Combinatorial Generalizations of Sieve Methods and Characterizing Hamiltonicity via Induced Subgraphs

by

Zishen Qu

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Mathematics
in
Combinatorics and Optimization

Waterloo, Ontario, Canada, 2022

Author's Declaration

This thesis consists of material all of which I authored or co-authored: see Statement of Contributions included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Statement of Contributions

Chapter 2 and 3 of this thesis are on work with Yu-Ru Liu. The results in Chapter 4 of this thesis are based on the paper "Minimal induced subgraphs of two classes of 2-connected non-Hamiltonian graphs." *Discrete Mathematics*, 345(7):112869, 2022, co-authored with Joseph Cheriyan, Sepehr Hajebi, and Sophie Spirkl. doi:10.1016/j.disc.2022.112869

Abstract

A sieve method is in effect an application of the inclusion-exclusion counting principle, and the estimation methods to avoid computing the explicit formula. Sieve methods have been used in number theory for over a hundred years. These methods have been modified to make use of the structure of integer-like objects; producing better estimates and providing more use cases. The first part of the thesis aims to analyze and use the analogues of number theoretic sieves in combinatorial contexts. This part consists of my work with Yu-Ru Liu in Chapters 2 and 3. We focus on two sieve methods: the Turán sieve (introduced by Liu and Murty in 2005) and the Selberg sieve (independently generalized by Wilson in 1969 and Chow in 1998 with slightly different formulations). Some comparisons and applications of these sieves are discussed. In particular, we apply the combinatorial Turán sieve to count labelled graphs and we apply the combinatorial Selberg sieve to count subspaces of finite spaces.

Finding sufficient conditions for Hamiltonicity in graphs is a classical topic, where the difficulty is bracketed by the NP-hardness of the associated decision problem. The second part of the thesis, consisting of Chapter 4, aims to characterize Hamiltonicity by means of induced subgraphs. The results in this chapter are based on the paper "Minimal induced subgraphs of two classes of 2-connected non-Hamiltonian graphs." *Discrete Mathematics*, 345(7):112869, 2022, co-authored with Joseph Cheriyan, Sepehr Hajebi, and Sophie Spirkl. We study induced subgraphs and conditions for Hamiltonicity. In particular, we characterize the minimal 2-connected non-Hamiltonian split graphs and the minimal 2-connected non-Hamiltonian triangle-free graphs.

Acknowledgements

I would like to thank my advisors Joseph Cheriyan and Yu-Ru Liu for the material direction and suggestions on the thesis. I would also like to thank my readers Kevin Hare and Sophie Spirkl for providing valuable revisions.

Thanks to Chris Godsil for helping me find some references for combinatorial identities. In addition to those named above, thanks to Aristotelis Chaniotis, Sepehr Hajebi, Patrick Hompe, Stephen Melczer, Benjamin Moore, Souradeep Purkayastha, Evelyne Smith-Roberge, Ronen Wdowinski, and Ali Zahabi for working with me during my Masters.

Dedication

Dedicated to Tessa

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

The following thesis comprises of two parts. The first part, consisting of my work with Yu-Ru Liu in Chapters 2 and 3, is to extend sieve methods from number theory to combinatorial contexts. The second part of the thesis consists of Chapter 4. The results in this chapter are based on the paper "Minimal induced subgraphs of two classes of 2-connected non-Hamiltonian graphs." *Discrete Mathematics*, 345(7):112869, 2022, co-authored with Joseph Cheriyan, Sepehr Hajebi, and Sophie Spirkl. Results are numbered in the order they appear in the non-introduction chapters. Summations are over a variable p or q denote the summation over the primes which satisfy the conditions given. Summations over other variables indicate a summation over the positive integers fulfilling the condition unless specified otherwise. Throughout the thesis log denotes the natural log unless specified otherwise.

For $f : \mathbb{N} \to \mathbb{R}$ and $g : \mathbb{N} \to \mathbb{R}^+$, we say that $f(n)$ is $O(g(n))$ if there exists some constant positive real number A such that

$$|f(n)| \leq Ag(n)$$

for all sufficiently large n. Writing

$$u(n) = h(n) + O(g(n))$$

means that there exists a function $f : \mathbb{N} \to \mathbb{R}$ such that $f(n)$ is $O(g(n))$ and

$$u(n) = h(n) + f(n).$$

We write $f(x) \ll g(x)$ or $g(x) \gg f(x)$ if $f(x)$ is $O(g(x))$. We say $f(n)$ is $o(g(n))$ if

$$\lim_{n\to\infty} \frac{f(n)}{g(n)} \to 0.$$

Writing

$$u(n) = h(n) + o(g(n))$$

means that there exists a function $f : \mathbb{N} \to \mathbb{R}$ such that $f(n)$ is $o(g(n))$ and

$$u(n) = h(n) + f(n).$$

We say that a certain property holds for *almost all* integers if for integers $n \leq N$, the number of exceptions is $o(N)$. We write $f(n) \sim g(n)$ if

$$\lim_{n\to\infty} \frac{f(n)}{g(n)} \to 1.$$

1.1 The Turán sieve

Here we give a general overview of the content of Chapter 2. Technical definitions are given in the chapter itself. Let $\omega(n)$ denote the number of distinct prime factors of n. Turán [29] showed the following result.

Theorem 8 (Turán [29]). *For positive real numbers N, we have*

$$\sum_{n\leq N} (\omega(n) - \log\log N)^2 \ll N \log\log N.$$

Turán uses the above result to produce the normal order of $\omega(n)$. One can view $\omega(n)$ as a random variable with $\log\log N$ as its expected value. Hence the sum in Theorem 8 can be viewed as a variance. This probabilistic approach leads the proof of Theorem 8 to a sieve formulation, of which a combinatorial version was developed by Liu and Murty [23], stated as follows.

Let (S, T) be a bipartite graph. For any $s \in S$ and any $t \in T$, we write $s \sim t$ if there is an edge between s and t. For any $s \in S$, let the degree of s be denoted by $\omega(s)$. Denote the degree of $t \in T$ with $\deg(t)$. Let $n(t_1, t_2)$ denote the number of common neighbours of $t_1, t_2 \in T$. Write $X = |S|$. Liu and Murty [23] show the following.

Corollary 12 (The Turán sieve, Liu and Murty [22]). *We have*

$$\#\{s \in S : \omega(s) = 0\} \leq X^2 \cdot \frac{\sum_{t_1,t_2\in T} n(t_1, t_2)}{\left(\sum_{t\in T} \deg(t)\right)^2} - X.$$

where the sum over t_1, t_2 is of ordered pairs, including pairs where $t_1 = t_2$.

In the same paper, Liu and Murty [23] provide an upper bound for the same context.

Theorem 13 (The simple sieve, Liu and Murty [22]). *We have*

$$\#\{s \in S : \omega(s) = 0\} \geq X - \sum_{t \in T} \deg(t).$$

This sieve counts the number of isolated vertices in one part of a bipartite graph. Note that if S is the set of positive integers up to a real number x and T is the set of prime numbers up to a real number z, then $\omega(s)$ is the number of distinct prime factors of s which are less than or equal to z. This allows us to use the Turán sieve to count the primes. In Section 2.1, we describe Turán's second moment method for the analysis of $\omega(n)$. In Section 2.2, we provide an exposition of the combinatorial Turán sieve developed by Liu and Murty. In Section 2.3, we will apply the Turán and simple sieves to count labelled graphs.

1.2 The Selberg sieve

We give a general overview of the content in Chapter 3, technical definitions are given in the chapter itself. For real numbers x and z with $x \geq z$, let $\Phi(x, z)$ be the number of integers less than or equal to x which are not divisible by primes less than z. We denote by (a, b) the greatest common divisor of a and b, and $[a, b]$ the least common multiple of a and b. It can be proved that

$$\Phi(x, z) = \sum_{n \leq x} \sum_{d | (n, P(z))} \mu(d),$$

where μ is the Möbius function and $P(z)$ is the product of all primes less than z. Selberg [27] noted that for $\lambda_1 = 1$ and λ_d arbitrary real numbers,

$$\sum_{d | m} \mu(d) \leq \left(\sum_{d | m} \lambda_d \right)^2,$$

for any positive integer m. Choosing appropriate λ_d and conducting a careful analysis leads to a good estimate of $\Phi(x, z)$. In particular, let ϕ be the *Euler phi function*, defined as

$$\phi(n) = \#\{k \in \mathbb{N} : 1 \leq k \leq n \text{ and } (k, n) = 1\}.$$

We have

$$\Phi(x,z) \leq \frac{x}{V(z)} + O\left(\sum_{d_1,d_2|P(z)} |\lambda_{d_1}||\lambda_{d_2}|\right),$$

where

$$V(z) = \sum_{d\leq z} \frac{\mu^2(d)}{\phi(d)}$$

and

$$\lambda_d = d\sum_{\substack{\delta\leq z\\ d|\delta}} \frac{\mu(\delta/d)\mu(\delta)}{\phi(\delta)V(z)}.$$

Let $\pi(x)$ denote the number of primes less than or equal to some real number x. One can make use of the above estimate to show the following.

Theorem 14 (Chebycheff 1850)**.** *We have that*

$$\pi(x) \ll \frac{x}{\log x}.$$

The derivation of this estimate as outlined above was conducted by Selberg [27] to demonstrate its potential as an alternative to the Brun sieve.

The derivation of the sieve can be generalized to work on a poset structure analogous to that of the divisibility poset. We will use $\preceq$ to denote the ordering relation of the poset. A few properties not found generally in posets are required for the extension. In particular, the existence of the meet (gcd) and join (lcm) is required. We also need the number of elements below a fixed element to be finite for the purposes of well-defined summation. Such a poset is called a locally finite lattice. Finally one needs an analogous Möbius function on a locally finite lattice, and such an extension was already found by Rota [25]. The *Möbius function* $\mu : L \times L \to \mathbb{Z}$ is defined as the function with the following three properties:

1. For $d \in L$, we have $\mu(d,d) = 1$.
2. Let $d, n \in L$. If $d \prec n$, then
$$\sum_{d\preceq e\preceq n} \mu(d,e) = 0.$$
3. If $d \not\preceq e$, then $\mu(d,e) = 0$.

Rota [25] showed that the above definition has many of the properties that we have for the Möbius function on the natural numbers. The generalization of the Selberg sieve was done by Wilson [32] (independently by Chow [7]) to produce the following result.

Theorem 20 (Selberg sieve on lattices, Wilson [32]). *Let $(L, \preceq)$ be a locally finite lattice. Let $S \subseteq L$ be a set of x distinct elements with the property that for any $d \in L$, the number of elements $n \in S$ satisfying $d \preceq n$ is of the form*

$$\frac{x}{f(d)} + R(d)$$

where f is a multiplicative function and R is some suitable error function. Let T be a set of atoms in $(L, \preceq)$ and write w to be the join of T. Let $M(S,T)$ be the set of elements in S for which the meet with w is 0 (the bottom of the lattice). Let $N : L \to \mathbb{N}$ be defined such that for lattice elements $d, e \in L$ with $d \prec e$, we have $N(d) < N(e)$. For any real number z, assume that the set $\{n \in L : N(n) < z\}$ is finite. Let

$$g(n) = \sum_{d \preceq n} \mu(d,n) f(d).$$

Define

$$V(z) = \sum_{\substack{N(d) \leq z \\ g(d) \neq 0}} \frac{\mu^2(0,d)}{|g(d)|}$$

and

$$\lambda(0,d) = \frac{f(d)}{V(z)} \sum_{\substack{d \preceq \delta \\ g(\delta) \neq 0 \\ N(\delta) \leq z}} \frac{\mu(d,\delta)\mu(0,\delta)}{|g(\delta)|}.$$

We have

$$|M(S,T)| \leq \frac{x}{V(z)} + \sum_{d_1, d_2 \preceq w} |\lambda(0,d_1)||\lambda(0,d_2)||R(d_1 \vee d_2)|,$$

where $d_1 \vee d_2$ is the join of d_1 and d_2.

To apply the combinatorial Selberg sieve, we need to translate the problem into a poset setting. In Section 3.1, we demonstrate the main ideas of the Selberg sieve by estimating the primes. In Section 3.2, we provide an exposition of the Selberg sieve on lattices given independently by Wilson and Chow. In Section 3.3, we will use the Selberg sieve on lattices to count finite subspaces.

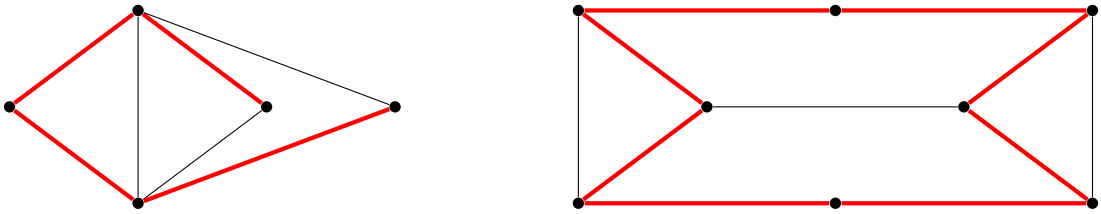


Figure 1.1: The graph on the left has a Hamiltonian path highlighted in red, and the graph on the right has a Hamiltonian cycle highlighted in red.

1.3 Hamiltonian graphs and induced subgraphs

Graphs in this chapter are finite and without loops or parallel edges. For a graph G and $X \subseteq V(G)$, $G[X]$ denotes the induced subgraph of G with vertex set X, and $G \setminus X$ denotes $G[V(G) \setminus X]$. A *Hamiltonian path* (resp. *Hamiltonian cycle*) in a graph G is a (not necessarily induced) subgraph H of G which is a path (resp. cycle), and $V(H) = V(G)$. A graph is *Hamiltonian* if it has a Hamiltonian cycle.

A *separating triangle* T in a graph G is a 3-cycle T where $G \setminus T$ is not connected. The approach of studying Hamiltonicity structurally was first exhibited by Whitney [31], who showed the following.

Theorem 1 (Whitney [31])**.** *Every planar triangulation with no separating triangle is Hamiltonian.*

Tutte [30] notes that Theorem 1 implies that all 4-connected planar triangulations are Hamiltonian. This result demonstrated the promise of finding sufficiency conditions that are based on structural properties. For us, a more specific class of structural conditions is of interest. The *claw* is the complete bipartite graph $K_{1,3}$. The *paw* is the graph composed of a triangle and one pendent leaf attached to the triangle. The *net* is the unique graph with degree sequence $(3, 3, 3, 1, 1, 1)$, and equivalently the graph with vertex set $\{a, b, c, a_1, b_1, c_1\}$ and edge set $\{ab, bc, ac, aa_1, bb_1, cc_1\}$. The *snare* is the graph obtained from a net by adding a vertex and making it adjacent to every vertex of the net. To demonstrate the type of desired result, consider the following theorem of Goodman and Hedetniemi.

Theorem 2 (Goodman and Hedetniemi [15])**.** *If G is 2-connected and contains no induced subgraph isomorphic to either claw or paw, then G is Hamiltonian.*

Proof. Suppose for a contradiction that G has no Hamiltonian cycle. Let C be the longest cycle in G and u be a vertex not in $V(C)$ which is adjacent to a vertex in C. Let v denote

the neighbour of u on C. Consider the vertices x, y which are adjacent to v on the cycle C. If $xu \in E(G)$, then we would have a longer cycle $(C \setminus xv) + xu + uv$. So $xu \notin E(G)$. Similarly $yu \notin E(G)$. Now consider the graph $G[\{u, v, x, y\}]$. If $xy \notin E(G)$, then this graph is an induced claw. If $xy \in E(G)$, then this graph is an induced paw. □

The following result was obtained by Duffus, Gould, and Jacobson.

Theorem 3 (Duffus, Gould, and Jacobson [13]; see also Shepherd [28])**.** *If G is connected and contains no induced subgraph isomorphic to the claw or the net, then G contains a Hamiltonian path.*

Duffus, Gould, and Jacobson also showed the following.

Theorem 4 (Duffus, Gould, and Jacobson [13]; see also Shepherd [28])**.** *If G is a 2-connected graph and contains no induced subgraph isomorphic to the claw or the net, then G is Hamiltonian.*

Shepherd [28] offers a simpler proof of the above two results by characterizing the structure of claw-free, net-free graphs. We say that $S \subseteq V(G)$ is a *separator* of G if $G \setminus S$ is not connected. We say that $S \subseteq V(G)$ is a *minimal separator* if no proper subset of S is a separator. Let G be a connected, claw-free, net-free graph. Shepherd shows the following. For any minimal separator S and any $v \in S$, the graph $G \setminus (S \setminus \{v\})$ has a cut vertex v, and $G \setminus S$ has two components. Moreover, the components are composed of vertex sets $S_1, \ldots, S_m$ and $T_1, \ldots, T_n$ respectively, where S_i is a clique of distance i from v in $G \setminus (S \setminus \{v\})$ and T_j is a clique of distance j from v in $G \setminus (S \setminus \{v\})$. There are no edges from S_i to T_j, and edges between different S_i only exist when the indices are consecutive. Similarly edges only exist between different T_j when the indices are consecutive. Moreover, if G is a graph with the above property for any minimal separator S and any vertex v in S, then G is connected, claw-free, and net-free. This characterization allows Shepherd to prove Theorem 3 and Theorem 4 by careful analysis of the minimal separators in the graph G.

We say that a graph H is an *HP-obstruction* if H is connected, has no Hamiltonian path, and every induced subgraph of H either equals H, or is not connected, or has a Hamiltonian path. The theorem of Duffus, Gould, and Jacobson can be viewed in this fashion as characterizing all HP-obstructions.

Theorem 5 (Duffus, Gould, and Jacobson [13]; see also Shepherd [28])**.** *There are exactly two HP-obstructions: the claw and the net.*

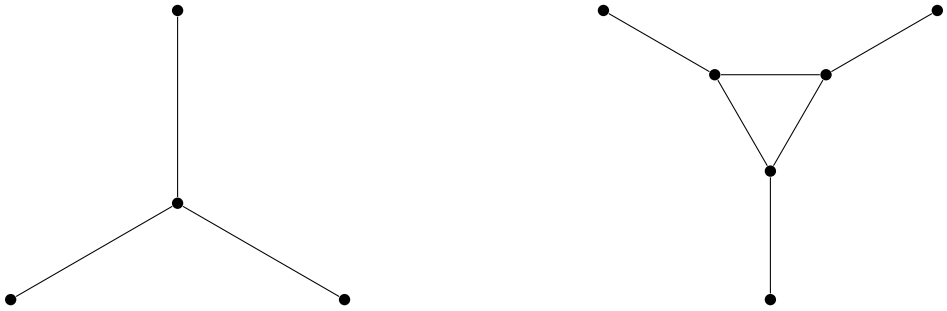


Figure 1.2: From left to right: a claw graph and a net graph.

A graph H is an *HC-obstruction* if H is 2-connected, has no Hamiltonian cycle, and every induced subgraph of H either equals H, or is not 2-connected, or has a Hamiltonian cycle. Following the same line of thought, we are interested in understanding HC-obstructions. In [3], Brousek gave a complete characterization of HC-obstructions that do not contain the claw as an induced subgraph. Let $\mathcal{P}$ be the class of graphs obtained by taking two disjoint triangles $\{u_1, u_2, u_3\}$ and $\{v_1, v_2, v_3\}$, joining u_i and v_i by a path with at least 3 vertices or a triangle for each i. Brousek showed the following.

Theorem 6 (Brousek [3])**.** *The graphs in $\mathcal{P}$ are HC-obstructions. Moreover, they are the only HC-obstructions which are claw-free.*

Chiba and Furuya [6] further studied induced subgraphs of non-minimal 2-connected non-Hamiltonian graphs. Let $N(v) = \{u \in V(G) : uv \in E(G)\}$. A vertex v is *locally connected* in G is $G[N(v)]$ is connected. We define the *closure* of a claw-free graph G as the graph obtained from G by recursively adding edges between vertices in a the neighbourhood of a locally connected vertex. This is well-defined as a result of Ryjáček [26]. We denote by $N_{3,1,1}$ the graph with the vertex set $\{a, b, c, a_1, b_1, c_1, a_2, a_3\}$ and edge set $\{ab, bc, ac, aa_1, bb_1, cc_1, a_1a_2, a_2a_3\}$. Chiba and Furuya showed the following.

Theorem 7 (Chiba and Furuya [6])**.** *Let G be a 2-connected $\{K_{1,3}, N_{3,1,1}\}$-free graph. Then G is Hamiltonian if and only if the closure of G is not one of the graphs depicted in Figure 1.3.*

Ding and Marshall [12] obtained a complete characterization in the case when "induced subgraph" is replaced by "induced minor" in the definition of an HC-obstruction.

Let us describe our main results. A *clique* in a graph G is a set K of pairwise adjacent vertices. A *stable set* in a graph G is a set S of pairwise non-adjacent vertices. A *split*

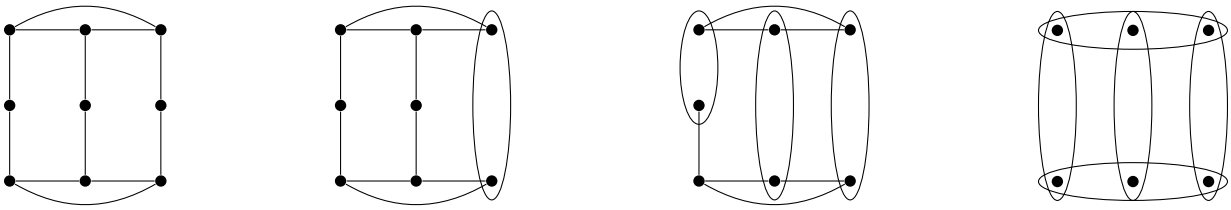


Figure 1.3: The four types of graphs which are the closures of the non-Hamiltonian 2-connected $\{K_{1,3}, N_{3,1,1}\}$-free graphs. The ellipses represent complete graphs of at least 3 vertices, including the explicitly drawn vertices.

Figure 1.4: From left to right: the snare, the 2-nova, a theta, and a triangle-free wheel. Squiggly edges represent paths of length at least one.

graph is a graph G with a partition (S, K) of $V(G)$ such that S is a stable set and K is a clique in G.

An *n-sun* is a graph obtained from a cycle C with $2n$ vertices $v_1, \ldots, v_{2n}$ that occur in this order along C by adding all edges $v_{2i}v_{2j}$ for distinct $i, j \in \{1, \ldots, n\}$. An *$n$-nova* is obtained from an n-sun by adding a vertex w and edges wv_{2i} for all $i \in \{1, \ldots, n\}$. Our first theorem, the following, gives a complete characterization of HC-obstructions that are split graphs.

Theorem 22 ([5])**.** *The snare and all n-novae for $n \geq 2$ are HC-obstructions. Moreover, these are the only HC-obstructions which are split graphs.*

A *theta* is a graph consisting of two non-adjacent vertices u and v and three paths P_1, P_2, P_3 from u to v and each of length at least two, such that the sets $V(P_1)\backslash\{u, v\}, V(P_2)\backslash\{u, v\}, V(P_3) \setminus \{u, v\}$ are disjoint and have no edges between them. The vertices u and v are the *ends* of the theta. A *closed theta* is a graph obtained from a theta with ends u, v by adding the edge uv.

A graph is *triangle-free* if it contains no three-vertex clique. A *wheel* is a pair (W, v) such that W is a cycle, and v is a vertex with at least three neighbours in W[1].

[1]In a standard definition of a wheel, the cycle W is required to be of length at least four. Note that this does not matter for our purposes as we are only concerned with triangle-free wheels.

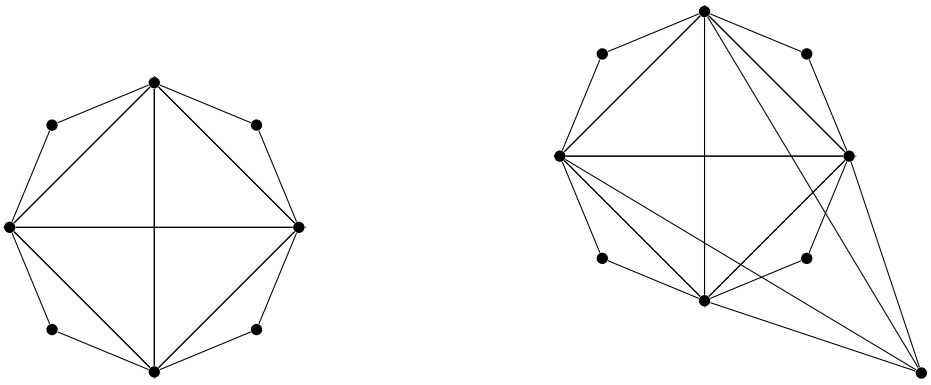


Figure 1.5: From left to right: a 4-sun and 4-nova.

Theorem 23 ([5])**.** *All thetas, triangle-free closed thetas, and triangle-free wheels are HC-obstructions, and they are the only HC-obstructions which are triangle-free.*

Chapter 2

The Turán sieve

In this Chapter we will first outline the proof of Turán, explain the extension of this method to sieve by Liu and Murty, and then apply the method to a labelled graph counting problem.

2.1 The normal order of the omega function

For a natural number n, let $\omega(n)$ denote the number of distinct prime factors of n. That is, for $n = p_1^{e_1} \cdots p_\ell^{e_\ell}$ with p_i distinct primes and e_i positive integers for all $1 \leq i \leq \ell$, we have $\omega(n) = \ell$. The asymptotic behaviour of ω is well-known. Indeed for any $\varepsilon > 0$, for almost all $n \in \mathbb{N}$, we have

$$(1-\varepsilon)\log\log n < \omega(n) < (1+\varepsilon)\log\log n.$$

This result was proved by Hardy and Ramanujan [19] in 1917.

In 1934, Turán [29] provided an alternative proof of the above result by viewing $\omega(n)$ as a random variable and computing a variance-like term. More precisely, Turán showed the following theorem.

Theorem 8 (Turán [29]). *For positive real numbers N, we have*

$$\sum_{n \leq N} (\omega(n) - \log\log N)^2 \ll N \log\log N.$$

We will demonstrate in Corollary 10 how Turán's theorem implies the result of Hardy and Ramanujan. To prove Theorem 8, we need an estimate on the reciprocals of primes.

Lemma 9 (Mertens, 1874)**.** *For positive real numbers N, we have*

$$\sum_{p\leq N}\frac{1}{p}=\log\log N+O(1).$$

A proof of Lemma 9 can be found in many standard analytic number theory books, for example, in *Introduction to Analytic Number Theory* by Apostol [1, Theorem 4.12]. We now prove Theorem 8 following Turán's original method.

Proof of Theorem 8. Consider the "variance" term

$$\sum_{n\leq N}(\omega(n)-\log\log N)^2=\sum_{n\leq N}\omega^2(n)-2\log\log N\sum_{n\leq N}\omega(n)+(\log\log N)^2\sum_{n\leq N}1. \tag{2.1}$$

For all $x\in\mathbb{R}$, we write $[x]$ to denote the largest integer which is at most x. The last summation on the right hand side of (2.1) becomes

$$\sum_{n\leq N}1=[N]=N+O(1).$$

By Lemma 9, the second summation in (2.1) becomes

$$\sum_{n\leq N}\omega(n)=\sum_{p\leq N}\left[\frac{N}{p}\right]=N\sum_{p\leq N}\frac{1}{p}+O(N)=N\log\log N+O(N),$$

where p is a prime. For the first summation in (2.1), write

$$\sum_{n\leq N}\omega^2(n)=\sum_{n\leq N}\sum_{p|n}\sum_{q|n}1=\sum_{p,q\leq N}\sum_{\substack{n\leq N\\ p|n,q|n}}1,$$

where p and q are primes. By considering the cases of $p=q$ and $p\neq q$, we can write

$$\sum_{n\leq N}\omega^2(n)=\sum_{\substack{p,q\leq N\\ p\neq q}}\left[\frac{N}{pq}\right]+\sum_{p\leq N}\left[\frac{N}{p}\right].$$

If $pq>N$, then $[N/pq]=0$. So we can write

$$\sum_{n\leq N}\omega^2(n)=\sum_{pq\leq N}\left[\frac{N}{pq}\right]-\sum_{p^2\leq N}\left[\frac{N}{p^2}\right]+\sum_{p\leq N}\left[\frac{N}{p}\right].$$

We remark that the first summation on the right includes cases of $p = q$. Since the series $\sum_p \frac{1}{p^2}$ converges, we have

$$\sum_{p^2 \le N} \left[\frac{N}{p^2}\right] = O(N).$$

Combining the above estimates with Lemma 9, we have

$$\sum_{n \le N} \omega^2(n) = \sum_{pq \le N} \left[\frac{N}{pq}\right] + O(N \log\log N). \tag{2.2}$$

To bound $\sum_{pq \le N} \left[\frac{N}{pq}\right]$, we note that

$$\left(\sum_{p \le \sqrt{N}} \frac{1}{p}\right)^2 \le \sum_{pq \le N} \frac{1}{pq} \le \left(\sum_{p \le N} \frac{1}{p}\right)^2. \tag{2.3}$$

By Lemma 9, we have

$$\sum_{p \le N} \frac{1}{p} = \log\log N + O(1).$$

Also we note that

$$\sum_{p \le \sqrt{N}} \frac{1}{p} = \log\log\sqrt{N} + O(1) = \log\left(\frac{1}{2}\log N\right) + O(1) = \log\log N + O(1).$$

By squaring the above two bounds, we obtain from (2.3) that

$$\sum_{pq \le N} \frac{1}{pq} = (\log\log N)^2 + O(\log\log N).$$

It follows from (2.2) that

$$\sum_{n \le N} \omega^2(n) = N(\log\log N)^2 + O(N \log\log N).$$

Combining the above estimates with (2.1), we have

$$\begin{aligned}
&\sum_{n \le N} (\omega(n) - \log\log N)^2 \\
=&\sum_{n \le N} \omega^2(n) - 2\log\log N \sum_{n \le N} \omega(n) + (\log\log N)^2 \sum_{n \le N} 1 \\
=&N(\log\log N)^2 - 2N(\log\log N)^2 + N(\log\log N)^2 + O(N \log\log N) \\
=&O(N \log\log N).
\end{aligned}$$

This completes the proof of Theorem 8. □

Corollary 10 (Hardy and Ramanujan [19]). *For any positive real numbers A and δ, the inequalities*

$$\log\log n - A(\log\log n)^{1/2+\delta} < \omega(n) < \log\log n + A(\log\log n)^{1/2+\delta}$$

hold for almost all $n \in \mathbb{N}$.

Proof. Fix a positive real number N. Consider positive integers n with $n \leq N$. Note that if

$$|\omega(n) - \log\log N| \geq (A/4)(\log\log N)^{\frac{1}{2}+\delta},$$

then

$$(\omega(n) - \log\log N)^2 \geq (A/4)^2(\log\log N)^{1+2\delta}.$$

Let

$$\epsilon(N) = \#\{n \leq N : |\omega(n) - \log\log N| \geq (A/4)(\log\log N)^{\frac{1}{2}+\delta}\}.$$

By Theorem 8, we have

$$\epsilon(N)(A/4)^2(\log\log N)^{1+2\delta} \leq \sum_{n\leq N}(\omega(n) - \log\log N)^2 \ll N\log\log N.$$

Therefore, we have

$$\epsilon(N) \ll \frac{N}{(A/4)^2(\log\log N)^{2\delta}},$$

which is $o(N)$. This proves that for almost all $n \leq N$, we have

$$\log\log N - (A/4)(\log\log N)^{1/2+\delta} < \omega(n) < \log\log N + (A/4)(\log\log N)^{1/2+\delta}.$$

Note that the number of positive integers $n \leq \sqrt{N}$ is $o(N)$. So it suffices to only consider positive integers n with $\sqrt{N} \leq n \leq N$. For $\sqrt{N} \leq n \leq N$ we have

$$\log\log n \leq \log\log N \leq \log\log n + \log 2.$$

By the above two inequalities, for almost all $n \leq N$, we have

$$\omega(n) > \log\log n - (A/4)(\log\log n + \log 2)^{1/2+\delta}, \tag{2.4}$$

and

$$\omega(n) < \log\log n + \log 2 + (A/4)(\log\log n + \log 2)^{1/2+\delta}. \tag{2.5}$$

Choose $N \in \mathbb{N}$ large enough such that for any $n \geq \sqrt{N}$ we have

$$2(\log\log n)^{\frac{1}{2}+\delta} > (\log\log n + \log 2)^{\frac{1}{2}+\delta}$$

and

$$(A/2)(\log\log n)^{\frac{1}{2}+\delta} > \log 2.$$

Using these inequalities with (2.4) and (2.5), for almost all $n \leq N$, we have

$$\log\log n - (A/2)(\log\log n)^{1/2+\delta} < \omega(n) < \log\log n + A(\log\log n)^{1/2+\delta}.$$

Hence the corollary follows. □

2.2 The Turán sieve

Turán's approach on the theorem of Hardy and Ramanujan concealed in it an elementary sieve method. We now follow in the approach of Liu and Murty [22] to formulate the sieve in a combinatorial setting.

Let (S, T) be a bipartite graph. For any $s \in S$ and any $t \in T$, we write $s \sim t$ if there is an edge between s and t. For any $s \in S$, let the *degree* of s, denoted by $\omega(s)$, be

$$\omega(s) = \#\{t \in T : s \sim t\},$$

Denote the degree of $t \in T$ with $\deg(t)$. So we have

$$\deg(t) = \#\{s \in S : s \sim t\}.$$

We are interested in counting the number of isolated vertices in S, i.e.,

$$\#\{s \in S : \omega(s) = 0\}.$$

Denote the size of S as X, so $|S| = X$. Note that we have

$$\sum_{s\in S} \omega(s) = \sum_{t\in T} \deg(t).$$

Hence the average degree of $s \in S$ is

$$\frac{1}{X}\sum_{t\in T} \deg(t).$$

Following the approach of Turán, we compute the "variance" as follows.

$$\begin{aligned}
&\sum_{s\in S}\left(\omega(s)-\frac{1}{X}\sum_{t\in T}\deg(t)\right)^2\\
=&\sum_{s\in S}\omega^2(s)-2\sum_{s\in S}\omega(s)\left(\frac{1}{X}\sum_{t\in T}\deg(t)\right)+\sum_{s\in S}\left(\frac{1}{X}\sum_{t\in T}\deg(t)\right)^2\\
=&\sum_{s\in S}\omega^2(s)-\frac{2}{X}\left(\sum_{t\in T}\deg(t)\right)^2+\frac{1}{X}\left(\sum_{t\in T}\deg(t)\right)^2\\
=&\sum_{s\in S}\omega^2(s)-\frac{1}{X}\left(\sum_{t\in T}\deg(t)\right)^2.
\end{aligned}$$

Let $n(t_1,t_2)$ denote the number of common neighbours of $t_1,t_2\in T$. Hence

$$n(t_1,t_2)=\#\{s\in S: s\sim t_1, s\sim t_2\}.$$

We have

$$\sum_{s\in S}\omega^2(s)=\sum_{s\in S}\sum_{\substack{t_1,t_2\in T\\ s\sim t_1\\ s\sim t_2}}1=\sum_{t_1,t_2\in T}\sum_{\substack{s\in S\\ s\sim t_1\\ s\sim t_2}}1=\sum_{t_1,t_2\in T}n(t_1,t_2).$$

Note that the summations are over the ordered pairs of t_1,t_2 including pairs where $t_1=t_2$.

Combining the above equations, we obtain the following result of Liu and Murty [22, Theorem 1].

Theorem 11 (Liu and Murty [22])**.** *We have*

$$\sum_{s\in S}\left(\omega(s)-\frac{1}{X}\sum_{t\in T}\deg(t)\right)^2=\sum_{t_1,t_2\in T}n(t_1,t_2)-\frac{1}{X}\left(\sum_{t\in T}\deg(t)\right)^2.$$

To estimate the number of isolated vertices in S, notice that we have

$$\#\{s\in S:\omega(s)=0\}\cdot\left(\frac{1}{X}\sum_{t\in T}\deg(t)\right)^2\le\sum_{s\in S}\left(\omega(s)-\frac{1}{X}\sum_{t\in T}\deg(t)\right)^2.$$

We combine the above inequality with Theorem 11 to obtain the following corollary of Liu and Murty [22, Corollary 1].

Corollary 12 (The Turán sieve, Liu and Murty [22]). *We have*

$$\#\{s \in S : \omega(s) = 0\} \leq X^2 \cdot \frac{\sum_{t_1,t_2 \in T} n(t_1, t_2)}{\left(\sum_{t \in T} \deg(t)\right)^2} - X.$$

where the sum over t_1, t_2 is of ordered pairs, including pairs where $t_1 = t_2$.

An easy lower bound is obtained by simply taking the size of S and subtracting the number of edges. This is stated in Liu and Murty [22, Proposition 1] as follows.

Theorem 13 (The simple sieve, Liu and Murty [22]). *We have*

$$\#\{s \in S : \omega(s) = 0\} \geq X - \sum_{t \in T} \deg(t).$$

Note that if S is the set of positive integers up to a real number x and T is the set of prime numbers up to a real number z, then $\omega(s)$ is the number of distinct prime factors of s which are less than or equal to z. This allows us to use the Turán sieve to count the primes.

2.3 Labelled graphs with no fixed size clique

An *undirected labelled graph* or a *graph* on n vertices is a pair $G = (V, E)$, where $V = \{1, 2, \ldots, n\}$ and E is a set of unordered pairs of V. We only consider the cases with no loops. In other words, E has no pairs of an element and itself. Consider all graphs on n vertices. Fix a positive integer r. A *r-clique* in a graph is a subgraph with r vertices $\{v_1, v_2, \ldots, v_r\}$ where every pair of vertices in $\{v_1, v_2, \ldots, v_r\}$ is an edge. We proceed to estimate the number of labelled graphs with no r-clique. The problem is uninteresting for $r = 1$ and $r = 2$, so we may assume $r \geq 3$.

Erdős, Kleitman, and Rothschild [14] showed that most triangle-free graphs on n vertices are bipartite and extended this idea to r-clique-free graphs. Using this, they determined the asymptotics of the number of n vertex graphs with no r-clique for the case of fixed r. They showed that if $G_r(n)$ is the number of n vertex graphs with no r-clique, then

$$\log_2(G_r(n)) = \frac{n^2}{2}\left(1 - \frac{1}{r-1}\right) + o(n^2).$$

Note that this implies that almost all n vertex graphs have an r-clique for fixed r.

Let the model $G(n,p)$ be graphs on n vertices where the edges are chosen independently and with probability p. Thus $G(n,1/2)$ are graphs on n vertices where any edge has $1/2$ chance of existing. To model $G(n,1/2)$, we note that every graph on n vertices has a $1/2^{\binom{n}{2}}$ chance of appearing, so we can simply multiply the number of graphs by $1/2^{\binom{n}{2}}$ to obtain the probability of an n vertex graph possessing the properties sieved for. In particular, Bollobás and Erdős [2] showed that for any $r < (2-\varepsilon)\log_2 n$, one can almost surely find an r-clique in a random graph on n vertices with probability $1/2$. It was shown in the same article that for any $r > (2+\varepsilon)\log_2 n$, there is almost surely no r-clique. However the method given relies on the Borel-Cantelli Lemma and thus provide no concrete bounds on the number (probability) of graphs on n vertices with no r-clique.

The following estimation uses techniques similar to those found in Kuo, Liu, Ribas, and Zhou [21]. We proceed to apply the Turán sieve to estimate the number of graphs on n vertices with no r-clique. Take S to be the set of all graphs on n vertices and T to be the set of all subsets of $\{1,2,\ldots,n\}$ with size r. For a graph $s \in S$ and a size r subset $t = \{v_1, v_2, \ldots, v_r\}$, we write $s \sim t$ if for all $1 \leq i < j \leq r$, the edge $v_i v_j$ is in the graph s. In other words, $s \sim t$ if and only if the r-clique formed by t is contained in the graph s. We have that $\omega(s) = 0$ if and only if s contains no r-clique.

We now apply the Turán sieve to estimate the number of r-clique free graphs. We have $\binom{n}{2}$ unordered pairs of $\{1,2,\ldots,n\}$, so

$$X = 2^{\binom{n}{2}}.$$

The size of T is the number of size r subsets in $\{1,2,\ldots,n\}$, so

$$|T| = \binom{n}{r}.$$

For a fixed $t \in T$, there are $\binom{r}{2}$ edges, so

$$\deg(t) = 2^{\binom{n}{2}-\binom{r}{2}}.$$

It follows that

$$\sum_{t \in T} \deg(t) = \binom{n}{r} 2^{\binom{n}{2}-\binom{r}{2}}.$$

We compute $n(t_1, t_2)$ with $t_1, t_2 \in T$. Suppose the subsets t_1 and t_2 have an intersection of size k. We have $\binom{r}{2}$ edges from each of t_1 and t_2 and $\binom{k}{2}$ edges in the intersection. We

take the convention that $\binom{k}{2} = 0$ when $k = 0$ or $k = 1$. This matches with the values obtained when counting edges. Therefore we have that

$$n(t_1, t_2) = 2^{\binom{n}{2} - 2\binom{r}{2} + \binom{k}{2}}.$$

There are $\binom{n}{r}\binom{r}{k}\binom{n-r}{r-k}$ choices for ordered pairs of (t_1, t_2) with size k intersection. Therefore

$$\sum_{t_1, t_2 \in T} n(t_1, t_2) = \sum_{k=0}^{r} \binom{n}{r}\binom{r}{k}\binom{n-r}{r-k} 2^{\binom{n}{2} - 2\binom{r}{2} + \binom{k}{2}}.$$

Combining the above equations with the Turán sieve stated in Corollary 12, we obtain

$$\begin{aligned}
\#\{s \in S : \omega(s) = 0\} &\leq X^2 \cdot \frac{\sum_{t_1, t_2 \in T} n(t_1, t_2)}{\left(\sum_{t \in T} \deg(t)\right)^2} - X \\
&= 2^{2\binom{n}{2}} \frac{\sum_{k=0}^{r} \binom{n}{r}\binom{r}{k}\binom{n-r}{r-k} 2^{\binom{n}{2} - 2\binom{r}{2} + \binom{k}{2}}}{\binom{n}{r}^2 2^{2\binom{n}{2} - 2\binom{r}{2}}} - 2^{\binom{n}{2}} \\
&= 2^{\binom{n}{2}} \left(\frac{\sum_{k=0}^{r} \binom{r}{k}\binom{n-r}{r-k} 2^{\binom{k}{2}}}{\binom{n}{r}} - 1 \right). \qquad (2.6)
\end{aligned}$$

By the simple sieve stated in Theorem 13, we have

$$\#\{s \in S : \omega(s) = 0\} \geq 2^{\binom{n}{2}} - \binom{n}{r} 2^{\binom{n}{2} - \binom{r}{2}} = 2^{\binom{n}{2}} \left(1 - \frac{\binom{n}{r}}{2^{\binom{r}{2}}}\right). \qquad (2.7)$$

We consider r as a function of n, say $r = r(n)$. Note that the simple sieve provides a nontrivial bound when

$$1 - \frac{\binom{n}{r(n)}}{2^{\binom{r(n)}{2}}} > 0.$$

In fact if we choose a function $r(n)$ such that

$$\lim_{n \to \infty} \frac{\binom{n}{r(n)}}{2^{\binom{r(n)}{2}}} = 0, \qquad (2.8)$$

then almost all n vertex graphs have no $r(n)$-clique.

We will use upper and lower bounds for the binomial coefficient $\binom{n}{k}$, which can be found in many texts, for example, in *The Art of Computer Programming, volume 1* by Knuth [20, Section 1.2.6, exercise 67]. Consider the bound

$$\binom{n}{r} \leq \left(\frac{en}{r}\right)^r,$$

where e is the base of the natural logarithm. It follows that

$$\frac{\binom{n}{r}}{2^{\binom{r}{2}}} \leq \frac{\left(\frac{en}{r}\right)^r}{2^{\frac{r(r-1)}{2}}} = \frac{1}{2^{\frac{r(r-1)}{2} - r\log_2 n + r\log_2\left(\frac{r}{e}\right)}} = \frac{1}{2^{r\left(\frac{r-1}{2} - \log_2 n + \log_2\left(\frac{r}{e}\right)\right)}}.$$

Choose $r(n) \geq 2\log_2 n + 1$, we obtain

$$\frac{\binom{n}{r}}{2^{\binom{r}{2}}} \leq \frac{1}{2^{r\left(\log_2 n - \log_2 n + \log_2\left(\frac{r}{e}\right)\right)}} = \frac{1}{2^{r\left(\log_2\left(\frac{r}{e}\right)\right)}}.$$

In this case, the limit condition of (2.8) holds. Therefore for $r(n) \geq 2\log_2 n + 1$, we conclude that almost all n-vertex graphs have no $r(n)$-clique. This is best possible when compared to the result stated earlier by Bollobás and Erdős [2].

A similar analysis can be conducted for the bound obtained by the Turán sieve in (2.6). We can see that if

$$\lim_{n\to\infty} \frac{\sum_{k=0}^{r} \binom{r}{k}\binom{n-r}{r-k}2^{\binom{k}{2}}}{\binom{n}{r}} - 1 = 0, \tag{2.9}$$

then almost all n vertex graphs have an $r(n)$-clique.

We proceed to compute $r = r(n)$ which fulfil the condition in (2.9). Since the simple sieve implies that almost all n vertex graphs have no $r(n)$-clique with $r(n) \geq 2\log_2 + 1$, we may assume that $0 \leq r(n) \leq 2\log_2 n + 1$. By Vandermonde's identity [16, Table 169], we have

$$\binom{n}{r} = \sum_{k=0}^{r} \binom{r}{k}\binom{n-r}{r-k}.$$

Hence we can write

$$\frac{\sum_{k=0}^{r} \binom{r}{k}\binom{n-r}{r-k}2^{\binom{k}{2}}}{\binom{n}{r}} - 1 = \frac{\sum_{k=0}^{r} \binom{r}{k}\binom{n-r}{r-k}2^{\binom{k}{2}}}{\sum_{k=0}^{r} \binom{r}{k}\binom{n-r}{r-k}} - 1 = \frac{\sum_{k=2}^{r} \binom{r}{k}\binom{n-r}{r-k}(2^{\binom{k}{2}} - 1)}{\sum_{k=0}^{r} \binom{r}{k}\binom{n-r}{r-k}}.$$

Therefore we can rewrite the limit in (2.9) as

$$\lim_{n\to\infty} \frac{\sum_{k=2}^{r} \binom{r}{k}\binom{n-r}{r-k}(2^{\binom{k}{2}} - 1)}{\binom{n}{r}} = 0.$$

Note that

$$\binom{n}{r} \geq \frac{n^r}{r^r}.$$

Also, we have that for $2 \leq k \leq r$,

$$\binom{r}{k} \leq 2^r \qquad \text{and} \qquad 2^{\binom{k}{2}} - 1 \leq 2^{\binom{r}{2}}.$$

For n sufficiently large, we have $r - 2 < n/2$. For $2 \leq k \leq r$, we have

$$\binom{n-r}{r-k} \leq \binom{n}{r-2}.$$

Combining the above inequalities, we obtain

$$\frac{\sum_{k=2}^{r} \binom{r}{k}\binom{n-r}{r-k}(2^{\binom{k}{2}} - 1)}{\binom{n}{r}} \leq \frac{(r-2)2^r\binom{n}{r-2}2^{\binom{r}{2}}}{r^{-r}n^r} \leq r^{r+1}2^{\binom{r}{2}+r}\binom{n}{r-2}n^{-r}.$$

We have

$$\binom{n}{r-2} \leq \left(\frac{en}{r-2}\right)^{r-2}.$$

We write

$$r^{r+1}2^{\binom{r}{2}+r}\binom{n}{r-2}n^{-r} \leq r^{r+1}2^{\binom{r}{2}+r}\left(\frac{en}{r-2}\right)^{r-2}n^{-r} = r^{r+1}2^{\binom{r}{2}+r}e^{r-2}n^{-2}(r-2)^{-(r-2)}.$$

Note that as $r \geq 4$, we have

$$r^{r+1} \leq (2r-4)^{r+1} = ((r-2)2)^{r+1},$$

so

$$\begin{aligned}
r^{r+1}2^{\binom{r}{2}+r}e^{r-2}n^{-2}(r-2)^{-(r-2)} &\leq (r-2)^{r+1}2^{\binom{r}{2}+2r+1}e^{r-2}n^{-2}(r-2)^{-(r-2)} \\
&= (r-2)^3 2^{\binom{r}{2}+2r+1}e^{r-2}n^{-2} \\
&= (r-2)^3 2^{\binom{r}{2}+2r+1}2^{(r-2)\log_2 e}n^{-2} \\
&\leq r^3 2^{r^2/2+3r/2+(r-2)\log_2 e+1}n^{-2} \\
&\leq r^3 2^{r^2/2+7r/2}n^{-2}
\end{aligned}$$

For $r \geq 7$ we have that

$$r^3 2^{r^2/2+7r/2} n^{-2} \leq r^3 2^{r^2} n^{-2}.$$

So for any fixed $\varepsilon > 0$, we can choose $r(n) = \sqrt{(2-\varepsilon)\log_2 n}$, and we obtain

$$r^3 2^{r^2} n^{-2} \leq (\sqrt{(2-\varepsilon)\log_2 n})^3 n^{2-\varepsilon} n^{-2} = \frac{(\sqrt{(2-\varepsilon)\log_2 n})^3}{n^{\varepsilon}}.$$

The last term clearly goes to 0 as n grows. Hence the limit (2.9) in approaches 0 in this case. Therefore almost all n vertex graphs have a clique of size $\sqrt{(2-\varepsilon)\log_2 n}$.

The Turán sieve is based on a probabilistic principle, and as a result many estimates in the asymptotic case can be obtained using other probabilistic methods. The advantage of the Turán sieve when compared to these probabilistic methods is that it provides concrete constructable bounds. When compared to the methods of Erdős, Kleitman, and Rothschild which require r to be fixed in relation to n, the Turán sieve allows for far more flexibility for varying r. Unfortunately we were unable to reach the best bounds asymptotically from below when compared to Bollobás and Erdős [2], being off by a square root. However, the method has the advantage of producing concrete bounds for given n and r.

Chapter 3

The Selberg sieve

To introduce the Selberg sieve, we give the following which are standard methods and applications; we present them as in the text *An Introduction to Sieve Methods and their Applications* by Cojocaru and Murty [9, Chapter 5, Chapter 7].

A fundamental function in number theory is the *prime counting function*. Let $\pi(x)$ denote the number of primes less than or equal to some real number x. In 1896, Hadamard [17] and de la Vallée Poussin [11] independently proved the prime number theorem, which states that

$$\pi(x) \sim \frac{x}{\log x}.$$

The prime numbers represent a prototypical object suitable for sieving. In this chapter, we will introduce the Selberg sieve by deriving an upper bound for $\pi(x)$. In particular, we will show the following.

Theorem 14 (Chebycheff 1850)**.** *We have that*

$$\pi(x) \ll \frac{x}{\log x}.$$

The method of conducting this estimate with the Selberg sieve was shown by Selberg [27] in 1947. Our exposition follows the textbook of Cojocaru and Murty in [9, Chapter 5, Chapter 7], which can also be found in Dalton [10].

3.1 Bounding the number of primes

Before we describe and apply the sieve method, we need to introduce some common notation. We denote the greatest common divisor of a and b as (a, b). Integers a, b with the property that $(a, b) = 1$ are called *coprime*. The least common multiple of two integers a and b is denoted by $[a, b]$. We often need to refer to the primes less than a number z, so we write $P(z)$ denote the product of primes less than z.

Let $\Phi(x, z)$ be the number of positive integers at most x which are not divisible by primes less than z. So we have

$$\Phi(x, z) = \#\{n \in \mathbb{N} : n \leq x \text{ and } (n, P(z)) = 1\}.$$

Note that $\Phi(x, z)$ is an upper bound for the number of primes between z and x.

A useful function when considering arithmetical properties of positive integers is the *Möbius function*, denoted $\mu : \mathbb{N} \to \{-1, 0, 1\}$. It is defined on the positive integers as follows. Let $n \in \mathbb{N}$. For $n = 1$, we take $\mu(1) = 1$. If $n > 1$ is *square-free*, that is $n = p_1 \cdots p_\ell$ for distinct primes $p_1, \ldots, p_\ell$, then $\mu(n) = (-1)^\ell$. Otherwise $\mu(n) = 0$.

Lemma 15. *For any $n \in \mathbb{N}$, we have*

$$\sum_{d|n} \mu(d) = \begin{cases} 1 & \text{if } n = 1, \\ 0 & \text{if } n > 1. \end{cases}$$

Proof. For $n = 1$ the equation holds by definition. Consider the case that $n > 1$ and write $n = p_1^{e_1} \cdots p_\ell^{e_\ell}$. Since no divisor with a squared prime factor contributes to the sum, we can replace n with a square-free $N = p_1 \cdots p_\ell$. Write

$$\sum_{d|n} \mu(d) = \sum_{d|N} \mu(d) = \sum_{d|p_1 \cdots p_\ell} \mu(d) = \sum_{i=0}^{\ell} \binom{\ell}{i} (-1)^i = (1-1)^\ell = 0.$$

The second to last equality is obtained by the binomial theorem. □

One item to note is that the Möbius function is *multiplicative*, that is, for positive integers n, m which are coprime, we have $\mu(n)\mu(m) = \mu(nm)$.

By Lemma 15, we observe that

$$\Phi(x, z) = \sum_{n \leq x} \sum_{d|(n, P(z))} \mu(d). \tag{3.1}$$

The equality follows since the inner sum only contributes when $(n, P(z)) = 1$, which is equivalent to n not having prime divisors less than z.

An important observation made by Atle Selberg [27] is that for $\lambda_1 = 1$ and λ_d arbitrary real numbers, we have

$$\sum_{d|m} \mu(d) \leq \left(\sum_{d|m} \lambda_d \right)^2$$

for any $m \in \mathbb{N}$. For convenience, we take $\lambda_d = 0$ for all $d > z$. Using Selberg's observation, we can choose λ_d to minimize the bound on $\Phi(x, z)$ in (3.1). This yields

$$\begin{aligned}
\Phi(x, z) &= \sum_{n \leq x} \sum_{d|(n, P(z))} \mu(d) \\
&\leq \sum_{n \leq x} \left(\sum_{d|(n, P(z))} \lambda_d \right)^2 \\
&= \sum_{n \leq x} \left(\sum_{d_1, d_2|(n, P(z))} \lambda_{d_1} \lambda_{d_2} \right) \\
&= \sum_{d_1, d_2|P(z)} \lambda_{d_1} \lambda_{d_2} \sum_{\substack{n \leq x \\ [d_1, d_2]|n}} 1.
\end{aligned}$$

This produces the estimate

$$\Phi(x, z) \leq x \sum_{d_1, d_2 \leq z} \frac{\lambda_{d_1} \lambda_{d_2}}{[d_1, d_2]} + O\left(\sum_{d_1, d_2|P(z)} |\lambda_{d_1}||\lambda_{d_2}| \right). \tag{3.2}$$

We will focus on minimizing the main term in the bound. Note that $[d_1, d_2](d_1, d_2) = d_1 d_2$. Hence we can write

$$\sum_{d_1, d_2 \leq z} \frac{\lambda_{d_1} \lambda_{d_2}}{[d_1, d_2]} = \sum_{d_1, d_2 \leq z} \frac{\lambda_{d_1} \lambda_{d_2}}{d_1 d_2} (d_1, d_2). \tag{3.3}$$

We want to optimize this sum by choosing the sequence λ_d, but the gcd would be difficult to deal with in the summation. Hence we look for a technique to eliminate this factor. To

do this, we introduce a new arithmetic function with the desired summation properties. For a positive integer n, we recall the *Euler phi function*, $\phi(n)$, is defined as

$$\phi(n) = \#\{k \in \mathbb{N} : 1 \leq k \leq n \text{ and } (k, n) = 1\}.$$

One can show that the Euler phi function is multiplicative, that is for $m, n \in \mathbb{N}$ with $(m, n) = 1$, we have that $\phi(mn) = \phi(m)\phi(n)$.

The desired summation property is the following.

Lemma 16. *For $n \in \mathbb{N}$, we have*

$$\sum_{\delta | n} \phi(\delta) = n.$$

Proof. Consider the list of fractions $1/n, 2/n, \ldots, n/n$ and write them in the lowest terms. Hence the numerator and denominator are coprime. The denominators contain all positive divisors of n. Note that for any divisor $d|n$, there are $\phi(d)$ terms in the list of fractions. This shows a bijection from $\{1, \ldots, n\}$ to $\{\phi(d) : d|n\}$. Therefore

$$\sum_{\delta | n} \phi(\delta) = n.$$

□

Note that using Lemma 16 and the property that ϕ is multiplicative, we can obtain the following formula for ϕ. For $n = p_1^{k_1} p_2^{k_2} \cdots p_\ell^{k_\ell}$ with $k_i \geq 1$ and p_i distinct primes, we have

$$\phi(n) = n \prod_{i=1}^{\ell} \left(1 - \frac{1}{p_i}\right). \tag{3.4}$$

We rewrite (3.3) as

$$\begin{aligned}
\sum_{d_1, d_2 \leq z} \frac{\lambda_{d_1} \lambda_{d_2}}{d_1 d_2} (d_1, d_2) &= \sum_{d_1, d_2 \leq z} \frac{\lambda_{d_1} \lambda_{d_2}}{d_1 d_2} \sum_{\delta | (d_1, d_2)} \phi(\delta) \\
&= \sum_{\delta \leq z} \phi(\delta) \sum_{\substack{d_1, d_2 \leq z \\ \delta | (d_1, d_2)}} \frac{\lambda_{d_1} \lambda_{d_2}}{d_1 d_2} \\
&= \sum_{\delta \leq z} \phi(\delta) \left(\sum_{\substack{d \leq z \\ \delta | d}} \frac{\lambda_d}{d} \right)^2.
\end{aligned}$$

Let

$$u_\delta = \sum_{\substack{d \leq z \\ \delta | d}} \frac{\lambda_d}{d}.$$

We look to optimize

$$\sum_{d_1, d_2 \leq z} \frac{\lambda_{d_1} \lambda_{d_2}}{d_1 d_2} (d_1, d_2) = \sum_{\delta \leq z} \phi(\delta) u_\delta^2. \tag{3.5}$$

The sequence u_δ is not unconstrained. It is subject to the original restrictions on λ_d, that is, the constraints $\lambda_1 = 1$ and $\lambda_d = 0$ for all $d > z$. Note that this implies $u_\delta = 0$ for any $\delta > z$.

To convert the other restriction we need to rewrite the relation between u_δ and λ_d. The following result allows us to do this. We say that $D \subseteq \mathbb{Z}$ is a *divisor closed set* if for every $n \in D$, we have that $d|n$ implies that $d \in D$.

Lemma 17. *(Dual Möbius inversion formula) Let $\mathcal{D}$ be a divisor closed set of natural numbers. Let f, g be complex valued functions on the natural numbers. We have*

$$f(n) = \sum_{\substack{n | d \\ d \in \mathcal{D}}} g(d)$$

if and only if

$$g(n) = \sum_{\substack{n | d \\ d \in \mathcal{D}}} \mu\left(\frac{d}{n}\right) f(d).$$

Given the additional assumption that the series are absolutely convergent.

Proof. Consider the forward direction first. We have

$$\begin{aligned}
\sum_{\substack{n | d \\ d \in \mathcal{D}}} \mu\left(\frac{d}{n}\right) f(d) &= \sum_{\substack{n | d \\ d \in \mathcal{D}}} \mu\left(\frac{d}{n}\right) \sum_{\substack{d | e \\ e \in \mathcal{D}}} g(e) \\
&= \sum_{\substack{s = d/n \\ n | d \\ d \in \mathcal{D}}} \sum_{\substack{d | e \\ e \in \mathcal{D}}} g(e) \mu(s) \\
&= \sum_{\substack{n | e \\ e \in \mathcal{D}}} g(e) \sum_{s | \frac{e}{n}} \mu(s) \\
&= g(n).
\end{aligned}$$

The last equality is obtained by Lemma 15.

For the backwards direction, we have

$$\begin{aligned}\sum_{\substack{n|d\\ d\in\mathcal{D}}} g(d) &= \sum_{\substack{n|d\\ d\in\mathcal{D}}}\sum_{\substack{d|e\\ e\in\mathcal{D}}} \mu\left(\frac{e}{d}\right) f(e)\\ &= \sum_{\substack{n|e\\ e\in\mathcal{D}}} f(e) \sum_{\substack{n|d\\ d|e}} \mu\left(\frac{e}{d}\right)\\ &= \sum_{\substack{n|e\\ e\in\mathcal{D}}} f(e) \sum_{\substack{\frac{e}{d}|\frac{e}{n}\\ d|e}} \mu\left(\frac{e}{d}\right)\\ &= \sum_{\substack{n|e\\ e\in\mathcal{D}}} f(e) \sum_{s|\frac{e}{n}} \mu\left(s\right)\\ &= f(n).\end{aligned}$$

The last equality is obtained by Lemma 15. □

Applying the Dual Möbius inversion formula for u_δ produces

$$\frac{\lambda_\delta}{\delta} = \sum_{\delta|d} \mu\left(\frac{d}{\delta}\right) u_d. \tag{3.6}$$

Thus we have the restriction that

$$\sum_{\delta\le z} \mu(\delta) u_\delta = 1. \tag{3.7}$$

To optimize the quadratic form in (3.5), let $V(z)$ be a function to be defined later. We write

$$\begin{aligned}\sum_{\delta\le z} \phi(\delta) u_\delta^2 &= \sum_{\delta\le z} \phi(\delta)\left(u_\delta - \frac{\mu(\delta)}{\phi(\delta)V(z)}\right)^2 + \frac{2}{V(z)}\sum_{\delta\le z}\mu(\delta)u_\delta - \frac{1}{V^2(z)}\sum_{\delta\le z}\frac{\mu^2(\delta)}{\phi(\delta)}\\ &= \sum_{\delta\le z} \phi(\delta)\left(u_\delta - \frac{\mu(\delta)}{\phi(\delta)V(z)}\right)^2 + \frac{2}{V(z)} - \frac{1}{V^2(z)}\sum_{\delta\le z}\frac{\mu^2(\delta)}{\phi(\delta)}.\end{aligned}$$

By choosing

$$V(z) = \sum_{d\le z} \frac{\mu^2(d)}{\phi(d)},$$

we obtain

$$\sum_{\delta\le z}\phi(\delta)u_\delta^2=\sum_{\delta\le z}\phi(\delta)\left(u_\delta-\frac{\mu(\delta)}{\phi(\delta)V(z)}\right)^2+\frac{1}{V(z)}. \tag{3.8}$$

Since the functions are real, the above form has a minimum when

$$u_\delta=\frac{\mu(\delta)}{\phi(\delta)V(z)}. \tag{3.9}$$

By (3.6) and (3.9), we obtain

$$\lambda_\delta=\delta\sum_{\delta|d}\mu\left(\frac{d}{\delta}\right)u_d=\delta\sum_{\substack{d\le z\\ \delta|d}}\frac{\mu(d/\delta)\mu(d)}{\phi(d)V(z)}.$$

Combining (3.2), (3.3), (3.5), and (3.8), we obtain the estimate

$$\Phi(x,z)\le\frac{x}{V(z)}+O\left(\sum_{d_1,d_2|P(z)}|\lambda_{d_1}||\lambda_{d_2}|\right). \tag{3.10}$$

It remains to bound the error term.

Our goal is to show that $|\lambda_\delta|\le 1$. This is true by definition for $\delta=1$. Note that

$$\begin{aligned}
V(z)\lambda_\delta&=\delta\sum_{\substack{d\le z\\ \delta|d}}\frac{\mu(d/\delta)\mu(d)}{\phi(d)}\\
&=\delta\sum_{t\le\frac{z}{\delta}}\frac{\mu(t)\mu(\delta t)}{\phi(\delta t)}\\
&=\delta\sum_{\substack{t\le\frac{z}{\delta}\\ (t,\delta)=1}}\frac{\mu^2(t)\mu(\delta)}{\phi(\delta)\phi(t)}\\
&=\mu(\delta)\prod_{p|\delta}\left(1+\frac{1}{p-1}\right)\sum_{\substack{t\le\frac{z}{\delta}\\ (t,\delta)=1}}\frac{\mu^2(t)}{\phi(t)}.
\end{aligned}$$

For δ not square-free, we have $\mu(\delta)=0$ and hence $V(z)\lambda_\delta=0$. Let $\delta=p_1\cdots p_n$ be

square-free. We have

$$\begin{aligned}
&\prod_{p|\delta}\left(1+\frac{1}{p-1}\right)\\
=&\prod_{i=1}^{n}\left(1+\frac{1}{p_i-1}\right)\\
=&1+\sum_{i=1}^{n}\frac{1}{p_i-1}+\sum_{1\le i<j\le n}\frac{1}{(p_i-1)(p_j-1)}+\sum_{1\le i<j<k\le n}\frac{1}{(p_i-1)(p_j-1)(p_k-1)}+\cdots\\
=&1+\sum_{i=1}^{n}\frac{1}{\phi(p_i)}\sum_{1\le i<j\le n}\frac{1}{\phi(p_ip_j)}+\sum_{1\le i<j<k\le n}\frac{1}{\phi(p_ip_jp_k)}+\cdots\\
=&\sum_{s|\delta}\frac{1}{\phi(s)}.
\end{aligned}$$

For $s|\delta$, we have

$$\left|\frac{1}{\phi(s)}\sum_{\substack{t\le\frac{z}{\delta}\\(t,\delta)=1}}\frac{\mu^2(t)}{\phi(t)}\right|=\left|\sum_{\substack{st\le\frac{zs}{\delta}\\(t,\delta)=1}}\frac{\mu^2(st)}{\phi(st)}\right|.$$

Combining the above, we obtain

$$\prod_{p|\delta}\left(1+\frac{1}{p-1}\right)\sum_{\substack{t\le\frac{z}{\delta}\\(t,\delta)=1}}\frac{\mu^2(t)}{\phi(t)}=\left|\left(\sum_{s|\delta}\frac{1}{\phi(s)}\right)\sum_{\substack{t\le\frac{z}{\delta}\\(t,\delta)=1}}\frac{\mu^2(t)}{\phi(t)}\right|=\left|\sum_{s|\delta}\sum_{\substack{st\le\frac{zs}{\delta}\\(t,\delta)=1}}\frac{\mu^2(st)}{\phi(st)}\right|.$$

Note that for distinct s_1, s_2 which are divisors of δ and for any $t_1, t_2 \le z/\delta$ with t_1, t_2 both coprime to δ, we have that $s_1t_1 \ne s_2t_2$. Thus we have

$$\left|\sum_{s|\delta}\sum_{\substack{st\le\frac{zs}{\delta}\\(t,\delta)=1}}\frac{\mu^2(st)}{\phi(st)}\right|\le\left|\sum_{d\le z}\frac{\mu^2(d)}{\phi(d)}\right|,$$

as the product st in the left sum is never in the sum more than once. Combining the above,

we obtain

$$\left|\mu(\delta)\prod_{p|\delta}\left(1+\frac{1}{p-1}\right)\sum_{\substack{t\leq\frac{z}{\delta}\\(t,\delta)=1}}\frac{\mu^2(t)}{\phi(t)}\right|\leq\left|\sum_{d\leq z}\frac{\mu^2(d)}{\phi(d)}\right|.$$

Thus we have $|V(z)||\lambda_\delta|\leq|V(z)|$. We conclude $|\lambda_\delta|\leq 1$ for any δ.

Combining the fact that $|\lambda_\delta|\leq 1$ with (3.10), we have

$$\Phi(x,z)\leq\frac{x}{V(z)}+O(z^2). \tag{3.11}$$

We now find an appropriate lower bound for $V(z)$. Note

$$V(z)=\sum_{d\leq z}\frac{\mu^2(d)}{\phi(d)}\geq\sum_{d\leq z}\frac{\mu^2(d)}{d}\geq\sum_{d\leq z}\frac{1}{d}-\sum_{\substack{d\leq z\\ d \text{ not square-free}}}\frac{1}{d}.$$

We can bound the summation of d which contains a square by

$$\sum_{\substack{d\leq z\\ d \text{ not square-free}}}\frac{1}{d}\leq\frac{1}{4}\sum_{d\leq\frac{z}{4}}\frac{1}{d}\leq\frac{1}{4}\sum_{d\leq z}\frac{1}{d}.$$

Since

$$\sum_{d\leq z}\frac{1}{d}\gg\log z,$$

we have

$$V(z)\geq\sum_{d\leq z}\frac{1}{d}-\sum_{\substack{d\leq z\\ d \text{ not square-free}}}\frac{1}{d}\geq\sum_{d\leq z}\frac{1}{d}-\frac{1}{4}\sum_{d\leq z}\frac{1}{d}\gg\log z.$$

Therefore using (3.11) we obtain

$$\Phi(x,z)\ll\frac{x}{\log z}+z^2.$$

Recall that $\Phi(x,z)$ is an upper bound for the number of primes between z and x, so we have that

$$\pi(x)\leq\Phi(x,z)+z.$$

It follows that

$$\pi(x) \ll \frac{x}{\log z} + z^2.$$

By choosing

$$z = \left(\frac{x}{\log x}\right)^{1/2},$$

we obtain the desired estimate of

$$\pi(x) \ll \frac{x}{\log x}.$$

3.2 The Selberg sieve in a lattice

In this section, we will extend the Selberg sieve in the previous section to a lattice setting. This work was conducted independently by Wilson [32] and Chow [7].

A *partially ordered set* or *poset* is a set P with a binary relation $\preceq$ with the following three properties:

1. For every $a \in P$, we have $a \preceq a$.
2. For every $a, b \in P$, we have that if $a \preceq b$ and $b \preceq a$, then $a = b$.
3. For every $a, b, c \in P$, if $a \preceq b$ and $b \preceq c$, then $a \preceq c$.

A lattice is a poset where meets and joins exist, as defined below. We write $(L, \preceq)$ to denote the set of elements and the relation of the lattice respectively. For $d, e \in L$, we write $d \prec e$ to denote $d \preceq e$ and $d \neq e$. We write $d \not\preceq e$ for the negation of $d \preceq e$. If we have both $d \not\preceq e$ and $e \not\preceq d$, then we say that e and d are *not comparable*. The *bottom* of the lattice $(L, \preceq)$, denoted 0, is the unique element for which $0 \preceq \ell$ for all $\ell \in L$. An element of the lattices a is called an *atom* if $0 \prec a$ and there is no element n in the lattice such that $0 \prec x \prec n$. Let n, m be elements of the lattice $(L, \preceq)$. The *meet* of n and m is the unique element $n \wedge m \in L$ such that $n \wedge m \preceq n$ and $n \wedge m \preceq m$, and for any $\ell \in L$ with $\ell \preceq n$ and $\ell \preceq m$, we have $\ell \preceq n \wedge m$. The *join* of n and m is the unique element $n \vee m \in L$ such that $n \preceq n \vee m$ and $m \preceq n \vee m$, and for any $\ell \in L$ with $n \preceq \ell$ and $m \preceq \ell$, we have $n \vee m \preceq \ell$. Let A be a subset of L. We write $\wedge A$ and $\vee A$ to be the meet and join of A respectively. We call a partially ordered set a *lattice* if the meet and join of any two elements exists. We call a lattice L *locally finite* if for any element e, the number of elements $d \in L$ such that $d \prec e$ is finite. Finally, a function $f : L \to \mathbb{R}$ is *multiplicative* on L if $f(n) \geq 1$ for all

$n \in L$ with equality holding only when $n = 0$ and $f(m \vee n)f(m \wedge n) = f(m)f(n)$ for all $m, n \in L$.

A useful example to keep in mind is the positive integers under the partial order of divisibility. More precisely, we consider the lattice $(\mathbb{N}, |)$. That is, the lattice elements defined as the positive integers with divisibility defining the partial order. For a prime p, there are no integers n with $n > 1$ and $n \neq p$ such that $1|n$ and $n|p$. So every prime is an atom in the division lattice of the integers. Note that $1|n$ for any positive integer n. Thus the integer 1 plays the role of the bottom element $0 \in L$ in the lattice of the positive integers. In this context the meet is the greatest common divisor, and the join is the least common multiple. In other words, for positive integers a and b, we have $a \wedge b = (a, b)$, the gcd of a and b, and we have $a \vee b = [a, b]$, the lcm of a and b.

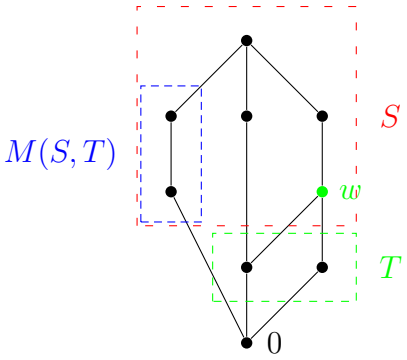


Figure 3.1: The Hasse diagram of an instance of a lattice with S, T selected, along with associated w in the lattice. The set $M(S, T)$ is labelled as well.

Fix S to be a subset of the lattice $(L, \preceq)$. Let T be a set of atoms and write w as the join of T. Let

$$M(S, T) = \{s \in S : \forall t \in T, t \not\preceq s\}.$$

This is equivalent to counting the elements in S for which the meet with w is 0 (the bottom element of L). Note that if we define $\omega(s)$ as the number of $t \in T$ such that $s \preceq t$, then we are counting the number of $s \in S$ such that $\omega(s) = 0$. Figure 3.1 provides an example of a lattice with the associated sets for this setting.

Consider the lattice formulation in the context of the natural number, where we count the number of primes in the divisor lattice. We have $S = \{1, \ldots, [x]\}$ and T is the set of the primes less than or equal to z. The integer w is the product of all of the primes in T and $M(S, T)$ is the set of integers coprime to w. As a numerical example, consider sieving

the set of primes $\{2, 3\}$ in the positive integers up to 24. We have

$$S = \{1, \ldots, 24\}, \quad T = \{2, 3\}, \quad w = 6.$$

We aim to count the number of integers in S for which the gcd (meet) with 6 is 1. This would be the set

$$M(S, T) = \{1, 5, 7, 11, 13, 17, 19, 23\}.$$

The *Möbius function* $\mu : L \times L \to \mathbb{Z}$ is defined as the function with the following three properties:

1. For $d \in L$, we have $\mu(d, d) = 1$.
2. Let $d, n \in L$. If $d \prec n$, then
$$\sum_{d \preceq e \preceq n} \mu(d, e) = 0.$$
3. If $d \not\preceq e$, then $\mu(d, e) = 0$.

In the context of the divisor lattice $(\mathbb{N}, |)$, consider $\mu(1, n)$ for $n \in \mathbb{N}$. We have that $\mu(1, 1) = 1$. By Property 2, we have

$$\mu(1, 1) + \mu(1, p) = 0$$

for any prime p, so $\mu(1, p) = -1$. We claim that $\mu(1, p_1 \cdots p_k) = (-1)^k$, where $p_1, \ldots, p_k$ are distinct prime numbers. We will show this by strong induction, so suppose that the formula is true for products of less than k distinct primes. Note by Property 2 that

$$\begin{aligned} 0 &= \sum_{1|e|p_1 \cdots p_k} \mu(1, e) \\ &= \sum_{1|e|p_1 \cdots p_{k-1}} \mu(1, e) + \sum_{p_k|e|p_1 \ldots p_k} \mu(1, e) \\ &= \sum_{1|e|p_1 \ldots p_{k-1}} \mu(1, ep_k), \end{aligned}$$

where the last equality holds as the first sum is 0 by Property 2 of the Möbius function. We separate the term where $e = p_1 \cdots p_{k-1}$ in this sum and apply induction to the remaining

terms to obtain

$$\begin{aligned}
0 &= \mu(1, p_1 \cdots p_k) + \sum_{i=0}^{k-2} \binom{k-1}{i} (-1)^{i+1} \\
&= \mu(1, p_1 \cdots p_k) - \sum_{i=0}^{k-2} \binom{k-1}{i} (-1)^{i} - (-1)^{k-1} + (-1)^{k-1} \\
&= \mu(1, p_1 \cdots p_k) - \sum_{i=0}^{k-1} \binom{k-1}{i} (-1)^{i} + (-1)^{k-1} \\
&= \mu(1, p_1 \cdots p_k) - (1-1)^{k-1} + (-1)^{k-1} \\
&= \mu(1, p_1 \cdots p_k) + (-1)^{k-1} = 0.
\end{aligned}$$

Note the second the last line is obtained by the binomial theorem. This completes the proof of the claim. A similar strong induction argument can be used to show that $\mu(1, n) = 0$ when $n \in \mathbb{N}$ is not square-free. Thus $\mu(1, n) = \mu(n)$ and we see that this definition of the Möbius function corresponds with the definition found in the integers.

The following result was first shown by Rota [25].

Lemma 18 (Möbius inversion formula)**.** *Let $(L, \preceq)$ be a locally finite lattice. If $f : L \to \mathbb{R}$ and g is defined by*

$$g(n) = \sum_{d \preceq n} \mu(d, n) f(d)$$

then

$$f(n) = \sum_{d \preceq n} g(d).$$

Proof. Note

$$\begin{aligned}
\sum_{d \preceq n} g(d) &= \sum_{d \preceq n} \sum_{e \preceq d} \mu(e, d) f(e) \\
&= \sum_{e \preceq n} f(e) \sum_{e \preceq d \preceq n} \mu(e, d) \\
&= f(n).
\end{aligned}$$

The last equality is obtained by the Property 2 of the Möbius function, which implies that the inner sum is 0 if $e \prec n$, and is 1 if $e = n$. □

We also have a dual statement for Möbius inversion first shown by Rota [25].

Lemma 19 (Dual Möbius inversion formula)**.** *Let $(L, \preceq)$ be a locally finite lattice. If $r : L \to \mathbb{R}$ and s is defined by*

$$s(d) = \sum_{d \preceq e} r(e),$$

then

$$r(d) = \sum_{d \preceq e} \mu(d, e) s(e).$$

Proof. Note

$$\begin{aligned} \sum_{d \preceq e} \mu(d, e) s(e) &= \sum_{d \preceq e} \mu(d, e) \sum_{e \preceq n} r(n) \\ &= \sum_{d \preceq e} \sum_{e \preceq n} \mu(d, e) r(n) \\ &= \sum_{d \preceq n} r(n) \sum_{d \preceq e \preceq n} \mu(d, e) = r(d). \end{aligned}$$

The last equality again obtained by property 2 of the Möbius function. □

Fix S, T where S is a set of x distinct elements in a lattice L, and T a finite set of atoms whose join is w. Assume that for $d \in L$, the number of elements of $n \in S$ satisfying $d \preceq n$ is of the form

$$\frac{x}{f(d)} + R(d)$$

where f is a multiplicative function on L, and $R(d)$ is some small error term. Since every element is above 0 in the lattice, we have $f(0) = 1$ and $R(0) = 0$.

Recall that $|M(S, T)|$ is the number of elements of S whose meet with w is 0. We aim to show that for any real number z, we have

$$|M(S, T)| \leq \frac{x}{V(z)} + E(z),$$

where the functions $V(z)$ and $E(z)$ will be defined later. The function $E(z)$ is intended to be a small error term dependent on z. In order to sieve up to z, we require some method of comparing lattice elements to the number z. This will be the norm function $N : L \to \mathbb{N}$. We require that the norm function be increasing, that is, for lattice elements $d, e \in L$ with

$d \prec e$, we have $N(d) < N(e)$. For any real number z, the set $\{n \in L : N(n) < z\}$ must also be finite. In the context of the integers, we can take the norm function to be the absolute value.

We now proceed to estimate $|M(S,T)|$. By the definition of the Möbius function, we observe that

$$|M(S,T)| = \sum_{\substack{n \in S \\ n \wedge w = 0}} 1 = \sum_{n \in S} \sum_{d \preceq n \wedge w} \mu(0,d). \tag{3.12}$$

Let $\lambda : L \times L \to \mathbb{R}$ be a function satisfying $\lambda(d,d) = 1$, $\lambda(d_1,d_2) = 0$ if $d_1 \not\preceq d_2$, and the remaining values arbitrary. We make the observation that

$$\sum_{d \preceq n} \mu(0,d) \leq \left(\sum_{d \preceq n} \lambda(0,d) \right)^2.$$

This produces Selberg's observation in this context. For convenience, take $\lambda(0,d) = 0$ if $N(d) > z$. Note that this implies that some of the following summations are over elements which are bounded in norm to less than or equal to z, as all remaining terms are 0. Write

$$\begin{aligned} |M(S,T)| &= \sum_{n \in S} \sum_{d \preceq n \wedge w} \mu(0,d) \\ &\leq \sum_{n \in S} \left(\sum_{d \preceq n \wedge w} \lambda(0,d) \right)^2 \\ &= \sum_{n \in S} \left(\sum_{d_1, d_2 \preceq n \wedge w} \lambda(0,d_1)\lambda(0,d_2) \right) \\ &= \sum_{d_1, d_2 \preceq w} \lambda(0,d_1)\lambda(0,d_2) \sum_{\substack{n \in S \\ d_1 \vee d_2 \preceq n}} 1 \end{aligned}$$

This produces the estimate

$$|M(S,T)| \leq x \sum_{d_1,d_2} \frac{\lambda(0,d_1)\lambda(0,d_2)}{f(d_1 \vee d_2)} + \sum_{d_1,d_2 \preceq w} |\lambda(0,d_1)||\lambda(0,d_2)||R(d_1 \vee d_2)|.$$

We aim to minimize the main term. Since f is multiplicative, we have

$$\sum_{d_1,d_2} \frac{\lambda(0,d_1)\lambda(0,d_2)}{f(d_1 \vee d_2)} = \sum_{d_1,d_2} \frac{\lambda(0,d_1)\lambda(0,d_2)}{f(d_1)f(d_2)} f(d_1 \wedge d_2).$$

Let g be the Möbius inversion of f defined as in Lemma 18. We obtain

$$\sum_{d_1,d_2} \frac{\lambda(0,d_1)\lambda(0,d_2)}{f(d_1)f(d_2)} f(d_1 \wedge d_2) = \sum_{d_1,d_2} \frac{\lambda(0,d_1)\lambda(0,d_2)}{f(d_1)f(d_2)} \sum_{\delta \preceq d_1 \wedge d_2} g(\delta)$$
$$\leq \sum_{\delta \preceq w} |g(\delta)| \left(\sum_{\delta \preceq d} \frac{\lambda(0,d)}{f(d)} \right)^2.$$

Let

$$u(\gamma, \delta) = \sum_{\gamma \preceq \delta \preceq d} \frac{\lambda(\gamma, d)}{f(d)}.$$

We look to optimize

$$\sum_{\delta \preceq w} |g(\delta)| u(0,\delta)^2. \tag{3.13}$$

Note that

$$\sum \mu(0,\delta) u(0,\delta) = \sum \mu(0,\delta) \sum_{\delta \preceq d} \frac{\lambda(0,d)}{f(d)}$$
$$= \sum_d \frac{\lambda(0,d)}{f(d)} \sum_{\delta \preceq d} \mu(0,\delta)$$
$$= \frac{\lambda(0,0)}{f(0)} = 1.$$

Therefore we can minimize (3.13) by writing

$$\sum_\delta |g(\delta)| u(0,\delta)^2 = \sum_{g(\delta)\neq 0} |g(\delta)| \left(u(0,\delta) - \frac{\mu(0,\delta)}{|g(\delta)|V(z)} \right)^2 + \frac{2}{V(z)} - \frac{1}{V^2(z)} \sum_{g(\delta)\neq 0} \frac{\mu^2(0,\delta)}{|g(\delta)|}$$

Let

$$V(z) = \sum_{\substack{N(d)\leq z \\ g(d)\neq 0}} \frac{\mu^2(0,d)}{|g(d)|}.$$

We obtain

$$\sum_\delta |g(\delta)| u(0,\delta)^2 = \sum_{g(\delta)\neq 0} |g(\delta)| \left(u(0,\delta) - \frac{\mu(0,\delta)}{|g(\delta)|V(z)} \right)^2 + \frac{1}{V(z)}.$$

This can be minimized by choosing

$$u(0,\delta) = \begin{cases} \dfrac{\mu(0,\delta)}{V(z)|g(\delta)|} & \text{if } g(\delta) \neq 0 \\ 0 & \text{if } g(\delta) = 0 \end{cases}.$$

Thus we have shown that the estimate is

$$|M(S,T)| \leq \frac{x}{V(z)} + \sum_{d_1,d_2 \preceq w} |\lambda(0,d_1)||\lambda(0,d_2)||R(d_1 \vee d_2)|.$$

Note the similarity of the form for this bound when compared to the integers. The main difference comes in the error term, which can be shown to be $O(z^2)$ in the integer case.

One can compute $\lambda(0,d)$ by using Lemma 19 (Dual Möbius inversion in lattices) on $n(0,\delta)$ and $\lambda(0,d)/f(d)$. We have

$$\frac{\lambda(0,d)}{f(d)} = \sum_{d \preceq \delta} \mu(d,\delta)u(0,\delta) = \sum_{\substack{d \preceq \delta \\ g(\delta)\neq 0 \\ N(\delta)\leq z}} \frac{\mu(d,\delta)\mu(0,\delta)}{V(z)|g(\delta)|}.$$

Therefore

$$\lambda(0,d) = \frac{f(d)}{V(z)} \sum_{\substack{d \preceq \delta \\ g(\delta)\neq 0 \\ N(\delta)\leq z}} \frac{\mu(d,\delta)\mu(0,\delta)}{|g(\delta)|}.$$

To summarize, we have shown the following the bound, which was first independently obtained by Chow [7] and Wilson [32].

Theorem 20 (Selberg sieve on lattices, Wilson [32])**.** *Let $(L,\preceq)$ be a locally finite lattice. Let $S \subseteq L$ be a set of x distinct elements with the property that for any $d \in L$, the number of elements $n \in S$ satisfying $d \preceq n$ is of the form*

$$\frac{x}{f(d)} + R(d)$$

where f is a multiplicative function and R is some suitable error function. Let T be a set of atoms in $(L,\preceq)$ and write w to be the join of T. Let $M(S,T)$ be the set of elements in S for which the meet with w is 0 (the bottom of the lattice). Let $N : L \to \mathbb{N}$ be defined such

that for lattice elements $d, e \in L$ with $d \prec e$, we have $N(d) < N(e)$. For any real number z, assume that the set $\{n \in L : N(n) < z\}$ is finite. Let

$$g(n) = \sum_{d \preceq n} \mu(d, n) f(d).$$

Define

$$V(z) = \sum_{\substack{N(d) \leq z \\ g(d) \neq 0}} \frac{\mu^2(0, d)}{|g(d)|}$$

and

$$\lambda(0, d) = \frac{f(d)}{V(z)} \sum_{\substack{d \preceq \delta \\ g(\delta) \neq 0 \\ N(\delta) \leq z}} \frac{\mu(d, \delta)\mu(0, \delta)}{|g(\delta)|}.$$

We have

$$|M(S, T)| \leq \frac{x}{V(z)} + \sum_{d_1, d_2 \preceq w} |\lambda(0, d_1)||\lambda(0, d_2)||R(d_1 \vee d_2)|,$$

where $d_1 \vee d_2$ is the join of d_1 and d_2.

3.3 Counting the subspaces of a finite vector space

It was noted by Wilson [32] that the lattice of subspaces of a finite vector space was a context where the Selberg sieve could be applied. Since Wilson only mentioned the possible use of the Selberg sieve without computation, we will give explicit estimates in this section. The lattice of subspaces certainly has some properties which make the application of the Selberg sieve possible, although we will have to make some assumptions to obtain numerical estimates.

Let $\mathcal{V}$ be an n-dimensional vector space over $\mathbb{F}_q$, a finite field with q elements. Let L be comprised of the set of all subspaces of $\mathcal{V}$, ordered by inclusion. For subspaces d and e of the vector space $\mathcal{V}$, we have $d \preceq e$ if d is a subspace of e. Thus L is a lattice where the bottom element 0 is the 0-dimensional subspace $\{0\}$, and the top element is $\mathcal{V}$.

Fix a set of vectors $\{t_1, \ldots, t_\ell\}$ in $\mathcal{V}$ where the t_i are pairwise linearly independent. We aim to compute the number of subspaces which do not contain any non-zero vector as a linear combination of $t_1, \ldots, t_\ell$. If we take S to be the set of all subspaces of $\mathcal{V}$ and T to be the set $\{\text{span}\{t_1\}, \ldots, \text{span}\{t_\ell\}\}$, then $M(S, T)$ is the set of subspaces with no non-zero

vector as a linear combination of $\{t_1, \ldots, t_\ell\}$. We aim to compute $|M(S,T)|$. For the sake of brevity, we identify a vector t_i with $\text{span}\{t_i\}$. Thus we write $T = \{t_1, \ldots, t_\ell\}$.

The *Gaussian binomial coefficient*

$$\binom{n}{k}_q$$

is the number of subspaces of dimension k in an n-dimensional vector space over $\mathbb{F}_q$. It is well-known that

$$\binom{n}{k}_q = \frac{(1+q)\cdots(1+q+\cdots+q^{n-1})}{(1+q)\cdots(1+q+\cdots+q^{k-1})(1+q)\cdots(1+q+\cdots+q^{n-k-1})}.$$

The Gaussian binomial theorem (for example, see [24]) states that

$$\sum_{k=0}^{n} \binom{n}{k}_q q^{\binom{k}{2}} y^k = \prod_{i=0}^{n-1}(1+q^i y). \tag{3.14}$$

We take the convention that $\binom{k}{2} = 0$ if k is 0 or 1. We may now define a Möbius function on L. The following result is well-known, for example, in Rota [25, Example 2], and by Hall [18, Section 2.7]. The proof is provided for the sake of completeness.

Lemma 21 (Möbius function on the lattice of subspaces, [18], [25])**.** *Let L be the lattice of subspaces of a vector space $\mathcal{V}$ over $\mathbb{F}_q$, ordered by inclusion. For arbitrary $d, e \in L$, we define*

$$\mu(d,e) = \begin{cases} (-1)^{\dim e - \dim d} q^{\binom{\dim e - \dim d}{2}}, & \text{if } d \preceq e, \\ 0, & \text{if } d \not\preceq e. \end{cases}$$

The function μ satisfies the three properties of Möbius functions as defined in Section 3.2.

Proof. Properties 1 and 3 of the Möbius function directly follow from the definition of μ. To show that μ satisfies Property 2 of the Möbius function, consider $d, c \in L$ with $d \preceq c$. We have

$$\begin{aligned} \sum_{d \preceq e \preceq c} \mu(d,e) &= \sum_{d \preceq e \preceq c} (-1)^{\dim e - \dim d} q^{\binom{\dim e - \dim d}{2}} \\ &= \sum_{k=0}^{\dim c - \dim d} \binom{\dim c - \dim d}{k}_q q^{\binom{k}{2}} (-1)^k \\ &= \prod_{i=0}^{\dim c - \dim d - 1} (1 + q^i(-1)). \end{aligned}$$

The last equality is obtained by (3.14). Note that the first term of the product on the right side is $(1 + q^0(-1)) = 0$. Thus we have

$$\sum_{d \preceq e \preceq c} \mu(d, e) = \begin{cases} 0, & \text{if } \dim c > \dim d, \\ 1, & \text{if } \dim c = \dim d. \end{cases}$$

The latter is only possible in the case that $c = d$. This proves that the function μ as defined satisfies all three properties, and is therefore the Möbius function on the lattice of subspaces. □

Now we apply the Selberg sieve to count $M(S, T)$. Recall that $S = L$ is the set of all subspaces of $\mathcal{V}$ with $x = |S|$. The set T is some subset of 1-dimensional subspaces with w as the join of T. Let $m = \dim w$. For $d \in L$, the number of elements $e \in L$ satisfying $d \preceq e$ is of the form

$$\frac{x}{f(d)} + R(d),$$

for some functions f and R. For $d \in L$, we take f to be

$$f(d) = x^{\frac{\dim d}{n}}.$$

Then

$$f(\mathcal{V}) = x^{\frac{\dim \mathcal{V}}{n}} = x^1.$$

Since $\mathcal{V}$ is the only element which is above $\mathcal{V}$ itself, we should have $x/f(\mathcal{V}) = 1$, which is the case. Likewise the 0-dimensional subspace gives

$$f(\{0\}) = x^{\frac{\dim\{0\}}{n}} = 1.$$

Note that lattice elements $d, e \in L$ are subspaces of $\mathcal{V}$. So $d \wedge e$, which denotes the largest subspace of $\mathcal{V}$ which is contained in d and e, is the subspace $d \cap e$. Similarly $d \vee e$ denotes the smallest subspace containing both d and e, which is the subspace $d + e$ spanned by d and e. Thus for $d, e \in L$, we have

$$\dim(d \wedge e) + \dim(d \vee e) = \dim(d) + \dim(e).$$

It follows that

$$x^{\frac{\dim(d \wedge e)}{n}} \cdot x^{\frac{\dim(d \vee e)}{n}} = x^{\frac{\dim d}{n}} \cdot x^{\frac{\dim e}{n}}.$$

This implies that f is multiplicative in the lattice of subspaces.

Using this function as f also provides the useful property that the function value is only dependent on the dimension of the lattice element as a subspace. Recall that g is the Möbius inversion function of f. We write

$$\begin{aligned} g(e) &= \sum_{d \preceq e} \mu(d,e) f(d) \\ &= \sum_{d \preceq e} (-1)^{\dim e - \dim d} q^{\binom{\dim e - \dim d}{2}} f(d) \\ &= \sum_{k=0}^{\dim e} \binom{\dim e}{k}_q (-1)^{\dim e - k} q^{\binom{\dim e - k}{2}} x^{\frac{k}{n}}. \end{aligned} \tag{3.15}$$

Next we compute $V(z)$. Choose a norm function N and a number z so that every subspace which is below w has norm at most z, and every other element in S has norm greater than z. Then we have

$$V(z) = \sum_{\substack{N(d) \leq z \\ g(d) \neq 0}} \frac{\mu^2(0,d)}{|g(d)|} = \sum_{\substack{d \preceq w \\ g(d) \neq 0}} \frac{q^{2\binom{\dim d}{2}}}{|g(d)|} = \sum_{k=0}^{m} \sum_{g(d) \neq 0 \text{ where } \dim d = k} \binom{m}{k}_q \frac{q^{2\binom{k}{2}}}{|g(d)|}.$$

Also, we have

$$\begin{aligned} \lambda(0,d) &= \frac{f(d)}{V(z)} \sum_{\substack{d \preceq \delta \\ g(\delta) \neq 0 \\ N(\delta) \leq z}} \frac{\mu(d,\delta)\mu(0,\delta)}{|g(\delta)|} \\ &= \frac{f(d)}{V(z)} \sum_{\substack{d \preceq \delta \\ g(\delta) \neq 0 \\ N(\delta) \leq z}} \frac{(-1)^{2\dim \delta - \dim d} q^{\binom{\dim \delta - \dim d}{2} + \binom{\dim \delta}{2}}}{|g(\delta)|} \\ &= (-1)^{\dim d} \frac{f(d)}{V(z)} \sum_{\substack{d \preceq \delta \\ g(\delta) \neq 0 \\ N(\delta) \leq z}} \frac{q^{\binom{\dim \delta - \dim d}{2} + \binom{\dim \delta}{2}}}{|g(\delta)|}. \end{aligned}$$

Combine the above estimates with Theorem 20 to obtain

$$|M(S,T)| \leq \frac{x}{V(z)} + \sum_{d_1, d_2 \preceq w} |\lambda(0,d_1)||\lambda(0,d_2)||R(d_1 \vee d_2)|.$$

In order to produce numerical estimates, we now make some additional assumptions. Recall that w is the join of T and that $m = \dim w$. We compute an illustrative example for with $m = 2$. Since w has dimension 2, we have that the number of 1-dimension subspaces of w is $\binom{2}{1}_q = q + 1$. We recall

$$V(z) = \sum_{\substack{k=0 \\ g(d)\neq 0 \text{ where } \dim d = k}}^{m} \binom{m}{k}_q \frac{q^{2\binom{k}{2}}}{|g(d)|}.$$

By noting that the function g is only dependent on the dimension of the input, we may simplify $V(z)$ as follows. Let d_1 be some 1-dimensional subspace of w. Recall that we choose z and the norm function N such that all subspaces below w are included, but no others. We have

$$\begin{aligned} V(z) &= \binom{2}{0}_q \frac{q^{2\binom{\dim\{0\}}{2}}}{|g(0)|} + \binom{2}{1}_q \frac{q^{2\binom{\dim d_1}{2}}}{|g(d_1)|} + \binom{2}{2}_q \frac{q^{2\binom{\dim w}{2}}}{|g(w)|} \\ &= \frac{q^{2\binom{0}{2}}}{|g(0)|} + (q+1)\frac{q^{2\binom{1}{2}}}{|g(d_1)|} + \frac{q^{2\binom{2}{2}}}{|g(w)|} \\ &= 1 + \frac{q+1}{|g(d_1)|} + \frac{q^2}{|g(w)|}. \end{aligned}$$

Since $\dim d_1 = 1$, by (3.15) we have

$$\begin{aligned} g(d_1) &= \sum_{k=0}^{1} \binom{1}{k}_q (-1)^{1-k} q^{\binom{1-k}{2}} x^{\frac{k}{n}} \\ &= \binom{1}{0}_q (-1)^1 q^{\binom{1}{2}} + \binom{1}{1}_q q^{\binom{0}{2}} x^{\frac{1}{n}} \\ &= -1 + x^{\frac{1}{n}}. \end{aligned}$$

Also as $\dim w = 2$, by (3.15) we have

$$\begin{aligned} g(w) &= \sum_{k=0}^{2} \binom{2}{k}_q (-1)^{2-k} q^{\binom{2-k}{2}} x^{\frac{k}{n}} \\ &= q - (q+1)x^{\frac{1}{n}} + x^{\frac{2}{n}}. \end{aligned}$$

Combining the above, we find that

$$V(z) = 1 + \frac{q+1}{|-1 + x^{\frac{1}{n}}|} + \frac{q^2}{|q - (q+1)x^{\frac{1}{n}} + x^{\frac{2}{n}}|}.$$

Since

$$x = \sum_{k=0}^{n} \binom{n}{k}_q,$$

we have

$$\frac{x}{V(z)} = \frac{\sum_{k=0}^{n} \binom{n}{k}_q}{1 + \frac{q+1}{\left|-1+\left(\sum_{k=0}^{n} \binom{n}{k}_q\right)^{\frac{1}{n}}\right|} + \frac{q^2}{\left|q-(q+1)\left(\sum_{k=0}^{n} \binom{n}{k}_q\right)^{\frac{1}{n}}+\left(\sum_{k=0}^{n} \binom{n}{k}_q\right)^{\frac{2}{n}}\right|}}. \tag{3.16}$$

This completes the estimate by the Selberg sieve in this setting.

Now we will see how well $x/V(z)$ approximates $|M(S,T)|$. The true value of $|M(S,T)|$ may be computed as follows. Consider the set of 1-dimensional subspaces of w. For any higher dimensional subspace with non-zero meet with w, we have that it must contain some 1-dimensional subspace of w. The converse is also true. Thus to compute the number of subspaces with non-zero meet with w, we only need to compute the number of subspaces containing one of the 1-dimensional subspaces of w.

Suppose $d \in L$ contains two distinct 1-dimensional subspaces of w. Then d contains w, and so it contains every 1-dimensional subspace of w. We conclude that every subspace d containing a 1-dimensional subspace of w either contains exactly one 1-dimensional subspace or all of the 1-dimensional subspaces of w. Note that for a single 1-dimensional subspace, there are exactly

$$\sum_{k=0}^{n-1} \binom{n-1}{k}_q$$

subspaces of $\mathcal{V}$ which contain it. Similarly there are

$$\sum_{k=0}^{n-2} \binom{n-2}{k}_q$$

subspaces of $\mathcal{V}$ which contain w. Therefore by inclusion-exclusion, the number of elements in L which contain at least one 1-dimensional subspace of w is

$$\binom{2}{1}_q \sum_{k=0}^{n-1} \binom{n-1}{k}_q - \left(\binom{2}{1}_q - 1\right) \sum_{k=0}^{n-2} \binom{n-2}{k}_q,$$

which is equal to

$$(q+1) \sum_{k=0}^{n-1} \binom{n-1}{k}_q - q \sum_{k=0}^{n-2} \binom{n-2}{k}_q.$$

Thus we may write

$$|M(S,T)| = \sum_{k=0}^{n} \binom{n}{k}_q - (q+1)\sum_{k=0}^{n-1} \binom{n-1}{k}_q + q\sum_{k=0}^{n-2} \binom{n-2}{k}_q \tag{3.17}$$

This is slightly more difficult to compute when compared to the estimate in (3.16) as there are three distinct sums of Gaussian binomial coefficients which need to be computed. For higher $m = \dim w$, we would need to compute $m+1$ such sums.

The ratio between $|M(S,T)|$ and $x/V(z)$ gets closer to 1 as n and q get larger. In particular, the computed values have errors less than 1% for $n \geq 45$. For $n \geq 50$ and $q \geq 7$, the estimate is accurate to the first ten digits. This decrease in error for larger n can be seen in Table 3.1 at the end of the section, which contains $(x/V(z))/|M(S,T)|$ for various values of n and q. All numbers in the table are rounded to the first ten digits, and then truncated to five decimal places.

It can be seen from the above that the Selberg sieve requires one to make many decisions on making the estimate. Even with the many assumptions and lengthy analysis above, we have only obtained an estimate for the $m = 2$ case. The ratios of the estimate to the true value does approach 1 asymptotically, but one should note that this is mostly the effect that the size of the lattice grows very quickly, instead of the difference between $x/V(z)$ and $|M(S,T)|$ being small. For example, the absolute error at $n = 16, q = 5$ is on the magnitude of 10^{42}.

$q\backslash n$	4	5	6	7	8	9	10	12
5	0.16199	0.42150	0.58826	0.70836	0.79578	0.85858	0.90302	0.95522
7	0.11336	0.44138	0.63881	0.76744	0.85239	0.90695	0.94198	0.97765
8	0.09855	0.45139	0.65870	0.78822	0.87069	0.92137	0.95276	0.98304
11	0.07077	0.47769	0.70394	0.83136	0.90610	0.94744	0.97108	0.99120
13	0.05957	0.49219	0.72604	0.85058	0.92072	0.95745	0.97764	0.99375
16	0.04813	0.51041	0.75172	0.87154	0.93579	0.96726	0.98374	0.99591

Table 3.1: The value $(x/V(z))/|M(S,T)|$ for various of n and q, where $m = 2$. Values computed by applying (3.17) and (3.16).

3.4 Subspaces of a vector space using the Turán sieve

We now apply the Turán sieve to the same problem as Section 3.3. Recall that the lattice $S = L$ is constructed to be all of the subspaces of $\mathcal{V}$ over $\mathbb{F}_q$, where the ordering is by inclusion. The set T is some subset of one-dimensional subspaces of S. We denote w as the join of T, and we assume that $m = \dim w = 2$. We also assume that for $d \in L$, the number of elements of $e \in S$ satisfying $d \preceq e$ is of the form

$$\frac{x}{f(d)} + R(d),$$

where $x = |S|$, f is a multiplicative function on L, and $R(d)$ is the error term. For $t \in T$ and $s \in S$, we take $t \sim s$ if t is a subspace of s. We define

$$\omega(s) = \#\{t \in T : t \preceq s\}.$$

To estimate the same set $M(S,T)$ as the Selberg sieve, we need the additional assumption that T is the set of all atoms which are below an element w in the lattice. Recall that the join of T is w. Combining the above two statements we have that every element $s \in S$ that has 0 meet with w is above some element in T. Therefore

$$|M(S,T)| = \#\{s \in S : \omega(s) = 0\}.$$

By the assumption above, we have

$$\deg(t) = \frac{x}{f(t)} + R(t).$$

Using the lattice structure, we may write

$$n(t_1,t_2) = \begin{cases} \frac{x}{f(t_1 \vee t_2)} + R(t_1 \vee t_2), & \text{if } t_1 \neq t_2, \\ \deg(t_1), & \text{if } t_1 = t_2. \end{cases}$$

Since f is multiplicative, we have $f(t_1 \vee t_2) = f(t_1)f(t_2)$.

We recall by the Turán sieve that

$$\#\{s \in S : \omega(s) = 0\} \leq x^2 \frac{\sum_{t_1,t_2 \in T} n(t_1,t_2)}{\left(\sum_{t \in T} \deg(t)\right)^2} - x,$$

where the sum is over all ordered pairs (t_1, t_2), including pairs where $t_1 = t_2$. We may write

$$\#\{s \in S : \omega(s) = 0\} \leq x \left(x \frac{\sum_{\substack{t_1,t_2 \in T \\ t_1 \neq t_2}} n(t_1, t_2) + \sum_{t \in T} \deg(t)}{\left(\sum_{t \in T} \deg(t)\right)^2} - 1 \right)$$
$$= x \left(x \frac{\sum_{\substack{t_1,t_2 \in T \\ t_1 \neq t_2}} \left(\frac{x}{f(t_1 \vee t_2)} + R(t_1 \vee t_2)\right) + \sum_{t \in T} \left(\frac{x}{f(t)} + R(t)\right)}{\left(\sum_{t \in T} \left(\frac{x}{f(t)} + R(t)\right)\right)^2} - 1 \right).$$

Now by choosing $f(d) = x^{\frac{\dim d}{n}}$ as before, we have

$$\#\{s \in S : \omega(s) = 0\} \leq x \left(x \frac{\sum_{\substack{t_1,t_2 \in T \\ t_1 \neq t_2}} \left(x^{1-\frac{2}{n}} + R(t_1 \vee t_2)\right) + \sum_{t \in T} \left(x^{1-\frac{1}{n}} + R(t)\right)}{\left(\sum_{t \in T} \left(x^{1-\frac{1}{n}} + R(t)\right)\right)^2} - 1 \right).$$

Lastly, we recall $\dim w = 2$. This implies $|T| = \binom{2}{1}_q = q + 1$. It follows that

$$\#\{(t_1, t_2) : t_1, t_2 \in T, t_1 \neq t_2\} = 2\binom{q+1}{2}.$$

Thus we produce the estimate

$$x \left(\frac{2\binom{q+1}{2} x^{2-\frac{2}{n}} + x \sum_{\substack{t_1,t_2 \in T \\ t_1 \neq t_2}} R(t_1 \vee t_2) + (q+1)x^{2-\frac{1}{n}} + x \sum_{t \in T} R(t)}{(q+1)^2 x^{2-\frac{2}{n}} + 2(q+1)x^{1-\frac{1}{n}} \sum_{t \in T} R(t) + \left(\sum_{t \in T} R(t)\right)^2} - 1 \right). \tag{3.18}$$

To compute a similar numerical estimate to the Selberg sieve estimate in Section 3.3, we ignore the error term and approximate (3.18) with

$$E(n, q) = x \left(\frac{2\binom{q+1}{2} x^{2-\frac{2}{n}} + (q+1)x^{2-\frac{1}{n}}}{(q+1)^2 x^{2-\frac{2}{n}}} - 1 \right). \tag{3.19}$$

Recall that

$$x = \sum_{k=0}^{n} \binom{n}{k}_q,$$

so equation (3.19) suffices for numerical estimates.

At the end of this section, we have Table 3.2, which contains $E(n,q)/|M(S,T)|$ for various values of n and q. All numbers in the table are rounded to the first ten digits, and then truncated to five decimal places. The estimate from (3.19) is superior to the Selberg sieve estimate for very small values of n. On the other hand, as n increases, the error of the Turán sieve estimate grows faster than the size of the lattice, so the fraction does not converge to 1. One can see that this should be the case by noting that

$$E(n,q) = x\left(\frac{2\binom{q+1}{2} + (q+1)x^{\frac{1}{n}}}{(q+1)^2} - 1\right).$$

As x becomes larger in relation to q, we should expect $E(n,q)/|M(S,T)| \geq E(n,q)/x$ to grow at the rate of $x^{1/n}$.

The greater advantage here is the ease of use. Construction of the Selberg estimate for other m requires the computation of

$$V(z) = \sum_{\substack{k=0 \\ g(d)\neq 0 \text{ where } \dim d = k}}^{m} \binom{m}{k}_q \frac{q^{2\binom{k}{2}}}{|g(d)|},$$

where

$$g(d) = \sum_{k=0}^{\dim d} \binom{\dim d}{k}_q (-1)^{\dim d - k} q^{\binom{\dim d - k}{2}} x^{\frac{k}{n}}.$$

Notably the derivation of $g(d)$ for larger dimensions requires successively more difficult computations of the Möbius inversion on f. For the Turán sieve, define $T(m,q)$ to be the size of T given $m = \dim w$ in $\mathbb{F}_q$. We obtain using the Gaussian binomial that

$$T(m,q) = \sum_{k=0}^{m-1} q^k.$$

With a minor extension of the method above, we find that the Turán sieve estimate with arbitrary m is

$$E(n,q,m) = x\left(\frac{2\binom{T(m,q)}{2}x^{2-\frac{2}{n}} + T(m,q)x^{2-\frac{1}{n}}}{T(m,q)^2 x^{2-\frac{2}{n}}} - 1\right).$$

For fixed n, q, m this estimate is relatively easy to derive and compute.

$q\backslash n$	4	5	6	7	8	9	10	12
5	1.15103	1.45675	2.03562	2.96445	4.39316	6.53832	9.76656	21.77040
7	1.10674	1.57346	2.45086	3.92193	6.37177	10.31796	16.80511	44.31948
8	1.09323	1.62517	2.63939	4.37478	7.36747	12.32577	20.78005	58.58711
11	1.06771	1.75667	3.14569	5.64880	10.36266	18.71488	34.28441	113.37373
13	1.05731	1.82916	3.44526	6.44175	12.36206	23.22709	44.46666	159.88985
16	1.04659	1.92195	3.85319	7.56666	15.36233	30.30719	61.31231	244.65714

Table 3.2: The value $E(n,q)/|M(S,T)|$ for various of n and q, where $m = 2$. Values computed by applying (3.17) and (3.19).

Chapter 4

Hamiltonian graphs and induced subgraphs

The results in this chapter are based on the paper "Minimal induced subgraphs of two classes of 2-connected non-Hamiltonian graphs." *Discrete Mathematics*, 345(7):112869, 2022, co-authored with Joseph Cheriyan, Sepehr Hajebi, and Sophie Spirkl. The aim of this chapter is to attack the problem of characterizing Hamiltonicity from the perspective of forbidden induced subgraphs. Of particular interest are results where classes of 2-connected graphs produced by excluding induced subgraphs are shown to be Hamiltonian. The main results of this chapter use the notion of HC-obstructions.

4.1 Introduction

Let us describe our main results. A *clique* in a graph G is a set K of pairwise adjacent vertices. A *stable set* in a graph G is a set S of pairwise non-adjacent vertices. A *split graph* is a graph G with a partition (S, K) of $V(G)$ such that S is a stable set and K is a clique in G. The left-most graph in Figure 4.1 is a split graph, with the four central vertices comprising the clique and the three outer vertices making the stable set. Recall that a graph H is an *HC-obstruction* if H is 2-connected, has no Hamiltonian cycle, and every induced subgraph of H either equals H, or is not 2-connected, or has a Hamiltonian cycle.

An *n-sun* is a graph obtained from a cycle C with $2n$ vertices $v_1, \dots, v_{2n}$ that occur in this order along C by adding all edges $v_{2i}v_{2j}$ for distinct $i, j \in \{1, \dots, n\}$. An *n-nova* is

Figure 4.1: From left to right: the snare, the 2-nova, a theta, and a triangle-free wheel. Squiggly edges represent paths of length at least one.

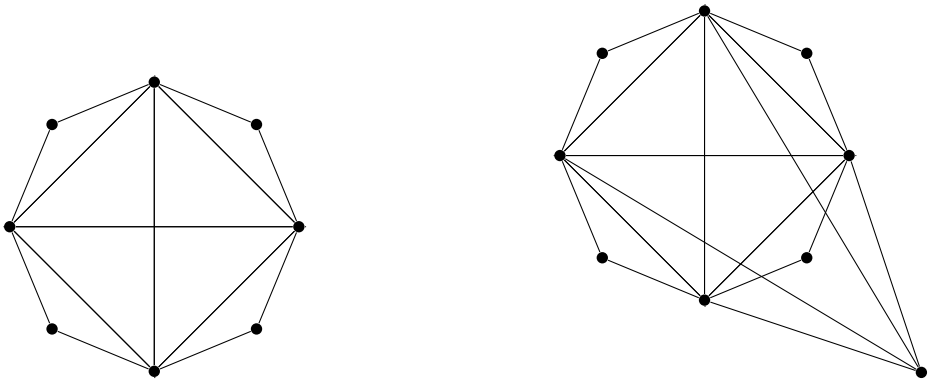


Figure 4.2: From left to right: a 4-sun and 4-nova.

obtained from an n-sun by adding a vertex w and edges wv_{2i} for all $i \in \{1, \ldots, n\}$. See Figure 4.2 for an illustration of a 4-sun and a 4-nova. The *net* is the unique graph with degree sequence $(3, 3, 3, 1, 1, 1)$, and equivalently the graph with vertex set $\{a, b, c, a_1, b_1, c_1\}$ and edge set $\{ab, bc, ac, aa_1, bb_1, cc_1\}$. The *snare* is the graph obtained from a net by adding a vertex and making it adjacent to every vertex of the net, this graph is depicted as the left-most graph of Figure 4.1. Our first theorem, the following, gives a complete characterization of HC-obstructions that are split graphs.

Theorem 22 ([5])**.** *The snare and all n-novae for $n \geq 2$ are HC-obstructions. Moreover, these are the only HC-obstructions which are split graphs.*

A *theta* is a graph consisting of two non-adjacent vertices u and v and three paths P_1, P_2, P_3 from u to v and each of length at least two, such that the sets $V(P_1)\backslash\{u, v\}, V(P_2)\backslash\{u, v\}, V(P_3) \setminus \{u, v\}$ are disjoint and have no edges between them. The vertices u and v are the *ends* of the theta. A *closed theta* is a graph obtained from a theta with ends u, v by adding the edge uv.

A graph is *triangle-free* if it contains no three-vertex clique. A *wheel* is a pair (W, v) such

that W is a cycle, and v is a vertex with at least three neighbours in W[1]. A triangle-free wheel is depicted as the right-most graph in Figure 4.1.

Theorem 23 ([5])**.** *All thetas, triangle-free closed thetas, and triangle-free wheels are HC-obstructions, and they are the only HC-obstructions which are triangle-free.*

4.2 Split graphs

In this section we prove Theorem 22. The following is well-known (see, for example, [8]):

Lemma 24. *Let G be a graph and $X \subseteq V(G)$. If $G \setminus X$ has more than $|X|$ connected components, then G has no Hamiltonian cycle.*

From this, we deduce:

Lemma 25. *The snare and all n-novae for $n \geq 2$ are 2-connected graphs with no Hamiltonian cycle.*

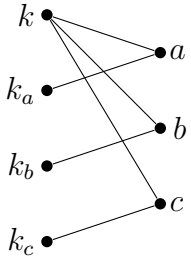
Proof. Clearly, these graphs are 2-connected. First, consider the snare. Suppose it has a Hamiltonian cycle; then, a Hamiltonian path of a net can be obtained by deleting one particular vertex of the snare. This is a contradiction. Next, consider an n-nova for $n \geq 2$. The graph is non-Hamiltonian, by Lemma 24 with $X = \{v_{2i} : i \in \{1, \ldots, n\}\}$, where the vertex labels are as in the definition. □

Note that the snare has the property that every 2-connected induced subgraph is Hamiltonian. Similarly, every n-nova for $n \geq 2$ has the property that every 2-connected induced subgraph is Hamiltonian. Combining with Lemma 25, we have that snares and n-novae are HC-obstructions. A notable example is a K_5 deleting three edges which form a triangle; this graph is a 2-nova. See the second figure from the left of Figure 4.1,

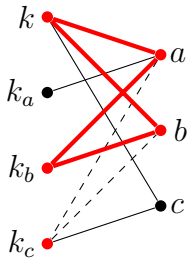
This naturally leads to the more focused question of resolving the problem for split graphs. In other words, which 2-connected split graphs are HC-obstructions? This is answered by the previously stated Theorem 22.

In view of Lemma 25, in order to prove Theorem 22, it suffices to show the following Lemma.

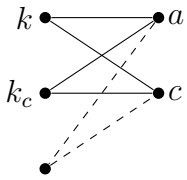
Lemma 26. *Let G be a 2-connected split graph (S, K) which is an HC-obstruction. We have that G contains a snare or an n-nova (for $n \geq 2$) as an induced subgraph.*



(a) The starting configuration of vertices. The vertices a, b, c are fixed and distinct. The vertices k_a, k_b, k_c are arbitrary and may not be distinct.



(b) The 2-nova formed by k, k_a, k_c, a, b if k_c is not adjacent to a or b, highlighted in red.



(c) The vertices which form a 2-nova.

Figure 4.3: Some 2-nova contradictions from the proof of Lemma 26. The convention taken in these diagrams is that the clique is on the left and the stable set is on the right. Dashed lines indicate non-adjacent vertex pairs.

Proof. Take G with K made as large as possible. In other words, if $s \in S$ is adjacent to all vertices in K, then we should add s to K instead.

Case 1: Consider the case that there exists a vertex $k \in K$ such that $|N(k) \cap S| \geq 3$, where there are at least 3 neighbours of k in S. Denote three of the neighbours of k in S as a, b, c. Write k_a, k_b, k_c to be an arbitrary set of neighbours of a, b, c respectively, not necessarily distinct and not the vertex k. Note that k_a, k_b, k_c must necessarily exist as a consequence of G being 2-connected. This is depicted in Figure 4.3a. In the case that k_a, k_b, k_c are all distinct and is each adjacent to only one of a, b, c, we have a snare. If it is the case that one of k_a, k_b, k_c is adjacent to all three of a, b, c, then we have an induced 2-nova. Since we chose k_a, k_b, k_c arbitrarily, we have that no neighbour of a, b, or c can be adjacent to all three of a, b, c. Therefore $N(a) \cap N(b) \cap N(c) = \{k\}$.

So we only have the case that one of the vertices of k_a, k_b, k_c is adjacent to two of a, b, c. Without loss of generality suppose that it is k_b, being adjacent to a and b. In this case consider the induced subgraph on vertices k, a, b, c, k_b, k_c. Note that a, k, b, k_b form an even cycle. If k_c is not adjacent to a or b, then a, k, b, k_b, k_c is a 2-nova. This 2-nova is depicted in Figure 4.3b. Thus it must be the case that k_c is adjacent to one of a, b. Without loss of generality suppose that it is adjacent to a.

[1]In a standard definition of a wheel, the cycle W is required to be of length at least four. Note that this does not matter for our purposes as we are only concerned with triangle-free wheels.

Note that there is also an even cycle on c, k_c, a, k, so any other vertex in K must be adjacent to a or both b and c, else we would obtain an induced 2-nova. This is depicted in Figure 4.3c. There must be vertices in K not adjacent to a, as we assumed K to be as large as possible. We therefore find $k' \in K$ that is adjacent to both b and c. Now there is an even cycle c, k, b, k'. Every other vertex in K now must be adjacent to two of a, b, c as otherwise we have an induced 2-nova. Note that the vertices k, k_b, k_c, k' are adjacent to at least two of a, b, c. Therefore we have

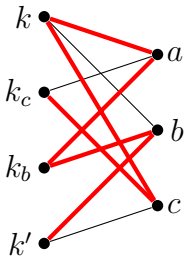
$$\begin{aligned} N(a) \setminus (N(b) \cup N(c)) &= \emptyset \\ N(b) \setminus (N(a) \cup N(c)) &= \emptyset \\ N(c) \setminus (N(a) \cup N(b)) &= \emptyset \\ N(a) \cup N(b) &= K \\ N(a) \cup N(c) &= K \\ N(b) \cup N(c) &= K \\ N(a) \cap N(b) \cap N(c) &= \{k\} \end{aligned}$$

Consider a fourth vertex $s \in S$. There must be one, as otherwise we have k_c, c, k, a, k_b, b, k' a path starting and ending in K with all of the vertices in S and this can to extended to a Hamiltonian cycle. This path is depicted in Figure 4.4a. There are at least two neighbours of s. The neighbours of s are in the intersection of two of the neighbourhoods of a, b, c. Suppose without loss of generality that one of the neighbours is $k_1 \in N(a) \cap N(b)$. Let k_2 be a neighbour of s distinct from k_1. If $k_2 \in N(a) \cap N(b)$, then the graph induced on a, k_1, b, k_2, s produces an induced 2-nova. This 2-nova is depicted in Figure 4.4b. Suppose without loss of generality that $k_2 \in N(b) \cap N(c)$. We have s, k_2, b, k_1 an even cycle with k_c which is an induced 2-nova.

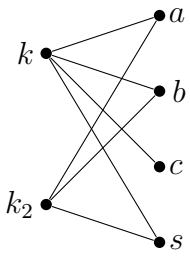
This ends the proof of the case where there exists $k \in K$ with $|N(k) \cap S| \geq 3$.

Case 2: Suppose that for all $k \in K$, we have $|N(k) \cap S| \leq 2$. For the sake of exposition, let us consider the simple case that each $k \in K$ has at most 1 neighbour in S. We can take two edges incident to each $s \in S$, and connect the ends of these paths in K appropriately to form a Hamiltonian cycle. This is depicted in Figure 4.4c.

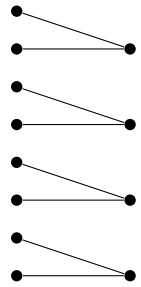
Now that each $k \in K$ has at most 2 neighbours in S, choose two edges e_s, f_s incident to each $s \in S$ and take the graph with these edges to be H. Observe that H has maximum degree 2. Choose H such that the graph has the least number of cycles. If H is acyclic, then we have that H is composed of paths $P_1, \ldots, P_t$ ending at vertices in K. In this case we can connect the ends of the paths in K as in the previous paragraph to form a Hamiltonian cycle.



(a) The path which can be extended to a Hamiltonian cycle if there are only three vertices in S, highlighted in red.

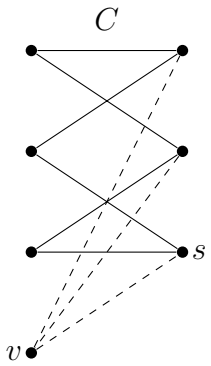


(b) A diagram for one case that a fourth vertex is in S. The 2-nova is induced by the vertices k, k_2, a, b, s.

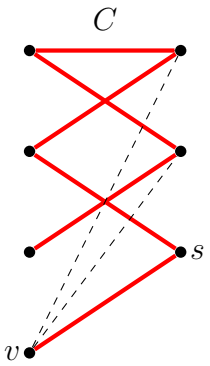


(c) A diagram on the case that $|N(k) \cap S| \leq 1$ for all $k \in K$. By connecting the ends of the shown paths in K with any remaining vertices, we construct a Hamiltonian cycle.

Figure 4.4: Some graphs used in the proof of Lemma 26. The convention taken in these diagrams is that the clique is on the left and the stable set is on the right.



(a) The vertices which form a n-nova from the cycle C in H.



(b) The choice of edges for H' given a $v \in K$ adjacent to a cycle C in H.

Figure 4.5: Some parts of the proof of Claim 2 in Lemma 26. Dashed lines indicate non-adjacent vertex pairs.

Consider one of the shortest cycles C in H. If C contains all vertices in K, then C is a Hamiltonian cycle. If there exists a vertex $v \in K$ that is not incident to any vertex of $C \cap S$ in G, then we have an n-nova. This is depicted in Figure 4.5a. If v is incident to a vertex $s \in C \cap S$, then we can switch one of the edges incident to s, say $e_s = st$ to the edge sv instead. Call this graph H'. This is depicted in Figure 4.5b. Now the edges of C no longer form a cycle. If no cycles are formed by this operation, then this contradicts the minimality of the number cycles. If a cycle C' is formed, then it must use the edge sv. But as H' is a graph of degree at most 2, the cycle must use all of the edges of $C \setminus \{st\}$. But the vertex t has degree 1 in H', so no cycle is formed. This contradicts the choice of H.

□

4.3 Triangle-free graphs

In this section, we prove Theorem 23.

Lemma 27. *Thetas, closed thetas, and triangle-free wheels are 2-connected graphs with no Hamiltonian cycle.*

Proof. Again, 2-connectivity can be checked easily. Thetas and closed thetas have no Hamiltonian cycles by Lemma 24, letting X be the set of ends of the (closed) theta. For a triangle-free wheel $H = (W, v)$, note that every edge e of W contains a vertex of degree two in H, and therefore every Hamiltonian cycle of H contains e. It follows that every Hamiltonian cycle contains all edges of W; but these edges form a cycle that does not contain v, and hence no Hamiltonian cycle exists. □

We assume that the reader is familiar with standard definitions for graph minors and planar graphs. A *model* of graph H in graph G is a collection of disjoint sets $(A_h)_{h \in V(H)}$ such that $G[A_h]$ is connected for all $h \in V(H)$, and for every edge $e = hh' \in E(H)$, there is at least one edge between A_h and $A_{h'}$ in G. We say that graph G contains H *as a minor* (or *contains an H-minor*) if G contains a model of H. A graph is *outerplanar* if it has a planar embedding with all vertices incident with the outer face.

Theorem 28 ([4])**.** *A graph is outerplanar if and only if it contains no subdivision of $K_{2,3}$ as a subgraph and no subdivision of K_4 as a subgraph.*

Lemma 29 ([4])**.** *Every 2-connected outerplanar graph is Hamiltonian.*

In view of Lemma 27, in order to prove Theorem 23, it suffices to prove the following.

Theorem 30. *Let G be a triangle-free 2-connected graph. If G is not Hamiltonian then it contains one of the following as an induced subgraph:*

1. *a triangle-free wheel;*
2. *a theta;*
3. *a closed theta.*

Proof. If G is outerplanar, then it is Hamiltonian by Lemma 29. We consider the cases where G has a subdivision of K_4 or $K_{2,3}$ as a subgraph.

Case 1: If G contains a subdivision of K_4 as a subgraph, then G contains K_4 as a minor. Consider the sets A_1, A_2, A_3, A_4 of a model of K_4 in G.

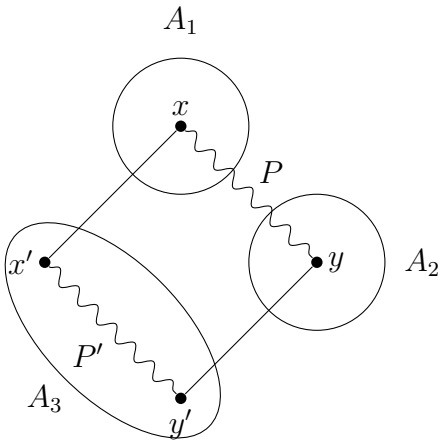
We claim that there exists an induced cycle in $G[A_1 \cup A_2 \cup A_3]$ with at least one vertex from each of A_1, A_2, A_3. Consider the following construction of a cycle. Choose $x \in A_1$ with a neighbour $x' \in A_3$ and $y \in A_2$ with a neighbour $y' \in A_3$. Let P be a shortest path from x to y in $G[A_1 \cup A_2]$. Let P' be a shortest path from x' to y' in $G[A_3]$. The cycle is formed by P, x, x', P', y', y. Let C be a shortest cycle constructed with the above method, and fix the associated vertices for the construction x, x', y, y'. If there is a chord in C, it is an edge between a vertex in P and P'. Suppose without loss of generality that it is between a vertex in $z \in A_1$ and a vertex in $z' \in A_3$, which is not the pair x and x'. We can construct a shorter cycle with z, z', y, y' and the associated paths. Therefore no such chord exists and C is an induced cycle. This is depicted in Figure 4.6a.

Note the following observation.

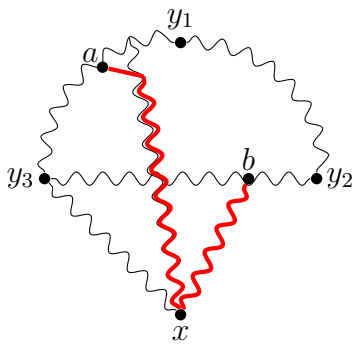
(1) *If any vertex in $G \setminus C$ has at least two neighbours in C, then we have a triangle-free wheel or a subdivision of $K_{2,3}$ (a theta).*

Let $x \in A_4$. Let P_i be an induced path in $G[A_i \cup A_4]$ from x to $y_i \in A_i \cap C$ for $i \in \{1, 2, 3\}$. Choose the P_i to be minimal. Note that the second to last vertex z_i of P_i has y_i as its only neighbour on C, as it cannot have more than one neighbour on C.

As y_1, y_2, y_3 cannot be a triangle, suppose that y_1 and y_2 are non-adjacent without loss of generality. Let Q be a shortest path between y_1 and y_2 contained in $G[P_1 \cup P_2]$. Choose a shortest path R which has endpoints a and b non-adjacent vertices on C and whose interior



(a) The cycle produced in the proof.



(b) The path R is depicted in red.

Figure 4.6: Some depictions of steps in the proof of Theorem 30

is a subpath of Q. If no vertex in the interior of R has a neighbour in $C \setminus \{a, b\}$, then the graph $G[R \cup C]$ is a theta. The path R is depicted in Figure 4.6b.

If there are at least two vertices $a', b' \in C \setminus \{a, b\}$ adjacent to the interior of R, then note that these vertices cannot be adjacent to $N(\{a, b\}) \cap R$. Otherwise there would be vertices in $G \setminus C$ with at least 2 neighbours in C and we obtain a contradiction by (1). Note that a, a', b' cannot form a triangle. Therefore we can choose a shorter R.

So there is at most one vertex in C adjacent to the interior of R. If a vertex in the interior of R has a neighbour $c \in C \setminus \{a, b\}$, then it must be adjacent to both a and b. Otherwise we could choose a shorter R with one endpoint being c. Consider the induced cycle $C' = G[(C \setminus \{c\}) \cup R]$ with the vertex c. As c has 3 neighbours in C', we have an induced triangle-free wheel.

This completes the proof in the case that G contains a K_4-minor.

Case 2: Suppose that G does not contain a K_4-minor. Since it is not outerplanar, it contains a subdivision of $K_{2,3}$ as a (not necessarily induced) subgraph. Choose such a subgraph H to have as few vertices as possible, and denote the two vertices of degree 3 as u and v. Let the three paths between u and v be denoted P_1, P_2, P_3.

Note that the paths P_1, P_2, P_3 are induced except possibly the edge uv. Otherwise we can shorten one of the paths, decreasing the number of vertices in H. There are no edges between $P_1 \setminus \{u, v\}, P_2 \setminus \{u, v\}, P_3 \setminus \{u, v\}$ as we obtain a K_4 minor if there is such an edge.

If uv is an edge, then $G[V(H)]$ a closed theta. If not, then $G[V(H)]$ is a theta. $\square$

Chapter 5

Conclusion

In the first part of the thesis we consider some generalizations of number theoretic sieves to combinatorics. In Chapter 2, we gave an exposition of Turán's [29] proof of the Hardy-Ramanujan theorem, demonstrating the idea of computing probabilistic parameters. This idea was extended to a combinatorial sieve method by Liu and Murty [22], and we outline the proof in the Chapter. Applying the Turán sieve to the problem of labelled graphs was relatively easy, difficulties arising only when looking for asymptotic implications. In Chapter 3, we start by outlining Selberg's [27] sieve estimate on the prime numbers, as explained by Cojocaru and Murty [9]. We give an exposition of Wilson's [32] (independently by Chow [7]) generalization of the sieve to lattices, following the flow and notation of the earlier exposition as much as possible. Lastly we apply the Selberg sieve to count subspaces of finite vector spaces. The examples demonstrate the differences clearly. In the Turán sieve, there is no need for the existence of many structural conditions, and no estimates on the Möbius function. The Selberg sieve allows for and requires many more choices in applying the estimate. This can lead to better results with the judicious application of analytic methods to choose estimates in reducing the errors.

The second part of the thesis is joint work with Joseph Cheriyan, Sepehr Hajebi, and Sophie Spirkl in Chapter 4. Here we define an minimal HC-obstruction to Hamiltonicity under the 2-connected induced subgraph relation. These HC-obstructions are then characterized for two classes of graphs, split graphs and triangle-free graphs.

References

[1] T. M. Apostol. *Introduction to Analytic Number Theory*. Undergraduate Texts in Mathematics. Springer New York, 1998.

[2] B. Bollobás and P. Erdős. Cliques in random graphs. *Mathematical Proceedings of the Cambridge Philosophical Society*, 80(3):419–427, 1976.

[3] J. Brousek. Minimal 2-connected non-hamiltonian claw-free graphs. *Discrete Mathematics*, 191(1-3):57–64, 1998.

[4] G. Chartrand and F. Harary. Planar permutation graphs. In *Annales de l'IHP Probabilités et statistiques*, volume 3 (4), pages 433–438, 1967.

[5] J. Cheriyan, S. Hajebi, Z. Qu, and S. Spirkl. Minimal induced subgraphs of two classes of 2-connected non-Hamiltonian graphs. *Discrete Mathematics*, 345(7):112869, 2022.

[6] S. Chiba and M. Furuya. A characterization of 2-connected $\{K_{1,3}, N_{3,1,1}\}$-free non-Hamiltonian graphs. *Discrete Mathematics*, 344(5):112321, 2021.

[7] T. Y. Chow. The combinatorics behind number-theoretic sieves. *Advances in Mathematics*, 138(2):293–305, 1998.

[8] V. Chvátal. Tough graphs and hamiltonian circuits. *Discrete Mathematics*, 5(3):215–228, 1973.

[9] A. C. Cojocaru and M. R. Murty. *An Introduction to Sieve Methods and Their Applications*. London Mathematical Society Student Texts. Cambridge University Press, 2005.

[10] J. Dalton. An exposition of Selberg's sieve. Master's thesis, University of Vermont, 2017.

[11] C.-J. É. G. N. de La Vallée Poussin. Recherches analytiques de la théorie des nombres premiers. *Annales de la Société scientifique de Bruxelles*, 21:351–368, 1896.

[12] G. Ding and E. Marshall. Minimal k-connected non-Hamiltonian graphs. *Graphs Comb.*, 34(2):289–312, 2018.

[13] D. Duffus, R. J. Gould, and M. S. Jacobson. Forbidden subgraphs and the Hamiltonian theme. *The Theory and Applications of Graphs (Kalamazoo, Mich. 1980, Wiley, New York, 1981)*, pages 297–316, 1981.

[14] P. Erdős, D. J. Kleitman, and B. L. Rothschild. Asymptotic enumeration of k_n-free graphs. In *Colloquio Internazionale sulle Teorie Combinatorie (Rome, 1973)*, Tomo II, Atti dei Convegni Lincei, No. 17, Accademia Nazionale dei Lincei, Rome, pages 19–27, 1976.

[15] S. Goodman and S. Hedetniemi. Sufficient conditions for a graph to be Hamiltonian. *Journal of Combinatorial Theory, Series B*, 16(2):175–180, 1974.

[16] R. L. Graham, D. E. Knuth, and O. Patashnik. *Concrete Mathematics: A Foundation for Computer Science*. Addison-Wesley, Reading, 1989.

[17] J. Hadamard. Sur la distribution des zéros de la fonction $\zeta(s)$ et ses conséquences arithmétiques. *Bulletin de la Société Mathématique de France*, 24:199–220, 1896.

[18] P. Hall. The Eulerian functions of a group. *The Quarterly Journal of Mathematics*, os-7(1):134–151, 1936.

[19] G. H. Hardy and S. Ramanujan. The normal number of prime factors of a number n. *The Quarterly Journal of Pure and Applied Mathematics*, 48:76–92, 1917.

[20] D. E. Knuth. *The Art of Computer Programming, Vol. 1: Fundamental Algorithms*. Addison-Wesley, Reading, Mass., third edition, 1997.

[21] W. Kuo, Y.-R. Liu, S. Ribas, and K. Zhou. The shifted Turán sieve method on tournaments. *Canadian Mathematical Bulletin*, 62(4):841–855, 2019.

[22] Y.-R. Liu and M. R. Murty. Sieve methods in combinatorics. *Journal of Combinatorial Theory, Series A*, 111(1):1–23, 2005.

[23] Y.-R. Liu and J. C. Saunders. Sieve methods in random graph theory, 2018. arXiv:1805.11153.

[24] G. Pólya and G. L. Alexanderson. Gaussian binomial coefficients. *Elemente der Mathematik*, 26:102–109, 1971.

[25] G.-C. Rota. On the foundations of combinatorial theory I. Theory of Möbius functions. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 2(4):340–368, 1964.

[26] Z. Ryjáček. On a closure concept in claw-free graphs. *Journal of Combinatorial Theory, Series B*, 70(2):217–224, 1997.

[27] A. Selberg. On an elementary method in the theory of primes. *Det Kongelige Norske Videnskabers Selskabs Forhandlinger*, 19:64–67, 1947.

[28] F. B. Shepherd. Hamiltonicity in claw-free graphs. *Journal of Combinatorial Theory, Series B*, 53(2):173–194, 1991.

[29] P. Turán. On a theorem of Hardy and Ramanujan. *Journal of the London Mathematical Society*, s1-9(4):274–276, 1934.

[30] W. T. Tutte. A theorem on planar graphs. *Transactions of the American Mathematical Society*, 82(1):99–116, 1956.

[31] H. Whitney. A theorem on graphs. *Annals of Mathematics*, 32(2):378–390, 1931.

[32] R. J. Wilson. The Selberg sieve for a lattice. In *Combinatorial Theory and Its Applications*, volume 4 of *Colloquia Mathematica Societatis János Bolyai*, pages 1141–1149, Amsterdam, 1969. North-Holland.