Regeneration of Elliptic Chains with Exceptional Linear Series

Dissertation Advisor: Joseph Harris Nathan K. Pflueger

Regeneration of Elliptic Chains with Exceptional Linear Series

Abstract

We study two dimension estimates regarding linear series on algebraic curves. First, we generalize the classical Brill-Noether theorem to many cases where the Brill-Noether number is negative. Second, we extend results of Eisenbud, Harris, and Komeda on the existence of Weierstrass points with certain semigroups, by refining their dimension estimate in light of combinatorial considerations. Both results are proved by constructing chains of elliptic curves, joined at pairs of points differed by carefully chosen orders of torsion, and smoothing these chains. These arguments lead to several combinatorial problems of separate interest.

Contents

Acknowledgements

I am very grateful to Joe Harris for his close guidance. His ideas can be found in every part of this thesis. I also owe a great debt to Sam Payne, who has been very generous to me with his support and suggestions. This work also benefitted greatly from conversations with Nathan Kaplan and Melody Chan, especially in its combinatorial aspects.

The community of graduate students at Harvard has been my most important resource in graduate school. I am also grateful for the resources and support of the math department, as well as the NSF graduate research fellowship program.

I dedicate this thesis to the memory of my father, who, despite his lack of mathematical training, inspired my first interest and enthusiastically edited my first writing in mathematics.

Notation and conventions

Throughout this thesis, we work over an algebraically closed field K of arbitrary characteristic, except in section 4, where we will assume that K is characteristic 0. A *curve* will always mean a projective curve over K, with at worst nodes as singularities. For convenience, we collect here some of the main notation.

Symbol	Definition	
$\rho(g,r,d)$	The Brill-Noether number.	$p.3$
$\rho(g,S)$	The once-pointed Brill-Noether number.	$p.7$
$\rho(g,S:T)$	The twice-pointed Brill-Noether number.	$p.24$
$\mathrm{wt}(S)$	The weight of a numerical semigroup.	$p.14$
$\varepsilon(S)$	The effective weight of a numerical semigroup.	$p.20$
$\mathfrak{g}^r_d$	Linear series of rank r and degree d.	$p.2$
$\lvert\mathcal{L}\rvert$	The complete linear series of a line bundle.	$p.2$
$a_i^L(p)$	ith vanishing order of a linear series.	$p.5$
$G^r_d(C)$	Scheme of $\mathfrak{g}^r_d$s on C.	$p.2$
$\mathcal{G}^r_{d,g}$	Moduli of curves with a $\mathfrak{g}^r_d$, over $\mathcal{M}_g$.	$p.2$
$\mathcal{G}^S_g$	Moduli of pointed curves with a $\mathfrak{g}^r_d$ with specification ramification.	$p.7$
$\mathfrak{G}^S_d(f)$	Once-pointed limit linear series over a family f.	$p.27$
$\mathcal{W}_S$	Moduli of Weierstrass points.	$p.13$
	(if $0 \notin S$) moduli of twisted Weierstrass points.	$p.78$
$\mathcal{B}^{S:T}_g$	Moduli of twice-pointed linear series.	$p.23$
$\mathfrak{B}^{S:T}_d(f)$	Twice-pointed limit linear series over a family f.	$p.27$
$d(P)$	Displacement distance of a partition P.	$p.90$
$\delta(P)$	Displacement difficulty of a partition P.	$p.90$

A tilde over a symbol will always denote an open version of a given space (for example, $\widetilde{\mathcal{W}}_S$ denotes the pointed curves with Weierstrass semigroup exactly S, not the closure of this locus). The word "sep" in a superscript will indicate the subspace parametrizing only *separable* objects, e.g. $G_d^{r,\mathrm{sep}}(C)$ denotes the scheme of separable $\mathfrak{g}_d^r$s on a curve C. This distinction does not exist in characteristic 0.

The word *stack* will always refer to a Deligne-Mumford stack. A *point* will mean a closed (geometric) point unless otherwise stated.

1. Introduction

The roots of our main results are classical. In the following two subsections, we summarize the background material and previous work. We also state our main results, theorems A, B and C.

This rest of this thesis is organized as follows. Section 2 states and proves the main lemma about regeneration of elliptic chains used to prove the main result. Sections 3 and 4 concern the existence and deformations of Weierstrass points: section 3 proves theorems B and C, while section 4 is largely speculative and explores extremal cases to which these theorems cannot be extended. This leads to some purely combinatorial conjectures. Section 5 gives the proof of theorem A, by introducing the notion of twisted Weierstrass points. The proof requires some combinatorial analysis, which is deferred to section 6. Finally, section 6 provides the necessary combinatorial analysis for theorem A, as well as exploring several cases of a combinatorial problem about partitions that naturally arises from it. Stronger results about this combinatorial problem would improve theorem A, and also give generalizations in the context of pointed curves.

1.1. **Brill-Noether theory.** We will begin with a brief introduction to Brill-Noether theory. This section will also serve to introduce notation that will be used throughout the thesis.

Brill-Noether theory is concerned with curves in projective space. There are two typical ways to study these curves. The most "extrinsic" approach is to study the Hilbert scheme. The second approach is more intrinsic: one studies an abstract algebraic curve, together with a choice of map to projective space.

The basic objects of study in Brill-Noether theory are *linear series*. A linear series is characterized by two integers, traditionally denoted r (called the *rank*) and d (called the *degree*). Given a smooth curve C, any two nonnegative integers d and r give rise to the following set.

$$\begin{aligned} G^r_d(C) &= \{(\mathcal{L}, V) : \ \mathcal{L} \text{ a degree } d \text{ line bundle,} \\ & \qquad V \subseteq H^0(\mathcal{L}) \text{ an } (r+1)\text{-dimensional subspace}\} \end{aligned}$$

A map $f : C \to \mathbf{P}^r$ has a naturally associated linear series, namely $(f^*\mathcal{O}_{\mathbf{P}^r}(1), V)$, where V is the image of the induced map $H^0(\mathcal{O}_{\mathbf{P}^r}(1)) \to H^0(f^*\mathcal{O}_{\mathbf{P}^r}(1))$. This linear series uniquely determines f up to automorphisms of $\mathbf{P}^r$. Not all linear series arise in this was, however; those that do arise are said to be base point free. Although it is tempting to view linear series with base points as pathological, we will see the possibility of base points in linear series is a feature, not a bug, as it makes many useful constructions possible.

Any line bundle determines a linear series $|\mathcal{L}|$, given by $(\mathcal{L}, H^0(\mathcal{L}))$. This is called the complete linear series of $\mathcal{L}$. A curve in projective space which is embedded by a complete linear series to usually called *linearly normal.*

Classically, an element of this set is called a $\mathfrak{g}^r_d$ on C. This set also has a natural scheme structure (see [ACGH]). This scheme structure globalizes nicely to a Deligne-Mumford stack $\mathcal{G}^r_{d,g} \to \mathcal{M}_g$, whose geometric points correspond to triples $(C, \mathcal{L}, V)$, where C is a genus g curve, $\mathcal{L} \in \mathrm{Pic}^d(C)$, and $V \subseteq H^0(\mathcal{L})$ has dimension $(r+1)$ (details can be found in [ACG]). We will also make use of the following open subset.

$$\widetilde{G}^r_d(C) = \{(\mathcal{L}, V) \in G^r_d(C) : \ V = H^0(\mathcal{L})\}$$

These are called *complete* linear series can be identified with line bundles, hence $\widetilde{G}^r_d(C)$ may be regarded as a locally closed subscheme of $\mathrm{Pic}^d(C)$.

Brill and Noether [BN] estimated the dimension of the scheme $G^r_d(C)$, and arrived at the following number, now called the *Brill-Noether number.* We will see one derivation of this estimate (and a proof that it is a local lower bound) in lemma 1.9.

$$\rho(g,d,r) = g - (r+1)(g-d+r)$$

This number is a lower bound for $\dim G^r_d(C)$, when $G^r_d(C)$ is nonempty. The estimate is sharp for general curves of genus g, as the following theorem shows.

Theorem 1.1 (Brill-Noether-Petri theorem)**.** *Suppose that g, r, d are fixed integers.*

- *(Existence) If $\rho(g,d,r) \geq 0$, then $G^r_d(C)$ is nonempty, and all components have dimension at least $\rho(g,d,r)$.*
- *(Non-existence) For C general, $G^r_d(C)$ has no components of dimension greater than $\rho(g,d,r)$.*
- *(Irreducibility) For C general, $G^r_d(C)$ is irreducible.*
- *(Smoothness) For C general, $G^r_d(C)$ is smooth.*

Despite being named for its classical conjecturers, this theorem is an amalgamation of several results from the 1970s and early 1980s. Brill and Noether proved the dimension bound in the existence theorem, assuming the (not yet proved) first part of the existence theorem. The first part of the existence theorem was proved independently by Kempf [K71] and Kleiman-Laksov [KL72, KL74]. The non-existence theorem was proved by Griffiths and Harris [GH80] (see also a recent combinatorial proof [CDPR] using tropical curves). The smoothness theorem was conjectured by Petri [P24] (who asserted in an offhand remark that it is "known to hold") and proved by Gieseker [G82] (a shorter proof, by theory of limit linear series, can be found in [EH83]). Fulton and Lazarsfeld [FL81] proved that for C general, $G^r_d(C)$ is connected, which, together with smoothness, implies the irreducibility theorem. More recently, Osserman [Oss11] gave a short proof of the existence and non-existence portions of this theorem using limit linear series. The papers mentioned so far study degenerations to singular curves; a non-degenerative proof of theorem 1.1 was later provided by Lazarsfeld [L86].

We will use the following terms to describe linear series, in terms of how unusual they are from the standpoint of Brill-Noether theory.

- A linear series is *non-special* if $\rho(g, d, r) = g$.
- A linear series is *special* if $\rho(g, d, r) < g$.
- A linear series is *exceptional* if $\rho(g, d, r) < 0$.

The Brill-Noether theorem asserts the existence (in great abundance) of special linear series; one of the main results of this thesis is the construction of exceptional linear series. More specifically, we will study exceptional linear series where the Brill-Noether dimension estimate is sharp. To preface our new result, observe that the $\rho \geq 0$ case of the Brill-Noether theorem can be re-stated in a global way as follows.

Definition 1.2. A geometric point $(C, L) \in \mathcal{G}^r_{d,g}$ is called *dimensionally proper* if the local dimension of $\mathcal{G}^r_{d,g}$ at this point is exactly $\dim \mathcal{M}_g + \rho(g, d, r)$. An irreducible component of $\mathcal{G}^r_{d,g}$ is called dimensionally proper if it has dimension $\dim \mathcal{M}_g + \rho(g, d, r)$.

Corollary 1.3 (Restatement of theorem 1.1 when $\rho \geq 0$)**.** *Suppose that g, d, r are fixed integers such that $\rho(g, d, r) \geq 0$.*

- *(Existence) $\mathcal{G}^r_{d,g}$ has an irreducible component that dominates $\mathcal{M}_g$.*
- *(Non-existence) No dimensionally improper component of $\mathcal{G}^r_{d,g}$ dominates $\mathcal{M}_g$.*
- *(Irreducibility) The component dominating $\mathcal{M}_g$ is unique.*
- *(Smoothness) The component dominating $\mathcal{M}_g$ has smooth general fiber.*

When $\rho < 0$, there can no longer be components dominating $\mathcal{M}_g$, so it is no longer sufficient to study general curves. Instead, we work globally, and construct dimensionally proper components of $\mathcal{G}^r_{d,g}$.

Theorem A. *Suppose that g, d, r are positive integers with $0 > \rho(g, d, r) \geq -\frac{r}{r+2}g + 3r - 3$. If $r = 1$ or $g - d + r = 1$, then $\mathcal{G}^r_{d,g}$ is empty. Otherwise $\mathcal{G}^r_{d,g}$ has an irreducible component of dimension $\dim \mathcal{M}_g + \rho$, whose image in $\mathcal{M}_g$ has codimension equal to $-\rho$, and whose general member is complete and separable.*

The lower bound on ρ here is almost certainly far from sharp. Indeed, it could be improved by a purely combinatorial analysis, as we will see is section 6. This result was proved in the case $\rho = -1$ be Eisenbud and Harris [EH89], and a slightly different statement was proved by Edidin [E93] in the case $\rho = -2$. Similar results (with different sorts of bounds) have been obtained by Sernesi [S84], Pareschi [P89] and Lopez [L91, L99] by a different method.

Actually, theorem A will be deduced from a stronger result (corollary 5.17) about *pointed* curves. For the rest of this subsection we will discuss the generalizations of the main notions of Brill-Noether theory to this setting. This more general setting allows more versatile arguments by induction, and is essential for studying limits at the boundary of $\mathcal{M}_g$. As a first demonstration of the utility of this perspective, we will derive the Brill-Noether dimension estimate and prove that it is a local lower bound on the dimension of $G^r_d(C)$.

Given a linear series $L = (\mathcal{L}, V)$ on C, and a marked point $p \in C$, the *vanishing orders* are the $r + 1$ nonnegative integers a such that

$$V \cap H^0(\mathcal{L}(-ap)) \neq V \cap H^0(\mathcal{L}(-(a+1)p)).$$

These numbers will be denoted in increasing order as $a^L_0(p), a^L_1(p), \cdots, a^L_r(p)$. We will almost always work with *separable* linear series, defined as follows.

Definition 1.4. A point p is called a *ramification point* for L if $a^L_r(p) > r$. A linear series L is *separable* if not every point $p \in C$ is a ramification point for L. For any moduli stack of linear series, a superscript "sep" indicates the substack where the linear series are separable, e.g. $G^{r,\mathrm{sep}}_d(C)$ denotes the moduli scheme of separable $\mathfrak{g}^r_d$s on C.

In characteristic 0, all linear series are separable. In characteristic p, separable linear series form an open substack of $\mathcal{G}^r_{d,g}$ (see [Oss06]).

For notational reasons (see the remark below) we will find it convenient in this thesis to consider the equivalent sequence $\{d - a_i^L(p)\}$. Therefore we introduce the following notation.

Definition 1.5. For a smooth pointed curve (C, p) of genus g and finite set S, let $\widetilde{G}_d^S(C, p)$ denote the subscheme of $G_d^{|S|-1}(C)$ parameterizing linear series L such that the vanishing orders of L at p are precisely $(d - S)$. Let $G_d^S(C, p)$ denote the scheme parameterizing linear series (separable or otherwise) with vanishing orders at least $(d - S)$.

Clearly $\widetilde{G}_d^S(C)$ is an open subscheme of $G_d^S(C)$.

Remark 1.6. The reason that it will be particularly convenient to denote the set S rather than the set of vanishing orders is that it is well-behaved under the addition of base points. More precisely, the map

$$\begin{array}{rccc} \iota: & G_d^r(C) & \hookrightarrow & G_{d+1}^r(C) \\ & (\mathcal{L}, V) & \mapsto & (\mathcal{L}(p), V) \end{array}$$

(by abuse of notation, sections of $\mathcal{L}$ may be identified with sections of $\mathcal{L}(p)$ which vanish at p) induces isomorphisms $G_d^S(C, p) \xrightarrow{\sim} G_{d+1}^S(C, p)$ and $\widetilde{G}_d^S(C, p) \xrightarrow{\sim} \widetilde{G}_{d+1}^S(C, p)$. These isomorphisms also preserve the respective separable loci. So in fact the number d is redundant, and we will often omit it.

Definition 1.7. The notation $G^S(C, p)$ denotes $G_d^S(C, p)$, where d is any integer greater than or equal to the largest element of S. The different possible choices of d are identified by the isomorphism described in the remark above.

The following example shows how the study of the schemes $G^S(C,p)$ subsumes the study of the schemes $G^r_d(C)$ from the non-pointed case. Therefore we lose nothing by focusing attention on the pointed situation.

Example 1.8. Fix positive integers g, r, d and a curve C. Let S be the set $\{d-r, d-r+1, \cdots, d\}$. Then for any point $p \in C$,

$$G^S(C,p) \cong G^r_d(C).$$

Under this isomorphism, $\widetilde{G}^S(C,p)$ consists of those $\mathfrak{g}^r_d$s which are unramified at p.

Assume now that $g-d+r \geq 1$ and $r \leq g-1$. Let T be the set $S \cup \{g+r+1, g+r+2, \cdots, 2g-1\}$. Then

$$G^T(C,p) \cong W^r_d(C),$$

where $W^r_d(C) \subseteq \mathrm{Pic}^d(C)$ is the locus of degree d line bundles $\mathcal{L}$ with $h^0(\mathcal{L}) \geq r+1$. Under this isomorphism, $\widetilde{G}^T(C,p)$ consists of those line bundles $\mathcal{L}$ for which p is not ramified in either $|\mathcal{L}|$ or $|\omega_C \otimes \mathcal{L}^\wedge|$. $\lhd$

The second case in the example above is the basis of our analysis in section 5.

This construction can also be relativized. The resulting stacks will be denoted $\mathcal{G}^S_g \to \mathcal{M}_{g,1}$ and $\widetilde{\mathcal{G}}^S_g \to \mathcal{M}_{g,1}$.

The Brill-Noether number has the following analog for pointed curves. Here, the elements of S are denoted $s_0 < s_1 < \cdots < s_{|S|-1}$.

$$\rho(g,S) := g + \sum_{i=0}^{|S|-1} (s_i - i - g)$$

To demonstrate the techniques we will use extensively later, we will give a proof that this gives a bound on the dimension of $\mathcal{G}^S_g$.

Lemma 1.9. *Let $\pi : \mathcal{C} \to B$ be a flat family of smooth curves of genus g over a scheme B, with a section $s : B \to \mathcal{C}$, and let $\mathcal{G}^S_g(\mathcal{C}, s) \to B$ denote the pullback of $\mathcal{G}^S_g$*

to B. Then the following inequality holds locally at every point.

$$\dim \mathcal{G}^S_d(\mathcal{C}) \geq \rho(S,g) + \dim B$$

Proof. We will describe $\mathcal{G}^S_g(\mathcal{C},s)$ as the pullback of a Schubert cell from a certain Grassmannian bundle over the relative Picard scheme of $\mathcal{C}$. The dimension estimate will follow from a computation of the dimension of this Schubert cell.

Fix an integer $N \geq 2g-1$ (so that all line bundles of degree N are nonspecial). Let $r = |S|-1$. Let $\mathcal{P}_N \to B$ denote the relative Picard scheme of degree N line bundles on $\mathcal{C}$, and let $\mathcal{M}$ denote a relative Poincaré line bundle on $\mathcal{P}_N \times_B \mathcal{C}$. Let $\pi : \mathcal{P}_N \times_B \mathcal{C} \to \mathcal{P}_N$ be the projection, and let $\mathcal{S}$ denote the pushforward $\pi_* \mathcal{M}$. Notice that the fiber of $\mathcal{S}$ over a point $\mathcal{L}$ of $\mathcal{P}_N$ is naturally identified with the space of global sections of $\mathcal{L}$. In particular, ever fiber is a vector space of dimension $N+1-g$ (since N was chosen large enough that $\mathcal{L}$ is guaranteed to be non-special).

By the theorem on cohomology and base change ([H77] theorem 12.11), $\mathcal{S}$ is a vector bundle of rank $N+1-g$. Let $P \subseteq \mathcal{C}$ be the image of the section s, and define the vector bundle $\mathcal{T} \to \mathcal{P}_N$ of "Taylor expansions" of sections at P, given by $\mathcal{T} = \pi_*(\mathcal{M}/\mathcal{M}(-(N+1)P))$. There is a natural map of vector bundles $t : \mathcal{S} \to \mathcal{T}$. The map t must in fact be an injection, since a section of a degree N line bundle on a curve vanishing to order $N+1$ at a point must vanish entirely. Therefore t induces a map τ on Grassmannian bundles:

$$\tau : \ \mathrm{Gr}(r+1, \mathcal{S}) \to \mathrm{Gr}(r+1, \mathcal{T}).$$

The vector bundle $\mathcal{T}$ has a natural filtration by sub bundles $\{0\} = \mathcal{T}_{-1} \subset \mathcal{T}_0 \subset \cdots \subset \mathcal{T}_N = \mathcal{T}$, where $\mathcal{T}_n = \pi_*(\mathcal{M}(-(N-n)P)/\mathcal{M}(-(N+1)P))$. This filtration defines Schubert cells $\Sigma(a_0, \cdots a_r) \subseteq \mathrm{Gr}(r+1, \mathcal{T})$ for any sequence $0 \leq a_0 < a_1 < \cdots < a_r \leq N$, defined by the condition that a section of $\Sigma(a_0, \cdots, a_r)$ over $\mathcal{P}_N$ (regarded as a sub-bundle of $\mathcal{T}$) must meet $\mathcal{T}_{a_n}$ in a sub-bundle of rank at least $n+1$, for all $n = 0, 1, \cdots, r$ (this index convention is slightly non-standard, but it will be convenient for our purposes). It is a standard fact that

$$\operatorname{codim}\left(\Sigma(a_0, \cdots, a_r) \subseteq \operatorname{Gr}(r+1, \mathcal{T})\right) = \sum_{n=0}^{r}(N - r + n - a_n).$$

Observe that $\mathcal{G}^S_{N,g}(\mathcal{C})$ is isomorphic to (indeed, the scheme structure can be defined to be) the inverse image of $\Sigma(s_0, s_1, \cdots, s_r)$ under τ. Therefore we conclude that:

$$\begin{aligned}
\dim \mathcal{G}^S_{N,g}(\mathcal{C}) &\geq \dim \operatorname{Gr}(r+1, \mathcal{M}) - \sum_{n=0}^{r}(N - r + n - s_n) \\
&= \dim \mathcal{P}_N + (r+1)(N - g - r) - \sum_{n=0}^{r}(N - r + n - s_n) \\
&= \dim B + g + \sum_{n=0}^{r}\left[(N - g - r) - (N - r + n - s_n)\right] \\
&= \dim B + g + \sum_{n=0}^{r}(s_n - n - g) \\
&= \dim B + \rho(g, S).
\end{aligned}$$

□

Throughout this thesis, we will be particularly interested in cases where this dimension estimate is sharp. Such linear series are particularly useful in smoothing arguments.

Definition 1.10. A geometric point (C, p, L) of $\mathcal{G}^S_g$ is called *dimensionally proper* if the local dimension of $\mathcal{G}^S_g$ at this point is exactly $\rho(g, S) + \dim \mathcal{M}_{g,1}$.

The Brill-Noether theorem has a several extensions for pointed curves, based on this notion of "dimensionally proper." One version is the following. A more general version (for arbitrarily many marked points) is true in characteristic 0 only, and was proved by Eisenbud and Harris ([EH86], theorem 4.5 and the subsequent remark).

In arbitrary characteristic, a version for up to two marked points (implying this statement) was proved by Osserman [Oss11].

Theorem 1.11 (Eisenbud,Harris,Osserman)**.** *Suppose that* g *is a positive integer,* S *is a finite set of nonnegative integers such that* $\max S \leq |S| + g - 1$*, and* $d \geq \max S$*. Then:*

- *(Existence) If* $\rho(g, S) \geq 0$*, then* $\mathcal{G}^S_{d,g}$ *has a component dominating* $\mathcal{M}_{g,1}$*.*
- *(Dimension) If* $\rho(g, S) \geq 0$*, at least one such component dominating* $\mathcal{M}_{g,1}$ *is dimensionally proper.*
- *(Non-existence) No dimensionally improper dominates* $\mathcal{M}_{g,1}$ *(in particular, if* $\rho(g < S) < 0$ *then no component dominates* $\mathcal{M}_{g,1}$*).*

Remark 1.12. By letting S be the sequence $\{d-r, d-r+1, \cdots, d-1, d\}$ and forgetting the marked point, theorem 1.11 implies most of the Brill-Noether theorem. It has nothing to say about irreducibility or smoothness, however.

Remark 1.13. The hypothesis $\max S \leq |S| + g - 1$ simply means that points of $\mathcal{G}^S_{d,g}$ have the possibility of being complete. It can be weakened, but not removed entirely; for example, if S contains both 0 and 1 and one extremely large element, it could occur that $\rho(g, S) \geq 0$ even though such a linear series is impossible on a curve of positive genus.

We will generalize this theorem as well (corollary 5.17); the statement requires combinatorial terminology that we introduce in section 5. The generalized form, together with some combinatorial analysis, will give theorem A.

1.2. **Weierstrass points.** The second main result of this thesis concerns the existence and deformations of Weierstrass points. We now briefly survey the background on this subject.

Throughout this thesis, a *numerical semigroup* is a set S of nonnegative integers such that $S + S = S$ (in particular, $0 \in S$) and $\mathbf{N}\backslash S$ is finite. The size of the complement will be called the *genus* of S (for reasons that will be clear shortly).

Weierstrass points are reflect a curious feature of algebraic curves: although every point of a smooth curve is (étale-)locally isomorphic, there are a finite number of distinguished points, called Weierstrass points. The fact that these points are distinguished makes them helpful props in proving that algebraic curves of genus $g \geq 2$ have finitely many automorphisms. For other background and early applications, see the expository article [dC].

Definition 1.14. Given a geometric point p on a smooth curve C of genus g, the *Weierstrass semigroup* is the set $S(C, p) = \{n : \ h^0(np) > h^0((n-1)p)\}$. The point p is called a *Weierstrass point* if $S(C, p) \neq \{0, g+1, g+2, \cdots\}$.

In other words, the Weierstrass semigroup is the set of pole orders at p of regular functions on $C\backslash\{p\}$. Because functions can be multiplied together, $S(C, p)$ is a numerical semigroup.

Remark 1.15. Definition 1.14 is slightly non-standard in characteristic p. In particular, it is possible for a curve over a field of positive characteristic to consist entirely of Weierstrass points (by this definition). Other authors define the Weierstrass points of such a curve to be those points whose Weierstrass semigroup differs from that of a general point on the curve. However, we will use the definition above in this thesis, since we are mainly interested in characteristic 0 anyway, and we will be working globally in $\mathcal{M}_{g,1}$, rather than focusing attention on a fixed curve.

Example 1.16. Let C be a hyperelliptic curve of genus $g \geq 2$. Then the double cover $\pi : C \to \mathbf{P}^1$ has $2g+2$ ramification points, by Riemann-Hurwitz. For each such point p, it follows that $h^0(2p) = 1$, hence $2 \in S(C, p)$. Since $S(C, p)$ is a semigroup, it must contain all positive even numbers. On the other hand, it can contain no odd numbers less than $2g$ (this follows from Clifford's inequality), and must contain all integers greater than $2g$ (by Riemann-Roch). It follows that at these ramification points,

$$S(C,p) = \{0, 2, 4, \cdots, 2g-2, 2g, 2g+1, 2g+2, \cdots\}.$$

This semigroup is called the *hyperelliptic semigroup of genus g.* As long as the canonical series is separable (for example, in characteristic 0), there are no other Weierstrass points (this can be shown, for example, using the Plücker formula 1.1). So a curve is hyperelliptic if and only if it has a point with the hyperelliptic semigroup, and once one Weierstrass point has this semigroup, all Weierstrass points must have the same semigroup. $\triangleleft$

Example 1.17. Let C be a non-hyperelliptic curve of genus 3. Then the canonical bundle embeds C as a smooth quartic curve in $\mathbf{P}^2$. For simplicity, consider the case where the canonical series is separable (or work in characteristic 0). It follows from the geometric Riemann-Roch formula that the Weierstrass points of C are precisely the inflection points of C in this embedding. For C general, there will be 24 ordinary flex points (where the tangent line meets with contact of order 3), which have Weierstrass semigroup $\mathbf{N}\backslash\{1,2,4\}$. Some curves C may instead have some number of hyperflexes, where the tangent line has contact of order 4. These have Weierstrass semigroup $\mathbf{N}\backslash\{1,2,5\}$. $\triangleleft$

An easy fact (which follows quickly from the Riemann-Roch formula) is that the complement of the Weierstrass semigroup always has exactly g elements, often called the *Weierstrass gaps.*

$$|\mathbf{N}\backslash S(C,p)| = g$$

In other words: the genus of the Weierstrass semigroup is equal to the genus of the curve[1]. Another fact, directly following from Riemann-Roch, is that $S(C,p) \supset \{2g, 2g+1, \cdots\}$. So S is determined by its intersection with $\{0,1,2,\cdots,2g-1\}$; it contains exactly half of the elements of this set (in fact an easy exercise in combinatorics is to prove that that the largest gap of any numerical semigroup of genus g is at most $2g-1$; the equality cases are called *symmetric semigroups* and are interesting special cases; see [S93, CS13] and [B13]).

A basic question is the classification of Weierstrass semigroups.

Question 1.18. *For which numerical semigroups S of genus g does there exist a pointed curve (C,p) with Weierstrass semigroup S?*

Theorem C makes substantial progress on this question, but most cases remain open. The only large family of semigroups such that $\widetilde{\mathcal{W}}_S = \emptyset$ was provided by Buchweitz [B80], but Kaplan and Ye [KY13] have shown that these counterexamples occupy a proportion 0 subset (asymptotically in the genus) of all numerical semigroups. On the other hand, Kaplan and Ye also proved that the families of semigroups such that $\widetilde{\mathcal{W}}_S$ is known to be nonempty also occupy a proportion 0 subset of all numerical semigroups.

Secondly, we will consider the *abundance* of particular semigroups, i.e. the dimension of the corresponding locus in moduli.

Definition 1.19. Denote be $\widetilde{\mathcal{W}}_S$ the (open) substack of $\mathcal{M}_{g,1}$ consisting[2] of pointed curves (C,p) such that $S(C,p) = S$.

[1] This appears to the be origin of the word "genus" in the context of numerical semigroups, although the author has heard the joke that the reason the number of gaps of a semigroup is called the genus is because it is "the number of holes" in the semigroup.

[2] A technical remark: this definition only states what the geometric points of $\widetilde{\mathcal{W}}_S$ should be. To define the stack structure, simply construct an appropriate map of vector bundles and pull back the appropriate (open) Schubert cell.

Question 1.20. *How many irreducible components does $\widetilde{\mathcal{W}}_S$ have? What are their dimensions?*

One easy dimension estimate comes from Brill-Noether theory. Notice that there is an isomorphism to a pointed Brill-Noether variety as follows.

$$\begin{array}{rcl} \widetilde{\mathcal{W}}_S & \xrightarrow{\sim} & \widetilde{\mathcal{G}}^{S\cap\{0,1,\cdots,2g-1\}}_{2g-1,g} \\ (C,p) & \mapsto & (C,p,|\mathcal{O}_C((2g-1)p)|) \end{array}$$

Therefore it follows from lemma 1.9 that the codimension of any component of $\widetilde{\mathcal{W}}_S$ is bounded by the following number.

$$\begin{aligned} -\rho(g, S\cap\{0,1,\cdots,2g-1\}) &= -g + \sum_{i=0}^{g-1}(g+i-s_i) \\ &= \sum_{i=1}^{g-1}(g+i-s_i) \\ &= \sum_{i=1}^{g-1}(\text{number of gaps larger than } s_i) \\ &= |\{(s,t) : \ s\in S, \ t\notin S, \ 0<s<t\}| \end{aligned}$$

This quantity is called the *weight* of S, and is denoted $\mathrm{wt}(S)$. Notice that if p is not a Weierstrass point, then all the positive gaps come before all of the positive semigroup elements, so the weight is 0. The weight is important partly because it serves as an appropriate notion of *multiplicity* for Weierstrass points. Indeed, in characteristic 0 we have the following Plücker formula (see [ACGH], exercise E-8 for a derivation). Here K_C is a canonical divisor on C.

$$\sum_{p\in C} \mathrm{wt}\, S(C,p)\cdot p \sim \binom{g+1}{2} K_C \tag{1.1}$$

In particular, $\sum_{p\in C} \mathrm{wt}\, S(C,p) = g^3 - g$. This is one way to show that (in characteristic 0) there are finitely many Weierstrass points on a given curve. In characteristic p, we have an inequality $\sum_{p\in C} \mathrm{wt}\, S(C,p) \leq g^3 - g$, as long as the canonical series is separable (see [Oss06], proposition 2.4).

A Weierstrass point is called *dimensionally proper* if the corresponding point in $\mathcal{G}_{2g-1,g}^{S\cap\{0,1,\cdots,2g-1\}}$ is dimensionally proper; in other words, if the local codimension of $\widetilde{\mathcal{W}}_S$ in $\mathcal{M}_{g,1}$ is exactly $\mathrm{wt}(S)$. Semigroups of low weight very frequently belong only to dimensionally proper points. However, the weight is not always equal to the codimension.

Example 1.21. Let S be the hyperelliptic semigroup of genus g. Then $\widetilde{\mathcal{W}}_S \to \mathcal{M}_g$ is an étale map of degree $2g+2$, whose image is the hyperelliptic locus. The hyperelliptic locus has codimension $g-2$ in $\mathcal{M}_g$. Therefore the codimension of $\widetilde{\mathcal{W}}_S$ in $\mathcal{M}_{g,1}$ is $g-1$. However, $\mathrm{wt}(S) = \binom{g}{2}$. Therefore if $g \geq 3$, $\widetilde{\mathcal{W}}_S$ has no dimensionally proper points. We will see that hyperelliptic Weierstrass points do satisfy a different, more broadly applicable condition: they are *effectively proper*. $\triangleleft$

To illustrate questions 1.18 and 1.20, we will enumerate all numerical semigroups of genus up to 4 and describe the corresponding locus in moduli. In these genera, the answer to question 1.18 is that all semigroups S correspond to nonempty loci $\widetilde{\mathcal{W}}_S$, and the answer to question 1.20 is that these loci are irreducible but not always dimensionally proper. The result is summarized in figure 1. This tree organizes all numerical semigroups of genus up to 4, by assigning to every semigroup its "parent," obtained by filling in the last gap of the semigroup.

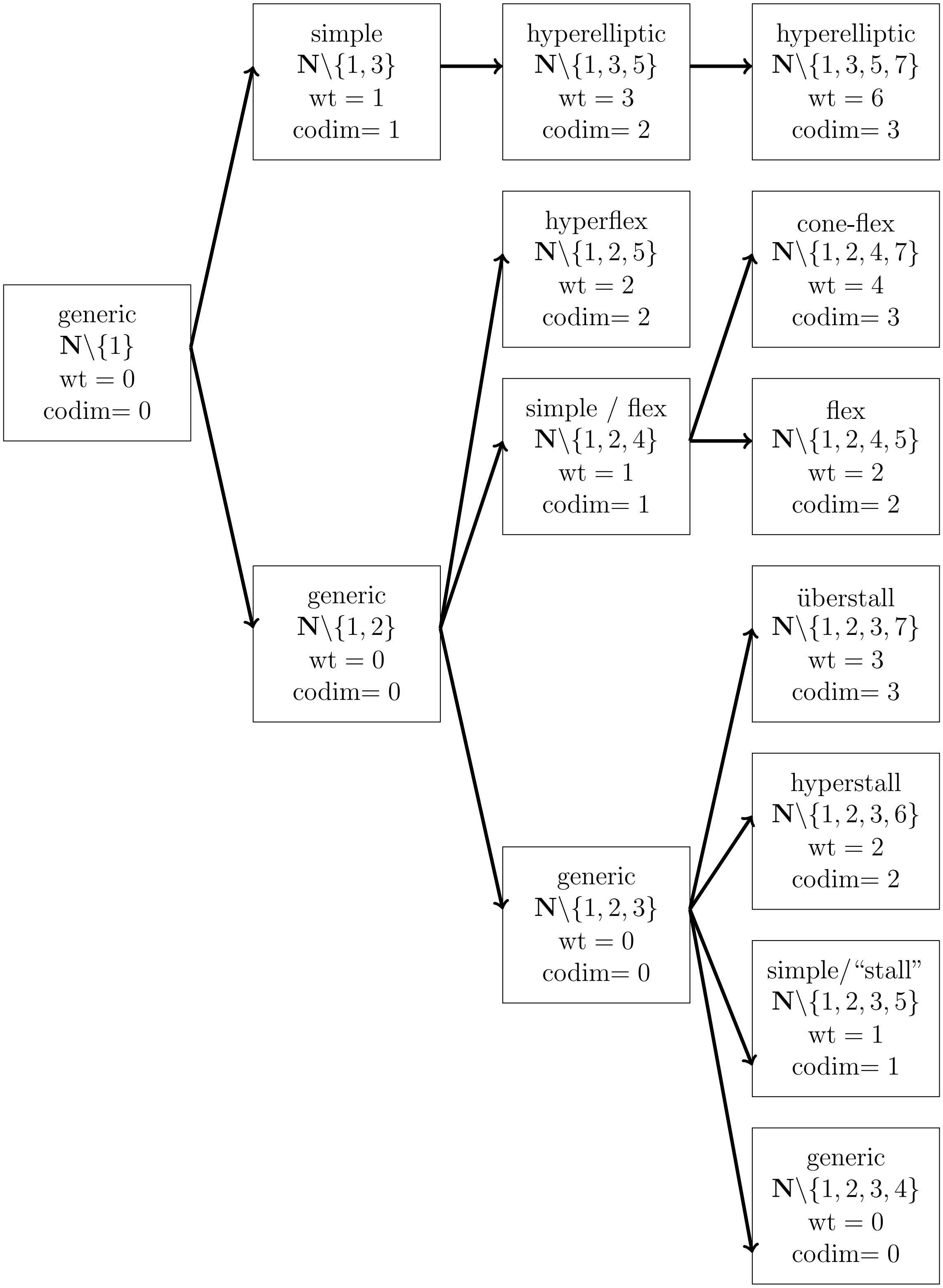


FIGURE 1. Semigroups of genus 1 through 4.

Remark 1.22. It is an easy combinatorial exercise to show that a semigroup S of genus g has exactly one child (of genus $g+1$) in this tree for each "effective generator," defined to be a generator of S (that is, a positive element that is not the sum of two other positive elements) that is larger than all gaps of S. One can attempt to enumerate numerical semigroups of genus g by studying the branching of this tree, i.e. by studying how the number of effective generators changes with each descent in the tree. Bras-Amorós [BA08] observed that this method might be used to determine the asymptotic growth rate of the number of numerical semigroups of genus g, which was successfully carried out by Zhai [Z13]. In fact, the number of numerical semigroups of genus g is asymptotic to a constant times the gth Fibonacci number.

In the casework below, we will assume that we are working in characteristic 0, to simplify special cases. We also remark that not all of these names are standard.

- **The generic semigroups**. These form the "spine" of the semigroup tree; they are semigroups of the form $\mathbf{N}\backslash\{1, 2, \cdots, g\}$. For S generic, $\widetilde{\mathcal{W}_S}$ is a dense open subset of $\mathcal{M}_{g,1}$, consisting of the non-Weierstrass points.
- **The hyperelliptic semigroups.** For every genus, there is a hyperelliptic semigroup $\mathbf{N}\backslash\{1, 3, \cdots, 2g-1\}$, which can exist only on hyperelliptic curves. Its weight is $\binom{g}{2}$, but the codimension of $\widetilde{\mathcal{W}_S}$ in $\mathcal{M}_{g,1}$ is $g-1$, by an easy calculation of the moduli of double-covers of $\mathbf{P}^1$. Therefore for all $g \geq 3$, the weight is not equal to the codimension.
- **The simple semigroups.** For every genus, there is a unique weight-1 semigroup $\mathbf{N}\backslash\{1, 2, \cdots, g-1, g+1\}$. This semigroup exists is codimension 1 in $\mathcal{M}_{g,1}$, and a general curve of genus g has only simple Weierstrass points.
- **Genus 3: flexes and hyperflexes.** In genus 3, non-hyperelliptic curves can have only two types of Weierstrass points, with semigroups $\mathbf{N}\backslash\{1, 2, 4\}$ or $\mathbf{N}\backslash\{1, 2, 5\}$. These have a nice geometric interpretation. If C is not hyperelliptic, its canonical series embeds it as a smooth plane quartic curve. By

Riemann-Roch, the number γ is a gap of the Weierstrass semigroup if and only if $\gamma - 1$ is the order of vanishing of a global 1-form, i.e. if and only if some line in $\mathbf{P}^2$ meets C to order $\gamma - 1$. This shows that $\mathbf{N}\backslash\{1,2,4\}$ (which is also the genus 3 "simple" semigroup) corresponds to points which have lines meeting them to orders $0, 1$ and 3, which are called flex points. On the other hand, those with semigroup $\mathbf{N}\backslash\{1,2,5\}$ have lines meeting them to orders $0, 1$, and 4; these are called *hyperflexes.* The Plücker formula implies that, if hyperflexes count as two flexes, there are 24 total flexes on a plane quartic (and there is a sextic curve, the Wronskian, which intersects C transversely at flexes and meets it to order 2 at hyperflexes). For both these semigroups, the weight is equal to the codimension, as some elementary dimension calculations in the plane easily show.

- **Genus 4: stalls, hyperstalls, and überstalls.** If a genus 4 curve C is not hyperelliptic, then its canonical model is a degree 6 curve in $\mathbf{P}^3$. The complement of the Weierstrass semigroup at $p \in C$ is given by the contact orders of hyperplanes with the curve at p, incremented by 1. Classically, a point $p \in C$ is called a *stall* if it is not an inflection point (i.e. the tangent line meets to order 2), but there is a plane that meets the curve to order 4. This plane is the osculating plane to C at p; the terminology presumably comes from the fact that (over $\mathbf{R}$) the osculating plane is the unique plane containing the tangent line such that the curve crosses from one side to the other (like an airplane gently ascending), while at a stall the curve meets the osculating plane to high order and then comes back down on the same side (its flight has "stalled"). An ordinary stall corresponds to the "simple" genus 4 semigroup $\mathbf{N}\backslash\{1,2,3,5\}$. There are two higher types of stalls, where the point still is not a flex point, but the osculating plane meets to even higher order: these correspond to the *hyperstall* semigroup $\mathbf{N}\backslash\{1,2,3,6\}$ and the *überstall* semigroup $\mathbf{N}\backslash\{1,2,3,7\}$ (the author has invented this word, as far as he knows). These three types of stalls have weight $1,2,3$ respectively, and in all cases the weight is equal to the codimension in $\mathcal{M}_{g,1}$.

- **Genus 4: flexes and cone-flexes.** There remain only two semigroups in genus 4, namely those that contain the number 3. These are $\mathbf{N}\backslash\{1,2,4,5\}$ and $\mathbf{N}\backslash\{1,2,4,7\}$, which we shall call the *flex* and *cone-flex* semigroups (for reasons to be explained). A curious feature, which we have not yet seen in non-hyperelliptic semigroups, is that these semigroups are consecutive children of a common parent in the semigroup tree, yet their weights differ by 2 rather than by 1. The reason for this is that their missing middle sibling $\mathbf{N}\backslash\{1,2,4,6\}$ is not a semigroup: 6 cannot be a gap once 3 is in the semigroup. In fact, this missing sibling makes itself felt in a geometric fashion as well. Once the gaps $\{1,2,4\}$ are fixed, the point $p \in C$ is necessarily an *inflection point*, meaning that its tangent line meets it to order 3. Now we use the geometry of canonical curves in $\mathbf{P}^3$: an easy dimension count shows that there is exactly one quadric surface containing C. For C to be inflected at p, the tangent line must lie entirely within the quadric: it must be a ruling line. Now, one can consider the tangent plane to the quadric at p; if the quadric is smooth, then this tangent plane meets the quadric as a union of two lines (one from each ruling); one must be the tangent line to C at p, and the other must be transverse; hence this plane meets C at p to order 4, and the Weierstrass semigroup must be the flex semigroup. On the other hand, if the quadric is singular, then the tangent plane meets it as the ruling line to multiplicity 2, hence meets the curve C to order 6 at p. As a result, an inflection point of C has the flex semigroup if the quadric is smooth, and the cone-flex semigroup if the quadric is in fact a quadric cone (hence our terminology). Therefore a single equation on the moduli of C (namely, the condition that the unique quadric containing it is singular, i.e. a quadric cone) forces the weight of the Weierstrass semigroup at a flex to jump by 2. Indeed, one may verify along these lines that the flex semigroup occurs as a codimension 2 locus in $\mathcal{M}_{4,1}$ (codimension is equal to weight), while the cone-flex semigroup occurs as a codimension 3 locus, despite being of weight 4.

Those low-genus examples suggest that when the combinatorics of the semigroup become nontrivial (which is seen in the semigroup tree when a semigroup has "missing" siblings which would not be closed under addition), the codimensions of the corresponding loci become smaller than the weight would predict. Using the hyperelliptic and cone-flex semigroups as our model, we use this idea to define a new quantity, called the *effective weight*.

Question 1.20 was given a substantial partial answer by Eisenbud and Harris [EH87], and Komeda [K91], stated in the following theorem. A semigroup is called *primitive* if the sum of any two positive elements exceeds the largest gap. It is easy to check that a semigroup is primitive if and only if its weight and effective weight coincide.

Theorem 1.23 (Eisenbud,Harris,Komeda)**.** *If S is a primitive semigroup with* $\mathrm{wt}(S) \leq g-1$*, then $\widetilde{\mathcal{W}}_S$ has a dimensionally proper component.*

Our main result on Weierstrass points is to calculate the exact minimum dimension of a component of $\widetilde{\mathcal{W}}_S$ for $\mathrm{wt}(S) \leq g-1$ (and in a wider range of cases as well). To state the result, we first define our notion of the effective weight.

Definition 1.24. A *generator* of a semigroup S is an element that is not a sum of two positive elements of S. The *effective weight* of a semigroup S, denoted $\varepsilon(S)$, is the following number.

$$\varepsilon(S) = |\{(s,t) :\ s \in S \text{ a generator of } S,\ t \notin S,\ 0 < s < t\}|$$

The author chose the term *effective weight* to refer to the combinatorial notion of *effective generators*. Specifically, the definition is chosen so that two consecutive children of a given semigroup in the semigroup tree (which correspond to two consecutive effective generators) always have consecutive effective weight (whereas their weights may differ by more than 1 if the effective generators are not consecutive).

Observe that if S is the hyperelliptic semigroup, then $\varepsilon(S) = g - 1$, which is the exact codimension of $\widetilde{\mathcal{W}}_S$ in $\mathcal{M}_g$ (see example 1.16). Our main theorems on Weierstrass points are the following.

Theorem B. *If X is any irreducible component of $\widetilde{\mathcal{W}}_S$, then*

$$\dim X \geq \dim \mathcal{M}_{g,1} - \varepsilon(S).$$

Theorem C. *If S is a numerical semigroup of genus g with $\varepsilon(S) \leq g - 1$, then $\widetilde{\mathcal{W}}_S$ has a component of dimension exactly $\dim \mathcal{M}_{g,1} - \varepsilon(S)$.*

The proofs of these theorems will occupy section 3. One special case of theorem C (in the non-primitive case) was studied by Bullock his his thesis [B13]: he considered the case where $2g - 1 \in S$ (such semigroups are called *symmetric* semigroups, and the points are sometimes called *subcanonical points*).

2. Elliptic chains

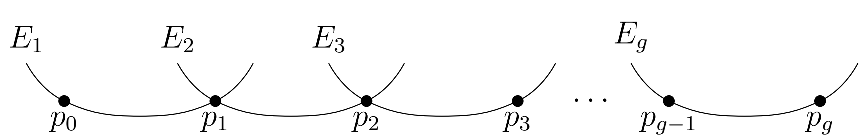
The main results of this thesis give results about smooth curves by constructing, and then smoothing, limit linear series on elliptic chains. Elliptic chains are very convenient curves to work with from the standpoint of limit linear series; their use as central fibers of degenerations was first leveraged by Welters [W85] to study Prym varieties, and their use was in the background of Eisenbud and Harris's work on Weierstrass points [EH87]. The main result of this section is lemma 2.10, which will drive our main results in subsequent sections.

2.1. **Limit linear series on elliptic chains.**

The theory of limit linear series was first proposed by Eisenbud and Harris in characteristic 0 (see [EH86] for a detailed treatment, or [HM] for an informal exposition of the underlying ideas). Recently, Osserman [Oss06] gave a more general and powerful construction, valid over an arbitrary algebraically closed field. The objective of limit linear series is to provide a partial compactification of the moduli space of linear series, allowing the underlying curves to become mildly singular (specifically, degenerations are allowed to curves of compact type).

We will give a rapid summary of the theory of limit linear series in this subsection. However, the only part of this analysis that is logically necessary for the rest of this thesis is lemma 2.10; the reader may therefore read the statement of this lemma and move on to the next subsection if desired.

While powerful and relatively elementary, the theory of limit linear series can be intimidating simply by virtue of its extremely cumbersome notation. For this reason, we will not make definitions in full generality, but restrict to the case of chains of elliptic curves.



Let X denoted the chain of elliptic curves shown above. The chain X consists of g elliptic curves $E_1, \cdots, E_g$, joined at points $p_1, \cdots, p_{g-1}$. We also mark two additional points $p_0 \in E_1$ and $p_g \in E_g$.

To describe limit linear series on X, we must parameterize linear series on each component E_i according to their ramification at *both* their marked points p_{i-1} and p_i. We will use a notation convention that appears bizarre at first, but will turn out to be very convenient, as follows. Informally speaking, we identify the ramification at p_{i-1} using order of zeros, and the ramification at p_i using orders of poles (where pole orders arise by regarding sections of a degree d line bundle $\mathcal{L}$ as rational sections of $\mathcal{L}(-dp_i)$ instead).

Definition 2.1. Suppose that S, T are two sets of nonnegative integers, with $|S| = |T| = r+1$. Let $\mathcal{B}^{S:T}_{d,g}$ denote the substack of $\mathcal{G}^r_{d,g} \times_{\mathcal{M}_g} \mathcal{M}_{g,2}$ parameterizing curves with two marked points and a linear series (C, p, q, L) such that the following vanishing conditions hold (here, $s_0 < \cdots < s_r$ and $t_0 < \cdots < t_r$ are the elements of S and T arranged in increasing order).

$$\begin{aligned} a_i^L(p) &\geq s_i \\ a_i^L(q) &\geq d - t_{r-i} \end{aligned}$$

Further, let $\widetilde{\mathcal{B}}^{S:T}_{d,g}$ denote the substack consisting of linear series where equality holds. Let $B^{S:T}_{d,g}(C, p, q)$ denote the fiber of $\mathcal{B}^{S:T}_{d,g}$ over $(C, p, q) \in \mathcal{M}_{g,2}$.

As usual, we will call a point of $\mathcal{B}^{S:T}_{d,g}$ *dimensionally proper* if the local dimension is given by the expected dimension from Brill-Noether theory. This expected dimension can be expressed as follows.

$$\begin{aligned} \text{expected dimension} &= \dim \mathcal{M}_{g,2} + \rho(g, S:T) \\ \text{where } \rho(g, S:T) &:= g + \sum_{i=0}^{r} (t_i - s_i - g) \end{aligned}$$

Let X be our elliptic chain. We first define *refined limit linear series* on X, which are the most important special case of limit linear series (and sufficient for our main theorems). Suppose that we have chosen a linear series L_i on each E_i individually. Suppose that its ramification at p_{i-1} and p_i is such that $(E_i, p_{i-1}, p_i, L_i) \in \widetilde{\mathcal{B}}_{d,g}^{S_i:T_i}$ (of course this will be true for a unique pair of sets S_i and T_i). We will say that the g-tuple $L = (L_1, \cdots, L_g)$ of linear series is a *refined limit linear series on* X if

$$T_i = S_{i+1} \text{ for all } i \in \{1, 2, \cdots, g-1\}.$$

The individual linear series L_i on the component E_i is called the E_i-aspect of L.

In other words, according to our convention of thinking of pole orders at the right side and zero orders at the left: a refined limit linear series is a set of linear series, one one each component, so that the pole orders on the left of any node equal the zero orders on the right of the node.

For obvious reasons, we will regard the elements of S_1 and T_g as the zero orders at p_0 and pole order at p_g of the limit linear series L. Therefore we make the following definition. Note that throughout this thesis, we will use fraktur letters for spaces of limit linear series.

Definition 2.2. Given a chain X of elliptic curves as above, any two sets of nonnegative integers S, T, and any integer $d \geq \max S \cup T$, define:

$$\widetilde{\mathfrak{B}}_d^{S:T}(X, p_0, p_g) = \coprod_{S_1, S_2, \cdots S_{g-1} \subseteq \mathbf{N}} \left(\prod_{i=1}^{g} \widetilde{B}_d^{S_{i-1}:S_i}(E_i, p_{i-1}, p_i) \right),$$

where $S_0 = S$ and $S_g = T$. Elements of this scheme are called *refined limit linear series with ramification* $(S : T)$ *at* (p_0, p_g). Given a refined limit linear series $L = (L_1, L_2, \cdots, L_g)$, the series $L_i \in B_d^{S_{i-1}:S_i}(E_i, p_{i-1}, p_i)$ is called the E_i-aspect of L.

Of course one may define a one-pointed version as well. It is easy to see, in fact, how this definition can be generalized to any number of marked points.

Definition 2.3. Given X, S, and d as above, define

$$\widetilde{\mathfrak{G}}_d^S(X, p_0, p_g) = \coprod_{S_1, S_2, \cdots S_{g-1} \subseteq \mathbf{N}} \left(G_d^{S_1}(E_1, p_1) \times \prod_{i=2}^{g} \widetilde{B}_d^{S_{i-1}:S_i}(E_i, p_{i-1}, p_i) \right),$$

where $S_g = S$. As before, the individual series are called the E_i-aspects.

Remark 2.4. The reader will observe that it was not really necessary for the components E_i to be genus 1; the same definitions can be used for any chain of positive-genus curves.

These schemes are obviously not proper. Eisenbud and Harris made use of the following compactification, although it has the substantial defect of failing to globalize well (this defect was rectified by Osserman, by a more intricate construction). A collection $L = (L_1, \cdots, L_g)$ is called an EH-coarse limit linear series if there are sets S_i, T_i such that $L_i \in \widetilde{B}_d^{S_i:T_i}(E_i, p_{i-1}, p)$, and

$$T_i \leq S_{i+1} \quad \text{for all } i \in \{1, 2, \cdots, g-1\}$$

Here, $T_i \leq S_{i+1}$ means that, once both sets are sorted into increasing order, the kth element of T_i is at most the kth element of S_i.

Definition 2.5. Define an *Eisenbud-Harris coarse limit linear series* on X with ramification at least $(S : T)$ at p_0, p_g to be an element of the following space.

$$\mathfrak{B}_d^{S:T,\mathrm{EH}}(X,p_0,p_g) = \bigcup_{S_1,S_2,\cdots S_{g-1}\subseteq \mathbf{N}} \left(\prod_{i=1}^{g} B_d^{S_{i-1}:S_i}(E_i,p_{i-1},p_i) \right),$$

regarded as a subscheme of $\displaystyle\prod_{i=1}^{g} G_d^r(E_i)$.

The scheme $\mathfrak{G}_d^{S,\mathrm{EH}}(X,p_g)$ can be defined similarly.

The Eisenbud-Harris spaces are easy to work with on the level of individual curves, but they are problematic from the standpoint of globalization, although Eisenbud and Harris were nevertheless able to prove very strong regeneration theorems in characteristic 0. Osserman's construction works in any characteristic and also gives a nice moduli scheme. However, we will not need the details of his construction, so we will content ourselves to state the following theorem without giving further details about the construction. The following theorem follows easily for the results of Osserman's [Oss06], although we state it in our notation and in our special case.

Black box. Osserman defines a certain type of family of pointed curves as a *smoothing family.* We will omit the definition of this term, but we will use it in the theorem below; the following lemma shows that the specific definition will not matter for us.

Lemma 2.6 (Osserman)**.** *Suppose that B is a regular connected scheme, and $B \to \overline{\mathcal{M}}_{g,n}$ is a morphism to the moduli stack of* 2*-pointed genus g stable curves, whose image is contained in the locus of curves that are chains of curves joined at nodes. Let $b \in B$ be any geometric point.*

There exists an étale map $B' \to B$, with b in its image, such that the composition $B' \to \mathcal{M}_{g,n}$ is a smoothing family.

Theorem 2.7 (Osserman)**.** *For any smoothing family $f: B \to \overline{\mathcal{M}}_{g,2}$, whose image consists of chains of smooth curves with the two marked points on opposite ends, and any S, T, d as above, there exists a scheme*

$$\mathfrak{B}_d^{S:T}(f) \to B,$$

proper over B and of dimension at least $\dim B + \rho(g, S : T)$ at all geometric points. These schemes are compatible with base change, and if the image of f is contained in $\mathcal{M}_{g,2}$, then $\mathfrak{B}_d^{S:T}(f) \to B$ is isomorphic to the scheme $\mathcal{B}_d^{S:T}(f)$ (the usual scheme of limit linear series with ramification). Also, $\mathfrak{B}_d^{S:T}(f)$ has an open subscheme $\widetilde{\mathfrak{B}}_d^{S:T}(f)$, whose geometric fibers are isomorphic to the spaces of refined limit linear series described above.

There is a corresponding statement about the globalization of $\mathfrak{G}_d^S$, and indeed similar statements for any number of marked points; the statement it completely analogous.

Although we will not need it in this thesis, we also mention the following result, which shows that analysis of Eisenbud-Harris coarse series (which are relatively easy to study directly) gives information about Osserman's coarse linear series.

Lemma 2.8 (Osserman)**.** *Let X be a chain of smooth curves with two marked points p, q. There is a set-theoretic map of geometric poitns*

$$e: \ \mathfrak{B}_d^{S:T}(X,p,q)(\overline{K}) \to \mathfrak{B}_d^{S:T,\mathrm{EH}}(X,p,q)(\overline{K})$$

such that for any limit linear series $L \in \mathfrak{B}_d^{S:T}(X,p,q)$, the local dimension of $\mathfrak{B}_d^{S:T}(X,p,q)$ at L satisfies

$$\dim_L \mathfrak{B}_d^{S:T}(X,p,q) \leq \dim_{e(L)} \mathfrak{B}_d^{S:T,\mathrm{EH}}(X,p,q) + \sum \rho(0, T_i : S_{i-1}),$$

where the sets S_i, T_i are the orders of zeros at the left endpoint and poles at the right endpoint of the aspects of $e(L)$ on the components of X.

One way to view this lemma is that Osserman's limit linear series constitute a sort of blow-up of the Eisenbud-Harris limit linear series, where the fibers have dimension bounded by the extent to which an Eisenbud-Harris coarse series fails to be a fine limit linear series. For a nice application of this last lemma, see Osserman's recent simple proof of the Brill-Noether theorem in arbitrary characteristic [Oss11]. The analysis in that paper essentially amounts to proving the following lemma.

Lemma 2.9. *If X is a chain of g elliptic curves as above, and on curve E_i the two points p_{i-1}, p_i do not differ by a torsion element of the Jacobian, then no scheme $\mathcal{B}_d^{S:T}(X, p_0, p_g)$ has a dimensionally improper component.*

This lemma can be proved by induction on g. That this lemma implies the Brill-Noether theorem 1.11 follows by taking a versal deformation of the curve and using theorem 2.7.

The upshot of all this, for our purposes, is the following lemma. After this section, we will not need to refer to limit linear series directly, but will instead be content to cite this fact.

Lemma 2.10 (Regeneration)**.** *Let S, T, U be arbitrary sets of integers of the same size, and d, g, h any positive integers.*

- *If $\widetilde{\mathcal{B}}_{d,g}^{S:T}$ and $\widetilde{\mathcal{B}}_{d,h}^{T:U}$ both have a dimensionally proper component, then so does $\widetilde{\mathcal{B}}_{d,g+h}^{S:U}$.*
- *If $\widetilde{\mathcal{G}}_{d,g}^{S}$ and $\widetilde{\mathcal{B}}_{d,h}^{S:T}$ both have dimensionally proper points, then so does $\widetilde{\mathcal{G}}_{d,g+h}^{T}$.*

Call a component of $\mathcal{B}_{d,g}^{S:T}$ full-rank *if it either dominates $\mathcal{M}_{g,2}$ or has finite fibers over it, and call a component of $\mathcal{G}_{d,g}^{S}$* full-rank *if it either dominates $\mathcal{M}_{g,1}$ or has finite fibers over it. Then in both of the above statements, if the two postulated components are full-rank, and either both have positive Brill-Noether numbers or both have negative Brill-Noether numbers, then the resulting component is also full-rank.*

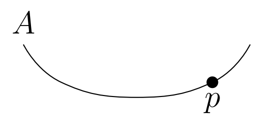
Proof. We will discuss a proof of the first bulleted statement; the second bulleted statement follows from a similar argument. Let (X_i, p_i, q_i, L_i) for $i = 1, 2$ be dimensionally proper geometric points of $\widetilde{\mathcal{B}}^{S:T}_{d,g}$ and $\widetilde{\mathcal{B}}^{T:U}_{d,h}$, respectively. Let $f_1 : B_1 \to \overline{\mathcal{M}}_{g,2}$ and $f_2 : B_2 \to \overline{\mathcal{M}}_{h,2}$, where B_1, B_2 are regular connected schemes, be versal deformations of (X_i, p_i, q_i), where $b_i \in B_i$ corresponds to (X_i, p_i, q_i). Let $f : B_1 \times B_2 \to \overline{\mathcal{M}}_{g+h,2}$ be obtained by composing $f_1 \times f_2 : B_1 \times B_2 \to \overline{\mathcal{M}}_{g,2} \times \overline{\mathcal{M}}_{h,2}$ with the gluing map that attaches q_1 to p_2. After an appropriate étale base change, $\widetilde{\mathcal{B}}^{S:T}_{d,g} \times \widetilde{\mathcal{B}}^{T:U}_{d,h}$ may be identified as an open subscheme of $\widetilde{\mathfrak{B}}^{S:U}_{d,g+h}(f)$. An elementary calculation shows that

$$\rho(g + h, S : U) = \rho(g, S : T) + \rho(h, T : U)$$

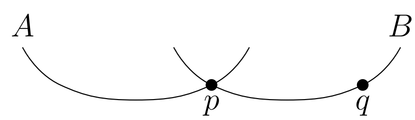
(Eisenbud and Harris refer to this fact in [EH86] as the "additivity of the Brill-Noether number"), and therefore $\widetilde{\mathfrak{B}}^{S:U}_{d,g+h}(f)$ has a dimensionally proper component. Perhaps after an additional étale base change, we may embed $B_1 \times B_2$ as a divisor in a versal family $f_3 : B_3 \to \overline{\mathcal{M}}_{g+h,2}$. Denote by C an irreducible component of $\mathfrak{B}^{S:U}_{d,g+h}(f_3)$ containing the (base change of) the product of our two dimensionally proper components of $\widetilde{\mathcal{B}}^{S:T}_{d,g}$ and $\widetilde{\mathcal{B}}^{T:U}_{d,h}$. Then the local dimension of C is at least $\dim B_3 + \rho(g + h, S : U)$, which is at least $\dim B_1 \times B_2 + \rho(g + h, S : U) + 1$; therefore since the part of C lying above $B_1 \times B_2$ has dimension equal to $\dim B_1 \times B_2 + \rho(g+h, S : U)$, it follow that C cannot lie entirely over $B_1 \times B_2$, and its dimension must be exactly $\dim B_3 + \rho(g + h, S : U) = \dim \mathcal{M}_{g+h,2} + \rho(g + h, S : U)$. Therefore, restricting to the smooth locus $\mathcal{M}_{g+h,2}$, we have proved that there is a dimensionally proper component of $\mathcal{B}^{S:U}_{d,g+h}$. In addition, since we started from a point in the open substack $\widetilde{\mathfrak{B}}^{S:U}_{d,g+h}(f_3)$, the general point of this component must lie in $\widetilde{\mathcal{B}}^{S:U}_{d,g+h}$. This establishes the claim in the first bullet point. The proof of the second bullet point is similar.

The last claim (about full rank components) follows from the upper semicontinuity of dimensions of fibers of $\mathfrak{B}^{S:U}_{d,g+h}$ over $\overline{\mathcal{M}}_{g,2}$ and of fibers of $\mathcal{G}^{S}_{d,g+h}$ over $\overline{\mathcal{M}}_{g,1}$. $\square$

2.2. **The displacement lemma.** The main result of this section is lemma 2.15, which is the essential geometric input in the proofs of theorems A and C. I refer to this lemma as a "displacement" lemma, because it concerns the following general situation. We have a linear series L^A on a smooth curve A of genus g, with known ramification at $p \in A$.



We wish to prove the existence of curves in higher genus, together with linear series with ramification at a marked point given by a slight modification of the ramification at p. To achieve this, we begin by constructing a nodal curve X of compact type by gluing a curve B to A at p, and marking a second point $q \in B$.



Now a refined limit linear series can be constructed on X with the desired ramification at q. In this situation, we will say that the B-aspect of this series *displaces* the ramification of L^A from the point p to the point q.

If S is a sequence of pole orders so that $(A, p, L^A) \in \widetilde{\mathcal{G}}^S_{d,g'}$, and T is the displaced sequence, so that we wish to construct a limit linear series L on X with $(X, q, L) \in \widetilde{\mathfrak{G}}^T_{d,g'+g}$, then it follows that we should choose (B, q, L^B) from $\widetilde{\mathcal{B}}^{S:T}_{d,g}$, defined as in the previous section. The letter $\mathcal{B}$ here is chosen to stand for "bridge."

Definition 2.11. A pair of sets $(S : T)$ is called g-valid if $\widetilde{\mathcal{B}}^{S:T}_{d,g}$ has a dimensionally proper point, for some $d \geq \max(S \cup T)$.

Notice that "for some d" is equivalent to "for all d," since $\mathcal{B}^{S:T}_{d,g} \cong \mathcal{B}^{S:T}_{d+1,g}$, by adding a base point at q. The reader may also verify that adding 1 to all the elements of S and T also results in an isomorphic space.

We will also make use of a slightly more refined notion, which essentially states that the image in $\mathcal{M}_{g,2}$ is as large as possible.

Definition 2.12. A pair of sets $(S:T)$ is called *strictly g-valid* if $\widetilde{\mathcal{B}}^{S:T}_{d,g}$ has a dimensionally proper point x such that either:

- $\rho(g,S:T) \geq 0$ and the fiber of x over $\mathcal{M}_{g,2}$ is $\rho(g,S:T)$-dimensional, or
- $\rho(g,S:T) < 0$ and x is an isolated point in its fiber over $\mathcal{M}_{g,2}$.

A large supply of valid pairs $(S:T)$ allows dimensionally proper linear series (with ramification) to be constructed by inductive arguments. The following lemma makes this precise.

Lemma 2.13. *Let S, T, U be three finite sets of integers of the same size.*

(1) If $\widetilde{\mathcal{G}}^{S}_{d,g}$ has a dimensionally proper point, and $(S:T)$ is h-valid, then $\widetilde{\mathcal{G}}^{T}_{d,g+h}$ has a dimensionally proper point.

(2) If $(S:T)$ is g-valid and $(T:U)$ is h-valid, then $(S:U)$ is $(g+h)$-valid.

Proof. Immediate from lemma 2.10. □

There is also a "strict" version of this lemma, taking fiber dimensions into account.

Lemma 2.14. *Let S, T, U be three finite sets of integers of the same size.*

(1) If $\widetilde{\mathcal{G}}^{S}_{d,g}$ has a dimensionally proper point x lying in a fiber whose local dimension at x is $\min(0,\rho(g,S))$, $(S:T)$ is h-valid, and the numbers $\rho(g,S)$ and $\rho(h,T)$ are either both nonnegative or both nonpositive, then $\widetilde{\mathcal{G}}^{T}_{d,g+h}$ has a dimensionally proper point y lying in a fiber whose local dimension at y is $\min(0,\rho(g,T))$

(2) If $(S:T)$ is strictly g-valid and $(T:U)$ is strictly h-valid, and the numbers $\rho(g,S:T)$ and $\rho(h,T:U)$ are either both nonnegative or both nonpositive, then $(S:U)$ is strictly $(g+h)$-valid.

Proof. Immediate from lemma 2.10. □

In other words, valid pairs have a sort of composition law, corresponding to gluing curves together. In particular, a large supply of 1-valid pairs will give a large supply of g-valid pairs for all positive g. Lemma 2.15 gives a versatile set of 1-valid pairs, sufficient for our applications but somewhat more general.

In the following statement, an *arithmetic progression* is a subset of $\mathbf{N}$, possibly empty or with only one element, such that $\Lambda - \Lambda$ is closed under addition.

Lemma 2.15 (The displacement lemma)**.** *Suppose that S, T are two finite sets of nonnegative integers, such that $s_i \leq t_i \leq s_{i+1}$ (where $s_0 < \cdots < s_r$ and $t_0 < \cdots < t_r$ are the elements of S and T). Let $\Lambda(S : T)$ be the arithmetic progression generated by $\{s_i : \ s_i = t_i\}$. Then $(S : T)$ is 1-valid if and only if the the following conditions hold.*

(1) There are at most 2 values of i such that $s_i = t_i$, and these are non-consecutive.
(2) If $t_i = s_i + 1$ and $t_i \in \Lambda(S : T)$, then $s_{i+1} = t_i$.
(3) If $t_i = s_i + 1$ and $s_i \in \Lambda(S : T)$, then $t_{i-1} = s_i$.

Proof. Fix $d \geq \max(S \cup T)$. The stack $\mathcal{B}^{S:T}_{d,1}$ has a map to $\mathcal{P}^d_{1,2}$ (the moduli stack of genus 1 curves with two marked points and a chosen degree d line bundle). Denote by $B(E, p, q, \mathcal{L})$ a fiber of this map. We will now compute explicitly the dimension of this fiber (which will depend on the choice of $(E, p, q, \mathcal{L})$). Since we are concerned only with dimension, we may work set-theoretically, and regard $B(E, p, q, \mathcal{L})$ as a subvariety of the Grassmannian $\mathrm{Gr}(r + 1, H^0(\mathcal{L}))$. For brevity (and to distinguish the variety from the scheme structure), denote the reduced structure of $B(E, p, q, \mathcal{L})$ simply by B (the data $E, p, q, \mathcal{L}$ will remain fixed). Denote the points of B by $[V]$, where $V \subseteq H^0(\mathcal{L})$. Denote by $\widetilde{B}$ the open subvariety where equality holds in all ramification conditions.

Let $\Lambda = \{n : \ \mathcal{L} \cong \mathcal{O}_E(n \cdot p + (d-n) \cdot q)\}$. This is an arithmetic progression, in the sense described before the statement of the lemma (possibly empty, possibly with only one element), and it is guaranteed to be a proper subset of $\mathbf{N}$. Every element $\lambda \in \Lambda \cap \{0, 1, \cdots, d\}$ corresponds to a nonzero section $\mu_\lambda \in H^0(\mathcal{L})$, unique up to scale, with divisor of zeros given by $(n \cdot p + (d-n) \cdot q)$. Choose one such section μ_λ for each $\lambda \in \Lambda \cap S \cap T$ (notice that if $\mu_\lambda \in V$, then certainly $\lambda \in S$ and $\lambda \in T$).

Any point $[V] \in \widetilde{B}$ has $\mu_\lambda \in V$ for some elements λ, but not for others. We will see that some of the elements μ_λ must be present, but any choice of a subset of the remaining elements of $\Lambda \cap S \cap T$ gives a distinct irreducible component of B, and all of these components have the same dimension. In order to study each possibility separately, let $M \subseteq \Lambda \cap S \cap T$ be any subset, and denote by B_M the locus in B consisting of those $[V]$ such that $V \cap \{\mu_\lambda\} = \{\mu_\lambda : \ \lambda \in M\}$. Again, we are working at the level of geometric points; therefore we will study B_M as a locally closed subvariety of $\mathrm{Gr}(r+1, H^0(\mathcal{L}))$. Not surprisingly, we will let $\widetilde{B}_M$ denote $B_M \cap \widetilde{B}$.

Suppose that $[V] \in \widetilde{B}_M$. We will analyze the structure of V by breaking into a direct sum in a particular way.

Consider subspaces of V of the following form. Here i, j are elements of $\{0, 1, \cdots, r\}$ with $i \le j$.

$$V_{i,j} = V \cap H^0(\mathcal{L}(-s_i p - (d - t_j)q)$$

By definition of vanishing sequences, the codimension of $V_{i,r}$ in V is exactly i, while the codimension of $V_{0,j}$ is exactly $(r-j)$. Therefore, since $V_{i,j} = V_{i,r} \cap V_{0,j}$, $\dim V_{i,j} \ge (j - i + 1)$. In fact, we will see shortly that equality holds in a broad range of cases.

The orders of vanishing at p of sections in $V_{i,j}$ must all be at least s_i (by definition) and at most t_j (since the orders of vanishing at q are at least $d - t_j$). Therefore these orders of vanishing are a subset of $\{s_i, s_{i+1}, \cdots, s_j, s_{j+1}\}$, where s_{j+1} occurs if and only if $s_{j+1} = t_j = \lambda \in M$. In particular, if j is selected so that either $s_{j+1} \neq t_j$ or $t_j \notin M$, then the orders of vanishing of sections in $V_{i,j}$ at p must be precisely $\{s_i, s_{i+1}, \cdots, s_j\}$. Similarly, if I is selected so that $s_i \neq t_{j-1}$ or $s_i \notin M$, then the order of vanishing at q are precisely $d - \{t_i, t_{i+1}, \cdots, t_j\}$. In either case, the number of orders of vanishing is $(j - i + 1)$, hence $\dim V_{i,j} = (j - i + 1)$.

In light of the above considerations, call two adjacent indices $i, i+1$ *linked* if $t_i = s_{i+1}$ and $t_i \in M$. We have shown that if i is not linked to $i-1$ and j is not linked to $j+1$, then $V_{i,j}$ has dimension equal to $(j-i+1)$, vanishing orders $\{s_i, \cdots s_j\}$ at p and vanishing orders $d - \{t_i, \cdots, t_j\}$ at q. Call the subspace $V_{i,j}$ a *linked block* if i is not linked to $i-1$, j is not linked to $j+1$, but for all $k \in \{i, i+1, \cdots, j-1\}$, k is linked to $k+1$.

Observe that any two linked blocks have disjoint sets of vanishing orders at p (or q), and that every possible order of vanishing (each value in $\{s_0, \cdots s_r\}$ at p, each value of $d - \{t_0, \cdots, t_r\}$ at q) occurs in some linked block. It follows that V is a direct sum of its linked blocks.

Suppose that there are ℓ linked blocks, and label them $V_{a_1,b_1}, V_{a_2,b_2}, \cdots, V_{a_\ell,b_\ell}$. In particular, $0 = a_1, b_\ell = r$, and $a_k = b_{k-1} + 1$. Notice that the values a_k, b_k do not depend on V: they are determined by M, S, and T.

Each block V_{a_k,b_k} must contain the sections μ_λ for all $\lambda \in \{t_{a_k} = s_{a_k+1}, t_{a_k+1} = s_{a_k+2}, \cdots, t_{b_k-1} = s_{b_k}\}$. Together, these sections span a subspace W_{a_k,b_k} of dimension $(b_k - a_k)$, which does not depend on V. The dimension of V_{a_k,b_k} is $(b_k - a_k + 1)$. Therefore the linked block V_{a_k,b_k} determines, and is determined by, a one-dimensional subspace of $Q_k := H^0(\mathcal{L}(-s_{a_k}p - (d - t_{b_k})q))/W_{a_k,b_k}$. The vector space Q_k, as well, does not depend on the choice of V, but only on M. So the block V_{a_k,b_k} determines, and is determined by, an element of $\mathbf{P}Q_k$. Note that the vanishing orders of sections in W_{a_k,b_k} are precisely $\{s_{a_k+1}, s_{a_k+2}, \cdots, s_{b_k}\}$ at p and $d - \{t_{a_k}, t_{a_k+1}, \cdots, t_{b_k-1}\}$ at q. So the element of $\mathbf{P}Q_k$ must be represented by a class of sections of $H^0(\mathcal{L}(-s_{a_k}p - (d - t_{b_k})q)$

which have vanishing order exactly s_{a_k} at p and t_{b_k} at q. Such elements form an open subset U_k of $\mathbf{P}Q_k$ (the complement of two linear subspaces).

To summarize: the choice of M determines a number ℓ and vector spaces $Q_1, \cdots, Q_\ell$. The element $[V] \in \widetilde{B}_M$ is determined uniquely by a choice of a single element from each of the sets $U_1, \cdots U_\ell$, where U_k is an open subset of $\mathbf{P}Q_k$.

Conversely, suppose that we have selected an element from each U_k (implicit in this assumption is that each U_k is nonempty; we will presently determine when this obtains). These choices produce a space $V \subseteq H^0(\mathcal{L})$ with vanishing orders precisely given by S at p and $d - T$ at q. So $[V] \in \widetilde{B}$. The point $[V]$ need not be in $\widetilde{B}_M$ specifically, but it is easy to check when this is the case. If $\lambda \in M$, then either $\lambda = t_i = s_{i+1}$ for some i or $\lambda = s_i = t_i$ for some i. In the former situation, i and $i+1$ are linked, and so μ_λ lies in some W_{a_k,b_k} and therefore lies in V. In the latter case, there must be a singleton block $a_k = b_k = s_i = t_i$. In this case, the only possible divisor of zeros for a section in this block is $s_i p + (d - t_i)q$, hence μ_λ spans this block and is guaranteed to be in V. So $\mu_\lambda \in V$ for all $\lambda \in M$. On the other hand, suppose that $\lambda \in (S \cap T \cap \Lambda) \backslash M$. From the construction of V, it is apparent that the only way that μ_λ could "accidentally" be present in V is if there is a singleton block $V_{i,i}$ such that $s_i = t_i$ (this is because the construction of V specifies a basis of V, all of whose members have distinct orders of zero at p and q, so the order of zero at p of a linear combination of these basis elements is equal to the lowest order of vanishing of any of the basis elements involved; since the orders of vanishing of μ_λ at p and q add up to d, the only way that μ_λ can be a linear combination of the basis elements is if it is a scalar multiple of one basis element). And indeed, if $s_i = t_i$ for some i, then the block $V_{i,i}$ must be precisely the span of μ_λ.

It follows from the discussion above that:

(1) If M does not contain all the elements $\{\lambda : s_i = t_i = \lambda \text{ for some } i\}$, then $\widetilde{B}_M$ is empty, and

(2) Otherwise, $\widetilde{B}_M$ is isomorphic to $\displaystyle\prod_{k=1}^{\ell} U_k$.

It remains to determine when all the open sets U_k are nonempty, and then to determine the dimension of each set U_k.

First, U_k is nonempty if and only if both of the following inequalities hold.

$$\begin{aligned} h^0(\mathcal{L}(-(s_{a_k}+1)p-(d-t_{b_k})q)) &< h^0(\mathcal{L}(-s_{a_k}p-(d-t_{b_k})q)) \\ h^0(\mathcal{L}(-s_{a_k}p-(d-t_{b_k}+1)q)) &< h^0(\mathcal{L}(-s_{a_k}p-(d-t_{b_k})q)) \end{aligned}$$

By Riemann-Roch, these inequalities fail to hold in precisely two situations.

(1) $s_{a_k} = t_{b_k}$ and $s_{a_k} \notin \Lambda$

(2) $s_{a_k} = t_{b_k} - 1$ and either $s_{a_k} \in \Lambda$ or $t_{b_k} \in \Lambda$

Notice that if situation 2 holds, then either $b_k = a_k$ or $b_k = a_k + 1$; but the latter case is impossible since it would imply that $s_{a_k} = t_{a_k}$ and thus a_k would not be linked to $a_k + 1$. Therefore, U_k is nonempty for all values of k if and only if the following three conditions hold.

(1) If $s_i = t_i$ then $s_i \in M$.

(2) If $s_i + 1 = t_i \in \Lambda$, then $t_i \in M$.

(3) If $s_i = t_i - 1 \in \Lambda$, then $s_i \in M$.

Note that implicit in (2) is that $s_{i+1} = t_i$, and implicit in (3) is that $t_{i-1} = s_i$.

If these three conditions hold, then it also follows that $\widetilde{B}_M \cong \prod_{k=1}^{\ell} U_k$, by earlier remarks. Each U_k is a dense open subset of a projective space. We can also find their dimensions explicitly. In the computation below, we use the fact that if $t_j = s_i$ for $j \geq i$, then $j = i$ and $s_i = t_i \in M$, so in particular $s_i \in \Lambda$.

$$
\begin{aligned}
\dim U_k &= \dim \mathbf{P} Q_k \\
&= \dim Q_k - 1 \\
&= h^0(\mathcal{L}(-s_{a_k} p - (d - t_{b_k})q)) - \dim W_{a_k,b_k} - 1 \\
&= \left\{ \begin{array}{ll} (t_{b_k} - s_{a_k}) & \text{if } t_{b_k} \neq s_{a_k} \\ 1 & \text{if } t_{b_k} = s_{a_k} \end{array} \right\} - (b_k - a_k) - 1 \\
&= (t_{b_k} - s_{a_k}) - (b_k - a_k) - 1 + \delta(t_{b_k} - s_{a_k})
\end{aligned}
$$

Here δ is the Dirac function. Adding these dimensions together, we obtain the dimension of $\widetilde{B}_M$.

$$
\begin{aligned}
\dim \widetilde{B}_M &= \sum_{k=1}^{\ell} (t_{b_k} - s_{a_k}) - \sum_{k=1}^{\ell} (b_k - a_k) - \ell + |\{i : \ s_i = t_i\}| \\
&= \sum_{k=1}^{\ell} (t_{b_k} - s_{b_k} + t_{b_k - 1} - s_{b_k - 1} + \cdots + t_{a_k} - s_{a_k}) \\
&\quad -(b_\ell - a_0) - \sum_{k=1}^{\ell-1} (b_k - s_{k+1}) - \ell + |\{i : \ s_i = t_i\}| \\
&= \sum_{i=0}^{r} (t_i - s_i) - (r - 0) - \sum_{k=1}^{\ell} (-1) - \ell + |\{i : \ s_i = t_i\}| \\
&= \sum_{i=0}^{r} (t_i - s_i - 1) + |\{i : \ s_i = t_i\}| \\
&= \rho(1, S : T) - 1 + |\{i : \ s_i = t_i\}|
\end{aligned}
$$

In particular, this dimension does not depend on the set M. It follows that if $\widetilde{B}$ is nonempty, then it is of pure dimension $\rho(1, S : T) - 1 + |\{i : \ s_i = t_i\}|$. It has one component for each possible choice of set M, each of which is a isomorphic to a dense open subset of a product of projective spaces.

Whether or not $\widetilde{B}$ is nonempty depends on the arithmetic progression Λ. The stack $\mathcal{P}^d_{1,2}$ of genus 1 curves with two marked points and a degree d line bundle is a disjoint union of locally closed substacks $\mathcal{P}^d_{1,2}(\Lambda)$, parameterizing those choices such that the arithmetic progression $\{n: \mathcal{L} \cong \mathcal{O}_E(np + (d-n)q)\}$ is equal to Λ. It is easy to see that the dimension of these substacks is given as follows.

$$\dim \mathcal{P}^d_{1,2}(\Lambda) = \begin{cases} 3 & \text{if } \Lambda = \emptyset \\ 2 & \text{if } |\Lambda| = 1 \\ 1 & \text{if } |\Lambda| = \infty \end{cases}$$

As we have seen, every fiber of $\mathcal{B}^{S:T}_{d,1}$ over a given substack $\mathcal{P}^d_{1,2}(\Lambda)$ is isomorphic. These fibers are nonempty if and only if Λ contains $\{s_i: s_i = t_i\}$ but does not contain any elements of $\{t_i: s_i + 1 = t_i \text{ and } s_{i+1} \neq t_i\}$ or $\{s_i: s_i + 1 = t_i \text{ and } t_{i-1} \neq s_i\}$, because these are precisely the cases in which a set M can be selected so that $\widetilde{B}_M$ is nonempty. The hypotheses of the lemma guarantee that there is some Λ such that the fibers over $\mathcal{P}^d_{1,2}(\Lambda)$ are nonempty. In fact, there is a minimal such Λ, generated by $\{s_i: s_i = t_i\}$, and the closure of $\mathcal{P}^d_{1,2}(\Lambda)$ contains all points with nonempty fibers above them. The size of Λ is 0 if $|\{i: s_i = t_i\}| = 0$, 1 if $|\{i: s_i = t_i\}| = 1$, or ∞ if $|\{i: s_i = t_i\}| \geq 2$. It follows that, for this minimal progression Λ,

$$\begin{aligned} \dim \widetilde{\mathcal{B}}^{S:T}_{d,1} &= \dim \mathcal{P}^d_{1,2}(\Lambda) + \rho(1, S:T) - 1 + |\{i: s_i = t_i\}| \\ &= 3 - \min(|\{i: s_i = t_i\}|, 2) + \rho(1, S:T) - 1 + |\{i: s_i = t_i\}| \\ &= 2 + \rho(1, S:T) + \max(0, |\{i: s_i = t_i\}| - 2) \\ &= \dim \mathcal{M}_{1,2} + \rho(1, S:T) + \max(0, |\{i: s_i = t_i\}| - 2). \end{aligned}$$

Therefore if the hypotheses are met, then $\widetilde{\mathcal{B}}^{S:T}_{d,1}$ has dimensionally proper points, i.e. $(S:T)$ is 1-valid.

Conversely, if hypothesis (2) or (3) fails, then there is no Λ such that the fibers over $\mathcal{P}^d_{1,2}$ are nonempty. If hypothesis (1) fails, then either $\Lambda = \mathbf{N}$ (in which case $\widetilde{\mathcal{B}}^{S:T}_{d,1}$ is empty) or else the dimension computation above shows that no component of $\widetilde{\mathcal{B}}^{S:T}_{d,1}$ is dimensionally proper. Therefore the hypotheses are both necessary and sufficient. $\square$

The proof of the displacement lemma, with a few modifications, also yields the following "strict" version.

Lemma 2.16. *Suppose that S, T are two finite sets of nonnegative integers such that $s_i \leq t_i \leq s_{i+1}$ (where $s_0 < \cdots < s_r$ and $t_0 < \cdots < t_r$ are the elements of S and T).*

- *If there are 0 or 1 indices i such that $s_i = t_i$, then $(S : T)$ is strictly 1-valid.*
- *If there are exactly 2 indices i such that $s_i = t_i$, then $(S : T)$ is strictly 1-valid if and only if*

  (1) These two values of i are not consecutive,

  (2) For all other indices i, $t_i = s_i + 1$,

  (3) If $t_i = s_i + 1$ and $t_i \in \Lambda(S : T)$, then $s_{i+1} = t_i$, and

  (4) If $t_i = s_i + 1$ and $s_i \in \Lambda(S : T)$, then $t_{i-1} = s_i$,

  where $\Lambda(S : T)$ is the arithmetic progression generated by the two values s_i such that $s_i = t_i$.

Proof. Observe in the proof of lemma 2.15 that, in the cases where $(S : T)$ is 1-valid, the map $\widetilde{\mathcal{B}}^{S:T}_{d,1} \to \mathcal{M}_{g,2}$ is either surjective or has image of codimension 1 (corresponding to twice-pointed curves with the points differing by some specific torsion order). The map is surjective if there are 0 or 1 indices with $s_i = t_i$ (in which case $\rho(1, S : T) \geq 0$), and has image of codimension 1 if there are two such indices (in which case $\rho(1, S : T) = -1$ if and only if $t_i = s_i + 1$ for all other i). Therefore there are no additional constraints on S and T when at most one index i gives $s_i = t_i$, but in the case where there are two such indices, the condition that $s_i = t_i + 1$ for all other indices must be added to insure strictness. $\square$

Example 2.17. Lemma 2.16 implies the existence and dimension portions of the generalized Brill-Noether theorem 1.11. These two statements amount to the following: if S is a finite set of nonnegative integers such thet $\max S \leq |S| + g - 1$ and $\rho(g, S) \geq 0$, then there is a dimensionally proper component of $\mathcal{G}^S_{d,g}$ that dominates $\mathcal{M}_{g,1}$. In fact, we can prove something slightly stronger: there is a dimensionally proper component of $\widetilde{\mathcal{G}}^{S,\mathrm{sep}}_g$ which dominates $\mathcal{M}_{g,1}$. For convenience, denote $|S| - 1$ by r and the set $\{0, 1, \cdots, r\}$ by $[r]$. What we shall show is that the pair $([r] : S)$ is strictly g-valid. Therefore there is a dimensionally proper component X of $\widetilde{\mathcal{B}}^{[r]:S}_g$. The forgetful map $X \to \widetilde{\mathcal{G}}^S_g$ has all fibers of dimension 1, since the forgotten point is unramified, hence the linear series is separable and thus all but finitely many points of the curve are unramified. Since X dominates $\mathcal{M}_{g,2}$, it must also dominate $\mathcal{M}_{g,1}$. Therefore it suffices to show that if $\rho(g, S) \geq 0$, then $([r] : S)$ is strictly g-valid.

This follows by induction on g. The base case $g = 1$ is easy: for given value of r, there are only two possible sets S, namely $\{1, 2, \cdots, r+1\}$ and $\{0, 2, \cdots, r+1\}$. In both cases $([r] : S)$ is strictly 1-valid. For the inductive step, construct a set S' from S be decreasing all element of S by one, except the minimum element $s \in S$ such that $(s+1) \notin S$. The assumptions that $\max S \leq r + g$ and $\rho(g, S) \geq 0$ imply that S' can contain no negative elements. It is easy to check that $\rho(g-1, S') = \rho(g, S)$. Also $\max S' = \max S - 1$, unless S consists of a contiguous block of integers, in which case $\max S' = \max S$. Therefore S' will satisfy the hypotheses of the claim, unless $S = \{g, g+1, \cdots, g+r\}$. In this latter case, simply take S' to be $S - 1$ instead. Lemma 2.16 implies that in either case, $(S' : S)$ is strictly 1-valid. By inductive hypothesis, $([r] : S')$ is strictly $(g-1)$-valid. Since $\rho(g-1, S')$ and $\rho(g, S' : S)$ are both nonnegative, lemma 2.14 implies that $([r] : S)$ is strictly g-valid, completing the induction. $\lhd$

Example 2.18. The following situation will be used to construct Weierstrass points. Suppose that $m > 1$ and that $S = \{0, m, s_2, s_3, \cdots, s_r\}$, where $s_k > m$ and $m \nmid s_k$ for $k \geq 2$. Define T to be $\{0, m, s_2', s_3', \cdots, s_r'\}$, where $s_k' = \begin{cases} s_k + 1 & \text{if } m \nmid (s_k + 1) \\ s_k + 2 & \text{otherwise} \end{cases}$. Then (in the notation of the statement of the displacement lemma) $\Lambda(S : T)$ consists of all multiples of m. There are exactly two places where $s_i = t_i$, and no other places where S or T meet $\Lambda(S : T)$. Therefore $(S : T)$ is 1-valid. Notice however that $(S : T)$ is *not strictly* 1-valid. Indeed, our constructions in section 3 will lead to linear series which are dimensionally proper, do not extend to all of $\mathcal{M}_{g,1}$, but always lie in components with positive-dimensional fibers over $\mathcal{M}_{g,1}$. $\triangleleft$

3. Effectively proper Weierstrass points

In section 1, we defined the effective weight $\varepsilon(S)$ of a numerical semigroup S to be the number of pairs $(s,t) \in S \times (\mathbf{N} \backslash S)$ such that s is a generator and $s < t$. In this section we prove the following two theorems about the effective weight.

Theorem B. *If X is any irreducible component of $\mathcal{W}_S$, then*

$$\dim X \geq \dim \mathcal{M}_{g,1} - \varepsilon(S).$$

Theorem C. *If S is a numerical semigroup of genus g with $\varepsilon(S) \leq g-1$, then $\widetilde{\mathcal{W}}_S$ has a component of dimension exactly* $\dim \mathcal{M}_{g,1} - \varepsilon(S)$.

The following terminology will be convenient.

Definition 3.1. A component of $\widetilde{\mathcal{W}}_S$ or $\mathcal{W}_S$ is called *effectively proper* if it has codimension exactly $\varepsilon(S)$ in $\mathcal{M}_{g,1}$.

3.1. Restricted Weierstrass sequences.

The proof of theorem C proceeds by deducing the existence of effectively proper Weierstrass points from the existence of dimensionally proper linear series of a different sort. In this subsection we prove that the effective weight is a bound on codimension (theorem B), and we also prove corollary 3.4, which will be used in the proof of theorem C.

The construction is as follows: beginning with the complete linear series of the line bundle $\mathcal{O}(Np)$ (for suitably large n), we can project down to a subseries that includes enough vanishing orders at p to completely determine the Weierstrass semigroup. More specifically, we shall consider points of $\mathcal{G}_g^T$, where T is taken to be a restricted Weierstrass sequences, as defined below.

Definition 3.2. Let S be a numerical semigroup, and $T \subset S$ a finite subset. Call T a *restricted Weierstrass sequence* of S if

- T contains 0 and all generators of S.

- T does not contain any non-generators of S that are less than the largest gap of S.

The following lemma demonstrates the close link between restricted Weierstrass sequences and the effective weight of a semigroup.

Lemma 3.3. *Suppose that S is a numerical semigroup of genus g, and $T \subset S$ is a restricted Weierstrass sequence.*

(1) $\widetilde{G}^T(C,p)$ is nonempty if any only if $(C,p) \in \widetilde{\mathcal{W}}_S$.

(2) If $(C,p) \in \widetilde{\mathcal{W}}_S$, then the reduced structure of $\widetilde{G}^T(C,p)$ is isomorphic to affine space of dimension $\rho(g,T) + \varepsilon(S)$.

Proof. Fix an integer d, at least as large as the maximum value in T, and regard $\widetilde{G}^S(C,p)$ as a subschema of $G_d^{|S|-1}(C)$.

First consider part 1. If $\widetilde{G}^T(C,p)$ contains a point $L = (\mathcal{L}, V)$, then since $0 \in T$, $\mathcal{L} \cong \mathcal{O}(dp)$. Hence for all $t \in T$, there is a rational function on C with pole of order exactly t at p, and no other poles. So the Weierstrass semigroup of (C,p) contains all the generators of S. Therefore the Weierstrass semigroup of (C,p) is contained in S; since both have genus g, they must be equal. So $(C,p) \in \widetilde{\mathcal{W}}_S$. Conversely, if the Weierstrass semigroup of (C,p) is S, then for all $t \in T$, there exists a rational function f_t on C with pole of order t at p and no other poles. Regarded as sections of $\mathcal{O}(dp)$, these sections are linearly independent (since they have different orders of vanishing at p), and their span gives a linear series $L = (\mathcal{O}(dp), V)$ in $\widetilde{G}^T(C,p)$.

Now consider part 2. Assume that $(C,p) \in \widetilde{\mathcal{W}}_S$. Since $0 \in T$, each L in $\widetilde{G}^T(C,p)$ is given by $(\mathcal{O}(dp), V)$, for some $V \subseteq H^0(\mathcal{O}(dp))$. The vector space $H^0(\mathcal{O}(dp))$ has a complete flag given by the spaces $H^0(\mathcal{O}(sp))$, where s is an element of $S \cap \{0, 1, \cdots, d\}$, and those subspaces $V \subseteq H^0(dp)$ such that $(\mathcal{O}(dp), V) \in \widetilde{G}^T(C,p)$ form an open Schubert cell in $\mathrm{Gr}(|T|, H^0(\mathcal{O}(dp)))$ with respect to this filtration, hence isomorphic to an affine space.

To facilitate computations, define the following notation: for any two sets A, B of integers, let $\lambda(A, B)$ be the number of pairs (a, b), where $a \in A, b \in B$, and $a < b$. Also, partition the set $\{0, 1, 2, \cdots, d\}$ into three sets, as follows.

$$\begin{aligned} S &= T \cup T' \cup G, \text{ where} \\ T' &:= (S \cap \{0, 1, 2, \cdots, d\}) \backslash T \\ G &= \{0, 1, 2, \cdots, d\} \backslash S \end{aligned}$$

The codimension in $\mathrm{Gr}(|T|, H^0(\mathcal{O}(dp)))$ of this Schubert cycle can be described as the number of pairs (t, s), where $t \in T$, $s \in S$ such that $s \leq d$ and $s \notin T$. In other words, this codimension is $\lambda(T, T')$. The dimension of $\mathrm{Gr}(|T|, H^0(\mathcal{O}(dp)))$ is $|T| \cdot |T'|$, so it follows that the dimension of $\widetilde{G}_d^T(C, p)$ is $\lambda(T', T)$. On the other hand, $\varepsilon(S)$ and $\rho(g, T)$ are expressed in this notation as follows.

$$\varepsilon(S) = \lambda(T, G)$$

$$\begin{aligned} \rho(g, T) &= \sum_{i=0}^{r} (t_i - i) - (r+1)g \\ &= \lambda(T' \cup G, T) - |T| \cdot |G| \\ &= \lambda(T', T) + \lambda(G, T) - |T| \cdot |G| \end{aligned}$$

$$\begin{aligned} \rho(g, T) + \varepsilon(S) &= \lambda(T', T) + \lambda(G, T) + \lambda(T, G) - |T| \cdot |G| \\ &= \lambda(T', T) \end{aligned}$$

Therefore $\dim \widetilde{G}_d^T(C,p) = \rho(g,T) + \varepsilon(S)$, as claimed. □

This lemma rapidly gives a proof of theorem B.

Proof of theorem B. Let X be any irreducible component of $\mathcal{W}_S$, and let $\widetilde{X}$ be the intersection with $\widetilde{\mathcal{W}_S}$ (which is dense and open). Let T be the set of generators of S, plus 0, and let d be the maximum element of T. Consider the forgetful map $\pi : \widetilde{\mathcal{G}}_g^T \to \mathcal{M}_{g,1}$. By lemma 3.3, the image of this map is precisely $\widetilde{\mathcal{W}_S}$, and all the fibers are irreducible of dimension $\rho(g,T) + \varepsilon(S)$. Hence $\pi^{-1}(\widetilde{X})$ is irreducible of dimension $\dim X + \rho(g,T) + \varepsilon(S)$. On the other hand, the dimension of $\pi^{-1}(\widetilde{X})$ must be at least $\dim \mathcal{M}_{g,1} + \rho(g,T)$. It follows that $\dim X \geq \dim \mathcal{M}_{g,1} - \varepsilon(S)$. □

Lemma 3.3 and the remarks in the proof of theorem B also give the following corollary, which will be used to establish the existence of effectively proper Weierstrass points.

Corollary 3.4. *A pointed curve $(C,p) \in \widetilde{\mathcal{W}_S}$ is effectively proper if and only if (C,p,L) is dimensionally proper for all $L \in \widetilde{G}^T(C,p)$. In particular, $\widetilde{\mathcal{W}_S}$ has a dimensionally proper point if and only if $\widetilde{\mathcal{G}}_g^T$ has a dimensionally proper point.*

Before discussing the proof of the existence of effectively proper points in general, we show how it will work on two examples of non-primitive semigroups. Of course the results of both examples are obvious by other methods, but the examples will demonstrate a more general approach.

Example 3.5. Consider the genus g hyperelliptic semigroup $S = \{2, 4, 6, \cdots, 2g, 2g+1, \cdots\}$. A reduced Weierstrass sequence for S is simply $T = \{2, 2g+1\}$. If we know of the existence of an effectively proper hyperelliptic Weierstrass point in genus g, then $\widetilde{\mathcal{G}}_g^T$ has a dimensionally proper point. Lemma 2.15 shows that for $T' = \{2, 2g+3\}$, the pair $(T : T')$ is 1-valid. Notice that T' is a restricted Weierstrass sequence for $S' = \{2, 4, \cdots, 2g+2, 2g+3, \cdots\}$, i.e. the hyperelliptic semigroup of genus $g+1$.

It follows from lemma 2.13 that $\widetilde{\mathcal{G}}_{g+1}^{T'}$ has a dimensionally proper point. Therefore by corollary 3.4, $\mathcal{W}_{S'}$ has an effectively proper point. By induction, the hyperelliptic semigroup of genus g has effectively proper points for all g. $\triangleleft$

Example 3.6. Consider the first example of a non-hyperelliptic semigroup that is not primitive: the cone-flex semigroup $S = \mathbf{N}\backslash\{1,2,4,7\} = \{0,3,5,6,8,9,\cdots\}$. Then $\varepsilon(S) = 3$ and $\mathrm{wt}(S) = 4$. One restricted Weierstrass sequence for S is $T = \{0,3,5,8,10\}$ (shorter sequences will also work).

To show that $\mathcal{W}_S$ has effectively proper points, we begin with the genus 3 hyperflex semigroup $S' = \mathbf{N}\backslash\{1,2,5\} = \{0,3,4,6,7,\cdots\}$. It has restricted a Weierstrass sequence $T' = \{0,3,4,7,8\}$. If we know that $\mathcal{W}_{S'}$ has effectively proper points, then $\widetilde{\mathcal{G}}_3^{T'}$ has dimensionally proper points. By lemma 2.15, the pair $(T' : T)$ is 1-valid, so by lemma 2.13, $\widetilde{\mathcal{G}}_4^{T}$ has dimensionally proper points. Therefore by corollary 3.4, $\mathcal{W}_S$ has effectively proper points. $\triangleleft$

3.2. **Semigroups of low effective weight.** This subsection collects some combinatorial properties of semigroups of low effective weight, needed in the proof of theorem C. Most of the arguments are fairly mechanical, so the reader may wish to skip this section and take the statements as black boxes.

The proof of theorem works by induction on the genus, using the fact that a genus g semigroup S with $\varepsilon(S) \leq g-1$ can be built up by a "displacement" process from a trivial semigroup, increasing the effective weight and genus by 1 at each step.

Recall that a *generator* of S is a positive element that is not a sum of two positive elements, a "gap" is an element of the complement, and the *effective weight* $\varepsilon(S)$ of a semigroup is defined to be the number of pairs (s,t), where s is a generator, t is a gap, and $s < t$. So $\varepsilon(S)$ measures how much the gaps intermingle with the generators.

First of all, semigroups with $\varepsilon(S) \leq g-1$ satisfy a combinatorial condition slightly weaker than primitivity. The author has invented the term "secundive" by replacing the Latin root *primus* ("first") in "primitive" with *secundus* ("second").

Definition 3.7. Let $S \subseteq \mathbf{N}$ be a numerical semigroup. Then S will be called

primitive	if	the largest gap is smaller than twice the smallest generator, and
secundive	if	the largest gap is smaller than the sum of the two smallest generators.

Example 3.8. The hyperelliptic semigroup $\{0, 2, 4, \cdots, 2g, 2g+1, \cdots\}$ is secundive, of effective weight $g-1$. $\triangleleft$

Lemma 3.9. *If $\varepsilon(S)$ is less than the genus of S, then S is secundive.*

Proof. Suppose that S is a non-secundive numerical semigroup. Let $E \subset S \times (\mathbf{N} \backslash S)$ denote the set $\{(x, y) : \ 0 < x < y, \ x \text{ is a generator } \}$. By definition, $|E| = \varepsilon(S)$. We will explicitly construct $g(S)$ distinct elements of E, which will show that $\varepsilon(S) \geq g$ and establish the lemma.

Let m and n be the smallest and second-smallest generators of S, respectively. Let f denote the largest gap of S. Since S is not secundive, $f > m+n$. Let $T = \{n, n+1, n+2, \cdots, n+m-1\} \backslash \{m \cdot \lceil \frac{n}{m} \rceil\}$. Observe that no element of T is a multiple of m, and each is smaller than $m+n$; therefore any element of $T \cap S$ cannot be a sum of two nonzero elements of S (n being the smallest non-multiple of m in S). In other words, every element of T is either a gap or a generator. For every element $t \in T$, define an element $e(t) \in E$ as follows.

$$e(t) = \begin{cases} (t, f) & \text{if } t \in S \\ (n, t) & \text{if } t \notin S \end{cases}$$

Notice that the pairs $e(t)$ are all distinct. Now, define two subsets of E as follows.

$$\begin{aligned} E_1 &= \{(m,x) : \ x > m, \ x \notin S\} \\ E_2 &= \{e(t) : t \in T\} \end{aligned}$$

These two sets are disjoint. Therefore

$$\begin{aligned} |E| &\geq |E_1| + |E_2| \\ &= (g(S) - (m-1)) + (m-1) \\ &= g(S). \end{aligned}$$

Therefore $\varepsilon(S) \geq g(S)$. □

Remark 3.10. The construction in the proof shows that this result is sharp – there exist non-secundive semigroups with $\varepsilon(S) = g$ for all $g \geq 6$, and it is easy to see from the proof how to enumerate them. For $g \leq 5$, all numerical semigroups are secundive.

Due to this lemma, we may confine our attention to secundive semigroups, which we will build up by an inductive procedure.

Definition 3.11. For S a subset of $\mathbf{N}$, and any nonnegative integer k, define the *upward and downward displacement across* k as follows. Here $0 \leq s_0 < s_1 < s_2 < \cdots$ denote the elements of S in sorted order.

$$D^+(S,k) = \{s_i^+\}$$

$$\text{where } s_i^+ = \begin{cases} s_i & \text{if } k | s_i \text{ and } (s_i - 1) \notin S \\ s_i + 1 & \text{otherwise} \end{cases}$$

$$D^-(S,k) = \{s_i^-\}$$

$$\text{where } s_i^- = \begin{cases} s_i & \text{if } k|s_i \text{ and } (s_i+1) \notin S \\ s_i - 1 & \text{otherwise} \end{cases}$$

Technically speaking, the downward displacement may no longer be a set of nonnegative integers (specifically, in case 0 and 1 are both in S). This is no obstacle, since the definition above is easily generalized to any set of integers.

From this definition, the following lemma is easy to verify.

Lemma 3.12. *If $S \subseteq \mathbf{N}$ contains all nonnegative multiples of k and has complement (in $\mathbf{N}$) of size g, then*

(1) $D^+(D^-(S,k),k) = S = D^-(D^+(S,k),k)$,

(2) $D^+(S,k)$ has complement of size $g+1$,

(3) If $1 \notin S$, then $D^-(S,k)$ has complement of size $g-1$.

Secundive semigroups behave well under displacement, as displayed in the following lemma.

Lemma 3.13. *Let S be a secundive semigroup, with smallest generator m. Then*

(1) $D^+(S,m)$ is a secundive semigroup, with genus $g+1$ and effective weight $\varepsilon(S)+1$, and

(2) if $(m+1) \notin S$, then $D^-(S,m)$ is a secundive semigroup with genus $g-1$ and effective weight $\varepsilon(S)-1$.

Proof. Let $u : \mathbf{N} \to (\mathbf{N} \backslash m\mathbf{N})$ be the order-preserving bijection between nonnegative integers and nonnegative integers not divisible by m (that is, $f(0) = 1, f(1) = 2, \cdots, f(m-2) = m-1, f(m-1) = m+1, f(m) = m+2$, and so forth). Let

$T = u^{-1}(S)$. Notice that all elements of T are at least $m-1$ and the difference between the last gap of T and the smallest element of T is less than $m-1$ (since $u(k+m-1) = u(k)+m$ for all nonnegative integers k).

Conversely, if $T \subset \mathbf{N}$ is any cofinite set such that $\min T \geq m-1$ and $\max(\mathbf{N}\backslash T) - \min T < m-1$, then $m\mathbf{N} \cup u(T)$ is a secundive semigroup with smallest generator m. So secundive semigroups with first generator m are in bijection with sets T such that $\min T \geq m-1$ and $\max(\mathbf{N}\backslash T) - \min T < m-1$. Notice that $g(S) = g(T)$ (since no multiples of m are gaps in S). Also $\varepsilon(S) = g(S) - (m-1) + \mathrm{wt}(T)$, because any pair (x, y) where $0 < x < y$, x is a generator, and $y \notin S$ either satisfies $x = m$ (there are $g(S) - (m-1)$ such pairs), or $m \nmid x$ and $m \nmid y$. All pairs of the latter sort occur in T, and all elements of S that are less than the last gap and not multiples of m are necessarily generators, so the pairs (x, y) where $x \neq m$ are precisely in bijection with pairs (x', y'), where $x' < y'$, $x' \in T$, and $y' \notin T$, of which there are precisely $\mathrm{wt}(T)$.

Now, observe that if $S = m\mathbf{N} \cup u(T)$ (where T is as above), then $D^+(S, m) = m\mathbf{N} \cup u(T+1)$ and $D^-(S, m) = m\mathbf{N} \cup (T-1)$. It follows from the previous paragraph that $D^+(S, m)$ is a secundive semigroup with smallest generator m, while the same is true of $D^-(S, m)$ if and only if $(m-1) \notin T$, i.e. $(m+1) \notin S$. Since $g(T \pm 1) = g(T) \pm 1$ and $\mathrm{wt}(T \pm 1) = \mathrm{wt}(T)$, the genus and effective weights of the displacements are as stated. $\square$

Under the stronger hypothesis that S is primitive, there is more flexibility in choosing the displacement. The following statement is sufficient for the application in this paper.

Lemma 3.14. *Let S be a primitive semigroup of genus g with smallest generator m such that $(m+1) \in S$ and $\varepsilon(S) > 0$. Let M be the largest element of S that is smaller than some gap of S.*

(1) *If* $\varepsilon(S) \leq g-2$*, then* $D^-(S,M)$ *is a primitive semigroup of genus* $g-1$ *and effective weight* $\varepsilon(S)-1$*.*

(2) *If* $\varepsilon(S) = g-1$*, then the same conclusion as above holds, unless* S *has the form* $\{0, m, m+1, 2m, 2m+1, 2m+2, \cdots\}$*.*

Proof. Suppose that $D^-(S,M)$ is not a primitive semigroup. Let f be the largest gap of S. Then the smallest positive element of $D^-(S,M)$ is $m-1$ and the largest gap is

$$f' = \begin{cases} f-2 & \text{if } f = M+1 \\ f-1 & \text{otherwise.} \end{cases}$$

Since $D^-(S,M)$ is a not a primitive semigroup, $2(m-1) \leq f'$. Since $2m > f$, this implies that $f' = f-1$ and $f = 2m-1$. Therefore $m, m+1 \in S$, while $2m-2, 2m-1 \notin S$, and $2m-2 > m+1$.

Let $T = S \cap \{m+2, m+3, \cdots 2m-3\}$. Observe that $g(S) = (m-1)+(m-4-|T|)+2 = 2m-3-|T|$. The weight of S (equal to $\varepsilon(S)$ since S is primitive) is the size of the set $W = \{(x,y) : \ 0 < x < y, \ x \in S, y \notin S\}$. This set contains the four elements $(m, 2m-2), (m, 2m-1), (m+1, 2m-2), (m+1, 2m-1)$. For each $z \in \{m+2, m+3, \cdots 2m-3\}$, either $z \in T$ and $(z, 2m-2), (z, 2m-1) \in W$, or else $z \notin T$ and $(m, z), (m+1, z) \in W$; either way z appears in two distinct elements of W. Taken together, this accounts for $4 + 2(m-4) = 2m-4$ distinct elements of W, hence $\varepsilon(S) \geq 2m-4 = g(S)-1-|T|$. Therefore $\varepsilon(S) \geq g-1$, with equality if and only if T is empty, in which case S is precisely the semigroup $\{0, m, m+1, 2m, 2m+1, 2m+2, \cdots\}$. $\square$

The exceptional case in part 2 of lemma 3.14 is the reason that Eisenbud and Harris originally proved their results in [EH87] only for the case $\mathrm{wt}(S) \leq g-2$.

3.3. **Existence of effectively proper points.** This subsection gives the proof of theorem C. The proof is by induction. Each inductive step replaces a semigroup S by a new semigroup $D^+(S,k)$ (in the notation of the previous subsection).

There will be two sorts of inductive steps, corresponding to the following two lemmas (we state them next to each other to show the similarity, and then prove each individually).

Lemma 3.15. *If S is secundive with smallest generator m and $\widetilde{\mathcal{W}}_S$ has an effectively proper component, then $\widetilde{\mathcal{W}}_{S'}$ has an effectively proper component, where $S' = D^+(S,m)$.*

Lemma 3.16. *If S is primitive and $\widetilde{\mathcal{W}}_S$ has an effectively proper component, then for any $k \in S$ such that $(k-1) \notin S$, $\widetilde{\mathcal{W}}_{S'}$ has an effectively proper component, where $S' = D^+(S,k)$.*

Proof of lemma 3.15. Let g be the genus of S. By lemmas 3.12 and 3.13, S' is a secundive semigroup of genus $g+1$. Let f be the largest gap of S. Let $T \subset S$ consist of all elements of S that are less than or equal to $f+m+1$ *except* multiples of m larger than m. Observe that T contains all generators of S (no generator can exceed $f+m$, since $n > f+m$ implies $n-m \in S$). Also, the only non-generators of S less than f must be multiples of m, since S is secundive. Therefore T is a restricted Weierstrass sequence of S.

Similarly, let f' be the largest gap of S'. Let $T' \subset S'$ consist of all elements of S' that are less than or equal to $f+m+2$ *except* multiples of m larger than m. Since $f' \leq f+2$, T' contains all generators of S' less than or equal to $f'+m$, therefore all generators of S'. The set S' is a secundive semigroup, so it follows that T' is a restricted Weierstrass sequence of S'.

Notice that if $T = \{t_0 = 0, t_1 = m, t_2, \cdots, t_r\}$ and $T' = \{t'_0 = 0, t'_1 = m, t'_2, \cdots, t'_r\}$ (in increasing order), then for $i \geq 2$,

$$t'_i = \begin{cases} t_i & \text{if } t_i = 0 \text{ or } t_i = m \\ t_i + 2 & \text{if } m | (t_i + 1) \\ t_i + 1 & \text{otherwise} \end{cases}.$$

It follows from the displacement lemma 2.15 that $(T : T')$ is 1-valid. By corollary 3.4, $\widetilde{\mathcal{G}}_g^T$ has a dimensionally proper point. It follows from lemma 2.15 that $\widetilde{\mathcal{G}}_{g+1}^{T'}$ has a dimensionally proper point. By corollary 3.4 again, $\widetilde{\mathcal{W}}_{S'}$ has an effectively proper point. □

Proof of lemma 3.16. Assume that k is not the smallest positive element of S; otherwise the result follows from lemma 3.15.

Let g be the genus of S. Then S' is a primitive semigroup of genus $g+1$ by lemma 3.14. Let $T \subset S$ consist of all elements less than or equal to N, and $T' \subset S'$ consist of all elements less than or equal to $N+1$, where N is chosen large enough that both T and T' contain all generators of S and S', respectively. Since S and S' are primitive, T and T' are restricted Weierstrass sequences for S and S', respectively.

Notice that if $T = \{t_0, t_1, \cdots, t_r\}$ and $T' = \{t'_0, t'_1, \cdots, t'_r\}$ (in increasing order), then

$$t'_i = \begin{cases} t_i & \text{if } t_i = 0 \text{ or } t_i = k \\ t_i + 1 & \text{otherwise} \end{cases}.$$

Since all multiples of k are at least two more than any gaps of S', it follows from the displacement lemma 2.15 that $(T : T')$ is 1-valid. By corollary 3.4, $\widetilde{\mathcal{G}}_g^T$ has a dimensionally proper point. It follows from lemma 2.15 that $\widetilde{\mathcal{G}}_{g+1}^{T'}$ has a dimensionally proper point. By corollary 3.4 again, $\widetilde{\mathcal{W}}_{S'}$ has an effectively proper point. □

Remark 3.17. Eisenbud and Harris [EH87] used a statement essentially equivalent to this lemma ([EH87], theorem 5.4) in their proof of the existence of primitive Weierstrass semigroups. The main difference is that they consider (complete) canonical series (whose ramification at p determines the Weierstrass semigroup), whereas we consider subseries of the complete linear series $|\mathcal{O}_C(Np)|$ for large integers N. These two methods are essentially dual to each other, via Serre duality.

Proof of theorem C. By induction on the genus of S. There are two base cases.

(1) For some m, $S = \{0, m, m+1, 2m, 2m+1, 2m+2, \cdots\}$. Then S is primitive, and Komeda [K91] proved that $\widetilde{\mathcal{W}}_S$ has an effectively proper component in this case.

(2) $\epsilon(S) = 0$. In this case, $\widetilde{\mathcal{W}}_S$ is dense in $\mathcal{M}_{g,1}$ (it includes all but finitely many points on a dense open substack of curves), and the result is obvious[3].

The inductive step has two cases.

In the first case, suppose that $(m+1) \notin S$, where m is the smallest generator of S. By lemma 3.9, S is secundive, so by lemma 3.13, the semigroup $S' = D^-(S, m)$ is secundive of genus $g-1$ and $\varepsilon(S') = \varepsilon(S) - 1 \leq (g-2)$. Also, by lemma 3.12, $D^+(S', m) = S$. By inductive hypothesis, $\widetilde{\mathcal{W}}_{S'}$ has an effectively proper component, so lemma 3.16 now implies that $\widetilde{\mathcal{W}}_S$ has an effectively proper component.

In the second case, suppose that $(m+1) \in S$. By lemma 3.9, S is secundive. By definition, this means that S contains all integers greater than or equal to $2m+1$; hence S is in fact primitive. We may assume that $S \neq \{0, m, m+1, 2m, 2m+1, \cdots\}$, since otherwise we could apply the first base case; hence by lemma 3.14 $S' = D^-(S, M)$ (where M is the largest element of S that is smaller than some gap) is a primitive semigroup of genus $g-1$ and with $\varepsilon(S') = \varepsilon(S) - 1$. Lemma 3.12 implies that $D^+(S', M) = S$. Of course $(M-1) \notin S'$, so we can apply lemma 3.16: by inductive

[3] In characteristic p, it is not the case that $\widetilde{\mathcal{W}}_S$ surjects onto $\mathcal{M}_g$, since some curves consist entirely of Weierstrass points. Nevertheless, non-Weierstrass points are still dense in $\mathcal{M}_{g,1}$.

hypothesis $\widetilde{\mathcal{W}}_{S'}$ has an effectively proper component, so in turn $\widetilde{\mathcal{W}}_S$ has an effectively proper component. $\square$

4. Semigroups of large effective weight

This section is largely experimental and speculative. The goal is to reveal certain patterns which may suggest the next direction in which theorems B and C might be extended.

Throughout this section we will assume that our field K has characteristic 0 (and we assume it is algebraically closed, as we do throughout the thesis). This assumption is necessary to apply Bertini's theorem to prove that certain curves are smooth.

We have shown elsewhere in this thesis (theorem B) that for any semigroup S of genus g such that $\mathcal{W}_S$ is nonempty, the codimension of $\mathcal{W}_S$ in $\mathcal{M}_{g,1}$ is bounded by the effective weight $\varepsilon(S)$, and that this bound holds exactly in many cases (especially when the effective weight is low). One of the main motivations for the definition of effective weight was the case of hyperelliptic semigroups: those which contain 2. There is one such semigroup for every genus g, and they are notable in that they are precisely the semigroups of maximal weight, given by $\mathrm{wt}(S) = \binom{g}{2}$. One reason that the new figure $\varepsilon(S)$ is so compelling is that it is precisely equal to the codimension $g-1$ in the case of hyperelliptic curves: the effective weight tames the semigroups of extremal *weight*.

With this in mind, we may naturally ask if we can tame the case of extremal *effective* weight. We will see that like the maximum weight in a given genus, the maximum effective weight in a given genus also grows quadratically. However, an examination of the maximal effective weights in genus up to 10 suggests that the classification is not as simple as the case of weight (see figure 2).

However, starting in genus 10, a simple pattern emerges in the semigroups of large effective weight; see figure 3.

On the basis of these data, we make the following combinatorial conjectures.

genus	$\max \varepsilon(S)$	gaps	generators
1	0	1	$\langle 2,3\rangle$
2	1	$1,3$	$\langle 2,5\rangle$
3	2	$1..5$	$\langle 3,4\rangle$
		$1,3,5$	$\langle 2,7\rangle$
4	3	$1..7$	$\langle 4..6\rangle$
		$1..4,7$	$\langle 3,5\rangle$
		$1,3,5,7$	$\langle 2,9\rangle$
5	4	$1..9$	$\langle 5..8\rangle$
		$1..5,9$	$\langle 4,6,7\rangle$
		$1..6,7$	$\langle 4..11\rangle$
		$1..4,5,8$	$\langle 3,7,11\rangle$
		$1,3,5,7,9$	$\langle 2,11\rangle$
6	6	$1..8,9$	$\langle 5..7\rangle$
		$1..5,7,11$	$\langle 4,6,9\rangle$
		$1..6,7,11$	$\langle 4,5\rangle$
		$1..4,5,8,11$	$\langle 3,7\rangle$
7	8	$1..10,11$	$\langle 6..9\rangle$
		$1..5,7,9,13$	$\langle 4,6,11\rangle$
8	10	$1..12,13$	$\langle 7..11\rangle$
		$1..5,7,9,11,15$	$\langle 4,6,13\rangle$
9	12	$1..14,15$	$\langle 8..13\rangle$
		$1..11,12,13$	$\langle 7..10\rangle$
		$1..9,10..17$	$\langle 6..8\rangle$
		$1..5,7,9,11,13,17$	$\langle 4,6,15\rangle$
10	15	$1..13,14,15$	$\langle 8..12\rangle$

FIGURE 2. Semigroups of maximal effective weight, up to genus 10.

Conjecture 4.1. *For all positive integers g, the maximum value of $\varepsilon(S)$, where S ranges over over all numerical semigroups of genus g, is* $\left\lfloor \dfrac{(g+1)^2}{8} \right\rfloor$.

genus	$\max \varepsilon(S)$	gaps	generators
11	18	$1..15, 16, 17$	$\langle 9..14\rangle$
12	21	$1..17, 18, 19$	$\langle 10..16\rangle$
13	24	$1..19, 20, 21$	$\langle 11..18\rangle$
		$1..16, 17..19$	$\langle 10..15\rangle$
14	28	$1..18, 19..21$	$\langle 11..17\rangle$
15	32	$1..20, 21..23$	$\langle 12..19\rangle$
16	36	$1..22, 23..25$	$\langle 13..21\rangle$
17	40	$1..24, 25..27$	$\langle 14..23\rangle$
		$1..21, 22..25$	$\langle 13..20\rangle$
18	45	$1..23, 24..27$	$\langle 14..22\rangle$
19	50	$1..25, 26..29$	$\langle 15..24\rangle$
20	55	$1..27, 28..31$	$\langle 16..26\rangle$
21	60	$1..29, 30..33$	$\langle 17..28\rangle$
		$1..26, 27..31$	$\langle 16..25\rangle$
22	66	$1..28, 29..33$	$\langle 17..27\rangle$
23	72	$1..30, 31..35$	$\langle 18..29\rangle$
24	78	$1..32, 33..37$	$\langle 19..31\rangle$
25	84	$1..34, 35..39$	$\langle 20..33\rangle$
		$1..31, 32..37$	$\langle 19..30\rangle$
26	91	$1..33, 34..39$	$\langle 20..32\rangle$
27	98	$1..35, 36..41$	$\langle 21..34\rangle$
28	105	$1..37, 38..43$	$\langle 22..36\rangle$
29	112	$1..39, 40..45$	$\langle 23..38\rangle$
		$1..36, 37..43$	$\langle 22..35\rangle$
30	120	$1..38, 39..45$	$\langle 23..37\rangle$

FIGURE 3. Semigroups of maximal effective weight, genus 11 through 30.

Conjecture 4.2. *If $g(S) \geq 10$, then $\varepsilon(S) = \left\lfloor \frac{(g+1)^2}{8} \right\rfloor$ if and only if S is one of the following semigroups.*

$$\begin{aligned} S &= \langle d-r+1, d-r+2, \cdots, d-1, d\rangle \\ &= \mathbf{N}\backslash\left(\{1,2,\cdots,d-r\} \cup \{d+1, d+2, \cdots, 2(d-r+1)-1\}\right) \end{aligned}$$

where $d = \frac{5}{4}g + \frac{1}{4} - \frac{3}{4}\eta$

and $r = \frac{1}{2}g + \frac{1}{2} - \frac{1}{2}\eta$

for some integer η such that $|\eta| \leq 2$.

Note in particular that in conjecture 4.2, η must be selected such that $\eta \equiv 3 - g \mod 4$ among the choices $\{-2, -1, 0, 1, 2\}$; this explains why there are two extremal semigroups when $g \equiv 3 \mod 4$, and there is only one otherwise. It is tedious but easy to check that the semigroups defined in this way do indeed have effective weight equal to $\left\lfloor \dfrac{(g+1)^2}{8} \right\rfloor$. In fact, they are also *primitive*, so their weight is equal to their effective weight (but of course they are not extremal with respect to weight). We will prove the following result about Weierstrass points with these extremal semigroups.

Proposition 4.3. *Suppose that $g \geq 6$ and S is one of the semigroups mentioned in conjecture 4.2. Then $\widetilde{\mathcal{W}}_S$ has codimension $\frac{7}{4}g - \frac{17+\eta}{4}$ in $\mathcal{M}_{g,1}$.*

This proposition will follow from corollary 4.13. It shows that the semigroups of maximal effective weight present a challenge: to find a new refinement of the effective weight that gives the correct codimension.

Question 4.4. *Is there a combinatorial quantity $\varepsilon'(S)$ associated to each semigroup which is an upper bound on the codimension of $\widetilde{\mathcal{W}}_S$ in $\mathcal{M}_{g,1}$, and which is equal to $\frac{7}{4}g - \frac{17+\eta}{4}$ for any of the semigroups S described in conjecture 4.2?*

The semigroups that appear as semigroups of maximal effective weight have a structure that turns out to have interesting geometric implications. The rest of this section is devoted to the study of a broad class of semigroups, including those mentioned in conjecture 4.2, which give interesting loci $\widetilde{\mathcal{W}}_S$ that can be studied rather explicitly. In particular, these semigroups provide a large class of examples of Weierstrass points which occur in larger dimensional families than what is predicted by the effective weight. They also provide a large class of semigroups for which $\widetilde{\mathcal{W}}_S$ is reducible, and even has impure dimension. We begin by defining the semigroups in this class.

Definition 4.5. Let d, r be positive integers such that $r \geq 2$ and $d \geq 2r + 1$. Define the semigroup $S(d, r)$ as follows.

$$S(d,r) = \langle d-r+1, d-r+2, \cdots, d\rangle$$

Such semigroups will be called *Castelnuovo semigroups.*

We name these semigroups after Castelnuovo because we shall see that they can only occur on Castelnuovo curves (which we will define and discuss presently; see chapter 3 of [H82] for a thorough discussion).

The gaps of $S(d,r)$ are arranged in several shrinking blocks of consecutive integers, beginning with a block of length $d-r$ and decreasing in size by $r-1$ at each stage (until they become empty). Therefore we can find the genus as follows.

$$\begin{aligned} g(S(d,r)) &= (d-r) + (d-r-(r-1)) + (d-r-2(r-1)) + \\ &\quad \cdots + (d-r+m(r-1)) \\ &= \binom{m}{2}(r-1) + me \\ \text{where} \qquad & m = \left\lfloor \frac{d-1}{r-1} \right\rfloor \text{ and } e = (d-1) - m(r-1). \end{aligned}$$

We can also easily compute the effective weight.

$$\varepsilon(S(d,r)) = r(g-d+r)$$

Conjectures 4.1 and 4.2 are not difficult to verify for Castelnuovo semigroups.

Proposition 4.6. *For any Castelnuovo semigroup $S = S(d,r)$ of genus g (with $r \geq 2$ and $d \geq 2r+1$ as usual), $\varepsilon(S) \leq \left\lfloor \frac{(g+1)^2}{8} \right\rfloor$. If $g \geq 10$, then equality holds only in the cases mentioned in the statement of conjecture 4.2.*

Proof. Fix positive integers d, e, and define integers m, e such that

$$(d-1) = m(r-1) + e,$$

where $e \in [0, r-2]$. Let $S = S(d,r)$. Then using the formulas above for the genus and effective weight of S, we can re-express $\varepsilon(S)$ as follows.

$$\begin{aligned}
\varepsilon(S) &= r\left(\binom{m}{2}(r-1) + me - m(r-1) - e + r - 1\right) \\
&= r\left(\binom{m-1}{2}(r-1) + (m-1)e\right) \\
&= \frac{m-1}{m} \cdot r\left(\frac{m(m-2)}{2}(r-1) + me\right) \\
&= \frac{m-1}{m} \cdot r\left(\binom{m}{2}(r-1) - \frac{m}{2}(r-1) + me\right) \\
&= \frac{m-1}{m} \cdot r\left(g + \frac{m}{2} - \frac{m}{2}r\right) \\
&= \frac{2(m-1)}{m^2} \cdot \left(\frac{m}{2}r\right)\left(g + \frac{m}{2} - \frac{m}{2}r\right)
\end{aligned}$$

Applying the arithmetic-geometric mean inequality, we obtain the following bound on the effective weight in terms of g and m.

$$\varepsilon(S) \leq \frac{m-1}{2m^2}\left(g + \frac{m}{2}\right)^2 \tag{4.1}$$

Consider first the case $m = 2$. Then taking floors of both sides, inequality 4.1 gives the desired bound $\varepsilon(S) \leq \left\lfloor \frac{(g+1)^2}{8} \right\rfloor$. Consider the equality case. Notice that $\left\lfloor \frac{(g+1)^2}{8} \right\rfloor = \frac{(g+1)^2}{8} - \zeta$, where ζ is one of $\frac{1}{8}, \frac{1}{2}$, or 0, according to whether g is even, $1 \mod 4$, or $3 \mod 4$ (respectively). Looking at the last step of the derivation of inequality 4.1 and setting $m = 2$, it follows that $\varepsilon(S) - \frac{(g+1)^2}{8}$ is given exactly by $\frac{1}{8}(g+1-2r)^2$. Therefore, $\varepsilon(S) = \left\lfloor \frac{(g+1)^2}{8} \right\rfloor$ if and only if $|r - \frac{1}{2}(g+1)| \leq 1$. Also, since

$g = (r-1) + 2e$, it follows that $g - r + 1$ must be even. Conversely, if g, r are chosen such that $|r - \frac{1}{2}(g+1)| \le 1$ and $r \equiv (g+1) \mod 2$, then letting $e = \frac{1}{2}(g+1-r)$ and $d = 2(r-1) + e$ gives a semigroup $S = S(d, r)$ such that $\varepsilon(S) = \left\lfloor \frac{(g+1)^2}{8} \right\rfloor$. Given g, the possible choices of r are precisely $r = \frac{1}{2}(g+1) + \frac{1}{2}\eta$, where η is an integer such that $|\eta| \le 2$ and $\eta \equiv (g+1) \mod 4$. Therefore, the equality cases such that $m = 2$ are precisely parameterized by

$$
\begin{aligned}
d &= \frac{5}{4}g + \frac{1}{4} = \frac{3}{4}\eta \\
r &= \frac{1}{2}g + \frac{1}{2} - \frac{1}{2}\eta \\
& \quad \eta \in \mathbf{Z}, \ |\eta| \le 2.
\end{aligned}
$$

Now consider the case $m \ge 3$. We know that $\varepsilon(S) \le \frac{m-1}{2m^2}\left(g + \frac{1}{2}m\right)^2$ (equation 4.1). First, we show that $\frac{m-1}{2m^2}\left(g + \frac{1}{2}m\right)^2 \le \frac{1}{9}(g + \frac{3}{2})^2$. To see this, fix g and consider the function $f(x) = \frac{x-1}{2x^2}(g + \frac{1}{2}x)^2$. Taking the logarithmic derivative,

$$
\begin{aligned}
\frac{f'(x)}{f(x)} &= \frac{1}{x-1} - \frac{2}{x} + \frac{2}{2g+x} \\
&= \frac{1}{x-1} - \frac{4g}{x(x+2g)} \\
&= \frac{x^2 - 2gx + 4g}{(x-1)x(x+2g)} \\
&= \frac{(x-g)^2 - (g^2 - 4g)}{(x-1)x(x+2g)}.
\end{aligned}
$$

It follows from this that as long as $g \ge 5$, $f(x)$ is a decreasing function on the interval $(g - \sqrt{g^2 - 4g}, g + \sqrt{g^2 - 4g})$, which contains the interval $[3, g]$. The formula $g = \binom{m}{2}(r-1) + me$ implies that $g \le \binom{m}{2} \le m$ (here we use $m \ge 3$), so m lies in this interval, and it follows that $\varepsilon(S) \le f(m) \le f(3) = \frac{1}{9}(g + \frac{3}{2})^2$, as claimed.

Next as long as $g \ge 11$, the following inequality holds.

$$\frac{1}{9}\left(g+\frac{3}{2}\right)^2 < \frac{1}{8}(g+1)^2 - \frac{1}{2}$$

Indeed, this inequality is equivalent to $(g-3)^2 > 54$, i.e. $|g-3| \geq 8$. It follows that for all $g \geq 11$, $\varepsilon(S) < \frac{(g+1)^2}{8} - \frac{1}{2} \leq \left\lfloor \frac{(g+1)^2}{8} \right\rfloor$.

Together with our enumeration of all semigroups of genus less than 11 (figure 2), this implies that $\varepsilon(S) \leq \left\lfloor \frac{(g+1)^2}{8} \right\rfloor$ for all Castelnuovo semigroups, and that if $g \geq 10$ then the only equality cases satisfy $m = 2$ and are precisely those described in conjecture 4.2. $\square$

The expression for the genus of $S(d, r)$ is intriguing due to the following theorem (which is our reason for naming these semigroups as we did).

Theorem 4.7 (Castelnuovo)**.** *Suppose that C is a reduced, irreducible, and non-degenerate curve of degree d in $\mathbf{P}^r$, where $d \geq 2r+1$. Let $m = \left\lfloor \frac{d-1}{r-1} \right\rfloor$ and let e be the remainder. Then the arithmetic genus of C is bounded by:*

$$p_a(C) \leq \binom{m}{2}(r-1) + me.$$

If equality holds in this expression, then the set-theoretic intersection of the quadric hypersurfaces containing C is an irreducible and non-degenerate surface of degree $r-1$.

The curves where equality holds in theorem 4.7 are called *Castelnuovo curves.*

Lemma 4.8. *Suppose that $S = S(d, r)$ is a Castelnuovo semigroup of genus g, and $(C, p) \in \mathcal{W}_S$ is a Weierstrass point. Then $\mathcal{O}_C(dp)$ is very ample, and embeds C in $\mathbf{P}^r$ as a Castelnuovo curve.*

Proof. By definition of the Weierstrass semigroup,

$$h^0(\mathcal{O}(dp)) = r+1, \text{ and}$$
$$h^0(\mathcal{O}((d-1)p)) = r.$$

Therefore $\mathcal{O}_C(dp)$ is base point free, and does indeed give a map $f: C \to \mathbf{P}^r$. Let N be an integer larger than $\frac{2g+1}{d}$, and let $\nu_N: \mathbf{P}^r \to \mathbf{P}^M$ be the Veronese map (here $M = \binom{r+N}{N} - 1$). Then the composition $\nu_N \circ f: C \to \mathbf{P}^M$ is given by the linear series $(\mathcal{O}_C(Ndp), V)$, where V is the image of $\mathrm{Sym}^N H^0(\mathcal{O}_C(dp)) \to H^0(\mathcal{O}_C(Ndp))$. Regard sections of $\mathcal{O}_C(dp)$ and $\mathcal{O}_C(Ndp)$ as rational functions on C. The pole orders of sections in $H^0(\mathcal{O}_C(dp))$ include all generators of the Weierstrass semigroup, and also 0. Therefore N-fold products of these sections give sections with all possible pole orders up to Nd. This implies that in fact V is all of $H^0(\mathcal{O}_C(Ndp))$. Since $Nd \geq 2g+1$, it follows that $\mathcal{O}_C(Ndp)$ is very ample, so $\nu_N \circ f$ is an embedding. Therefore f must also be an embedding. The fact that the image in $\mathbf{P}^r$ is a Castelnuovo curve follows from the calculation of the genus of $S(d,r)$ (which is also the genus of C). $\square$

The reason that Castelnuovo curves are particularly easy to study explicitly is that such a curve can be studied using the geometry of the surface cut out by the quadrics containing the curve. In fact, this surface is of a very simple kind, due to the following theorem. A proof of a more general statement, and a nice discussion of the surrounding ideas, can be found in [EH85].

Theorem 4.9 (Bertini)**.** *If S is a reduced, irreducible, and non-degenerate surface of degree $r-1$ in $\mathbf{P}^r$, then S must either be the Veronese surface in $\mathbf{P}^5$, or a rational normal surface scroll.*

We will now compute the dimension of $\widetilde{\mathcal{W}}_S$, when S is a Castelnuovo semigroup. This computation will reveal that these semigroups give a large supply of semigroups whose points are not effectively proper. We begin from the following fact, which is a relatively straightforward calculation (the necessary formulas can be found in the following two subsections). We simplify the statement slightly by restricting to the case $r \geq 6$, but the full statement simply has a few additional cases.

Proposition 4.10 ([H82] corollary 3.12)**.** *Let d, r be integers such that $d \geq 2r + 1$ and $r \geq 6$, define m, e as above, and let $g = \binom{m}{2}(r-1) + me$. Let $\mathcal{H}$ denote the union of the (open) components of the Hilbert scheme whose members correspond to smooth, irreducible, and non-degenerate curves of degree d and genus g in $\mathbf{P}^r$. Then:*

(1) There is a component of $\mathcal{H}$ of dimension

$$C(d,r) = \left(\binom{m+1}{2} + r + 2\right)(r-1) + (m+2)(e+2) - 4.$$

(2) If $e = 0$, then $\mathcal{H}$ has a second component of dimension $C(d,r) + r - 3$.

These are the only components of $\mathcal{H}$ (in particular, $\mathcal{H}$ is irreducible if $e = 0$).

Remark 4.11. Eisenbud and Harris prove in [H82] that the under weakened inequalities on the genus, most of Castelnuovo's argument remains valid, and it is still possible to draw conclusions about the geometry of the curves in question (most notably, about the rational normal scrolls they sit on). It would be interesting to try to apply these results to a broader class of numerical semigroups than the Castelnuovo semigroups.

If we denote by $\mathcal{H}'$ the incidence correspondence parameterizing points $[C] \in \mathcal{H}$, points $p \in C \subseteq \mathbf{P}^r$, and hyperplanes H containing p, an easy calculation shows that $\dim \mathcal{H}' = \dim \mathcal{H} + r$. We can consider the locus $\mathcal{H}'' \subseteq \mathcal{H}'$ corresponding to $\{(C, p, H) : C \cap H = dp\}$ (where this is equality as divisors). Since every Castelnuovo curve lies on a unique surface scroll in $\mathbf{P}^r$, there is a map from $\mathcal{H}'$ to the parameter space of triples (S, H, p), where S is a surface scroll, H is a hyperplane section, and $p \in H$ is a point. It follows from a dimension count in the following section that this parameter space has dimension $(r+1)(r+2) - 7$. We will show in theorem 4.20 that, for a *general* triple (S, H, p), the fiber of $\mathcal{H}''$ over (S, H, p) has codimension exactly d in the fiber of $\mathcal{H}'$ over (S, H, p). From this we obtain the following lemma.

Lemma 4.12. *The locus $\mathcal{H}''$ (as defined above) has a component of dimension exactly $\dim \mathcal{H} - (d - r) + 1$.*

Notice that there is a surjective map from $\mathcal{H}''$ to $\widetilde{\mathcal{W}}_S$, all of whose fibers are orbits by the action of PGL_{r+1}. Therefore an easy calculation immediately shows the following.

Corollary 4.13. *For any Castelnuovo semigroup S as above, with $r \geq 6$, $\widetilde{\mathcal{W}}_S$ has a component of codimension $2g - 2m - e$ in $\mathcal{W}_{g,1}$. In case $e = 0$, there is a second component of codimension $2g - 2m - e - r + 3$.*

Because the effective weight of these semigroups is unbounded, whereas these codimension are all less than $2g$, this shows that Castelnuovo semigroups are a large class of semigroups whose codimensions in $\mathcal{M}_{g,1}$ are much smaller than their effective weight. Another intriguing consequence of corollary 4.13 is that it shows $\widetilde{\mathcal{W}}_S$ is sometimes reducible, and sometimes does not have pure dimension.

The rest of this section contains an analysis of curves on rational normal surface scrolls. Much of this analysis can be found in [H82], but we reproduce it here for our convenience. The author believes that a more careful analysis of arbitrary fibers of the map from $\mathcal{H}'$ to the space of triples (S, H, p) can likely be used to show that the components of $\mathcal{W}_S$ described above are, in fact, the only components.

4.1. **Background on rational normal scrolls.** We recall some basic properties of rational normal surface scrolls, needed in the analysis above. None of this is original, but we have collected it in one place for convenience.

Let M be a $2 \times (r-1)$ matrix of linear forms on $\mathbf{P}^r$.

$$M = \begin{pmatrix} L_{11} & L_{12} & \cdots & L_{1,r-1} \\ L_{21} & L_{22} & \cdots & L_{2,r-1} \end{pmatrix} \quad L_{i,j} \in H^0(\mathcal{O}_{\mathbf{P}^r}(1))$$

We will study the scheme $S(M)$ cut out by the 2×2 minors of M. In particular, $S(M)$ remains unchanged if row or column operations are performed on M. If M is 1-generic (defined below), $S(M)$ will be called a *rational normal surface scroll.*

Call M 1-*generic* if there do not exist vectors $v \in K^2$, and $w \in K^{r-1}$ such that $v^T M w = 0$. If M is 1-generic, then by multiplying by invertible matrices (with coefficients in K) on the left and right, it can be arranged that $L_{2,i} = L_{1,i+1}$ for all values of i in $\{1, 2, \cdots, r-2\}$ except possibly one. Call this exception a (or let $a = r - 1$ if there is no exception). That M is 1-generic implies that these equalities span all K-linear relations between the forms $L_{i,j}$. Therefore there is a basis $X_0, X_1, \cdots, X_a, Y_0, Y_1, \cdots, Y_b$ (where $a + b = r - 1$) for $H^0(\mathcal{O}_{\mathbf{P}^r})$ such that M can be written in the following form.

$$M = \begin{pmatrix} X_0 & X_1 & \cdots & X_{a-1} & Y_0 & Y_1 & \cdots & Y_{b-1} \\ X_1 & X_2 & \cdots & X_a & Y_1 & Y_2 & \cdots & Y_b \end{pmatrix}$$

Without loss of generality, also assume that $a \leq b$. An easy calculation shows that $S(M)$ is nonsingular if and only if $b > 0$, whereas if $b = 0$ then $S(M)$ is a cone over a rational normal curve, with vertex given by $\{X_i = 0\}$. Define $\delta = b - a$; this measures the extent to which the scrolls fails to be "balanced."

Notice that, due to this parameterization, rational normal surface scrolls form a dense open subset of a component of the Hilbert scheme, of dimension $(r+1)^2 - 7$ (this can be computed by parametrizing all ways of choosing $L_{i,j}$, then adjusting for scaling, row operations and column operations). Within this component, the number δ is upper semicontinuous, hence the general member has $\delta \leq 1$ (note that δ is always congruent to $r - 1$ modulo 2). We will not need it here, but it is not difficult to determine that the locus of scrolls with invariant $\delta > 1$ has codimension δ.

Let S denote the Hirzebruch surface $\mathbf{P}(\mathcal{O}_{\mathbf{P}^1}(a) \oplus \mathcal{O}_{\mathbf{P}^1}(b))$ over $\mathbf{P}^1$. Denote the tautological line bundle of S by $\mathcal{O}_S(H)$; then $H^0(\mathcal{O}_S(H)) \cong H^0(\mathcal{O}_{\mathbf{P}^1}(a)) \oplus H^0(\mathcal{O}_{\mathbf{P}^1}(b))$.

Let $\{W, Z\}$ be a basis for $H^0(\mathcal{O}_{\mathbf{P}^1}(1))$, and define an isomorphism $H^0(\mathcal{O}_S(H)) \to H^0(\mathcal{O}_{\mathbf{P}^r})$ as follows.

$$\begin{aligned}
H^0(\mathcal{O}_{\mathbf{P}^1}(a)) &\to \operatorname{span}\langle X_0, X_1, \cdots, X_a\rangle \\
W^i Z^{a-i} &\mapsto X_i \\
H^0(\mathcal{O}_{\mathbf{P}^1}(b)) &\to \operatorname{span}\langle Y_0, Y_1, \cdots, Y_b\rangle \\
W^i Z^{b-i} &\mapsto Y_i
\end{aligned}$$

One can verify that the image of the resulting map $S \to \mathbf{P}^r$ is precisely $S(M)$. Indeed, the induced map $\iota : S \to S(M)$ is an isomorphism if $b > 0$, while otherwise it collapses the section $[0, 1]$ (the directrix) to a single point, resulting in a cone over a rational normal curve.

The Picard group of S has two generators: the directrix section D (given by $[0, 1]$, and unique if and only if $\delta > 0$) and the fiber class F. The intersection pairing is as follows.

$$\begin{aligned}
D^2 &= -\delta \\
F^2 &= 0 \\
D \cdot F &= 1
\end{aligned}$$

The canonical divisor is given as follows.

$$K = -2D - (\delta + 2)F$$

Therefore the Euler characteristics and arithmetic genera of divisors on S are given as follows.

$$
\begin{aligned}
\chi(\alpha D + \beta F) &= (\alpha+1)(\beta+1) - \binom{\alpha+1}{2}\delta \\
p_a(\alpha D + \beta F) &= (\alpha-1)(\beta-1) - \binom{\alpha}{2}\delta
\end{aligned}
$$

Note in particular that the smooth members of $|D + \beta F|$ are all rational: these arise as hyperplane sections of S when it is embedded as a rational normal surface scroll. We will see shortly that the series $|D+\beta F|$ has a smooth member if and only if $\beta = 0$ or $\beta \geq \delta$.

For convenience, we will compute the dimension of the first cohomology of all effective divisors on S. Together with Serre duality, this will make it possible to compute the dimension of all linear systems on S. We begin by determining which divisor classes are effective.

Lemma 4.14. *The divisor $\alpha D + \beta F$ satisfies $h^0(\alpha D + \beta F) > 0$ if and only if $\alpha \geq 0$ and $\beta \geq 0$.*

Proof. Since D and F are effective themselves, one direction is obvious. Conversely, suppose that $\alpha D + \beta F$ is effective. Notice that both of the classes F and $D + \delta F$ are base point free: the first because $|F|$ contains all fibers of the map $S \to \mathbf{P}^1$, the second because $|D + \delta F|$ contains the images of all of the sections $\mathbf{P}^1 \to S$ given by $[z, w] \mapsto [1, p(z, w)]$, where p is a degree δ homogeneous polynomial. Therefore

$$
\begin{aligned}
(\alpha D + \beta F) \cdot F &= \alpha \\
(\alpha D + \beta F) \cdot (D + \delta F) &= \beta
\end{aligned}
$$

Therefore $\alpha \geq 0$ and $\beta \geq 0$. $\square$

To state the computation of the first cohomology numbers, the following notation will be useful.

Definition 4.15. Let $f_\delta(n)$ denote the following piecewise-linear function.

$$
\begin{aligned}
f_\delta(n) &= \begin{cases} 0 & \text{if } n < 0 \\ \binom{q+1}{2}\delta + (q+1)r & \text{otherwise} \end{cases} \\
\text{where} \qquad & q = \left\lfloor \frac{x}{\delta} \right\rfloor \text{ and } x = q\delta + r.
\end{aligned}
$$

Another way to describe the function $f_\delta(n)$ is that $f_\delta(0) = 0$, and f is linear with slope $q+1$ on the interval $[q\delta, (q+1)\delta]$. In fact, we have met this function elsewhere: the value $f_{r-1}(d-r)$ computes the maximum genus of a smooth curve of degree d in $\mathbf{P}^r$. This appears to be a coincidence.

Lemma 4.16. *Suppose that $\alpha, \beta \geq 0$. Then $h^1(\alpha D + \beta F) = f_\delta(\alpha\delta - \beta - 1)$. In particular, $h^1(\alpha D + \beta F) = 0$ if and only if $\beta > \alpha\delta$.*

Proof. Proceed by induction on α. If $\alpha = 0$, then the complete system $|\beta F|$ consists of the pullback of the complete linear system of β points on $\mathbf{P}^1$, hence $h^0(\beta F) = \beta + 1 = \chi(\beta F)$, and the result follows. Now assume that $\alpha > 0$, and the result holds for all smaller values of α.

We will make use of the following exact sequence of sheaves.

$$0 \to \mathcal{O}_S((\alpha - 1)D + \beta F) \to \mathcal{O}_S(\alpha D + \beta F) \to \mathcal{O}_D(\beta - \alpha\delta) \to 0$$

By forming the associated long exact sequence in cohomology, we can draw two conclusions.

- If $\beta - \alpha\delta \leq -1$ then $h^0((\alpha - 1)D + \beta F) = h^0(\alpha D + \beta F)$.
- If $\beta - \alpha\delta \geq -1$ and $h^1((\alpha - 1)D + \beta F) = 0$, then $h^1(\alpha D + \beta F) = 0$.

If $\beta \geq \alpha\delta - 1$, then the result follows immediately from conclusion 2 and the induction hypothesis. Otherwise, we apply hypothesis 1 to deduce that

$$\begin{aligned} h^1(\alpha D + \beta F) &= h^1((\alpha-1)D + \beta F) + \chi((\alpha-1)D + \beta F) - \chi(\alpha D + \beta F) \\ &= h^1((\alpha-1)D + \beta F) + \alpha\delta - \beta - 1 \end{aligned}$$

Now let $(\alpha\delta - \beta - 1) = q\delta + r$, where q, r are integers with $0 \leq r < \delta$. Then if $q = 0$, it follows that $h^1((\alpha-1)D+\beta F) = 0$, hence $h^1(\alpha D+\beta R) = r$. Otherwise, the inductive hypothesis implies that $h^1((\alpha-1)D+\beta F) = \binom{q}{2}\delta+rq$, hence $h^1(\alpha D+\beta F) = \left(\binom{q}{2} + q\right)\delta + r(q+1)$, which is the desired result. This completes the induction. $\square$

Corollary 4.17. *The divisor class $C = \alpha D + \beta F$ (with $\alpha, \beta \geq 0$) contains a smooth member if and only if either $\beta \geq \alpha\delta$ or $\beta - \alpha\delta = -\delta$. The series $|C|$ cuts out a complete linear series on the directrix D if and only if $\beta \geq \alpha\delta$; otherwise all member of $|C|$ contain D as a component.*

Proof. If $C \cdot D \geq 0$, then $\beta \geq \alpha\delta$, and therefore $h^1(C - D) = 0$, so the map $H^0(\mathcal{O}_S(C)) \to H^0(\mathcal{O}_D(C \cdot D))$ is surjective. In particular, $|C|$ has no base points on D. Since $|C|$ includes all curves of the form αD plus any β fibers, the only possible base points of $|C|$ lie on D. Thus the general member of $|C|$ is smooth, by Bertini's theorem, and the series cuts out a complete linear series on D.

On the other hand, if $C \cdot D < 0$, then any member of $|C|$ contains D as a component. Hence this member can be smooth only if it is a union of D and a curve in the class $C - D$ that doesn't meet D. This is possible if and only if $(C - D) \cdot D = 0$ and the general member of $C - D$ does not contain D; by the previous case, this is true if and only if $C \cdot D = -\delta$. In this case, the general member of $|C|$ consist of the union of D and a disjoint smooth curve. $\square$

Corollary 4.18. *Suppose that H is a smooth curve of class $D+\gamma F$ with $\gamma > 0$, and C is any effective divisor class on S. Then $|C|$ cuts out a complete linear series on H if and only if either $\gamma = \delta$ or $\beta - \alpha\delta + 1 \geq \gamma - \delta$.*

Proof. By the previous corollary, in fact γ must be at least δ. It follows that $H \cdot C = -\alpha\delta + \alpha\gamma + \beta \geq -1$. Since $p_a(H) = 0$, it follows that $h^1(\mathcal{O}_H(H \cdot C)) = 0$. Therefore $H^0(\mathcal{O}_S(C)) \to H^0(\mathcal{O}_H(H \cdot C))$ is surjective if and only if $h^1(\mathcal{O}_S(C-H)) = h^1(\mathcal{O}_S(C))$. Lemma 4.16 now gives the result. $\square$

4.2. **Castelnuovo Weierstrass points on balanced scrolls.** We will now consider the case of interest in the previous section: those curves on S which become Castelnuovo curves under some embedding of S as a rational normal scroll.

Fix a surface S with a map $\iota : S \to S(M) \subseteq \mathbf{P}^r$ sending S to a rational normal scroll, as in the previous subsection. Let the integers a, b, δ be as described there. Recall that they satisfy the following relations.

$$\begin{aligned} 0 \leq \quad a \quad &\leq b \\ a + b \quad &= \quad r - 1 \\ b - a \quad &= \quad \delta \end{aligned}$$

Recall that the map $\iota : S \to S(M)$ sends S to a rational normal scroll (possibly contracting the directrix in the process) and that the hyperplane class in $\mathbf{P}^r$ pulls back to a class $H = D + bF$.

Throughout this subsection, we will sometimes assume that $S(M)$ is a *balanced* scroll, i.e. that it belongs to the generic isomorphism type among surface scrolls in $\mathbf{P}^r$. This amounts to assuming that $\delta \leq 1$, so that $a = \left\lfloor \frac{r-1}{2} \right\rfloor$ and $b = \left\lceil \frac{r-1}{2} \right\rceil$.

Fix an integer $d \geq 2r+1$. Define numbers m, e as before: $m = \left\lfloor \frac{d-1}{r-1} \right\rfloor$ and $e = (d-1) - m(r-1)$. Then a Castelnuovo curve on $S(M)$ is a smooth and irreducible curve C of degree d and genus g, where

$$g = \binom{m}{2}(r-1) + me.$$

More intrinsically, a Castelnuovo curve C on S is a curve such that

$$\begin{aligned} C \cdot H &= d, \text{ and} \\ p_a(C) &= g. \end{aligned}$$

First, we observe that there is at least one divisor class, and at most two of them, whose smooth members are Castelnuovo curves.

Lemma 4.19. *Let C be the divisor class $\alpha D + \beta F$. Then the smooth members of $|C|$ are Castelnuovo curves (of degree d in $\mathbf{P}^r$) if and only if the following conditions hold.*

$$\begin{aligned} \beta &= d - \alpha \cdot a \\ \alpha &\in \{m, m+1\} \\ \textit{where} \qquad & \alpha = m \textit{ is allowed if and only if } e = 0. \end{aligned}$$

Proof. The equation $\beta = d - \alpha \cdot a$ is equivalent to $C \cdot H = d$. Assume that it holds; it suffices to determine the possible values of α. From the genus formula,

$$
\begin{aligned}
p_a(\alpha D + \beta F) &= (\alpha-1)(\beta-1) - \binom{\alpha}{2}\delta \\
&= (\alpha-1)(d-\alpha a-1) - \frac{1}{2}(\alpha-1)\alpha\delta \\
&= \frac{1}{2}(\alpha-1)(2d-2-\alpha(2a+\delta)) \\
&= \frac{1}{2}(\alpha-1)(2d-2-\alpha(r-1))
\end{aligned}
$$

Regarding this expression as a quadratic function of a real variable α, notice that the maximum occurs at $\alpha = \frac{d-1}{r-1} + \frac{1}{2} = m + \frac{e}{r-1} + \frac{1}{2}$. Restricting α to integer values, the maximum occurs at the nearest integer, or possibly occurs twice if there is a tie for the nearest integer. Since $0 \le e < r-1$, we see that the maximum value occurs at $\alpha = m+1$ for certain; if $e = 0$ then the maximum occurs a second time at $\alpha = m$. In either case, the maximum genus is

$$
\begin{aligned}
\frac{1}{2}m(2d-2-(m+1)(r-1)) &= m(d-1) - \binom{m+1}{2}(r-1) \\
&= m(m(r-1)+e) - \binom{m+1}{2}(r-1) \\
&= \binom{m}{2}(r-1) + me
\end{aligned}
$$

This is the desired genus g, and therefore this value of $p_a(\alpha D + \beta F)$ occurs when $\beta = d - \alpha \cdot a$ and $\alpha = m+1$, or $e = 0$ and $\alpha = m$, and in no other cases. $\square$

We now state a theorem about the existence of Castelnuovo Weierstrass points. For simplicity, we assume that the scroll is balanced, but similar analysis would work for scrolls of sufficiently low values of δ, given specific values of d and r.

Theorem 4.20. *Suppose that d, r are integers such that $d \geq 2r+1$ and $r \geq 3$. Let S be a balanced scroll, i.e. with $\delta \in \{0,1\}$, $\delta \equiv r-1 \mod 2$. Let H be any smooth member of the linear series $|D + \frac{r-1+\delta}{2}F|$, and $p \in H$ any point.*

For each divisor class C giving rise to Castelnuovo curves of degree d in $\mathbf{P}^r$, there is a codimension d subseries of $|C|$ on S, whose general member is a smooth curve meeting H to degree d at p.

Proof. First consider the case where $r-1$ does not divide $d-1$. Then by lemma 4.19, any such curve C has class $\alpha D + \beta F$, where α and β are as follows.

$$\begin{aligned} \alpha &= m+1 \\ \beta &= d-(m+1)a \end{aligned}$$

Therefore we can calculate that

$$\begin{aligned} \beta + 1 - \alpha\delta &= d + 1 - (m+1)b \\ &= 2 - \delta + e + (m-1)a \end{aligned}$$

Now, consider the map $H^0(\alpha D + \beta F) \to H^0(\mathcal{O}_H(d))$. By corollary 4.18 in the case $\gamma = b$, it follows that this map is surjective if $2 - \delta + e + (m-1)a \geq a$, which is certainly true if $\delta \leq 1$. So $|C|$ cuts out a complete linear series on H. In particular, a codimension d linear series within $|C|$ meets H to order d at p, and the general curve in this series does not contain all of H. Call this linear series L.

It remains to show that the general member of L is smooth. To see this, notice that L contains a sub series given by taking the complete linear series $|C-H|$ and adding H to it. Corollary 4.17 implies that $|C-H|$ is base point free (since any base points would have to lie on the directrix D), hence any base points of L must lie on H. In particular, since L also contains members that meet H only at p, the only base point of L is the point p. Since the general member of $|C - H|$ does not meet p and p is a smooth point of H, it follows that the generic member of L is smooth at p. Since p is the only base point, it follows by Bertini's theorem that the general member of L is smooth.

Now consider the case where $r - 1$ divides $d - 1$. In this case there is a second divisor class to consider, where $\alpha = m$. In this case,

$$\begin{aligned} \beta + 1 - \alpha\delta &= d + 1 - mb \\ &= 2 + e + ma \end{aligned}$$

The rest of the analysis is analogous to the previous case. $\square$

5. Twisted Weierstrass points

Our object in this section is the proof of theorem A. To prove it, we will instead prove a somewhat more general result, corollary 5.17, which will imply theorem A when combined with a combinatorial argument (lemma 6.20, proved in section 6). Corollary 5.17 is a generalized form of the "pointed" Brill-Noether theorem 1.11 to negative values of ρ. The strength of the resulting bound depends substantially on some nontrivial combinatorics, which we study separately in section 6.

For this section, we change our vocabulary slightly (compared to section 1) and discuss "twisted Weierstrass points" rather than linear series. The virtue of this perspective is that displacement arguments will be easier to visualize (used extensively in section 6), will immediately give linear series that are guaranteed to be *complete*, and will possess an intriguing duality property (which we do not exploit in this thesis, but will be useful in forthcoming work).

As discussed in section 1, every point p on a smooth curve C determines a numerical semigroup called the *Weierstrass semigroup* of the point; it consists of those integers n such that C has a rational function of degree n whose only pole is at p. For all but finitely many points on a given curve C, this semigroup is $S = \{0, g+1, g+2, \cdots\}$; the other points are called *Weierstrass points*. We will consider a generalization of this concept, allowing sequences S that do not necessarily contain 0 (the special case $0 \in S$ will be precisely the same thing as a Weierstrass semigroup).

Let C be a smooth curve, $\mathcal{L}$ a degree 0 line bundle on C, and $p \in C$ a point. The *twisted Weierstrass sequence* of the triple $(C, \mathcal{L}, p)$ is the following set of nonnegative integers.

$$S(C, \mathcal{L}, p) = \{n \in \mathbf{Z}_{\geq 0} : \ h^0(\mathcal{L}(np)) > h^0(\mathcal{L}((n-1)p))\}$$

In other words, the twisted Weierstrass sequence is the set of possible pole orders at p of rational sections of $\mathcal{L}$ that are regular away from p. In the special case $\mathcal{L} = \mathcal{O}_C$, the twisted Weierstrass sequence is the classical Weierstrass semigroup. By the Riemann-Roch formula, the complement of S has precisely g elements, where g is the genus of C. If twisted Weierstrass sequences are given the obvious partial ordering, then they are upper semi-continuous in families; therefore the *general* twisted Weierstrass sequence is simply

$$S = \{g, g+1, g+2, \cdots\}.$$

A triple $(C, \mathcal{L}, p)$ with a different sequence is called a *twisted Weierstrass point.* We generalize the notation $\mathcal{W}_S$ (denoting the moduli of Weierstrass points) as follows.

Definition 5.1. Given a set $S \subseteq \mathbf{N}$ with $|\mathbf{N} \backslash S| = g$, let $\widetilde{\mathcal{W}_S}$ denote the moduli stack of triples $(C, \mathcal{L}, p)$, where C is a smooth curve, $\mathcal{L}$ is a line bundle of degree 0 and $p \in C$, such that $S(C, \mathcal{L}, p) = S$. Let $\mathcal{W}_S$ denote the module of triples $(C, \mathcal{L}, p)$ where $S(C, \mathcal{L}, p) \leq S$ (element by element, when the elements of both are sorted).

This is a mild abuse of notation since, in the case $0 \in S$, we previously defined $\mathcal{W}_S$ as a substack of $\mathcal{M}_{g,1}$, not of the moduli stack of pointed curves with choice of line bundle. However, the two definitions give stacks which are isomorphic (in the latter definition, the line bundle would always be $\mathcal{O}_C$), so this will not create any issues.

We can also identify $\mathcal{W}_S$ with a Brill-Noether variety, by truncating the sequence S at some degree $d \geq 2g-1$.

$$\begin{aligned} \mathcal{W}_S &\cong \mathcal{G}_g^{S \cap \{0,1,2,\cdots,d\}} \\ (C, \mathcal{L}, p) &\mapsto (C, p, |\mathcal{L}(dp)|) \end{aligned}$$

Remark 5.2. Going in reverse, a Brill-Noether variety $\mathcal{G}_g^T$ can be identified with a locus of twisted Weierstrass points as long as $d \geq 2g-1$ and $|T| = d+1-g$. This is because these constraints guarantee that $(C, p, L) \in \mathcal{G}_g^T$ is a *complete* linear series, and for all $n \geq 0$ the complete linear series of $\mathcal{L}(np)$ has no gaps in its vanishing sequence besides those already visible in T.

We can and will describe a twisted Weierstrass sequence using the (equivalent) data of a partition. Namely, the *twisted Weierstrass partition* $P(C, \mathcal{L}, p)$ is given by the multiset $\{(n+g) - s_n\}$ (restricted to positive entries), where the twisted Weierstrass sequence is $s_0 < s_1 < s_2 < \cdots$. It will often be convenient to fix the partition P and vary the genus g (producing a family of Weierstrass sequences, each a translation of any other). For this reason, we introduce the following notation.

Definition 5.3. If P is a partition, the elements of P will be denoted $P_0, P_1, P_2, \cdots, P_\ell$, where $P_0 \geq P_1 \geq \cdots \geq P_\ell \geq 1$. For all $k > \ell$, P_k will be defined to be 0. Let $|P|$ denote $\sum_{k=0}^{\infty} P_k$.

Definition 5.4. If P is a partition, and g is any integer with $g \geq P_0$, let $S_{P,g}$ denote the set of integers

$$S_{P,g} = \{g + k - P_k : \ k \geq 0\},$$

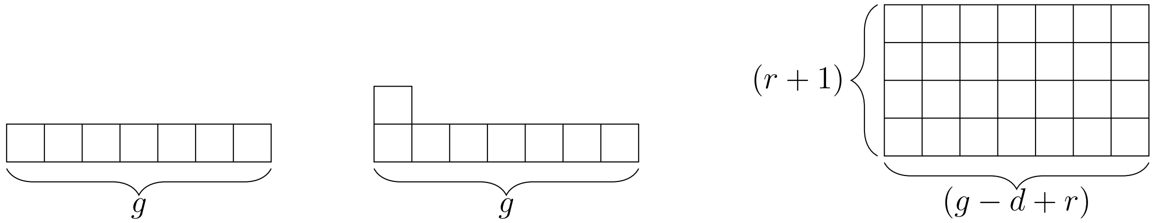
and let

$$\mathcal{W}_g(P) = \mathcal{W}_{S_{P,g}}.$$

One convenient aspect of working with partitions rather than sequences is the Brill-Noether dimension estimate, given by identifying $\mathcal{W}_S$ with a Brill-Noether variety, takes a particularly simple form.

$$\dim_{(C,\mathcal{L},p)} \widetilde{\mathcal{W}}_g(P) \geq (4g-2) - |P| \tag{5.1}$$

As usual, a point $(C, \mathcal{L}, p) \in \widetilde{\mathcal{W}}_g(P)$ where equality holds in 5.1 is called a *dimensionally proper point.*



$\mathcal{W}_g(P) \cong \mathcal{M}_{g,1}$ $\qquad$ $\mathcal{W}_g(P) \cong \{\text{Weierstrass points}\}$ $\qquad$ $\mathcal{W}_g(P) \cong \mathcal{W}^r_g \times_{\mathcal{M}_g} \mathcal{M}_{g,1}$

FIGURE 4. Three examples of partitions and the geometric interpretation of $\mathcal{W}_g(P)$.

Example 5.5. Let $P = (g)$. Then $0 \in S_{P,g}$, and $(C, \mathcal{L}, p) \in \widetilde{\mathcal{W}}_g(P)$ if and only if $h^0(\mathcal{L}) = 1$ and $h^0(\mathcal{L}(gp) = 1)$. This is true if and only if $\mathcal{L} = \mathcal{O}_C$ and p is not a Weierstrass point. So $\widetilde{\mathcal{W}}_g(P)$ is isomorphic to the complement in $\mathcal{M}_{g,1}$ of the locus of Weierstrass points, and $\mathcal{W}_g(P) \cong \mathcal{M}_{g,1}$. Therefore the local dimension at each point is $(3g-2) = (4g-2) - |P|$, so every point is dimensionally proper. $\triangleleft$

Example 5.6. Let $P = (g\ 1)$. Then $\widetilde{\mathcal{W}}_g(P)$ consists of triples $(C, \mathcal{L}, p)$ such that $h^0(\mathcal{L}) = 1, h^0((g-1)p) = 1$, and $h^0(gp) = 2$. In other words, this is the locus in $\mathcal{M}_{g,1}$ of simple Weierstrass points. This is étale-locally isomorphic to $\mathcal{M}_g$, so every point has local dimension $(3g-3) = (4g-2) - |P|$, so all points are dimensionally proper. $\triangleleft$

We will now study twisted Weierstrass points with the particular type of partition that will be relevant to theorem A. Let $P = (m^n)$ (i.e. the number m occurs n times). This partition corresponds to the following twisted Weierstrass sequence.

$$S = \{g-m, g-m+1, \cdots, g-m+n-2, g-m+n-1, g+n, g+n+1, g+n+2, \cdots\}$$

Then $(C, \mathcal{L}, p) \in \widetilde{\mathcal{W}}_g(P)$ if and only if the following conditions hold.

- $h^0(\mathcal{L}((g-m-1)p)) < h^0(\mathcal{L}(g-m)p) = 1$
- $h^0(\mathcal{L}(g-m+n-1)p) = h^0(\mathcal{L}(g+n-1)p) = n$

These conditions are equivalent to saying that $\mathcal{L}' = \mathcal{L}((g+n-1)p)$ is a line bundle of degree $(g-m+n-1)$ with $h^0(\mathcal{L}') = n$, such that p is not a ramification point for either the complete linear series $|\mathcal{L}'|$ or its dual $|\omega_C \otimes \mathcal{L}'^{\wedge}|$. In particular, both $|\mathcal{L}'|$ and $|\omega_c \otimes \mathcal{L}'^{\wedge}|$ are separable linear series.

Since any separable linear series has a finite number of ramification points, this means that for any line bundle $\mathcal{M}$ on C of degree $d = (g-m+n-1)$ and rank $r = (n-1)$, the triple $(C, \mathcal{M}(-dp), p)$ is a point of $\widetilde{\mathcal{W}}_g(P)$ for all but finitely many points $p \in C$. The upshot of this is the following.

Lemma 5.7. *Let g, d, r be integers, and P be the partition $((g-d+r)^{r+1})$. Then there is a map*

$$\begin{aligned} f : \widetilde{\mathcal{W}}_g(P) &\to \widetilde{\mathcal{G}}_g^{r,\mathrm{sep}} \\ (C, \mathcal{L}, p) &\mapsto (C, |\mathcal{L}(dp)|) \end{aligned}$$

which is surjective onto the open set $\{|\mathcal{L}| : \mathcal{L}^{\wedge} \text{ is also separable}\}$, and whose fiber over any point $(C, \mathcal{L}) \in \widetilde{\mathcal{G}}_g^r$ is isomorphic to C with finitely many punctures.

Notice that, in the notion of the lemma, $|P| = (r+1)(g-d+r) = g - \rho(g,d,r)$. It follows from this that the map f in the lemma sends dimensionally proper points to dimensionally proper points.

Corollary 5.8. *The stack $\widetilde{\mathcal{G}}_g^{r,\mathrm{sep}}$ has a dimensionally proper component if and only if $\widetilde{\mathcal{W}}_g(P)$ has a dimensionally proper component.*

Remark 5.9. Notice that twisted Weierstrass points have a duality property: namely if P^* is the dual partition of P (that is, $P_n^* = |\{m : P_m > n|)$, then $\widetilde{\mathcal{W}}_g(P) \cong \widetilde{\mathcal{W}}_g(P^*)$, via the map $(C, \mathcal{L}, p) \mapsto (C, \omega_C(-(2g-2)p) \otimes \mathcal{L}^{\wedge}, p)$. This duality is reflected, for example, in the two perspectives by which one typically studies classical Weierstrass points: in terms of pole order of rational functions or in terms of ramification of the canonical series.

Question 5.10. *Let $\mu(P, g)$ be the maximum codimension of a component of $\mathcal{W}_g(P)$ (or $-\infty$ if there are none). When is $\mu(P, g) < |P|$? Is there a purely combinatorial description of which partitions P and integers g give strict inequality?*

Another version of this question, which will be slightly more convenient for our purposes, is the following.

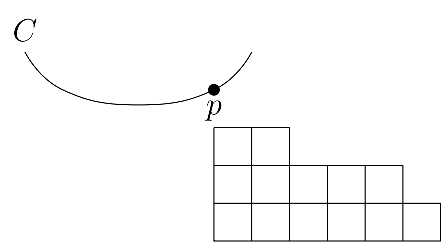
Question 5.11. *Let $\gamma(P)$ denote the minimum genus g such that $\mathcal{W}_g(P)$ has a dimensionally proper point. Can $\gamma(P)$ be determined (or bounded) by a combinatorial procedure?*

We will define in section 6 a function $\delta(P)$ of partitions such that $g \geq \frac{1}{2}(|P|+\delta(P))$. Bounding this function will give bounds on the function $\gamma(P)$ described above. In particular, suitable bounds of $\gamma(P)$ for "box-shaped" partitions will give theorem A.

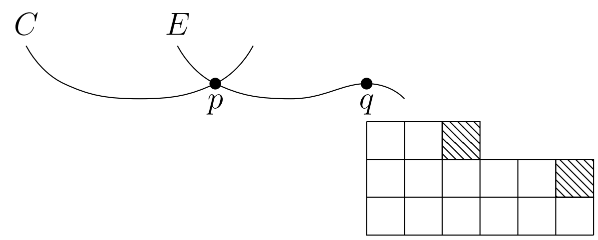
5.1. **Displacement of twisted Weierstrass sequences.** We will prove theorem A using the displacement lemma 2.15. The proof will inductively construct dimensionally proper twisted Weierstrass points, with the twisted Weierstrass partition growing at each step until eventually reaching a box-shaped partition.

The inductive steps will consist of displacing a twisted Weierstrass point of genus g across an elliptic curve to obtain a twisted Weierstrass point of genus $g+1$. We illustrate this process in an example.

Example 5.12. Suppose that $P = (6, 5, 2)$, and we have a dimensionally proper point $(C, p, \mathcal{L})$ of $\widetilde{\mathcal{W}}_g(P)$, for some g. This corresponds to a point of $x \in \widetilde{\mathcal{G}}_g^S$, where d is chosen to be at least $2g$, and $S = S_{P,g} \cap \{0, 1, \cdots, d\}$.



An elliptic curve with two marked points (one of which is identified with p) can be glued to C to produce a genus $g+1$ curve of compact type. Under certain conditions (analyzed by the displacement lemma), a dimensionally proper limit linear series can be constructed on the new reducible curve such that, when the result is smoothed, it will correspond to a point of $\widetilde{\mathcal{W}}_{g+1}(P')$ for some modified partition P'.



The new partition shown in this picture, namely $(6,6,3)$, can indeed by obtained from $(6,5,2)$ from this procedure. This follows from the displacement lemma 2.15. $\triangleleft$

The following definition will make precise which sorts of modifications of partitions can be accomplished in this way. Recall that $\widetilde{\mathcal{W}}_g(P)$ can be identified with a Brill-Noether variety $\mathcal{G}_g^{S(P,g,d)}$ for any $d \geq 2g-1$, where $S(P,g,d) = \{g+k-P_k : k = 0,1,\cdots,d-g\}$. Meanwhile, $\widetilde{\mathcal{W}}_{g+1}(P')$ can be identified with $\mathcal{G}_g^{S(P',g+1,d)}$ for any $d \geq 2g+1$. Observe that $|S(P,g,d)| = d-g+1$, while $|S(P',g+1,d)| = d-g$. Then, in the language of section 2, the construction in the example above will successfully give a dimensionally proper point of $\mathcal{W}_{g+1}(P')$ as long as, for some $d \geq 2g$, $(S(P,g,d) : S(P',g+1,d+1))$ is 1-valid. We will actually required that the pair is *strictly* 1-valid, to obtain slightly stronger results. Lemma 2.15 actually shows that assuming $P \leq P'$, the value of g does not matter, as long as $g \geq P_0$ and $g \geq P'_0$ (so that all elements of both sequences are nonnegative). Therefore the following is well-defined.

Definition 5.13. Let P, P' be two partitions with $P \leq P'$ and $|P'| \leq |P|+2$. Say that P is *linked* to P' if for all $g \geq \max\{P_0, P'_0\}$, there exists arbitrarily large integers d such that $(S(P,g,d) : S(P',g+1,d+1))$ is strictly 1-valid. In addition, if $|P'| - |P| = k$, we say that P and P' are k-linked.

By lemma 2.16, there are precisely three situations where P can be linked to P', given that $P \leq P'$. These are equivalent, respectively, to being $0-, 1-$, or 2-linked.

- 0-linked: $P = P'$.
- 1-linked: $|P'| = |P| + 1$.
- 2-linked: There are exactly two indices i such that $P'_i = P_i + 1$. For all other indices j, $P'_j = P_j + 1$. The arithmetic progression generated by $\{P_i - i : P'_i = P_i + 1\}$ is disjoint from the set $\{P_j - j : P_j = P'_j \text{ and } P_j < P_{j-1}\}$ and meets the set $\{P_j - j - 1 : P_j = P'_j \text{ and } P_j > P_{j+1}\}$ only at the generators of the arithmetic progression.

We will give a slightly different definition of "linked" in section 6; it is easy to verify that the characterization above is equivalent to the definition given there.

The following lemma follows immediately from definition 5.13.

Lemma 5.14. *Suppose that $P \leq P'$ and P is k-linked to P', where $k \leq 2$. For any g, if $\widetilde{\mathcal{W}}_g(P)$ has a dimensionally proper point, then $\widetilde{\mathcal{W}}_{g+1}(P')$ also has a dimensionally proper point.*

In addition, suppose that $\widetilde{\mathcal{W}}_g(P)$ has a dimensionally proper point, with local dimension equal to $\min(0, g - |P|)$ in its fiber over $\mathcal{M}_{g,1}$. If either

- *$g - |P| \leq 0$ and P is either 1-linked or 2-linked to P', or*
- *$g - |P| \geq 0$ and P is either 0-linked or 1-linked to P',*

then $\widetilde{\mathcal{W}}_{g+1}(P')$ has a dimensionally proper point, with local dimension equal to the minimum of 0 and $g - |P'|$ in its fiber over $\mathcal{M}_{g,1}$.

Definition 5.15. Let P be any partition. Define the *displacement distance* of P, denoted $d(P)$, to be the minimum value d such that there is an increasing sequence of partitions $0 = P_0 < P_1 < \cdots < P_d = P$ such that any two adjacent partitions in this sequence are k-linked for some $k \le 2$.

It is obvious from the definition that

$$\frac{1}{2}|P| \le d(P) \le |P|.$$

In fact, we will see in section 6 that in many cases, the value of $d(P)$ is much closer to the lower bound. Therefore we make the following definition.

Definition 5.16. The *displacement difficulty* $\delta(P)$ of a partition is

$$\delta(P) = 2d(P) - |P|.$$

This terminology allows us to state the following corollary, which we will use to prove theorem A.

Corollary 5.17. *Let P be any partition. For all integer $g \ge \frac{1}{2}(|P| + \delta(P))$, the stack $\widetilde{\mathcal{W}}_g(P)$ has a dimensionally proper component X, such that X*

- *dominates $\mathcal{M}_{g,1}$ if $g - |P| \ge 0$, and*
- *is generically finite over $\mathcal{M}_{g,1}$ if $g - |P| \le 0$.*

Proof. Immediate from induction on lemma 5.14. □

The machinery above can now be used to prove the first main theorem stated in section 1. We reproduce the statement here for convenience.

Theorem A. *Suppose that g, d, r are positive integers with*

$$0 > \rho(g, d, r) \ge -\frac{r}{r+2} g + 3r - 3.$$

If $r = 1$ or $g - d + r = 1$, then $\mathcal{G}^r_g$ is empty. Otherwise $\mathcal{G}^r_g$ has an irreducible component of dimension $\dim \mathcal{M}_g + \rho$, whose image in $\mathcal{M}_g$ has codimension equal to $-\rho$, and whose general member is complete and separable.

Proof. The cases $r \leq 1$ and $g - d + r \leq 2$ are easy. Therefore assume that $r \geq 2$ and $g - d + r \geq 3$.

By corollary 5.8, it suffices to prove that $\widetilde{\mathcal{W}}_g(P)$ has a dimensionally proper component generically finite over $\mathcal{M}_{g,1}$, where $P = ((g - d + r)^{r+1})$. By corollary 5.17, it suffices to prove that $g \geq \frac{1}{2}((r + 1)(g - d + r) + \delta(P))$. For convenience, denote $(g - d + r)$ by a and $(r + 1)$ by b. Then the desired inequality is $g \geq \frac{1}{2}(ab + \delta(a^b))$. We will prove the following lemma in section 6.

Lemma 6.20. *Let P be the partition (a^b), i.e. the partition of the number ab into b equal parts, where $a, b \geq 2$. Then $\delta(P) \leq a + 3b - 5$.*

Therefore is suffices to show that $g \geq \frac{1}{2}(ab + a + 3b - 5)$. We have assuming the following.

$$
\begin{aligned}
\rho(g, d, r) &\geq -\frac{r}{r+2}g + 3r - 3 \\
g - ab &\geq -\frac{b-1}{b+1}g + 3b - 6 \\
\frac{2b}{b+1}g &\geq ab + 3b - 6 \\
g &\geq \frac{1}{2}(a+3)(b+1) - 6 \cdot \frac{b+1}{2b} \\
2g &\geq (a+3)(b+1) - 6 - \frac{6}{b} \\
&= ab + a + 3b - 3 - \frac{6}{b}
\end{aligned}
$$

The result now follows, since we assumed that $b \geq 3$. $\square$

6. Displacement difficulty of partitions

This section is almost entirely self-contained and combinatorial. The purpose is to provide the necessary combinatorial analysis for theorem A, and to formulate and discuss a purely combinatorial problem (problem 6.6) that can generalize that theorem to the case of pointed Brill-Noether theory (that is, moduli of linear series with specified ramification) as well as improve the lower bound on ρ in theorem A. After formulating problem 6.6, the next four subsections will study it in four special cases to illustrate some of the interesting behavior that occurs. Of these, only the fourth subsection (on rectangular partitions) is needed for theorem A; the others serve mainly to illustrate problem 6.6, as well as to serve as a foundation for future work that may improve theorem A.

We will use the following convention in this section: an *arithmetic progression* will mean a proper subset $\Lambda \subset \mathbf{Z}$ such that $\Lambda - \Lambda$ is closed under addition. In particular, Λ may be empty or have a single element, but it cannot be all of $\mathbf{Z}$. Also, we adopt the following notational conventions: for a partition P, the partition elements are denoted $P_0 \geq P_1 \geq \cdots \geq P_n$, and P_k is defined to be 0 for $k > n$ and ∞ for $k < 0$.

Definition 6.1. Let P be a partition and Λ an arithmetic progression. Then define the *upward displacement* P^+_Λ and *downward displacement* P^-_Λ of P with respect to Λ as follows. For all $i \geq 0$,

$$(P^+_\Lambda)_i = \begin{cases} P_i + 1 & \text{if } (P_i - i) \in \Lambda \text{ and } P_{i-1} > P_i \\ P_i & \text{otherwise} \end{cases}$$

$$(P^-_\Lambda)_i = \begin{cases} P_i - 1 & \text{if } (P_i - i - 1) \in \Lambda \text{ and } P_{i+1} < P_i \\ P_i & \text{otherwise} \end{cases}$$

This definition is much easier to understand visually; it is illustrated in figure 5. Here the partition P is represented by its Young diagram, and the arithmetic

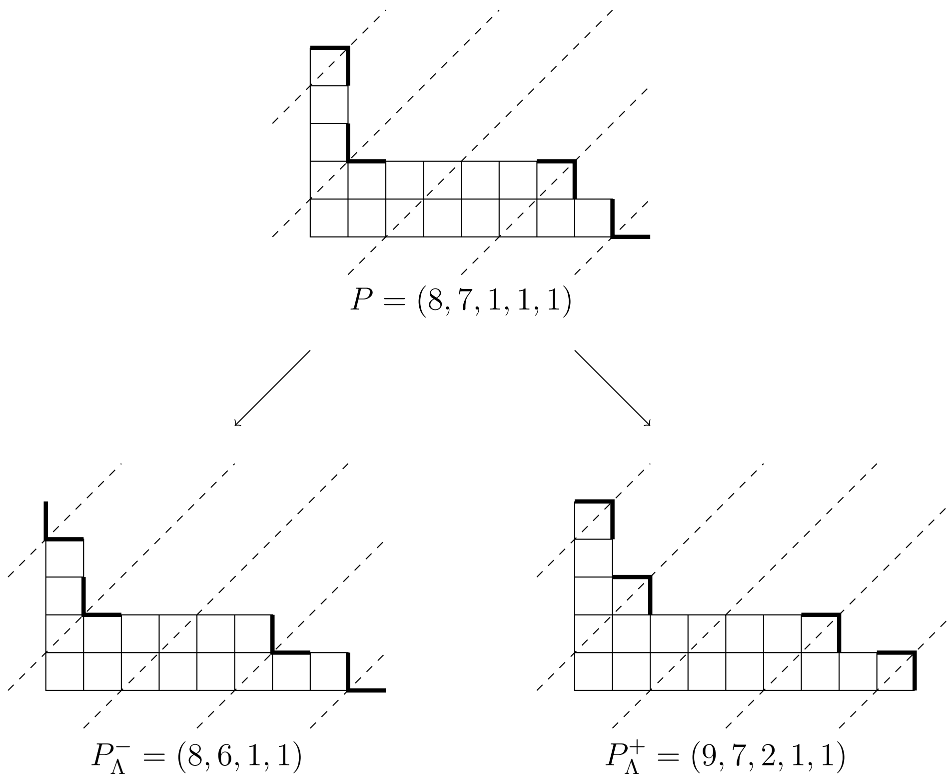


FIGURE 5. An example illustrating the definition of displacement. Here $\Lambda = \{2 \mod 3\}$.

progression Λ is represented by an evenly spaced sequence of diagonal lines, whose x-intercepts correspond to the values of Λ. Then the two displacements are obtained by finding all places where the line of Λ meet the corners of P, and either "turning the corners out" (in the case of P^+_Λ) or "turning the corners in" (in the case of P^-_Λ).

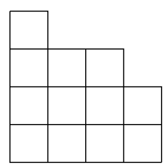
Note that, by these definitions, there are always two inward corners that are not immediately visible in the Young diagram: one at the end of the first row, and one at the end of the first column.

Observe that if P' is any other partition such that $P^-_\Lambda \leq P' \leq P^+_\Lambda$, then the upward and downward displacements of P' are the same as those of P (with respect to Λ). So displacement can be regarded as a sort of projection to the nearest partition that is stable with respect to the given arithmetic progression.

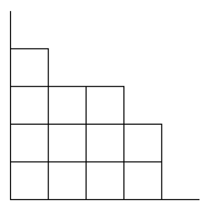
We will say that a partition P_1 *links to* a partition P_2 (or that P_1 and P_2 are *linked*, where it is understood that the smallest partition links to the larger) if there is an arithmetic progression Λ (proper but possibly empty or singleton) such that P_2 is the upward displacement of P_1 and P_1 is the downward displacement of P_2. Note that this implies that P_1 is its own downward displacement and P_2 is its own upward displacement. Say that P_1 and P_2 are *k-linked* if they are linked and $|P_2| - |P_1| = k$.

It is easy to verify that if P_1, P_2 are any two partitions with $P_1 \le P_2$, then P_1 can be connected to P_2 by a sequence of 1-linked partitions. Indeed, the arithmetic progressions can be taken to be singletons.

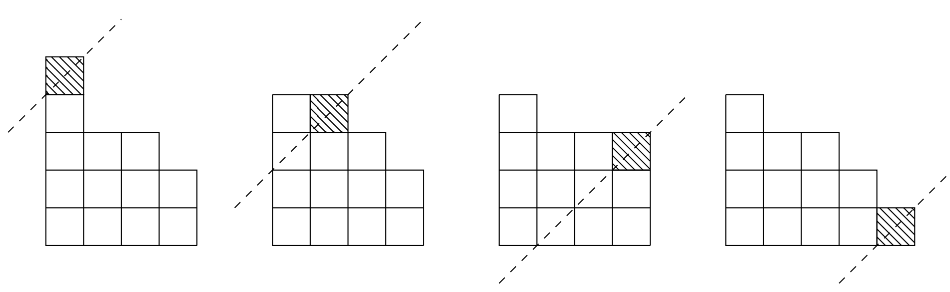
Example 6.2. Consider the partition $(4, 4, 3, 1)$, shown by the following Young diagram.



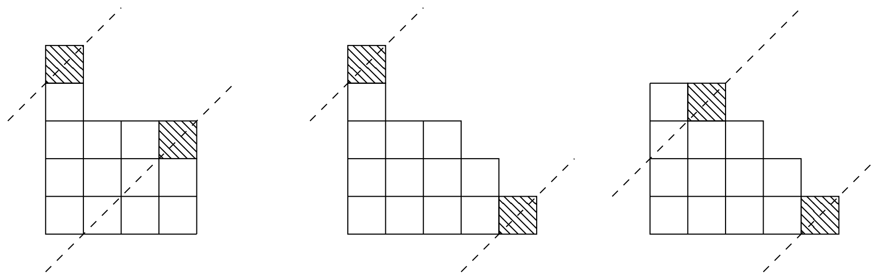
Remember that we consider the first row and first column to have inward corners at their ends, so it may be helpful to extend the axes when drawing the Young diagram to make these visible.



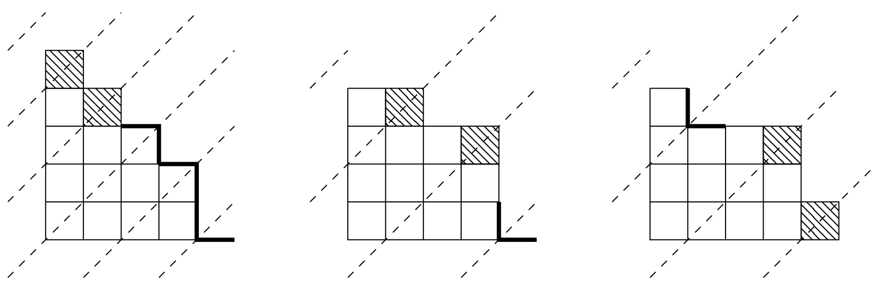
This partition is 0-linked to itself. It is 1-linked to 4 partitions (one for each inward corner).



There are therefore $\binom{4}{2}$ possible choices of two boxes to add. Of these, three of them give 2-linked partitions.



The other three do not give valid displacements, because the arithmetic progression of slope-one lines determined by the added squares meets at least one other corner of the partition (shown with bold lines).



$\triangleleft$

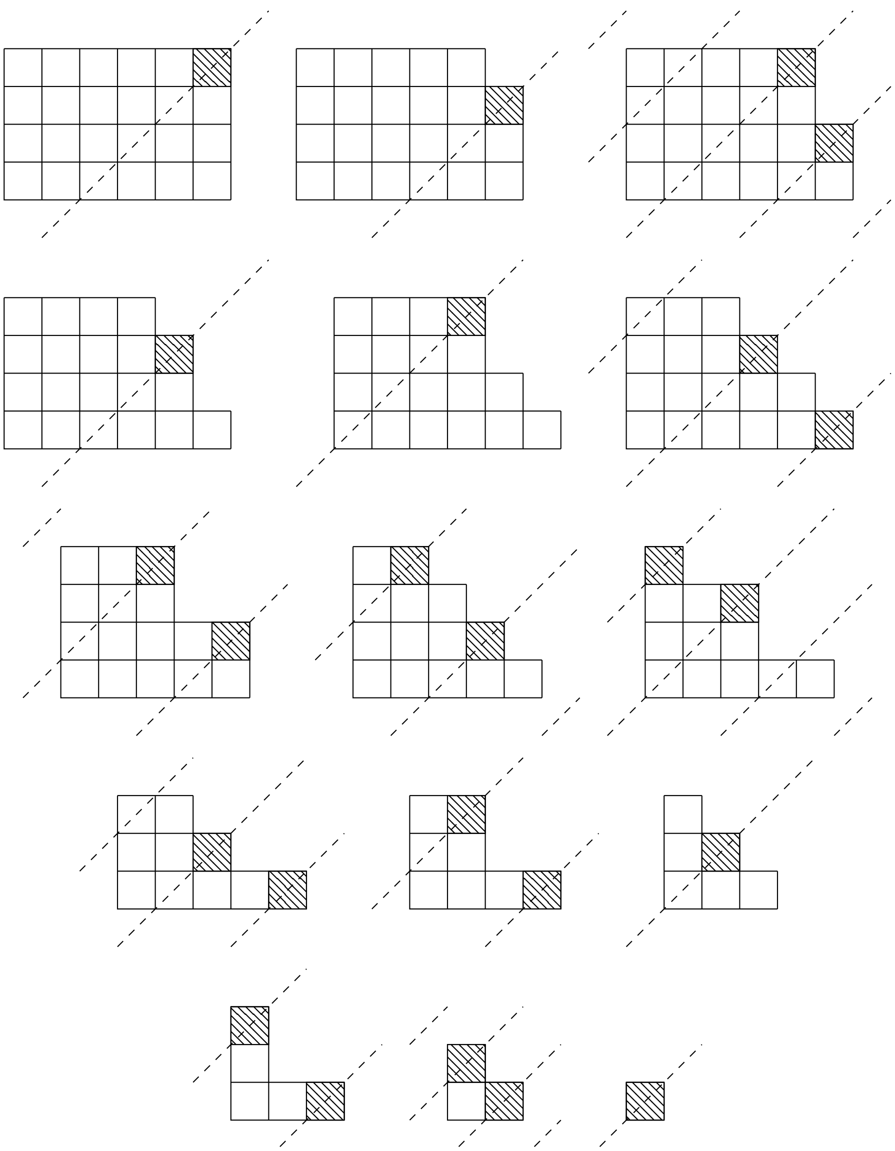
For reasons related to our intended application, we are particularly interested in 2-linked partitions. More specifically, we are interested in partitions that can be joined by a path of 1-linked and 2-linked pairs, using as few 1-linked pairs as possible. Therefore make the following definition, which corresponds to definition 5.16 from section 5.

Definition 6.3. Call a sequence of partitions of increasing sum *valid* if any two adjacent partitions in the sequence are 1-linked or 2-linked. Define the displacement distance $d(P)$ to the smallest value d such that there is a sequence $0 = P_0 < P_1 < \cdots < P_d = P$ of $1-$ or $2-$linked partitions.

Define the *difficulty* $\delta(P)$ of a partition P to be $2d(P) - |P|$.

The difficulty $\delta(P)$ quantifies how much the easy lower bound $\frac{1}{2}|P| \leq d(P)$ fails to be sharp. In practice, $\delta(P)$ tends to be quite low for many types of partitions.

Example 6.4. Consider the 4 by 6 rectangular partition. A computer search reveals that $\delta(P) = 6$. We computed this in less than a second using the Java code provided in the appendix (in particular, by running the method `difficulty`, which computes the shorted valid path by dynamic programming). One possible valid sequence, with only 6 one-links, is shown below. There are, in fact, 882 such paths (this can be found with the `numPaths` method in the code in the appendix).



$\lhd$

Remark 6.5. It is apparent from the definition that $\delta(P) = \delta(P^*)$ where P^* is the conjugate partition. This is not surprising, in light of the duality $\mathcal{W}_g(P) \cong \widetilde{\mathcal{W}}_g(P^*)$ (see remark 5.9).

The function δ is our object of study throughout this section.

Problem 6.6. *Find upper bounds for $\delta(P)$, given basic information about the shape of P.*

The particular case needed for our application is that of a "box-shaped" partition, but good bounds on difficulty of boxes will necessarily depend on bounds for the intermediate partitions. Therefore we will discuss some special classes of partitions to illustrate some of the phenomena that occur. The following subsections will each deal with a particular "shape" of partition, as summarized below.

- Subsection 6.1 considers those partitions with a very long bottom row (i.e. with one very large element). More specifically, we give an almost complete classification of those partitions which satisfy $d(P) = P_0$. These partitions turn out to correspond to primitive numerical semigroups.
- Subsection 6.2 computes the displacement difficulty of partitions with only two parts. This analysis is related to certain combinatorial games. We show when the two rows are nearly equal in size (more precisely, when the difference in their sizes is smaller than the square root of the length of the shorter row) then the displacement difficulty is asymptotic to $\sqrt{|P|}$, while when the difference of the sizes of the two rows is large enough, this difference gives the main term of an expression for $\delta(P)$.
- Subsection 6.3 considers the displacement difficulty of staircase partitions $(n, n-1, \cdots, 2, 1)$. These all have displacement difficulty at most 3.

- Subsection 6.4 considers the partitions that are relevant to our main theorem A, namely rectangular partitions. We give a relatively weak upper bound on the difficulty of such partitions, and provide experimental evidence that this bound should be able to be substantially improved.

6.1. **Partitions with a long bottom row.** The partitions we consider in this section are those which come up (implicitly) in Eisenbud and Harris's work on primitive Weierstrass points [EH87]. Indeed, their theorem will follow from the main result of this section, plus lemma 5.17.

We begin with the following easy observation.

$$d(P) \geq P_0 \tag{6.1}$$

This follows because the total number of terms in a valid sequence is at least P_0 (the bottom row can shrink by at most 1 square at at time).

In this subsection, we will give an almost complete classification of those partitions where equality holds, i.e. where each link removes exactly one square from the bottom row of the young diagram (and succeeds in removing all other squares in the process).

Recall that a numerical semigroup is a cofinite subset $S \subseteq \mathbf{N}$ containing 0, and a semigroup is *primitive* if the sum of any two positive elements exceeds the largest element of the complement. The *genus* of S is $|\mathbf{N}\backslash S|$, and the *weight* of S is $\sum(\mathbf{N}\backslash S) - \binom{g+1}{2}$. A semigroup $S = \{0 = s_0, s_1, s_2, \cdots\}$ gives rise to a partition given by $P_k = (g + k - s_k)$, and P uniquely determines S (since $g = P_0$).

For convenience, we introduce the following terminology.

Definition 6.7. A partition P is *primitive* if $P_0 - P_0^* \geq 2P_1 - 2$. The *weight* of a partition is $\mathrm{wt}(P) = |P| - P_0$.

Figure 6 gives a visual explanation of the condition that a partition is primitive.

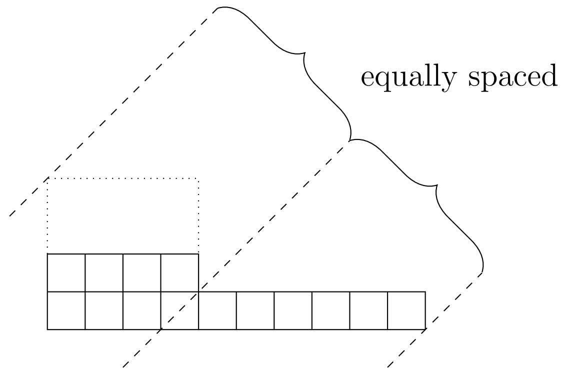


FIGURE 6. Visual representation of the primitivity condition. Draw diagonals through the corners at the ends of the first two rows of the Young diagram, and also a third diagonal, equally spaced from the second. Then the rest of the Young diagram must fit below this third diagonal.

The reason for these terms is that a partition P arises from a primitive semigroup if and only if P is primitive, and the weight of a semigroup is always equal to the weight of the partition that it determines.

The partial classification we obtain is the following.

Proposition 6.8. *Suppose that P is a partition.*

(1) If $d(P) = P_0$, then P arises from a primitive semigroup of genus g and weight at most $g-1$.

(2) If P arises form a primitive semigroup of weight at most $g-2$, then $d(P) = P_0$.

This fact was implicitly proved and used by Eisenbud and Harris [EH87] (using different terminology), who showed the existence of dimensionally proper Weierstrass points with primitive semigroups of weight less than $g-2$. The discrepancy between $g-1$ and $g-2$ (which makes this proposition not a complete classification of the equality cases $d(P) = P_0$) is why the existence of weight $g-1$ Weierstrass semigroups was unknown until the work of Komeda [K91].

Proof. The result is obvious if $P_1 = 0$, since any such partition is primitive, has weight 0, and has $d(P) = P_0 = |P|$. Next consider the case $P_1 = 1$. In this case the partition is simply a "hook," and it is obvious in this case that the following conditions are equivalent.

- $d(P) = P_0$.
- $P_0^* \le P_0$.
- P is primitive.
- The weight of P is at most $P_0 - 1$.

It remains the prove the proposition in the case $P_1 \ge 2$. We proceed by induction on $|P|$. The base cases for this induction will be the cases where $P_1 \le 1$.

Suppose P is a partition with $P_1 \ge 2$, and that the proposition holds for all partitions P' with $|P'| < |P|$.

First we shall prove part 1 of the proposition. Therefore assume that $d(P) = P_0$. We must show that P is primitive of weight at most $P_0 - 1$. The fact that the weight of P is at most $P_0 - 1$ is easy to see directly: each displacement in a valid path to P must add a square to the bottom row, and can add at most one square above the bottom row. The first displacement (from the empty partition to a single square) adds nothing above the bottom row, hence $P_0 = d(P) \ge 1 + \mathrm{wt}(P)$. It remains to show that P is primitive. Suppose for contradiction that P is not primitive. By the inductive assumption, there must exist an arithmetic progression Λ such that $P_\Lambda^+ = P$, P_Λ^- is a primitive partition with weight at most $(P_\Lambda^-)_0 - 1$, and $(P_\Lambda^-)_0 = P_0 - 1$. Consider the quantity $(P_0 - P_0^* - 2P_1 + 2) - ((P_\Lambda^-)_0 - (P_\Lambda^-)_0^* - 2(P_\Lambda^-)_1 + 2)$, which is equal to $1 + (P_\Lambda^-)_0^* - P_0^* + 2(P_\Lambda^-)_1 - 2P_1$. If this quantity is nonnegative, then it follows immediately that P is primitive, from the fact that P_Λ^- is primitive. The only way that this expression can be negative is if $P_1 = (P_\Lambda^-)_1 + 1$, in which case it is equal to -1. Therefore the only way that P could fail to be primitive is if $(P_\Lambda^-)_0 - (P_\Lambda^-)_0^* - 2(P_\Lambda^-)_1 + 2$ is exactly equal to 0 (so $P_0 - P_0^* - 2P_1 + 2 = -1$), and P is obtained from P_Λ^- by adding one square to each of the first two rows. In this case, the progression Λ must be generated by $P_0 - 1$ and $P_1 - 2$. Therefore Λ contains $2(P_1 - 2) - (P_0 - 1) = 2P_1 - P_0 - 3$,

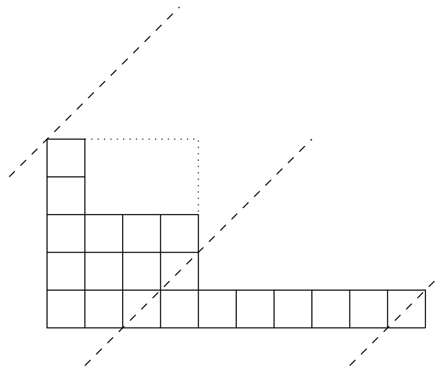


FIGURE 7. The only situation where primitivity may fail after displacement, in the proof of theorem 6.8

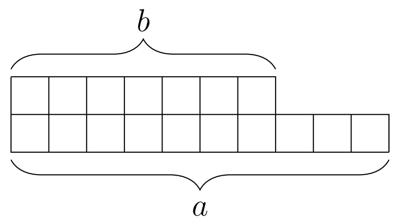
which is equal to $-P_0^*$. It follows that for $i = P_0^*$, $P_i - i = -P_0^* \in \Lambda$, hence $(P_\Lambda^+)_i = P_i + 1 = 1$. This contradicts the assumption that $P_\Lambda^+ = P$. Therefore P must in fact be primitive.

Now we shall prove that part 2 of the proposition holds for P. Therefore assume that P is primitive of weight at most $P_0 - 2$; we must show that $d(P) = P_0$.

Let $k \geq 1$ be the largest integer such that $P_k = P_1$. Then let Λ be the arithmetic progression generated by $P_0 - 1$ and $P_k - k - 1$. The corresponding diagonal lines meet the Young diagram of P at only two corners, both outward, at the ends of rows 0 and k (since P is primitive, the next element to the left of Λ, namely $2(P_k - k - 1) - (P_0 - 1)$, is strictly less than $-P_0^*$, and hence meets no corners of P; see figure 6). Thus $P_\Lambda^+ = P$ and P_Λ^- differs in exactly two places from P: P_0 and P_k are both decreased by 1. Now, it is immediate that $|P_\Lambda^-| \leq 2(P_\Lambda^-)_0 - 2$. It remains to show that $(P_\Lambda^-)_0 - (P_\Lambda^-)_0^* \geq 2(P_\Lambda^-)_1 - 2$. Since P_0 decreased by 1 under the displacement, the only way that this inequality could fail is if P_0^* is unchanged, P_1 is unchanged, and the inequality was sharp before, i.e. $P_0 - P_0^* = 2P_1 - 2$.

This would mean that $P_1 = P_2$ and the Young diagram meets the third diagonal in figure 6; see figure 7. But in this case, we would have $|P| \geq P_0 + 2P_1 + (P_0^* - 3) = 2P_0 - 1$, which contradicts the assumption that $|P| \leq 2P_0 - 2$. Hence P_Λ^- satisfies the hypotheses of the lemma. Also, it is clear that $2P_0 - |P|$ is unchanged and $d(P) \leq 1 + d(P_\Lambda^-) = P_0$, hence $d(P) = P_0$, completing the induction. $\square$

6.2. **Two-row partitions.** A second class of partitions for which $\delta(P)$ may be computed quite efficiently (and given an exact asymptotic in closed form) are partitions with only two rows. Throughout this section, we will use the abbreviated notation $\delta(a, b)$ to denote the difficulty of the partition (a, b), where $a \geq b$.



The simplest such partitions, namely $2 \times a$ rectangles, display an intriguing behavior in the limit.

Proposition 6.9. *The displacement difficulty of* $2 \times a$ *rectangles satisfies*

$$\lim_{a \to \infty} \frac{\delta(a, a)}{\sqrt{a}} = 2\sqrt{\pi}.$$

This proposition will be an immediate consequence of proposition 6.12 and theorem 6.14. We also obtain the following more specific (but less asymptotically precise) statement. It will follow from lemmas 6.15 and 6.16, together with proposition 6.12.

Proposition 6.10. *For any positive integers* a, b *with* $a \geq b$*,*

$$\delta(a, b) = \begin{cases} 2 \cdot \left\lceil \frac{b+1}{a-b+1} \right\rceil + (a - b) - 2 & \text{if } a - b \geq \sqrt{b+1} - 1 \\ C(a, b) \cdot \sqrt{b+1} - (a - b) - 4 & \text{if } a - b < \sqrt{b+1} - 1 \end{cases}$$

where $C(a, b)$ *is a number that satisfies* $3.26 \leq C(a, b) \leq 4$*.*

$\rho_{n,m}$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	3	2	0	0	0	0	0	0	0	0	0	0	0	0	0
3	5	4	3	0	0	0	0	0	0	0	0	0	0	0	0
4	9	8	6	4	0	0	0	0	0	0	0	0	0	0	0
5	11	10	9	8	5	0	0	0	0	0	0	0	0	0	0
6	17	16	15	12	10	6	0	0	0	0	0	0	0	0	0
7	21	20	18	16	15	12	7	0	0	0	0	0	0	0	0
8	29	28	27	24	20	18	14	8	0	0	0	0	0	0	0
9	33	32	30	28	25	24	21	16	9	0	0	0	0	0	0
10	41	40	39	36	35	30	28	24	18	10	0	0	0	0	0
11	47	46	45	44	40	36	35	32	27	20	11	0	0	0	0
12	57	56	54	52	50	48	42	40	36	30	22	12	0	0	0
13	59	58	57	56	55	54	49	48	45	40	33	24	13	0	0
14	77	76	75	72	70	66	63	56	54	50	44	36	26	14	0
15	81	80	78	76	75	72	70	64	63	60	55	48	39	28	15
16	101	100	99	96	95	90	84	80	72	70	66	60	52	42	30
17	107	106	105	104	100	96	91	88	81	80	77	72	65	56	45
18	117	116	114	112	110	108	105	104	99	90	88	84	78	70	60
19	131	130	129	128	125	120	119	112	108	100	99	96	91	84	75
20	149	148	147	144	140	138	133	128	126	120	110	108	104	98	90

FIGURE 8. Some values of $\rho_{n,m}$. The number n is the vertical axis.

More precisely, the lower bound for $C(a,b)$ that we will obtain is $\sqrt{32/3}$. By proposition 6.9, the limit of $C(a,a)$ as a grows to infinity is $2\sqrt{\pi}$, which is approximately 3.55.

We can state the exact values of $\delta(a,b)$ using the following notation.

Definition 6.11. For any two integers m, n, define an integer $\rho_{m,n}$ recursively as follows.

$$\rho_{n,m} = \begin{cases} 0 & \text{if } m > n \\ m & \text{if } m = n \\ m \cdot \left\lceil \dfrac{\rho_{n,m+1} + 1}{m} \right\rceil & \text{if } m < n \end{cases}$$

For example, figure 8 shows the values $\rho_{n,m}$ for $n \leq 20$ and $m \leq 15$.

Using this notation, we can exactly express the difficulty of 2-row partitions as follows.

Proposition 6.12. *The difficulty of any two-part partition* (a, b) *is*

$$\delta(a, b) = 2 \cdot \min\{n : \ \rho_{n,a-b+1} \geq b + 1\} - (a - b) - 2.$$

Before proving proposition 6.12, we briefly remark on the ideas behind the proof. One can prove that if one is looking for the shortest valid sequence from (a, b) to $(0, 0)$, then

- It is always better to take a two-link if it is available.
- If a two-link is not available, then it is better to remove a square from the top row rather than the bottom row.

It would be possible to write a proof of proposition 6.12 along these lines. For the sake of brevity, however, we give a proof by induction instead, which may seem more opaque. Nevertheless, the inductive proof does reveal these two "strategic" facts on a close reading. The proof of proposition 6.12 uses the following lemma.

Lemma 6.13. *For any positive integer* m *and nonnegative integer* k*, let* $\mu(m, k) = \min\{n : \ \rho_{n,m} \geq k\}$.

(1) For all k*,* $\mu(2, k) = \mu(1, k + 1)$.

(2) For all m, k*,* $\mu(m, k) + 1 \geq \mu(m + 1, k + 1)$.

(3) For all m, k *such that* m *divides* k*,* $\mu(m, k + 1) = \mu(m + 1, k)$.

(4) For all m, k *such that* m *does not divide* k*,* $\mu(m, k) = \mu(m, k + 1)$.

(5) For all m, k*,* $\mu(m, k + 1) \leq \mu(m + 1, k)$.

Proof. We consider each claim in turn.

(1) For all n, $\rho_{n,1} = \rho_{n,2} + 1$. The result follows immediately.

(2) This statement is equivalent to saying that if $\rho_{n,m} \geq k$ then $\rho_{n+1,m+1} \geq k+1$. In other words, this amounts to saying that $\rho_{n,m} + 1 \leq \rho_{n+1,m+1}$. This is obvious for $m > n$, so assume $m \leq n$. We can prove this statement by induction on $n - m$. For $n = m$ it follows from the definition of $\rho_{n,n}$. Now

suppose that $m < n$ and $\rho_{n,m+1} + 1 \leq \rho_{n+1,m+2}$. Then $\rho_{n,m}$ is the unique multiple of m in the set $\{\rho_{n,m+1}+1, \rho_{n,m+1}+2, \cdots, \rho_{n,m+1}+m\}$. On the other hand, $\rho_{n+1,m+1}$ is a multiple of $(m+1)$ than is strictly larger than $\rho_{n,m+1}$. Since $\rho_{n,m+1}$ itself is a multiple of $(m+1)$, it follows that $\rho_{n+1,m+1} \geq \rho_{n,m+1}+(m+1)$. It follows that $\rho_{n+1,m+1} \geq \rho_{n,m} + 1$, completing the induction.

(3) The statement that $\mu(m, k+1) = \mu(m+1, k)$ is equivalent to the statement that for all n, $\rho_{n,m} \geq k+1$ if and only if $\rho_{n,m+1} \geq k$. On the one hand, $\rho_{n,m}$ is always strictly greater than $\rho_{n,m+1}$. Conversely, if $\rho_{n,m+1} < k$, then k is a multiple of m strictly larger than $\rho_{n,m+1}$, hence $\rho_{n,m} \leq k$. The desired result follows by contrapositive.

(4) This follows since $\rho_{n,m}$ is always a multiple of m, therefore $\rho_{n,m} \geq k$ if and only if $\rho_{n,m} \geq k+1$.

(5) This follows because if $\rho_{n,m+1} \geq k$, then $\rho_{n,m} \geq k+1$.

□

Proof of proposition 6.12. We will use the notation $\mu(m, k)$ defined in lemma 6.13. Therefore we wish to show that

$$\delta(a, b) = 2\mu(a - b + 1, b + 1) - a + b - 2.$$

First, consider the case $b = 0$. In this case, $\delta(a, 0) = a$, and $\mu(a - b + 1, b + 1) = \mu(a + 1, 1) = a + 1$. The result follows.

Now assume that $b > 0$. We will proceed by induction on $a+b$. The base cases are provided by the case $b = 0$. Therefore assume that the result holds for all a', b' such that $a' + b' < a + b$. There are three possible partitions smaller than (a, b) that could link to (a, b): these are $(a - 1, b)$ (possible if and only if $a > b$), $(a, b - 1)$ (possible if and only if $b > 0$, which we have assumed) and $(a - 1, b - 1)$ (possible if and only if

$0 < b < a$ and $(a-b+1) \nmid b$). By the inductive hypothesis, these partitions (if they exist) have the following difficulties.

$$\begin{aligned}
\delta(a-1,b) &= 2\mu(a-b,b+1)-a+b-1\\
\delta(a,b-1) &= 2\mu(a-b+2,b)-a+b-3\\
\delta(a-1,b-1) &= 2\mu(a-b+1,b)-a+b-2
\end{aligned}$$

We consider three cases.

First, suppose that $a = b$. In this case, there is only one possible downward displacement; we have $\delta(a,a) = \delta(a,a-1)+1 = 2\mu(2,a)-2$. By lemma 6.13 part 1, this is equal to $2\mu(1,a+1)-2$, as desired.

Next, suppose that $a > b$ and that $(a-b+1)$ divides b. Then it is not possible to displace down to $(a-1,b-1)$. Therefore

$$\begin{aligned}
\delta(a,b) &= \min\{\delta(a-1,b),\delta(a,b-1)\}+1\\
&= \min\{2\mu(a-b,b+1)-a+b,2\mu(a-b+2,b)-a+b-2\}\\
&= 2\cdot\min\{\mu(a-b,b+1)+1,\mu(a-b+2,b)\}-a+b-2.
\end{aligned}$$

By the third part of lemma 6.13, $\mu(a-b+1,b+1) = \mu(a-b+2,b)$. Therefore

$$\delta(a,b) = 2\min\{\mu(a-b,b+1)+1,\mu(a-b+1,b+1)\}-a+b-2.$$

By the second part of lemma 6.13, $\mu(a-b,b+1)+1 \geq \mu(a-b+1,b+2)$, and $\mu(a-b+1,b+2) \geq \mu(a-b+1,b+1)$ by definition. Hence in fact $\delta(a,b) = 2\mu(a-b+1,b+1)-a+b-2$, as desired.

Finally, consider the case that $a > b > 0$ and $(a-b+1) \nmid b$. Then all three displacements are available, and therefore

$$\begin{aligned}\delta(a,b) &= \min\{\delta(a-1,b)+1, \delta(a,b-1)+1, \delta(a-1,b-1)\} \\ &= 2\cdot\min\{\mu(a-b,b+1)+1, \mu(a-b+2,b), \mu(a-b+1,b)\} - a + b - 2.\end{aligned}$$

Therefore it suffices to show that

$$\min\{\mu(a-b,b+1)+1, \mu(a-b+2,b), \mu(a-b+1,b)\} = \mu(a-b+1,b+1).$$

By the fourth part of lemma 6.13, $\mu(a-b+1,b) = \mu(a-b+1,b+1)$. By the fifth part of lemma 6.13, $\mu(a-b+2,b) \geq \mu(a-b+1,b+1)$. Finally, $\mu(a-b,b+1)+1 \geq \mu(a-b+1,b+2)$, and this in turn is at least $\mu(a-b+1,b+1)$. It follow that this minimum is indeed equal to $\mu(a-b+1,b+1)$, and the result follows. □

For example, to compute the difficulties $\delta(a,a)$ of two-row rectangles, it suffices to compute the first column of figure 8. The numbers $\rho_{n,1}$ (the first column of figure 8) form a sequence $1, 3, 5, 9, 11, 17, \cdots$ that determines the displacement difficulty $\delta(a,a)$: the number $\rho_{n,1}$ is the minimal number a such that $\delta(a,a) \geq 2n$. The online encyclopedia of integer sequences [OEIS] reveals that the sequence $\{\rho_{n,1}\}_{n\geq 1}$ is called the *Sieve of Tchoukaillon,* or the *Smarandache consecutive sieve*. More details can be found in [BL95]. This sequence was studied in various guises by several authors; the strongest asymptotic result is the following.

Theorem 6.14 (Broline and Loeb [BL95])**.** *As $n \to \infty$, $\rho_{n,1} = n^2/\pi + O(n)$.*

This result was previous obtained with a weaker error term by Erdös and Jabotinsky [EJ58], who proved that $\rho_{n,1} = n^2/\pi + O(n^{4/3})$ and conjectured the stronger result proved by Broline and Loeb. We observe that Broline and Loeb study a sequence defined in a completely different way (in terms of a certain solitaire game) that is not obviously equivalent to the definition we give above, but it is an elementary exercise to show that they are the same sequence.

A careful study of Broline and Loeb's analysis would give more precise information about $\rho_{n,1}$ for specific values of n (and hence about $\delta(a, a)$ for specific values of a), and it has not escaped our notice that a similarly explicit analysis could be carried out on the table $\rho_{n,m}$ as a whole. Rather that doing this precise analysis, however, we will content ourselves for the time being with some bounds sufficient for proposition 6.10.

Lemma 6.15. *For any integers n, m such that $m \geq \frac{n+1}{2}$,*

$$\rho_{n,m} = m(n - m + 1).$$

Proof. Let $r(n, m) = m(n-m+1) = (\frac{n+1}{2})^2 - (m - \frac{n+1}{2})^2$. Clearly $r(n, n) = n$. Also, m divides $r(n, m)$ for all m, n, and $r(n, m) - r(n, m + 1) = 2m - n$. So as long as $\frac{n+1}{2} \leq m \leq n$, $1 \leq r(n, m) - r(n, m + 1) \leq m$. The result follows. $\square$

Lemma 6.16. *For any positive integers n, m such that $m \leq \frac{n+1}{2}$,*

$$\frac{1}{4}(n + 1)^2 \leq \rho_{n,m} \leq \frac{3}{8}(n + 1)^2.$$

Note that these bounds do not depend on m. They could be improved to depend on m, but it will be slightly cleaner this way.

Proof. Let $\ell = \left\lceil \frac{n+1}{2} \right\rceil$. The previous lemma shows that $\rho_{n,\ell} = \left\lfloor \frac{(n+1)^2}{4} \right\rfloor$. We have $m \leq \ell$, therefore $\rho_{n,m} \geq \rho_{n,\ell}$. In case n is even, we have strict inequality $m < \ell$, and thus $\rho_{n,m} < \rho_{n,\ell}$. This gives the first inequality. For the second, observe that $\rho_{n,m} \leq \rho_{n,1} \leq \rho_{n,\ell} + (\ell - 1) + (\ell - 2) + \cdots + 1 = \rho_{n,\ell} + \binom{\ell}{2}$. If n is even, this upper bound is $\frac{3}{8}(n^2 + 2n)$, while if n is odd this upper bound is $\frac{3}{8}n^2 + \frac{1}{2}n + \frac{1}{8}$. In both cases, we can conclude that $\rho_{n,1} \leq \frac{3}{8}(n + 1)^2$ to obtain the second inequality. $\square$

Proposition 6.10 follows immediately from these two lemmas.

6.3. **Staircase partitions.** A natural class of partitions, which turn out to have extremely low displacement difficulty, are *staircase partitions*, of the form

$$\mathrm{St}^n = (n, n-1, \cdots, 2, 1).$$

For example, the width-11 staircase has displacement difficulty equal to 2 (which will follow from our results). Figure 9 shows one valid path with only 2 one-links (there are many others).

Staircase partitions are potentially of interest as intermediate stages in displacement to other partitions, especially given their very low displacement difficulty.

One can find by explicit enumeration the displacement difficulties for staircases with bottom row up to size 11. These numbers are shown in figure 10. These difficulties were computed using the `difficulty` method from the Java code in the appendix; the computation took several minutes on a MacBook Pro (2.7 GHz processor).

In fact, the displacement difficulty of staircases can be determined explicitly.

Proposition 6.17. *If $n = 1$ or $n = 2$, then $\delta(\mathrm{St}^n) = 1$. Otherwise,*

$$\delta(\mathrm{St}^n) = \begin{cases} 2 & \textit{if } n \equiv 0 \textit{ or } 3 \mod 4 \\ 3 & \textit{if } n \equiv 1 \textit{ or } 2 \mod 4 \end{cases}$$

Proof. First, notice that the only partitions for which $\delta(P) \leq 1$ are either empty or have the form $P = (n, 1, 1, \cdots, 1)$ (where the number of 1's is $n-1$). Therefore for all $n \geq 3$, $\delta(\mathrm{St}^n) \geq 2$. Also, it is always the case that $\delta(P) \equiv |P| \mod 2$. Since $|\mathrm{St}^n| = \binom{n+1}{2}$, it follows that $\delta(\mathrm{St}^n)$ is at least 2 when $n \equiv 0$ or $3 \mod 4$, and is at least 3 when $n \equiv 1$ or $2 \mod 4$. Therefore to prove the proposition, it suffices to show that $\delta(\mathrm{St}^n) \leq 3$ for all n. Since we have seen that this holds for $n \leq 5$ by brute enumeration (figure 10), it suffices to establish the following lemma.

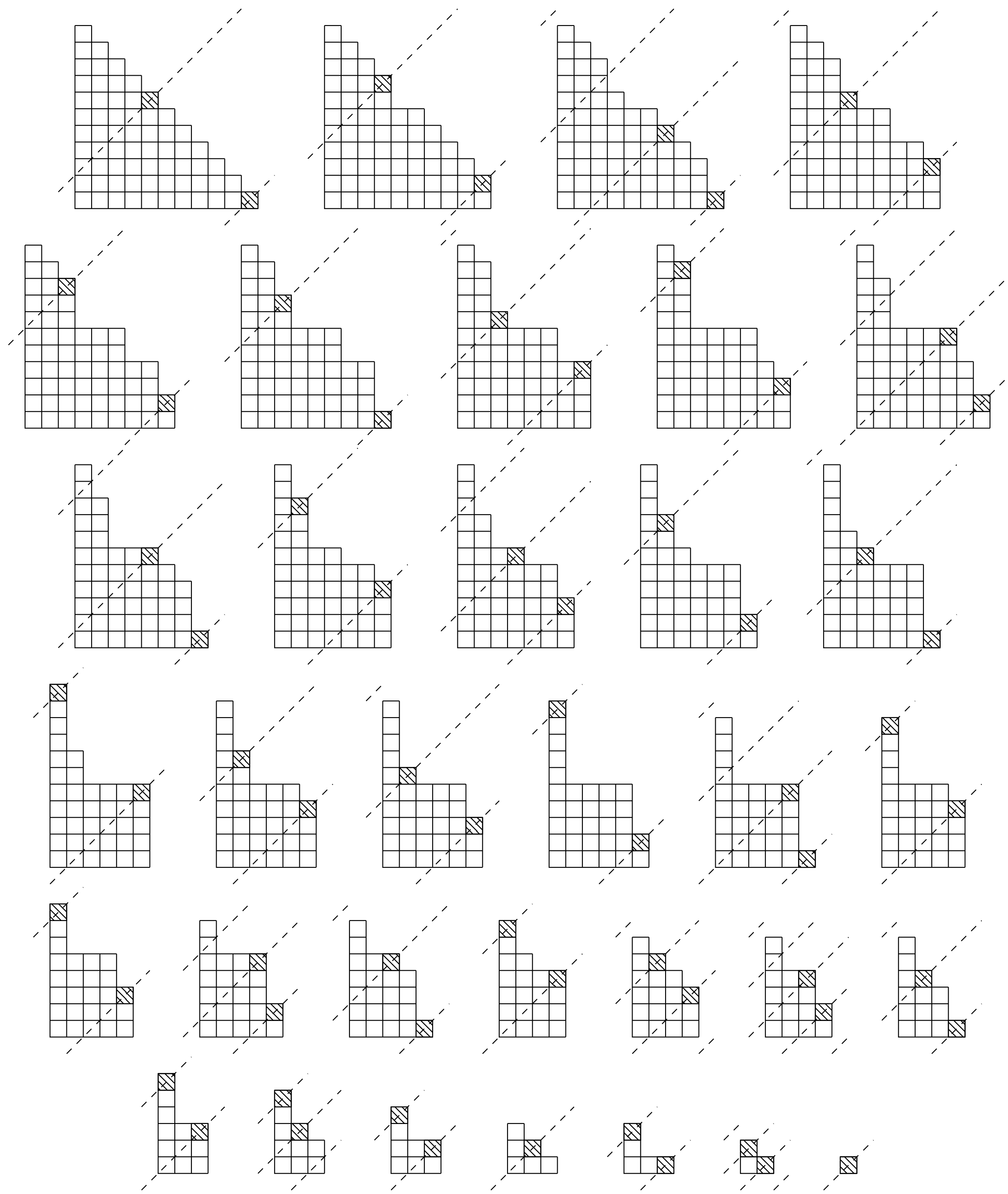


FIGURE 9. An optimal valid path for a width-11 staircase. There are only two one-links.

Lemma 6.18. *For all even n, $\delta(\mathrm{St}^n) \leq \delta(\mathrm{St}^{n-1})$. For all odd $n \geq 7$, $\delta(\mathrm{St}^n) \leq \delta(\mathrm{St}^{n-4})$.*

n	$\delta(\mathrm{St}^n)$	Number of optimal valid paths
1	1	1
2	1	1
3	2	2
4	2	4
5	3	84
6	3	$1,276$
7	2	$1,072$
8	2	$499,076$
9	3	$49,006,368,136$
10	3	$958,752,905,866,440$
11	2	$842,920,611,868,327,240$

FIGURE 10. Displacement difficulties, and the number of optimal valid paths, for small staircases.

For n even, the n boxes in St^n that are not in St^{n-1} can be added in $n/2$ pairs, all spaced the same distance apart. It is easy to verify that the arithmetic progression of slope-1 lines through each pair does not meet the Young diagram in any other points. This gives the first claim.

The second claim requires a more subtle construction. Figure 11 indicates the order that boxes should be added to St^7 to obtain St^{11}.

This figure is meant to be interpreted as follows: the first three displacements from St^7 consists of adding one $1a$ block and one $1b$ block (paired so that the distance between the two elements of any pair is the same); the order that this is done does not matter. The next displacement adds $2a$ and $2b$. The next adds $3a$ and $3b$. The next three add one $4a$ and one $4b$ each: again, they are paired so that each pair has equal distance between its members, and the order does not matter. Continue in this manner until all the indicated blocks are added. The reader may verify that each of these displacements is valid.

To generalize this construction to all odd integers $n \geq 7$, simply adjust the lengths of the sequences of 1s, 4s, 6s, and 9s. In general, there with be $\frac{1}{2}(n-5)$ pairs of 1s and pairs of 4s, and $\frac{1}{2}(n-3)$ pairs of 6s and pairs of 9s. There will always be exactly

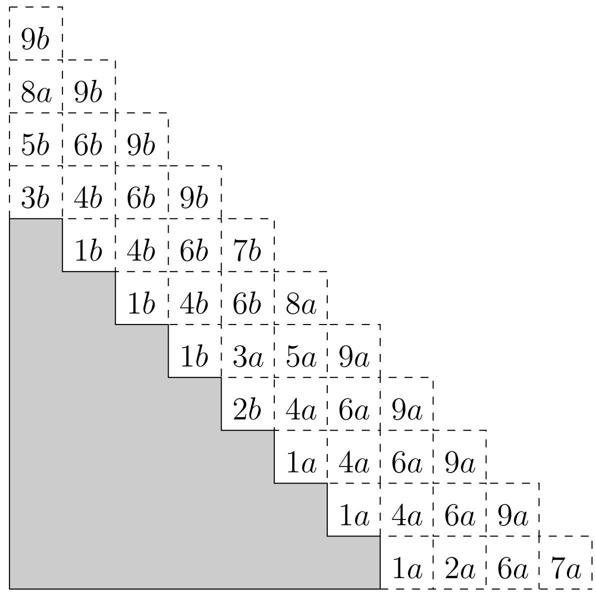


FIGURE 11. The displacement order to obtain one odd-width staircase from an another of width four smaller, demonstrated in the case $n = 11$.

one pair of each of $2, 3, 5, 7$, and 8. The reader may verify that this construction will always give a valid sequence of 2-links from St^{n-4} to St^n. $\square$

6.4. **Rectangular partitions.** Now that we have seen some cases where $\delta(P)$ can be determined exactly, we proceed to the type of partitions that are relevant to theorem A. We do not yet have a sharp result for these partitions (or even an exact asymptotic result), so we will content ourselves with an easy bound. A sharper result should be possible, and would improve theorem A.

Theorem A requires a bound of the difficulty of "rectangular partitions" $((a)^b)$. Experimental evidence suggests that such partitions have rather small difficulty when a and b are both at least 3. Figure 12 shows the displacement difficulties of some small rectangles, computed by dynamic programming using the `difficulty` method from the Java code in the appendix. These values took approximately 5 minutes to compute using a server in the Harvard math department.

In fact, we also have evidence that it is possible to get very short valid paths by simply choosing *random* downward displacements. In figure 13 we show some upper

	2	3	4	5	6	7	8	9	10
2	2	4	4	6	6	6	6	8	8
3	4	5	6	7	6	7	8	7	6
4	4	6	4	6	6	8	4	6	6
5	6	7	6	7	6	5	6	5	4
6	6	6	6	6	6	4	4	4	4
7	6	7	8	5	4	7	4	5	6
8	6	8	4	6	4	4	4	4	4
9	8	7	6	5	4	5	4	5	4
10	8	6	6	4	4	6	4	4	4
11	10	7	6	5	4	5	6	5	4
12	10	6	4	6	4	6	4	4	6
13	10	7	6	5	6	5	6	5	6
14	10	6	4	6	4	4	4	4	4
15	10	5	6	5	4	5	4	5	4

FIGURE 12. Displacement difficulties of small rectangles.

bounds on the displacement difficulty of larger rectangles. These lower bounds were found by finding 100 random valid paths from each partition to the empty partition and taking the minimum length among these, where the random path is chosen by choosing a random 2-link downward at each step (if one exists), or a random 1-link if not, until the partition has size at most 10, after which an optimal valid path is found by brute force enumeration. These values were computed by running the method `depthCharge` from the appendix 100 times on each partition, supplying the argument `getCautious=10`. The computation took several minutes on a Harvard math department server.

On the basis of these data, it appears that box-shaped partitions have very low displacement difficulty, when both sides are at least 3 units long. I previously conjectured that such difficulties are bounded by a constant, but this is not the case.

Proposition 6.19 (David Speyer, personal communication). *The numbers $\delta(P)$, where P ranges over all partitions (a^b) where $a, b \geq 3$, are unbounded.*

Nevertheless, we suspect that the growth rate is quite small.

	2	3	4	5	6	7	8	9	10	11	12	13	14	15
2	2	4	4	6	6	6	6	8	8	10	10	10	10	10
3	4	5	6	7	6	7	8	7	6	7	8	7	8	7
4	4	6	6	8	6	8	6	8	8	8	6	8	6	8
5	6	7	8	9	6	5	8	5	6	7	6	5	8	7
6	6	6	6	6	6	6	6	6	6	6	6	6	6	6
7	6	7	8	5	4	7	8	5	8	7	8	7	8	5
8	6	8	6	8	4	6	10	6	6	8	4	6	6	6
9	8	7	8	5	6	5	8	7	6	7	6	7	8	7
10	8	6	6	6	8	6	6	6	8	6	8	6	4	8
11	10	7	8	5	6	7	6	7	6	9	6	7	6	5
12	10	8	6	6	6	6	4	6	8	6	6	6	6	8
13	10	9	8	5	6	7	6	7	8	7	8	9	6	7
14	10	8	6	6	6	8	8	6	6	8	8	8	6	8
15	10	7	8	7	6	7	6	7	8	7	8	7	8	5
16	10	8	6	6	6	6	6	4	6	6	4	6	4	6
17	12	9	8	7	8	5	6	7	8	5	6	5	6	7
18	12	8	8	8	6	6	8	4	6	8	4	8	4	6
19	12	7	8	7	6	7	6	7	8	5	6	7	8	5
20	12	8	8	8	6	8	4	6	6	8	6	8	6	6
21	14	7	8	7	6	5	8	5	4	7	6	9	6	5
22	14	8	8	6	6	8	6	6	6	6	8	6	4	4
23	14	9	10	7	6	7	6	7	6	7	6	7	6	5
24	14	10	6	8	6	8	6	4	6	6	4	8	6	6
25	14	9	8	7	6	9	8	7	6	7	8	7	6	7

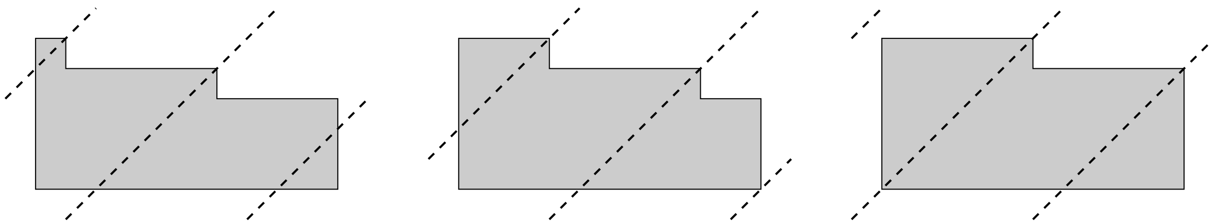
FIGURE 13. *Upper bounds* on displacement difficulties of rectangular partitions, found by computing lengths of 100 randomly chosen valid paths from each Partition.

The following lemma gives a very weak bound on difficulty of boxes, but it is nevertheless strong enough for our theorem A. This bound can certainly be improved.

Lemma 6.20. *Let P be the partition (a^b), i.e. the partition of the number ab into b equal parts, where $a, b \geq 2$. Then $\delta(P) \leq a + 3b - 5$.*

Proof of lemma 6.20. The proof will be by explicit construction of a sequence of partitions. First consider the case where a **is even**.

Define the following intermediate partitions: $P_{k,i} = (a^k\ (i + \frac{1}{2}a)\ i)$ (see figure 14), for $k \geq 0$ and $i \in \{0, 1, \cdots, \frac{1}{2}a\}$.



$P_{k,1}$ $P_{k,i}$ $P_{k,\frac{1}{2}a}$

FIGURE 14. The intermediate partitions $P_{k,i}$ used in the proof of lemma 6.20, together with the progressions $\Lambda_{k,i}$. The partition is it's own upward displacement for all values of i except possibly one (shown in the middle).

Let $\Lambda_{k,i}$ denote the arithmetic progression generated by the two diagonals meeting the outward corners of $P_{k,i}$ at the ends of rows k and $k+1$ (see figure 14). That is, $\Lambda_{k,i}$ is generated by the elements $P_k - k - 1 = i + \frac{1}{2}a - k - 1$ and $P_{k+1} - k - 2 = i - k - 2$. Then $\Lambda_{k,i} = \{n : \ n \equiv i - k - 2 \mod (\frac{1}{2}a + 1)\}$. Observe that if $1 \leq i \leq \frac{1}{2}a$, then $\Lambda_{k,i}$ does not meet the other outward-facing corner of the Young diagram (because $P_{k-1} - (k-1) - 1 = a - k \notin \Lambda_{k,i}$, since the next largest element of $\Lambda_{k,i}$ after $i + \frac{1}{2}a - k - 1$ is $i + a - k$ and we are assuming that $i \geq 1$), so it follows that

$$(P_{k,i})^-_{\Lambda_{k,i}} = P_{k,i-1} \text{ when } i > 0.$$

Now consider the upward displacement. The only inward-turned corner that $\Lambda_{k,i}$ can meet is the one at the end of the first row of the Young diagram; this corresponds to the value $P_0 - 0 = a$. From this we can conclude that

$$(P_{k,i})^+_{\Lambda_{k,i}} = P_{k,i} \text{ unless } k > 0 \text{ and } a \equiv i - k - 2 \mod (\frac{1}{2}a + 1).$$

For a fixed positive value of k, there is at most one value $i \in \{1, 2, \cdots, \frac{1}{2}a\}$ such that the congruence above holds. Therefore the sequence of partitions

$$P_{k,0} < P_{k,1} < \cdots < P_{k,\frac{1}{2}a}$$

is nearly a valid sequence of partitions; at most one adjacent pair is invalid. By inserting an intermediate partition at that place (if necessary), we obtain a valid sequence of partitions with at most two steps increasing the sum by only 1. Therefore $\delta(P_{k,\frac{1}{2}a}) \leq 2+\delta(P_{k,0})$. For $k = 0$, the original sequence is already valid, so $\delta(P_{0,\frac{1}{2}a}) \leq \delta(P_{0,0})$.

Since $P_{k,\frac{1}{2}a} = P_{k+1,0}$, it follows from this analysis that

$$\delta(P_{b-1,0}) \leq 2(b-2) + \delta(P_{0,0}).$$

Now, $P_{b-1,0} \leq (a^b)$ with $|(a^b)| - |P_{b-1,0}| = \frac{1}{2}a$ and $|P_{0,0}| = \frac{1}{2}a$. From this it follows (by a sequence of displacements along singleton progressions) that

$$\delta((a^b)) \leq a + 2b - 4 \text{ when } a \text{ is even.}$$

Now, if a is odd, then $\delta(((a-1)^b)) \leq a + 2b - 5$, and $((a-1)^b)$ can be linked to (a^b) by a length b sequence of length b. Therefore

$$\delta((a^b)) \leq a + 3b - 5 \text{ when } a \text{ is odd.}$$

So whether a is even or odd, $\delta((a^b)) \leq a + 3b - 5$. □

Appendix: source code

The following java code was used to compute displacement difficulties of partitions, cited elsewhere in this thesis. This code defines a class `Verbatim`, with methods which compute the displacement difficulty via dynamic programming (more precisely, via recursion with caching).

```
import java.util.*;

public class Partition {

	//Positive partition elements, arranged in nonincreasing order.
	private int[] parts;
	//Arrays of inward and outward corners, from right to left.
	private Corner[] iC,oC;
	private int size;

	//Constructs a partition from its elements.
	//NOTE this constructor is very rarely called directly. Instead,
	   the lookup method is used to find the "standard" object (to
	   prevent repetition).
	private Partition(int[] parts) {
		parts = truncate(parts); //remove terminal zeros
		this.parts = parts;
		computeCorners(); //Initialize the corner arrays
		size = 0;
		for (int n : parts) size += n;
	}
```

```
//Either constructs a new Partition object from its parts, or else
   returns the existing one.
//ALWAYS USE THIS to get Partitions from arrays of their parts.
public static Partition lookUp(int[] parts) {
       PNode n = PNode.getNode(parts);
       if (n.p != null) return n.p;
       n.p = new Partition(parts);
       return n.p;
}



//Modify the partition by turning corners in or out. Returns null
   if the input is not the right kind of corner.
public Partition turnOut(Corner c) {
       int[] newParts;
       if (c.y == parts.length && c.x == 0) {
              newParts = new int[parts.length+1];
              newParts[parts.length] = 0;
       }
       else {
              newParts = new int[parts.length];
       }
       for (int i=0; i<parts.length; i++) newParts[i] = parts[i];
       if (c.y >= 0 && c.y < newParts.length && newParts[c.y] == c.
          x && (c.y == 0 || c.x < newParts[c.y-1])) {
              newParts[c.y]++;
              return lookUp(newParts);
       } else {
```

```
            System.out.println("Tried to turn out (" + c.x + ","
                + c.y + ") in " + toString());
            return null; //Not actually an inward corner
        }
}
public Partition turnIn(Corner c) {
        int[] newParts = Arrays.copyOf(parts, parts.length);
        if (c.y >= 0 && c.y < parts.length && parts[c.y] == c.x+1 &&
             (c.y == parts.length-1 || parts[c.y+1] < parts[c.y])) {
            newParts[c.y]--;
            return lookUp(newParts);
        } else {
            System.out.println("Tried to turn in (" + c.x + "," +
                 c.y + ") in " + toString());
            return null; //Not actually an outward corner
        }
}

//Compute displacement along an arithmetic progression a mod m.
//The case m=0 means the singleton progression {a}.
public Partition displaceUp(int a, int m) {
        Partition result = this;
        for (Corner c : iC) {
            if (APContains(a,m,c.x-c.y)) result = result.turnOut(
                c);
        }
        return result;
}
```

```
public Partition displaceDown(int a, int m) {
        Partition result = this;
        for (Corner c : oC) {
                if (APContains(a,m,c.x-c.y)) result = result.turnIn(c
                    );
        }
        return result;
}
public Partition displaceUp(Corner c1, Corner c2) {
        return displaceUp(c1.x-c1.y,c1.x-c1.y-c2.x+c2.y);
}
public Partition displaceDown(Corner c1, Corner c2) {
        return displaceDown(c1.x-c1.y,c1.x-c1.y-c2.x+c2.y);
}
public static boolean APContains(int a, int m, int n) {
        if (m == 0) return a == n;
        else return ((a-n)%m == 0);
}

//Recursively compute difficulty. Caches results of previous calls.
private int diff = -1; //Set to -1 until computed, then the true
    value is cached.
public int difficulty() {
        if (diff != -1) return diff;

        int bestFound = size; //Larger than difficulty could
            possibly be.
```

```
        //Try all possible 2-links
        for (int i=0; i<oC.length; i++) {
                for (int j=0; j<i; j++) {
                        Partition parent = displaceDown(oC[i],oC[j]);
                        Partition child = displaceUp(oC[i],oC[j]);
                        if (parent.size == size-2 && equals(child)) {
                            //Make sure it's really a link
                                int candidate = parent.difficulty();
                                if (candidate < bestFound) bestFound =
                                    candidate;
                        }
                }
        }

        //Try all 1-links
        for (Corner c : oC) { //Try all 1-links
                Partition parent = turnIn(c);
                int candidate = parent.difficulty()+1;
                if (candidate < bestFound) bestFound = candidate;
        }

        diff = bestFound; //We've found the difficulty, so cache it.
        return diff;
}

//Computes the number of optimal paths
public int numPaths() {
        int diff = difficulty();
```

```
        if (size == 0) return 1; //Base case.
        int sum = 0;

        //Try all possible 2-links
        for (int i=0; i<oC.length; i++) {
                for (int j=0; j<i; j++) {
                        Partition parent = displaceDown(oC[i],oC[j]);
                        Partition child = displaceUp(oC[i],oC[j]);
                        if (parent.size == size-2 && equals(child)) {
                            //Make sure it's really a link
                                int candidate = parent.difficulty();
                                if (candidate == diff) sum += parent.
                                    numPaths();
                        }
                }
        }

        //Try all 1-links
        for (Corner c : oC) { //Try all 1-links
                Partition parent = turnIn(c);
                int candidate = parent.difficulty()+1;
                if (candidate == diff) sum += parent.numPaths();
        }
        return sum;
}

//Prints some optimal path
public void printPath() {
```

```
        int diff = difficulty();
        if (size == 0) return; //Base case.
        Partition par = null; //Will find a parent

        //Try all possible 2-links
        for (int i=0; i<oC.length; i++) {
                for (int j=0; j<i && par == null; j++) {
                        Partition parent = displaceDown(oC[i],oC[j]);
                        Partition child = displaceUp(oC[i],oC[j]);
                        if (parent.size == size-2 && equals(child)) {
                            //Make sure it's really a link
                                int candidate = parent.difficulty();
                                if (candidate == diff) par = parent;
                        }
                }
        }

        //Try all 1-links
        for (Corner c : oC) { //Try all 1-links
                if (par != null) break;
                Partition parent = turnIn(c);
                int candidate = parent.difficulty()+1;
                if (candidate == diff) par = parent;
        }
        System.out.println(this.toString());
        par.printPath();
}
```

```
//Computes an upper bound on difficulty, using randomization.
//Randomly displaces the partition down (by 2-links when possible)
   to a partition of the chosen size, then computes difficulty
   exactly.
public int depthCharge(int getCautious) {
        if (size <= getCautious) return difficulty();
        ArrayList<Partition> links = new ArrayList<Partition>();
        for (int i=0; i<oC.length; i++) {
                for (int j=0; j<i; j++) {
                        Partition parent = displaceDown(oC[i],oC[j]);
                        Partition child = displaceUp(oC[i],oC[j]);
                        if (parent.size == size-2 && equals(child)) {
                           //Make sure it's really a link
                              links.add(parent);
                        }
                }
        }
        if (links.size() > 0) {
                int i = r.nextInt(links.size());
                return links.get(i).depthCharge(getCautious);
        }
        //Otherwise, no 2-links; choose a random 1-link.
        int i = r.nextInt(oC.length);
        return 1+turnIn(oC[i]).depthCharge(getCautious);
}
public static Random r = new Random();
```

```
//Computes the arrays of inward and outward corners. Invoked in the
     constructor only.
private void computeCorners() {
        ArrayList<Corner> inList = new ArrayList<Corner>();
        ArrayList<Corner> outList = new ArrayList<Corner>();
        for (int i=0; i<parts.length; i++) {
                if (i==0 || parts[i] < parts[i-1]) inList.add(new
                    Corner(parts[i],i));
                if (i==parts.length-1 || parts[i] > parts[i+1])
                    outList.add(new Corner(parts[i]-1,i));
        }
        inList.add(new Corner(0,parts.length));
        iC = new Corner[inList.size()];
        iC = inList.toArray(iC);
        oC = new Corner[outList.size()];
        oC = outList.toArray(oC);
}

//Just a container for two integers. These refer to the coordinates
     of the lower-left corner of the square in question.
private static class Corner {
        public int x;
        public int y;
        public Corner(int x, int y) {
                this.x = x;
                this.y = y;
        }
```

```
}

//Function for convenience: removes any zeros from the end of an
    array.
//Assumes that arr is in nonincreasing order.
public static int[] truncate(int[] arr) {
        int l = arr.length;
        if (l == 0 || arr[l-1] > 0) return arr;
        while (l>0 && arr[l-1] == 0) l--;
        int[] result = new int[l];
        for (int i=0; i<l; i++) result[i] = arr[i];
        return result;
}

public String toString() {
        String result = "(";
        for (int i=0; i<parts.length; i++) {
                result += parts[i];
                if (i<parts.length-1) result += ",";
        }
        result += ")";
        return result;
}
public boolean equals(Partition p) {
        return (lookUp(parts) == lookUp(p.parts));
}
```

```
//Inner class used to index the Partition cache for the lookUp
   method.
private static class PNode {
        private Partition p;
        private int[] parts;
        private PNode[] children;
        private PNode(int[] parts) {
                p = null;
                this.parts = truncate(parts);
                children = new PNode[(parts.length>0)? parts[parts.
                    length-1]+1 : 0]; //The node of the empty
                    partition has no children.
                for (int i=0; i<parts[parts.length-1]; i++) children[
                    i] = null;
        }

        private static ArrayList<PNode> roots = new ArrayList<PNode
            >(); //Nodes for one-row partitions, at the bottom.
        public static PNode getNode(int[] parts) {
                int bottom = (parts.length>0)? parts[0] : 0;
                PNode root = getRoot(bottom);
                return getNodeRec(parts,root,1);
        }
        private static PNode getRoot(int bottom) {
                if (roots.size() > bottom) return roots.get(bottom);
                else {
                        for (int n=roots.size(); n<=bottom; n++) {
                                int[] parts = new int[1];
```

```
                                parts[0] = n;
                                roots.add(new PNode(parts));
                        }
                        return roots.get(bottom);
                }
        }
        private static PNode getNodeRec(int[] parts, PNode curr, int
             startIndex) {
                if (startIndex >= parts.length) return curr;
                PNode nextNode;
                int nextPart = parts[startIndex];
                if (curr.children[nextPart] != null) {
                        nextNode = curr.children[nextPart];
                } else {
                        int[] newParts = Arrays.copyOf(parts,
                            startIndex+1);
                        curr.children[nextPart] = new PNode(newParts);
                        nextNode = curr.children[nextPart];
                }
                return getNodeRec(parts,nextNode,startIndex+1);
        }
    }
}
```

References

[ACGH] E. Arbarello, M. Cornalba, P. Griffiths, J. Harris: *Geometry of Algebraic Curves, vol. 1.* Springer-Verlag (1985).

[ACG] E. Arbarello, M. Cornalba, P. Griffiths: *Geometry of Algebraic Curves, vol. 2.* Springer-Verlag (2011).

[B80] R.-O. Buchweitz: *On deformations of monomial curves.* Lecture Notes in Mathematics 777, Springer, Berlin, 205–220 (1980).

[BA08] M. Bras-Amorós: *Fibonacci-like behavior of numerical semigroups of a given genus.* Semigroup Forum, vol. 7, no. 2, 379-384 (2008).

[BN] A. Brill, M. Noether: *Über die algebraischen Functionen und ihre Anwendungen in der Geometrie.* Math. Ann. 7, 269–310 (1873).

[BL95] D. M. Broline, D. E. Loeb: *The combinatorics of Mancala-type games: Ayo, Tchoukaillon and* $1/\pi$. J. Undergrad. Math. Applic., vol. 16, 21–36 (1995).

[B13] E. Bullock: *Subcanonical points on algebraic curves.* Transactions of the American Mathematical Society 365.1, 99–122 (2013).

[dC] A. del Centina: *Weierstrass points and their impact in the study of algebraic curves: a historical account from the "Lückensatz" to the 1970s.* Ann. Univ. Ferrara 54, 37–59 (2008).

[CS13] A. Contiero, K.-O. Stöhr: *Upper bounds for the dimension of moduli spaces of curves with symmetric Weierstrass semigroups.* J. London Math. Soc. 2:88, 580–598 (2013).

[CDPR] F. Cools, J. Draisma, S. Payne, E. Robeva: *A tropical proof of the Brill-Noether Theorem.* Adv. Math. 230 no. 2, 759–776 (2012).

[E93] D. Edidin: *Brill-Noether theory in codimension-two.* J. Algebraic Geom. 2(1), 25–67 (1993).

[EH83] D. Eisenbud, J. Harris: *A simpler proof of the Gieseker-Petri theorem on special divisors.* Invent. Math. 74, 269–280 (1983).

[EH85] D. Eisenbud, J. Harris: *On varieties of minimal degree (a centennial account).* Algebraic geometry, Bowdoin (Brunswick, Maine), 3–13 (1985).

[EH87] D. Eisenbud, J. Harris: *Existence, decomposition, and limits of certain Weierstrass points.* Invent. Math. 87, 495–515 (1987).

[EH89] D. Eisenbud, J. Harris: *Irreducibility of some families of linear series with Brill-Noether number* -1. Annales scientifiques de l'É.N.S. 4^e serie 22:1, 33–53 (1989).

[EH86] D. Eisenbud, J. Harris: *Limit linear series: basic theory.* Invent. Math. 85, 337–371 (1986).

[EJ58] P. Erdős, E. Jabotinsky: *On Sequences of Integers Generated by a Sieving Process.* Nederl. Akad. Wetensch. Proc. Ser. A 61, and Indag. Math. 20, 115–128 (1958).

[FL81] W. Fulton, R. Lazarsfeld: *On the connectedness of degeneracy loci and special divisors.* Acta Math. 146, 271–283 (1981).

[G82] D. Gieseker: *Stable curves and special divisors.* Invent. Math. 66, 251–275 (1982).

[GH80] P. Griffiths, J. Harris: *On the variety of special linear systems on a general algebraic curve.* Duke Math. Journal 47:1, 233–272 (1980).

[GH78] P. Griffiths, J. Harris: *Principles of algebraic geometry.* Wiley-Interscience (1978).

[H09] J. Harris: *Brill-Noether theory.* Surveys in Differential Geometry XIV, 131–143 (2009).

[H82] J. Harris (with the collaboration of D. Eisenbud): *Curves in projective space.* Séminaire de Mathématiques Supérieures 85. Presses de l'Université de Montréal, Montreal, Que. (1982).

[HM] J. Harris, I. Morrison: *Moduli of Curves.* Springer-Verlag 1991.

[H77] R. Hartshorne: *Algebraic Geometry.* Springer-Verlag (1977).

[KY13] N. Kaplan, L. Ye: *The proportion of Weierstrass semigroups.* J. Algebra 373, 377–391 (2013).

[K71] G. Kempf: *Schubert methods with an application to algebraic curves.* Publ. Math. Centrum, Amsterdam, 1971.

[K91] J. Komeda: *On primitive Schubert indices of genus g and weight $g-1$.* J. Math. Soc. Japan 43:3, 437–445 (1991).

[KL72] S. Kleiman, D. Laksov: *On the existence of special divisors.* Amer. J. Math. 94, 431–436 (1972).

[KL74] S. Kleiman, D. Laksov: *Another proof of the existence of special divisors.* Acta Math. 132, 163–176 (1974).

[L86] R. Lazarsfeld: *Brill-Noether-Petri without degenerations.* J. Differential Geom. 23, 299–307 (1986).

[L91] A. F. Lopez: *On the existence of components of the Hilbert scheme with the expected number of moduli.* Math. Ann. 289, no. 3, 517–528 (1991).

[L99] A. F. Lopez: *On the existence of components of the Hilbert scheme with the expected number of moduli II.* Comm. Algebra 27, no. 7, 3485–3493 (1999).

[Oss06] B. Osserman: *A limit linear series moduli scheme.* Ann. Inst. Fourier (Grenoble) 56:4, 1165–1205 (2006).

[Oss11] B. Osserman: *A simple characteristic-free proof of the Brill-Noether theorem.* Preprint: arxiv.org/abs/1108.4967v1

[P89] G. Pareschi: *Components of the Hilbert scheme of smooth space curves with the expected number of moduli.* Manuscripta Math. 63, no. 1, 1–16 (1989).

[P24] K. Petri: *Über Spezialkurven I.* Math. Ann. 93, 182–209 (1924).

[S84] E. Sernesi: *On the existence of certain families of curves.* Invent. Math. 75, no. 1, 25–57 (1984).

[OEIS] N. Sloane: *The on-line encyclopedia of integer sequences.*

[S93] K.-O. Stöhr. *On the moduli spaces of Gorenstein curves with symmetric Weierstrass semigroups.* J. reine angew. Math. 441, 189–213 (1993).

[W85] G. Welters: *A theorem of Gieseker-Petri type for Prym varieties.* Annales Scientifiques de l'École Normale Supérieure 18, no. 4, 671–683 (1985).

[Z13] A. Zhai: *Fibonacci-like growth of numerical semigroups of a given genus.* Semigroup Forum, vol. 86, no. 3, 634–662 (2013).