BOLETIM
DA SOCIEDADE BRASILEIRA DE MATEMÁTICA

Bol. Soc. Bras. Mat., Vol. 21, No. 1, 11-50

Some Arithmetic Properties of Weierstrass Points: Hyperelliptic Curves

Joseph H. Silverman

Abstract. The set of Weierstrass points for pluricanonical linear systems on an algebraic curve C have been extensively studied from a geometric viewpoint. If the curve is defined over a number field k, then these n th order Weierstrass points are defined over an algebraic extension k_n of k, and it is an interesting question to ask for the arithmetic properties of the points and the extension that they generate. In this paper we begin the study of the arithmetic properties of higher order Weierstrass points in the special case of hyperelliptic curves. We give an upper bound for the average height of these points, and we show that for sufficiently large primes p, the first order Weierstrass points and the n th order Weierstrass points remain distinct modulo p. This limits to some extent the ramification that can occur in the extension k_n/k. We also present two numerical examples which indicate that a complete description of the ramification is likely to be complicated.

0. Introduction

Let C be a smooth projective curve of genus $g \geq 2$ defined over a field k of characteristic 0. If P is any point of C, then the Riemann-Roch theorem [[4], IV.1.3] says that there exists a non-constant function on C which has a pole of order at most $g+1$ at P and no other poles. For most points, the pole of such a function will have order exactly $g+1$. The few remaining points are called *Weierstrass points*. Equivalently, by the Riemann-Roch theorem, a point $P \in C$ is a Weierstrass point if

$$\ell(K_C - g(P)) \geq 1,$$

where K_C is a canonical divisor on C.

A curve has only finitely many Weierstrass points. More precisely, it is possible to assign weights to the Weierstrass points; and then the total weight of the Weierstrass points is $g^3 - g$. The set of Weierstrass points is defined over

Received 22 May 1990.
Research partially supported by NSF DMS-8842154 and a Sloan Foundation Fellowship.

the field k, and it is an interesting arithmetic question to study the field generated by these points. Some computations have been done for some special families of curves, such as Fermat curves and modular curves (cf. [11],) but in general little seems to be known.

One can also define higher order Weierstrass points by replacing K_C with some other divisor D, choosing $s \geq 1$ so that for most points $\ell(D - s(P)) = 0$, and then asking for which points is $\ell(D - s(P)) \geq 1$? Classically, one takes $D = nK_C$ for $n = 1, 2, \ldots$ and defines the set of n^{th} *order Weierstrass points of* C to be the set

$$C[n] = \{P \in C : \ell(nK_C - s(P)) \geq 1\},$$

where

$$s = \begin{cases} g & \text{if } n = 1, \\ (2n-1)(g-1) & \text{if } n \geq 2. \end{cases}$$

Again one can assign weights (cf. Sections 1, 3) in such a way that (for $n \geq 2$) the total weight of the n^{th} order Weierstrass points is $s^2 g$.

Mumford [[8], page 11] has suggested that the sets $C[n]$ are the analogue, for curves of genus $g \geq 2$, of the sets $E[n]$ of points of order n on curves of genus 1 (elliptic curves.) For example, one can use Weierstrass points to rigidify the space of curves and construct moduli spaces [[8], Appendix 7C], much as is done with torsion points on elliptic curves and abelian varieties. Further, generalizing an unpublished result of Mumford, Neeman [9] has shown that the sets $C[n]$ as $n \to \infty$ are uniformly distributed in $C(\mathbb{C})$ relative to the Bergman measure on C.

On the other hand, in characteristic p the sets $C[n]$ do not seem to behave as nicely as the corresponding $E[n]$'s. (See, e.g., [6],[10],[12].) This indicates that the arithmetic theory of the $C[n]$'s is likely to be complicated. However, in view of the vast richness of the arithmetic theory of torsion points on elliptic curves, it seems worthwhile to pursue Mumford's analogy and study the arithmetic theory of higher order Weierstrass points. The purpose of this article is to begin such a study, concentrating principally on hyperelliptic curves and using mostly elementary, but computationally rather intricate, methods. Our hope is that this will provide a foundation on which to build a general theory.

We now briefly summarize the sorts of results we obtain. All results are for hyperelliptic curves C defined over a number field k. For precise statements of the main theorems and some numerical examples, see Sections 1 and 2.

The arithmetic complexity of an algebraic point on C is measured by its height, say relative to a rational function $x : C \to \mathbf{P}^1$. We show (Theorem 1.1) that the average of the heights $h(x(P))$ for $P \in C[n]$ is bounded above by $O_C(\log n)$ as $n \to \infty$. Using our height estimate, we give a weak lower bound for the degree of the field generated by $C[n]$ (Corollary 1.2). Unfortunately, for a fixed curve C, this lower bound does not go to infinity with n.

For general curves (of genus ≥ 4), the $C[n]$'s for differing n's are undoubtedly disjoint; but for hyperelliptic curves one always has $C[1] \subset C[n]$. Let v be a place of $\bar{k}$ of characteristic p. Suppose that $p > 2n(g-1)$ and that C has good reduction at v. We then show that if $P \in C[n]$ and if the reduction $\tilde{P} \pmod{v}$ lies in $\widetilde{C[1]} \pmod{v}$, then $P \in C[1]$. This limits (to some extent) the ramification in $k(C[n])$.

Of course, the question one really wants to answer is for which v does the reduction map $C[n] \to \tilde{C} \pmod{v}$ fail to be injective. In Section 2 we present some numerical data for the curves $y^2 = x^6 + 1$ and $y^2 = x^5 + 1$ which shows that the set of such v behaves rather irregularly, but at the same time consists only of comparatively small primes. At present, we are unable to give a theoretical explanation for this behavior.

1. The main theorems

We begin by setting some notation, which will remain fixed throughout this paper.

k	a number field.
C/k	a hyperelliptic curve defined by $y^2 = f(x)$, where $f(x) \in k[x]$ is a monic polynomial with $\mathrm{Disc}(f) \neq 0$.
$g \geq 2$	the genus of C.
$n \geq 2$	an integer.
s	$= (2n-1)(g-1)$.
I	$= n(g-1) + 1$.
J	$= (n-1)(g-1) - 1 = s - I$.

$C[n]$ the set of n^{th} *order Weierstrass points* of C, defined as

$$C[n] = \{P \in C : \ell(nK_C - s(P)) \geq 1\}.$$

We note that for hyperelliptic curves one always has the inclusion $C[1] \subset C[n]$. The set $C[1]$ consists of the points with $y = 0$, together with one point "at infinity" if $\deg(f)$ is odd.

$\mathrm{wt}(P)$ the *weight* of a Weierstrass point $P \in C[n]$. See Section 3 for the definition of $\mathrm{wt}(P)$. We note that the total weight $\sum_{P \in C[n]} \mathrm{wt}(P)$ equals $s^2 g$, while the total weight excluding the points in $C[1]$ is

$$\sum_{\substack{P \in C[n] \\ P \notin C[1]}} \mathrm{wt}(P) = 4gIJ.$$

Our first result gives an upper bound for the average height of the points in $C[n]$.

Theorem 1.1. *Let $h: \bar{\mathbb{Q}} \to [0, \infty]$ be the absolute logarithmic height function. (See, e.g., [5].) Let $h(f)$ denote the height of the projective point defined by the coefficients of the polynomial $f(x)$. Then*

$$\frac{1}{4gIJ} \sum_{\substack{P \in C[n] \\ P \notin C[1]}} \mathrm{wt}(P) \cdot h(x(P)) \leq \frac{h(f) + 3 \log n}{g} + O(1).$$

The $O(1)$ is an absolute constant.

We can use Theorem 1.1 to give a weak estimate for the degree of the field generated by $C[n]$.

Corollary 1.2.

$$[k(C[n] \smallsetminus C[1]) : k] \gg \sqrt{\frac{\log ng + O(1)}{\dfrac{h(f) + \log n}{g} + O(1)}}.$$

The $\gg$ constant and the $O(1)$ constants are absolute.

As a particular example of Corollary 1.2, consider the curves

$$C_d : y^2 = x(x-1)(x-2)\cdots(x-d).$$

Then

$$[\mathbb{Q}(C_d[n]) : \mathbb{Q}] \gg \frac{\log n}{\log d} \quad \text{as } n \to \infty \text{ subject to } \frac{\log n}{d \log d} \ll 1.$$

Next we look at the behavior of the Weierstrass points "modulo p." We set a bit more notation:

v a place of $\bar{k}$ of characteristic $p > 0$.

ord_v the extension to $\bar{k}^*$ of the usual valuation $\mathrm{ord}_p : \mathbb{Q}^* \to \mathbb{Z}$.

$$\Phi_n(x) = \prod_{\substack{P \in C[n] \\ P \notin C[1]}} (x - x(P))^{\mathrm{wt}(P)/2}.$$

Notice that in the definition of $\Phi_n(x)$ we have used half the weight in the exponent. We do this because for the hyperelliptic curve $y^2 = f(x)$, if $P = (x, y)$ is in $C[n]$, then $(x, -y)$ is also in $C[n]$. So as P runs over $C[n] \smallsetminus C[1]$, the $x(P)$'s each appear twice. So $\Phi_n(x)$ will be a polynomial in $k[x]$.

Theorem 1.3. *Assume that $\deg f = 2g + 1$ is odd.*

(a) $\mathrm{Resultant}(\Phi_n, f) = \pm \mathrm{Disc}(f)^{IJ}$.

(b) *Assume that $p > 2n(g-1)$ and that the coefficients of f are v-integral. Then the coefficients of $\Phi_n(x)$ are v-integral.*

We can use this estimate to limit when points of $C[n]$ and $C[1]$ come together mod v.

Corollary 1.4. *Assume that $p > 2n(g-1)$ and that C has good reduction at v. If $P \in C[n]$ has the property that*

$$\tilde{P} \pmod{v} \in \widetilde{C[1]} \pmod{v},$$

then $P \in C[1]$.

A fancier way to restate Corollary 1.4 is as follows. First embed C in its Jacobian $\mu : C \hookrightarrow J$ by sending one of the points in $C[1]$ to the origin. As usual, let $J_1(\bar{k}_v)$ be the kernel of the reduction map $J(\bar{k}) \to \tilde{J} \pmod{v}$. Corollary 1.4 says that if $p > 2n(g-1)$ and if C has good reduction at v, then

$$\mu(C[n]) \cap J_1(\bar{k}_v) = \{0\}. \tag{1}$$

Notice that (1) remains true if we replace $C[n]$ by $J[n]$, the set of n-torsion points of J, provided C has good reduction and $p \nmid n$. This fact, of course, is central to the arithmetic study of Jacobians and abelian varieties. For example,

it is one of the crucial facts needed for the proof of the Mordell-Weil theorem. However, since $J[n]$ is a group, equation (1) for $J[n]$ tells us that $J[n]$ injects into the reduction $\tilde{J} \pmod{v}$. Unfortunately, $C[n]$ is not a group; so what we should really look at is the extent to which the map $C[n] \to \tilde{C} \pmod{v}$ fails to be one-to-one. Ignoring the points in $C[1]$, which we already understand, this means we want to know which primes divide the discriminant of the polynomial Φ_n. In the next section we present two examples which show that the answer must be fairly complicated.

2. Some Numerical Examples

In this section we present some numerical data. We begin with the curve

$$C : y^2 = x^6 + 1. \tag{2}$$

From the definition of Weierstrass points it is clear that the set $C[n]$ is invariant under any automorphism of C. In particular, for the curve (2), any point $P = (x, y) \in C[n]$ with $xy \neq 0$ gives rise to twelve points $(\zeta x, \pm y)$, where ζ is any sixth root of unity. It follows that up to a power of x, the polynomial $\Phi_n(x)$ defined in Section 1 is actually a polynomial in x^6. So we will work instead with the polynomial

$$\Psi_n(x) = c\frac{\Phi_n(x^{1/6})}{x^{a/6}} = c \prod_{\substack{P \in C[n]/Aut(C) \\ P \notin C[1],\, x(P) \neq 0}} (x - x(P))^{\mathrm{wt}(P)}.$$

Here $a = \operatorname{ord}_0 \Phi_n(x)$, and $c \in \mathbb{Z}$ is chosen so that $\Psi_n(x)$ has relatively prime integral coefficients. (I.e. $\Psi_n(x)$ is a primitive polynomial in $\mathbb{Z}[x]$.)

Our first table (Table 2.1) gives the Weierstrass polynomial $\Psi_n(x)$ for the curve (2) and $3 \leq n \leq 6$. Note that for curves of genus 2 such as (2), one has $C[2] = C[1]$, which explains why our table starts with $n = 3$.

A glance at Table 2.1 shows that the polynomials are "palindromic." This is easily explained by the fact that the curve (2) has the additional automorphism $(x, y) \to (x^{-1}, x^{-3}y)$. So if $P = (x, y) \in C[n]$, then $C[n]$ also contains a point whose first coordinate is x^{-1}.

It is also worth mentioning that in all cases listed in Table 2.1, the polynomial $\Psi_n(x)$ is irreducible in $\mathbb{Q}[x]$.

n	$\Psi_n(x)$
3	$4 - 19x + 4x^2$
4	$224 + 14952x - 528x^2 + 107269x^3 - 528x^4 + 14952x^5 + 224x^6$
5	$2048 - 1841664x + 61254528x^2 + 866502560x^3 - 136067400x^4 - 2150998344x^5 - 8234360583x^6 - 2150998344x^7 - 136067400x^8 + 866502560x^9 + 61254528x^{10} - 1841664x^{11} + 2048x^{12}$
6	$524288 - 38338560x + 12633292800x^2 + 2079068160x^3 + 2278700043264x^4 + 14040133779456x^5 + 11330384520000x^6 + 21378405927600x^7 + 88533623365668x^8 + 171597510043663x^9 + 88533623365668x^{10} + 21378405927600x^{11} + 11330384520000x^{12} + 14040133779456x^{13} + 2278700043264x^{14} + 2079068160x^{15} + 12633292800x^{16} - 38338560x^{17} + 524288x^{18}$

The Polynomial $\Psi_n(x)$ for $y^2 = x^6 + 1$

Table 2.1

Corollary 1.4 characterizes those primes for which points of $C[n]$ and points of $C[1]$ may coalesce modulo p. We now address the question of when two points in $C[n]$ not in $C[1]$ may come together modulo p. Equivalently, we ask which primes divide the discriminant of $\Psi_n(x)$? As one expects, this often happens for primes $p < 2n(g-1)$. However, Table 2 shows that there are also larger primes which occur. At present we have no simple way to describe these primes, but that will not prevent us from giving them a name.

Definition. Let C/k be a smooth curve of genus $g \geq 2$. A prime p is called *anomalous* for $C[n]$ if $p > 2n(g-1)$, and if there is some prime v of $\bar{k}$ lying over p such that C has good reduction at v and the reduction map $C(\bar{k}) \to \tilde{C} \pmod{v}$ is not one-to-one on the set of n-Weierstrass points $C[n]$.

Although the anomalous primes appear somewhat irregular, one feature immediately apparent from Table 2.2 is that they seem to be fairly small. If one were to take a random polynomial whose coefficients were roughly the same size as those of Ψ_5 or Ψ_6, one would not expect the discriminant to factor with such small primes.

n	$\mathrm{Disc}(\Psi)$	Factorization of $\mathrm{Disc}(\Psi)$
3	297	$3^3 \cdot 11$
4	$\approx 2.07 \cdot 10^{46}$	$2^{20} \cdot 3^{45} \cdot 5^3 \cdot 7^3 \cdot 11 \cdot 13 \cdot 17^2 \cdot 23^2 \cdot 191 \cdot 193^2$
5	$\approx 3.11 \cdot 10^{190}$	$2^{104} \cdot 3^{221} \cdot 5^{11} \cdot 11^6 \cdot 13 \cdot 17^2 \cdot 23^2 \cdot 29 \cdot 41^2 \cdot 59 \cdot 191 \cdot 193^2 \cdot 271^2 \cdot 421^2 \cdot 2161^2 \cdot 2579$
6	$\approx 1.52 \cdot 10^{417}$	$2^{296} \cdot 3^{467} \cdot 5^{14} \cdot 7^3 \cdot 23^5 \cdot 29 \cdot 31^3 \cdot 41^2 \cdot 43^2 \cdot 47^2 \cdot 59 \cdot 71^2 \cdot 83^2 \cdot 127^2 \cdot 157^2 \cdot 167^2 \cdot 271^2 \cdot 421^2 \cdot 433^2 \cdot 601^2 \cdot 887^2 \cdot 2161^2 \cdot 2579 \cdot 4057 \cdot 5449^2$

The Discriminant of $\Psi_n(x)$ for $y^2 = x^6 + 1$

Table 2.2

m	n	$\mathrm{Resultant}(\Psi_m, \Psi_n)$	Factorization of Resultant
3	4	$\approx 7.68 \cdot 10^{14}$	$2^{14} \cdot 3^{18} \cdot 11^2$
3	5	$\approx 1.12 \cdot 10^{29}$	$2^{22} \cdot 3^{36} \cdot 421^2$
3	6	$\approx 1.13 \cdot 10^{43}$	$2^{40} \cdot 3^{54} \cdot 421^2$
4	5	$\approx 5.11 \cdot 10^{100}$	$2^{76} \cdot 3^{108} \cdot 13^2 \cdot 17^4 \cdot 23^4 \cdot 191^2 \cdot 193^4$
4	6	$\approx 7.91 \cdot 10^{147}$	$2^{94} \cdot 3^{162} \cdot 5^6 \cdot 17^2 \cdot 47^2 \cdot \kappa^*$
5	6	$\approx 1.23 \cdot 10^{289}$	$2^{214} \cdot 3^{324} \cdot 5^{24} \cdot 29^2 \cdot 41^4 \cdot 59^2 \cdot 271^4 \cdot 421^4 \cdot 2161^4 \cdot 2579^2$

*κ=203558536134798796850303533660249 is composite, but is not divisible by "small" primes.

The Resultant of $\Psi_n(x)$ and $\Psi_m(x)$ for $y^2 = x^6 + 1$

Table 2.3

A closer look at Table 2.2 reveals another surprising fact. It appears that an anomalous prime for $C[n]$ is always also anomalous for either $C[n-1]$ or for $C[n+1]$. This suggests that we look for primes such that points in $C[n]$ and $C[n+1]$ come together modulo p; or, equivalently, primes dividing the resultant of Ψ_n and Ψ_{n+1}. The results are compiled in Table 2.3.

Table 2.3 confirms what we half expected. If p is anomalous for n and $n+1$, then it seems to divide the resultant of Ψ_n and Ψ_{n+1}. This suggests that if p is anomalous for n, then at least one of the double roots of $\Psi_n \pmod p$ will also be a double root of either $\Psi_{n-1} \pmod p$ or of $\Psi_{n+1} \pmod p$.

To describe this more intrinsically, we define

$$\Delta_n(p) = \{\tilde{P} \in \tilde{C}(\mathbb{F}_p) : C[n] \to \tilde{C} \pmod p \text{ is not one-to-one over } \tilde{P}\}.$$

Then Table 2.3 leads us to ask if

$$\Delta_n(p) \subset \Delta_{n-1}(p) \cup \Delta_{n+1}(p). \tag{3}$$

Note this is much stronger than merely saying that the anomalous primes match up.

Table 2.4 verifies (3) for the curve (2) and integers $3 \le n \le 5$. In all cases the quantities

$$\frac{\Psi_n}{\gcd(\Psi_n, \Psi_{n+1})} \pmod p$$

and

$$\frac{\Psi_{n+1}}{\gcd(\Psi_n, \Psi_{n+1})} \pmod p$$

are square-free in $\mathbb{F}_p[x]$. Note that the variable x used in Table 2.4 is really the quantity x^6 on the curve $y^2 = x^6 + 1$.

We next give some numerical data gathered for the curve

$$C : y^2 = x^5 + 1. \tag{4}$$

As above, any point $P = (x, y) \in C[n]$ with $x \neq 0$ gives rise to ten points $(\zeta x, \pm y)$, where ζ is any fifth root of unity. It follows that the polynomial $\Phi_n(x)$ defined in Section 1 is essentially a polynomial in x^5, at least if we ignore powers of x. So for the remainder of this section we let

$$\Psi_n(x) = c\frac{\Phi(x^{1/5})}{x^{a/5}} = c \prod_{\substack{P \in C[n]/Aut(C) \\ P \notin C[1],\, x(P) \neq 0}} (x - x(P))^{\mathrm{wt}(P)}.$$

Again $a = \operatorname{ord}_0 \Phi(x)$, and $c \in \mathbb{Z}$ is chosen so that $\Psi_n(x)$ has relatively prime integral coefficients.

n	$n+1$	p	$\gcd(\Psi_n, \Psi_{n+1}) \pmod p$
3	4	11	$(x-1)^2$
4	5	13	$(x-1)^2$
4	5	17	$(x^2-x+1)^2$
4	5	23	$(x+2)^2(x+12)^2$
4	5	191	$(x-1)^2$
4	5	193	$(x+50)^2(x-27)^2$
5	6	29	$(x-1)^2$
5	6	41	$(x^2-4x+1)^2$
5	6	59	$(x-1)^2$
5	6	271	$(x+117)^2(x-44)^2$
5	6	421	$(x+142)^2(x+169)^2$
5	6	2161	$(x+498)^2(x+959)^2$
5	6	2579	$(x-1)^2$

Points Common to $\Delta_n(p)$ and $\Delta_{n+1}(p)$

Table 2.4

The first table (Table 2.5) gives the Weierstrass polynomial $\Psi_n(x)$ for the curve (4) and $3 \le n \le 6$. Notice that the points with $x = 4$ have weight 2 when $n = 5, 6$. (Remember these are the points with $x^5 = 4$ on C.) Also, $x = 4$ gives 3^{rd} order Weierstrass points on C, although only with weight 1; but the points with $x = 4$ are not 4^{th} order Weierstrass points.

Since Ψ_5 and Ψ_6 have a multiple root, their discriminants are zero. So in Table 2.6 we have computed their discriminants after dividing by $(4-x)^2$. We again see in Table 2.6 that the anomalous primes reappear for consecutive n's, and that they are not too large (although for $n = 6$ they are considerably larger than the anomalous primes for the curve $y^2 = x^6 + 1$.)

n	$\Psi_n(x)$
3	$(4-x)(16-108x+x^2)$
4	$-32768-4767744x-3272704x^2-57028608x^3+170240x^4-9451008x^5+224896x^6-22344x^7+7x^8$
5	$(4-x)^2(-25165824+956301312x+20634402816x^2-3177185280x^3-137167994880x^4-471787716608x^5-345653886976x^6-157336670208x^7-18198906880x^8-1658593280x^9-4458384x^{10}-36688x^{11}+x^{12})$
6	$(4-x)^2(12094627905536+2585226714808320x-9399793625333760x^2+581389708311920640x^3+4975917911056056320x^4+6825760756013727744x^5+2413988213807906816 0x^6+10448053128388411392 0x^7+21577075075288399872 0x^8+19292143242996350976 0x^9+10504890768012450201 6x^{10}+4416351499339431936 0x^{11}+19543192767179653120x^{12}+4633662928140779520x^{13}+539433596351938560x^{14}+8318978192722944x^{15}-228260548980480x^{16}-4736403754560x^{17}-4687088560x^{18}-2849880x^{19}+11x^{20})$

The Polynomial $\Psi_n(x)$ for $y^2 = x^5 + 1$

Table 2.5

n	Disc(Ψ)*	Factorization of Disc(Ψ)
3	$\approx 1.86 \cdot 10^9$	$2^{12} \cdot 5^6 \cdot 29$
4	$\approx -4.80 \cdot 10^{93}$	$2^{112} \cdot 3^4 \cdot 5^{71} \cdot 7^6 \cdot 29 \cdot 79$
5	$\approx 4.58 \cdot 10^{214}$	$2^{264} \cdot 3^{12} \cdot 5^{168} \cdot 19^2 \cdot 29^2 \cdot 59 \cdot 79 \cdot 769$
6	$\approx 1.14 \cdot 10^{643}$	$2^{768} \cdot 3^{12} \cdot 5^{479} \cdot 11^{16} \cdot 19^2 \cdot 29^3 \cdot 59^3 \cdot 479 \cdot 769 \cdot 1019 \cdot 2879 \cdot 6841 \cdot 36479 \cdot 42839 \cdot 144899 \cdot 443701 \cdot 3508619$

*For n=5,6, this is the discriminant of $\Psi/(4-x)^2$.

The Discriminant of $\Psi_n(x)$ for $y^2 = x^5 + 1$

Table 2.6

3. Preliminaries

For any curve C of genus $g \geq 2$, the Riemann-Roch theorem tells us that $\ell(nK_C) = s = (2n-1)(g-1)$ provided $n \geq 2$. Choose a basis $\omega_1, \ldots, \omega_s$ for $H^0(C, \mathcal{L}(nK_C))$. Let $z \in k(C)$ be a non-constant function on C, and write $\omega_i(z) = \phi_i(z)\,(dz)^n$, where the ϕ_i's are regular functions on the Zariski open set

$$U_z = \{P \in C : z \text{ is regular at } P \text{ and } \operatorname{ord}_P(z - z(P)) = 1\}.$$

The *n^{th}-order Wronskian matrix* relative to the basis $\{\omega_i\}$ and the parameter z is

$$W_n(z) = \begin{pmatrix} f_1 & f_2 & \cdots & f_s \\ \frac{df_1}{dz} & \frac{df_2}{dz} & \cdots & \frac{df_s}{dz} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{d^{s-1}f_1}{dz^{s-1}} & \frac{d^{s-1}f_2}{dz^{s-1}} & \cdots & \frac{d^{s-1}f_s}{dz^{s-1}} \end{pmatrix}.$$

Then one easily checks that a point $P \in U_z$ is an n^{th}-order Weierstrass point if and only if $(\det W_n(z))(P) = 0$. Further, the order of vanishing of $\det W_n(z)$ at P is independent of the choice of the basis $\{\omega_i\}$ and the parameter z, subject of course to P being in U_z. We define the *(n^{th}-order) weight of P* to be

$$\operatorname{wt}(P) = \operatorname{wt}_n(P) = \operatorname{ord}_P \det W_n(z) \tag{5}$$

for any z such that $z - z(P)$ is a uniformizer at P. Thus P is in $C[n]$ if $\operatorname{wt}_n(P) \geq 1$. The following facts are well-known, but for the convenience of the reader we will sketch the proof of one of them.

Proposition 3.1. *Let C be a smooth projective curve of genus $g \geq 2$ (not necessarily hyperelliptic.)*

(a) *The total weight of the n^{th}-order Weierstrass points is*

$$\sum_{P \in C[n]} \operatorname{wt}_n(P) = s^2 g.$$

(b) *For every $P \in C$,*

$$\operatorname{wt}_n(P) \leq \tfrac{1}{2}g(g+1).$$

Further, equality holds if and only if C is hyperelliptic and P is a branch point of the double cover $C \to \mathbb{P}^1$. (I.e. $P \in C[1]$.)

(c) *Suppose C is hyperelliptic. Then $C[1] \subset C[n]$ for all n, and*

$$\sum_{\substack{P \in C[n] \\ P \notin C[1]}} \operatorname{wt}(P) = 4gIJ.$$

Proof. (a) See [3] or [6].

(b) The idea of this proof is due to Segre. For a generalization, see [[1], Theorem 3.1].

Fix a point $P \in C$. Let $1 \leq b_1 < b_2 < \cdots < b_g \leq s+g$ be the uniquely determined integers b in the indicated interval such that the divisor class of $b(P) - (n-1)K_C$ is effective. Let $1 \leq a_1 < a_2 < \cdots < a_s \leq s+g$ be the remaining integers in that interval. (The a_i's are classically called the *gap values for P*.) Then an alternative description of the weight of P is

$$\operatorname{wt}(P) = \sum_{i=1}^{s} a_i - i.$$

(See [3] or [6].) Since $\{a_i\} \cup \{b_i\} = \{1, 2, \ldots, s+g\}$, this gives the formula

$$\operatorname{wt}(P) = \sum_{i=1}^{g} s + i - b_i. \tag{6}$$

Next, the definition of the b_i's implies that the divisors $b_i(P) - (n-1)K_C$ are special, so Clifford's theorem [[4], IV.5.4] says that

$$\ell(b_i(P) - (n-1)K_C) \leq \tfrac{1}{2}\deg(b_i(P) - (n-1)K_C) + 1 = \tfrac{1}{2}(b_i - s + g + 1). \tag{7}$$

Further, from the definition of the b_i's, the dimension $\ell(b_i(P) - (n-1)K_C)$ is strictly increasing with i, so

$$\ell(b_i(P) - (n-1)K_C) \geq i \qquad \text{for all } 1 \leq i \leq g. \tag{8}$$

Combining (7) and (8), we find

$$b_i \geq s - g - 1 + 2i \qquad \text{and so} \qquad s + i - b_i \leq g + 1 - i. \tag{9}$$

Now substituting (9) into (6) gives the desired bound

$$\operatorname{wt}(P) \leq \sum_{i=1}^{g} g + 1 - i = \tfrac{1}{2}g(g+1). \tag{10}$$

Finally, Clifford's theorem (7) is a strict inequality unless the divisor is 0, K_C, or a multiple of the g^1_2 on a hyperelliptic curve. If C is not hyperelliptic, then $g \geq$

3, so there are at least three b_i's. It follows that (7) cannot be an equality for every b_i, so the inequality (10) is strict. On the other hand, if C is hyperelliptic, then $K_C = (g-1)\cdot g_2^1$ is a multiple of the g_2^1 on C. So in order for (10) to be an equality, we see that $b_i(P)$ is a multiple of the g_2^1 for every b_i. In particular, the b_i's are all even, and $2(P)$ is a g_2^1.

(c) From (b) the weight of any point $P \in C[1]$ satisfies $\mathrm{wt}_n(P) = g(g+1)/2 > 0$, so $P \in C[n]$. Since the set $C[1]$ on a hyperelliptic curve consists of $2g+2$ distinct points, the desired result follows immediately from (a), (b), and a little algebra. □

We now turn to the case that C is a hyperelliptic curve given by an equation

$$C : y^2 = f(x)$$

as described in Section 1. To compute $C[n]$ we use the following collection of differential forms.

Lemma 3.2. *Let $(x)_\infty$ be the polar divisor of x, so $K_C = (g-1)(x)_\infty$ is a canonical divisor on C. Then the set of differentials*

$$\left\{ \frac{x^i\,(dx)^n}{y^n} : 0 \le i < n(g-1)+1 \right\} \cup \left\{ \frac{x^j\,(dx)^n}{y^{n-1}} : 0 \le j < (n-1)(g-1)-1 \right\}$$

is a basis for $H^0(C, \mathcal{L}(nK_C))$. (If $n = g = 2$, the second set should be omitted.)

Proof. Riemann-Roch tells us that $\ell(nK_C) = s = (2n-1)(g-1)$ (remember we assume $n \ge 2$,) so the indicated set has the correct number of elements. Further, the indicated differentials are clearly linearly independent. It remains to check that they are all holomorphic on C, an exercise which we will leave for the reader. □

The Wronskian matrix relative to the parameter x and the basis described in Lemma 3.2 is thus

$$W_n(x) = \left(\frac{d^\ell x^i y^{-n}}{dx^\ell} \;\middle|\; \frac{d^\ell x^j y^{-(n-1)}}{dx^\ell} \right), \tag{11}$$

where i, j, ℓ run over the intervals

$$\begin{aligned} 0 \le i < I &= n(g-1)+1, \\ 0 \le j < J &= (n-1)(g-1)-1, \\ 0 \le \ell < s &= (2n-1)(g-1). \end{aligned} \tag{12}$$

Since $x - x(P)$ is a uniformizer at P for every point of C except for the branch points (i.e. points in $C[1]$) and the point(s) at infinity, the zeros of $\det W_n(x)$ describe the set $C[n] \setminus C[1]$. (Except possibly for the points at infinity if $\deg(f)$ is even. We will generally ignore this issue without further comment.) We now give a polynomial in x whose zeros match those of $\det W_n(x)$.

Theorem 3.3. *Define a doubly indexed sequence of polynomials $L_\ell^m(x)$ by the recursion*

$$L_0^m = 1, \qquad \text{and} \qquad L_{\ell+1}^m = 2\cdot f \cdot \frac{dL_\ell^m}{dx} - (m+2\ell)\cdot \frac{df}{dx}\cdot L_\ell^m.$$

Let $\mathbf{W}_n(x)$ be the matrix

$$\mathbf{W}_n(x) = \left(\binom{\ell}{i} L_{\ell-i}^n \;\middle|\; \binom{\ell}{j} L_{\ell-j}^{n-1} \right),$$

where $0 \le i < I$ and $0 \le j < J$ index the columns on the left and right sides respectively, and $0 \le \ell < s$ indexes the rows. Then

$$\det W_n(x) = (2y^2)^{-IJ}\cdot y^{-nI-(n-1)J}\cdot \Big(\prod_{i=0}^{I-1} i!\Big)\cdot \Big(\prod_{j=0}^{J-1} j!\Big)\cdot \det \mathbf{W}_n(x).$$

In particular, if $P \in C$ is a point with $y(P) \ne 0, \infty$, then

$$\mathrm{wt}_n(P) = \mathrm{ord}_P \det W_n(x) = \mathrm{ord}_P \det \mathbf{W}_n(x).$$

Proof. To ease notation, we will use D to denote differentiation with respect to x. In particular, differentiating the relation $y^2 = f(x)$, we find $2y\cdot Dy = Df$. So for any polynomial $F(x) \in k[x]$ and any integer $k \ge 1$, a quick calculation shows that

$$D\frac{F(x)}{y^k} = \frac{y^2 DF - FkyDy}{y^{k+2}} = \frac{2f\cdot DF - k\cdot Df\cdot F}{2y^{k+2}}. \tag{13}$$

Using (13) and the definition of the L_ℓ^m's, an easy induction on ℓ yields

$$D^\ell\left(\frac{1}{y^m}\right) = \frac{L_\ell^m}{2^\ell y^{m+2\ell}}. \tag{14}$$

In order to simplify the Wronskian matrix, we begin by proving the following formula, valid for any rational function $z = z(x)$:

$$D^\ell(x^i z) = i!\binom{\ell}{i} D^{\ell-i} z + (-1)^{i+1} \sum_{k=0}^{i-1} (-1)^k \binom{i}{k} x^{i-k} D^\ell(x^k z). \qquad (15)$$

There are two key points to note about the sum in (15). First, it is a linear combination of $D^\ell(x^k z)$'s with k strictly less than i. Second, the coefficients $(-1)^k\binom{i}{k}x^{i-k}$ are in $\mathbb{Z}[x]$ and are independent of ℓ.

To prove (15), we start with the sum

$$\begin{aligned}
\sum_{k=0}^{i-1}(-1)^k\binom{i}{k}x^{i-k}D^\ell(x^k z) &= \sum_{k=0}^{i-1}(-1)^k\binom{i}{k}x^{i-k}\sum_{r=0}^{k}\binom{\ell}{r}D^r x^k \cdot D^{\ell-r}z \\
&= \sum_{r=0}^{i-1}\binom{\ell}{r}D^{\ell-r}z\sum_{k=r}^{i-1}(-1)^k\binom{i}{k}x^{i-k}D^r x^k \\
&= \sum_{r=0}^{i-1}\binom{\ell}{r}D^{\ell-r}z\sum_{k=r}^{i-1}(-1)^k\frac{i!}{(i-k)!(k-r)!}x^{i-r} \\
&= \sum_{r=0}^{i-1}\binom{\ell}{r}D^{\ell-r}z\sum_{k=r}^{i-1}(-1)^k\binom{i-r}{i-k}\cdot D^r x^i \\
&= \sum_{r=0}^{i-1}\binom{\ell}{r}D^{\ell-r}z\cdot(-1)^{i+1}D^r x^i.
\end{aligned}$$

Therefore the quantity on the right-hand side of (15) is

$$\begin{aligned}
&i!\binom{\ell}{i}D^{\ell-i}z + (-1)^{i+1}\sum_{k=0}^{i-1}(-1)^k\binom{i}{k}x^{i-k}D^\ell(x^k z) \\
&\qquad = \binom{\ell}{i}D^i x^i\cdot D^{\ell-i}z + \sum_{r=0}^{i-1}\binom{\ell}{r}D^r x^i\cdot D^{\ell-r}z \\
&\qquad = \sum_{r=0}^{i}\binom{\ell}{r}D^r x^i\cdot D^{\ell-r}z \\
&\qquad = D^\ell(x^i z).
\end{aligned}$$

For the last equality, note that $D^r x^i = 0$ for $r > i$, and $\binom{\ell}{r} = 0$ for $r > \ell$, so the sum in the penultimate line may be replaced by the sum from $r = 0$ to $r = \ell$. This completes the proof of the formula (15).

Recall from above that the Wronskian matrix is given by

$$W_n(x) = \left(M_I(y^{-n}) \;\middle|\; M_J(y^{-(n-1)})\right),$$

where in general we write $M_R(z)$ to denote the rectangular matrix

$$M_R(z) = \left(D^\ell(x^r z)\right)_{\substack{0\le \ell < s \\ 0 \le r < R}}.$$

We are going to perform elementary column operations on $M_R(z)$, and then do these operations on each half of $W_n(x)$, thereby simplifying $W_n(x)$ without altering its determinant.

Let us denote the r^{th} column of $M_R(z)$ by $\vec{m}_r$, $0 \le r < R$. We form a new matrix $M'_R(z)$ with columns $\vec{m}'_r$ by replacing the r^{th} column of $M_R(x)$ with

$$\vec{m}'_r = \vec{m}_r + (-1)^r \sum_{k=0}^{r-1}\binom{r}{k}x^{r-k}\vec{m}_k.$$

Note the $\vec{m}'_r$'s are computed successively starting with $r = 1$ and increasing to $r = R - 1$. Using (15), we find the ℓ^{th} entry in the r^{th} column of M'_R is

$$D^\ell(x^r z) + (-1)^r\sum_{k=0}^{r-1}(-1)^k\binom{r}{k}x^{i-k}D^\ell(x^k z) = r!\binom{\ell}{r}D^{\ell-r}z.$$

Since these column operations will not change the determinant of $W_n(x)$, we conclude that

$$\begin{aligned}
\det W_n(x) &= \det\left(M'_I(y^{-n}) \;\middle|\; M'_J(y^{-(n-1)})\right) \\
&= \det\left(i!\binom{\ell}{i}D^{\ell-i}(y^{-n}) \;\middle|\; j!\binom{\ell}{j}D^{\ell-j}(y^{-(n-1)})\right).
\end{aligned}$$

Above we derived a formula (14) for $D^\ell(y^{-m})$ in terms of L^m_ℓ. Using this formula, we find that

$$\det W_n(x) = \det\left(\frac{i!\binom{\ell}{i}L^n_{\ell-i}}{(2y^2)^\ell(2y^2)^{-i}y^n} \;\middle|\; \frac{j!\binom{\ell}{j}L^{n-1}_{\ell-j}}{(2y^2)^\ell(2y^2)^{-j}y^{n-1}}\right)$$

It remains to pull out common factors from rows and columns. We can pull out a $(2y^2)^{-\ell}$ from the ℓ^{th} row, giving $(2y^2)^{-s(s-1)/2}$. We can pull out $i!(2y^2)^i y^{-n}$ from the i^{th} column on the left-hand-side, giving $(\prod i!)\,(2y^2)^{I(I-1)/2}y^{-In}$. And we can pull out $j!(2y^2)^j y^{-(n-1)}$ from the j^{th} column on the right-hand-side,

giving $(\prod j!)\,(2y^2)^{J(J-1)/2}y^{-J(n-1)}$. Looking again at the definition of $\mathbf{W}_n(x)$, we find that

$$\det W_n(x) = (2y^2)^{(\text{some power})}y^{(\text{some power})}\prod_{i=0}^{I-1} i!\cdot\prod_{j=0}^{J-1} j!\cdot\det \mathbf{W}_n(x).$$

A little bit of algebra to compute the powers of $2y^2$ and y then completes the proof of Theorem 3.3. □

4. Degree and height estimates

In view of Theorem 3.3, the polynomials L_ℓ^m described in that theorem tell us a great deal about the Weierstrass points of C. Our first result in this section begins our description of the L_ℓ^m's.

Proposition 4.1. *For any polynomial $F(x)\in k[x]$, let $\lambda(F)$ denote the leading coefficient of F.*

(a)
$$\deg L_\ell^m = (\deg f - 1)\ell.$$

(b)
$$\lambda(L_\ell^m) = (-1)^\ell(md)(md+2)(md+4)\cdots(md+2(\ell-1)).$$

Proof. Let $d=\deg f$. We prove (a) and (b) simultaneously by induction on ℓ. For $\ell=0$ we have $L_0^m=1$, so both parts are clearly true. Assume now that they are true for ℓ.

The recursive definition of L_ℓ^m given in Theorem 3.3 says that

$$L_{\ell+1}^m = 2\cdot f\cdot DL_\ell^m - (m+2\ell)\cdot Df\cdot L_\ell^m, \tag{16}$$

where as usual we are writing D for differentiation with respect to x. By the induction hypothesis we have

$$\deg(2\cdot f\cdot DL_\ell^m) = d+\deg L_\ell^m - 1 = (d-1)(\ell+1)$$

and

$$\deg((m+2\ell)\cdot Df\cdot L_\ell^m) = d-1+\deg L_\ell^m = (d-1)(\ell+1).$$

So assuming that the leading coefficients of the two terms in (16) do not cancel, we will have proven (a). We now check that the difference of those leading coefficients is as given in (b), which will simultaneously complete the proofs of (a) and (b).

We recall that the polynomial $f(x)$ defining C is assumed to be monic. Using the induction hypothesis, we compute

$$\begin{aligned}
&\lambda(2\cdot f\cdot DL_\ell^m) - \lambda((m+2\ell)\cdot Df\cdot L_\ell^m)\\
&\quad = \{2(\deg L_\ell^m) - (m+2\ell)d\}\lambda(L_\ell^m)\\
&\quad = \{2(d-1)\ell - (m+2\ell)d\}\cdot(-1)^\ell(md)(md+2)\cdots(md+2(\ell-1))\\
&\quad = (-1)^{\ell+1}(md)(md+2)\cdots(md+2\ell).
\end{aligned}$$

This completes the proof of Proposition 4.1. □

Next we estimate the size of the L_ℓ^m's. We recall that the height of a polynomial, such as L_ℓ^m, is defined to be the height of the projective point defined by its coefficients.

Proposition 4.2.

$$h(L_\ell^m) \le \ell(h(f)+2\log\deg f) + \sum_{k=0}^{\ell-1}\log(4k+m).$$

Proof. As above, we let $d=\deg f$. Also, it is convenient to use the multiplicative height H, where $h=\log H$. Note that H has the elementary properties:

$$H(FG)\le(\deg F+1)H(F)H(G) \qquad \text{and} \qquad H(DF)\le(\deg F)H(F).$$

Using these properties, the recursive definition of L_ℓ^m, and Proposition 4.1, we compute

$$\begin{aligned}
H(L_{\ell+1}^m) &= H(2\cdot f\cdot DL_\ell^m - (m+2\ell)\cdot Df\cdot L_\ell^m)\\
&\le 2H(f\cdot DL_\ell^m) + (m+2\ell)H(Df\cdot L_\ell^m)\\
&\le 2(d+1)H(f)H(DL_\ell^m) + (m+2\ell)dH(Df)H(L_\ell^m)\\
&\le 2(d+1)H(f)(d-1)\ell H(L_\ell^m) + (m+2\ell)d^2H(f)H(L_\ell^m)\\
&\le d^2(4\ell+m)H(f)H(L_\ell^m).
\end{aligned}$$

Using this formula repeatedly gives the upper bound

$$H(L_\ell^m)\le d^{2\ell}H(f)^\ell m(m+4)\cdots(m+4(\ell-1))H(L_0^m).$$

Since $L_0^m=1$, taking logarithms gives the desired result. □

We are now ready to begin estimating the degree and the height of the polynomial $\det\mathbf{W}_n(x)$ whose roots include $C[n]$.

Theorem 4.3.

(a)

$$\deg \det \mathbf{W}_n(x) \le (\deg f - 1)IJ.$$

(Later we will be able to show that the degree is exactly equal to $2gIJ$; but for now this rough estimate will suffice.)

(b)

$$h(\det \mathbf{W}_n(x)) \le 2IJ(h(f) + \log(gs) + O(1)).$$

Proof. Looking at Theorem 3.3, we see that $\det \mathbf{W}_n$ is a sum of $s!$ terms of the form

$$w(\sigma) = \prod_{i=0}^{I-1} \binom{\ell_i}{i} L^n_{\ell_i - i} \times \prod_{j=0}^{J-1} \binom{\ell'_j}{j} L^{n-1}_{\ell'_j - j}, \tag{17}$$

where

$$\sigma = (\ell_0, \ell_1, \dots, \ell_{I-1}, \ell'_0, \dots, \ell'_{J-1})$$

is some permutation of $0, 1, \dots, s-1$. As usual, we let $d = \deg f$.

(a) Using (17) and Proposition 4.1(a), we find

$$\begin{aligned}
\deg \det \mathbf{W}_n &\le \max_\sigma \deg w(\sigma) \\
&= \max_\sigma \Big\{ \sum_{i=0}^{I-1} \deg L^n_{\ell_i - i} + \sum_{j=0}^{J-1} \deg L^{n-1}_{\ell'_j - j} \Big\} \\
&= \max_\sigma \Big\{ \sum_{i=0}^{I-1} (d-1)(\ell_i - i) + \sum_{j=0}^{J-1} (d-1)(\ell'_j - j) \Big\} \\
&= (d-1)\Big\{ \sum_{\ell=0}^{s-1} \ell - \sum_{i=0}^{I-1} i - \sum_{j=0}^{J-1} j \Big\} \\
&= (d-1)IJ.
\end{aligned}$$

(b) For any polynomials $F_1, \dots, F_N \in \bar{\mathbb{Q}}[x]$, the height satisfies the elementary estimate:

$$H(F_1 F_2 \cdots F_N) \le \prod_{i=1}^{N} \{(1 + \deg F_i) H(F_i)\}.$$

We use this and Propositions 4.1 and 4.2 to estimate the height of $w(\sigma)$.

$$\begin{aligned}
H(w(\sigma)) &\le \prod_{i=0}^{I-1} (1 + \deg L^n_{\ell_i - i}) H\left(\binom{\ell_i}{i} L^n_{\ell_i - i} \right) \\
&\quad \times \prod_{j=0}^{J-1} (1 + \deg L^{n-1}_{\ell'_j - j}) H\left(\binom{\ell'_j}{j} L^{n-1}_{\ell'_j - j} \right) \\
&\le \prod_{i=0}^{I-1} \left\{ (1 + (d-1)(\ell_i - i)) \binom{\ell_i}{i} H(f)^{\ell_i - i} \prod_{k=1}^{\ell_i - i} d^2(4k+n) \right\} \\
&\quad \times \prod_{j=0}^{J-1} \left\{ (1 + (d-1)(\ell'_j - j)) \binom{\ell'_j}{j} H(f)^{\ell'_j - j} \prod_{k=1}^{\ell'_j - j} d^2(4k+n-1) \right\}
\end{aligned}$$

In this last estimate we can pull out $H(f)$ raised to the power

$$\sum_{\ell=0}^{s-1} \ell - \sum_{i=0}^{I-1} i - \sum_{j=0}^{J-1} j = \tfrac{1}{2}s(s-1) - \tfrac{1}{2}I(I-1) - \tfrac{1}{2}J(J-1) = IJ.$$

We will also use the fact that in the innermost products we always have $k \le s$, so $4k + n \le 5s$. This gives

$$\begin{aligned}
H(w(\sigma)) &\le H(f)^{IJ} 2^{s(s-1)/2} \prod_{\ell=0}^{s-1} (1 + d\ell) \times \prod_{i=0}^{I-1} (d^2 5s)^{\ell_i - i} \\
&\quad \times \prod_{j=0}^{J-1} (d^2 5s)^{\ell'_j - j} \qquad (18) \\
&\le H(f)^{IJ} 2^{s(s-1)/2} (2ds)^s (5d^2 s)^{IJ} \\
&\le (c_1 g^2 s H(f))^{IJ} (c_2 g s^2)^s,
\end{aligned}$$

where c_1, c_2 are absolute constants, and we have used that $s^2 \gg\ll IJ$ and $d \le 2g + 2$.

Since $\det \mathbf{W}_n$ is a sum of $s!$ of the $w(\sigma)$'s, we would like to say that $H(\det \mathbf{W}_n)$ is no larger than $s!$ times the maximum of the heights of the $w(\sigma)$'s. This would be true if the $w(\sigma)$'s had integral coordinates, but unfortunately for general polynomials $F_1, \dots, F_N$ we only have $H(\sum F_i) \le N \prod H(F_i)$. So we need to be a bit more careful about the denominators of the coefficients.

For any polynomial $F(x) \in \bar{\mathbb{Q}}[x]$, we let $\delta(F)$ denote the smallest positive integer such that $\delta(F)F(x)$ has integral coefficients. (I.e. Its coefficients are in

the integral closure of $\mathbb{Z}$ in $\bar{\mathbb{Q}}$.) We claim that

$$\delta(w(\sigma)) \text{ divides } \delta(f)^{IJ}. \tag{19}$$

To see this, we note that the recursive definition of L_ℓ^m, the fact that $\delta(DF)$ always divides $\delta(F)$, and induction on ℓ immediately imply that $\delta(L_\ell^m)$ divides $\delta(f)^\ell$. Then the definition of $w(\sigma)$ as a product of L_ℓ^m's gives (19).

We resume the proof of Theorem 4.3(b). Note that if $F_1, \dots, F_N \in \bar{\mathbb{Q}}[x]$ have integral coefficients, then it is true that $H(\sum F_i) \le N \max H(F_i)$. So if we let $\delta = \delta(F_1, \dots, F_N)$ denote the least common multiple of the $\delta(F_i)$'s, then for arbitrary polynomials we have

$$H\left(\sum F_i\right) \le \delta \cdot H\left(\sum \delta F_i\right) \le \delta N \max H(F_i).$$

We apply this to $\det \mathbf{W}_n$, which is a sum of $s!$ terms of the form $w(\sigma)$. From (19) we know that every $w(\sigma)$ has denominator $\delta(w(\sigma))$ dividing $\delta(f)^{IJ}$, so we find

$$H(\det \mathbf{W}_n) \le \delta(f)^{IJ} \cdot s! \cdot \max H(w(\sigma)).$$

Note that $\delta(f) \le H(f)$ and $s! < s^s$. Since (18) provides a bound for $H(w(\sigma))$, we obtain

$$H(\det \mathbf{W}_n) \le (c_1 g^2 s H(f)^2)^{IJ} (c_2 g s^2)^s \le (c_3 g s H(f))^{2IJ}.$$

Taking logarithms completes the proof of Theorem 4.3(b). □

It is a standard matter to relate the height of a polynomial to the height of its roots. We will use the following estimate, which follows from [[5], Chapter 3, Proposition 2.1 and Theorem 2.8].

Lemma 4.4. *Let*

$$F(x) = \sum_{i=0}^{d} a_i x^i = a_d \prod_{i=1}^{d} (x - \alpha_i) \in \bar{\mathbb{Q}}[x].$$

Then

$$\sum_{i=1}^{d} h(\alpha_i) \le h(F) + \tfrac{1}{2} \log(d+1).$$

We apply this lemma to the polynomial $\det \mathbf{W}_n(x)$ to complete the proof of Theorem 1.1.

Proof (of Theorem 1.1.). From Theorem 3.3, the points in $C[n]$ that are not in $C[1]$ are precisely the points P such that $x(P)$ is a root of $\det \mathbf{W}_n(x)$; and

further, the weight of P is the multiplicity of $x(P)$ as a root. Of course, for every root of $\det \mathbf{W}_n(x)$ there are two points in $C[n]$, corresponding to the two values for y. So if we factor $\det \mathbf{W}_n(x)$ as

$$\det \mathbf{W}_n(x) = c \prod_{i=1}^{N} (x - \alpha_i),$$

then Lemma 4.4 and Theorem 4.3 allow us to estimate

$$\begin{aligned} \sum_{\substack{P \in C[n] \\ P \notin C[1]}} \mathrm{wt}(P) \cdot h(x(P)) = 2 \sum_{i=1}^{N} h(\alpha_i) &\le \\ &\le 2h(\det \mathbf{W}_n(x)) + \log(\deg \det \mathbf{W}_n(x) + 1) \\ &\le 4IJ(h(f) + \log(gs) + O(1)) + \log((d-1)IJ + 1) \\ &\le 4IJh(f) + 12IJ \log(s) + O(IJ). \end{aligned}$$

Finally, we divide by $4gIJ$ and note that $\log(s)/g = \log(n)/g + O(1)$, which completes the proof of Theorem 1.1. □

The proof of Corollary 1.2 depends on the well-known fact that there are only finitely many algebraic numbers of bounded degree and height. In principle, Theorem 1.1 says that many of the points in $C[n]$ have small height, from which we can deduce that they generate a field of large degree. To quantify these ideas, we start by estimating the number of algebraic numbers of bounded degree and height.

Lemma 4.5. *For all real numbers* $B, D \ge 1$,

$$\#\{\alpha \in \bar{\mathbb{Q}} : H(\alpha) \le B \quad \textit{and} \quad [\mathbb{Q}(\alpha) : \mathbb{Q}] \le D\} \le (3B)^{D^2}.$$

Proof. Let $S(B, D)$ be the set in question. For $\alpha \in S(B, D)$, let $\{\alpha_1, \dots, \alpha_D\}$ be the conjugates of α. (If there are fewer than D of them, fill in the set with zeros.) Consider the map

$$\begin{array}{ccc} \phi : S(B,D) & \longrightarrow & \mathbf{P}^D(\mathbb{Q}) \\ \alpha & \longmapsto & [1, \sigma_1(\alpha_1, \dots, \alpha_D), \dots, \sigma_D(\alpha_1, \dots, \alpha_D)] \end{array}$$

where σ_i is the i^{th} elementary symmetric polynomial. Notice that ϕ is at most D-to-1, since

$$\phi^{-1}([1, a_1, \dots, a_D]) \subset \{\text{roots of } T^D + a_1 T^{D-1} + \cdots + a_D = 0.\}.$$

Further, $H(\phi(\alpha)) \le 2^D H(\alpha)^D$, so

$$H(\alpha) \le B \Longrightarrow H(\phi(\alpha)) \le (2B)^D.$$

Hence

$$\begin{aligned}\#S(B,D) &\le D\cdot\#\{P\in\mathbb{P}^D(\mathbb{Q}) : H(P)\le (2B)^D\}\\ &\le D(2(2B)^D+1)^D\\ &\le (3B)^{D^2}.\end{aligned}$$

□

Theorem 1.1 gives an upper bound for a weighted sum of heights, while Proposition 3.1 gives us the sum of the weights and an upper bound for the individual weights. We want to conclude that there are many points of small height. The following elementary counting lemma is what is needed.

Lemma 4.6. *Let $w_1,\dots,w_r \ge 1$ and $h_1,\dots,h_r \ge 0$ be real numbers, and let*

$$W=\max_{1\le i\le r}\{w_i\},\qquad N=\sum_{i=1}^r w_i,\qquad A=\frac{1}{N}\sum_{i=1}^r w_ih_i.$$

Then for any $0<\varepsilon\le 1$,

$$\#\{i : h_i\le \varepsilon^{-1}A\}\ge (1-\varepsilon)\frac{N}{W}.$$

For example, taking $\varepsilon=\frac12$, there are at least $N/2W$ of the h_i's which are larger than half the weighted average of the h_i's.

Proof. First we bound how many h_i's can be large.

$$\#\{i : h_i>\varepsilon^{-1}A\}=\sum_{h_i>A/\varepsilon}1\le\sum_{h_i>A/\varepsilon}\frac{h_i}{\varepsilon^{-1}A}\le\frac{\varepsilon}{W}\cdot\frac{1}{A}\sum_{h_i>A/\varepsilon}w_ih_i\le\frac{N\varepsilon}{W}$$

On the other hand,

$$N=\sum_{i=1}^r w_i\le rW,\qquad\text{so}\quad r\ge\frac{N}{W}.$$

Hence

$$\#\{i : h_i\le\varepsilon^{-1}A\}\ge r-\frac{N\varepsilon}{W}\ge(1-\varepsilon)\frac{N}{W}.$$

□

Proof (of Corollary 1.2.). Let $C[n]\setminus C[1]=\{P_1,\dots,P_r\}$. We apply

Lemma 4.6 with $\varepsilon=\frac12$ and

$$\begin{aligned}w_i&=\mathrm{wt}(P_i),\qquad h_i=h(x(P_i)),\\ W&=\max_{1\le i\le r}\{w_i\}\le\frac{g(g+1)}{2},\\ N&=\sum_{i=1}^r w_i=4gIJ,\\ A&=\frac{1}{N}\sum_{i=1}^r w_ih_i\le\frac{h(f)+3\log n}{g}+O(1).\end{aligned}$$

(The bound for W and the value for N come from Proposition 3.1(b,c), and the bound for A is Theorem 1.1.) We obtain the estimate

$$\#\{i : h(x(P_i))\le\frac{2h(f)+6\log n}{g}+O(1)\}\ge\frac{4gIJ}{g(g+1)}\gg gn^2.$$

So there are at least $O(gn^2)$ points in $C[n]\setminus C[1]$ with height bounded by some fixed multiple of $1+(h(f)+\log n)/g$.

Let $K[n]$ be the field generated by $C[n]\setminus C[1]$, and let $d[n]$ be the degree of $K[n]$. We apply Lemma 4.5 with

$$D=d[n]\qquad\text{and}\qquad B=O\left(1+\frac{h(f)+\log n}{g}\right).$$

We have just shown that $K[n]$ has $O(gn^2)$ elements with height less than B, and of course every element of $K[n]$ has degree at most $d[n]$, so we find

$$O(gn^2)\le(3e^B)^{d[n]^2}.$$

[Note Lemma 4.5 uses the multiplicative height, whence e^B in place of B.] Taking logarithms and solving for $d[n]$ gives the desired result

$$d[n]^2\gg\frac{\log gn+O(1)}{B+O(1)}\gg\frac{\log gn+O(1)}{\dfrac{h(f)+\log n}{g}+O(1)}.$$

□

5. A Resultant Computation

In this section we ask for which primes v is it true that points in $C[n]\setminus C[1]$ and points in $C[1]$ can coincide modulo v. In essence, we want to determine which

primes divide the resultant of $f(x)$ and $\Phi_n(x)$, where $\Phi_n(x)$ is the polynomial

$$\Phi_n(x) = \prod_{\substack{P \in C[n] \\ P \notin C[1]}} (x - x(P))^{\mathrm{wt}(P)/2}$$

defined in Section 1. Of course, this reasoning is only valid for primes v such that Φ_n has v-integral coefficients.

Our plan is as follows. First we compute the related quantity

$$\text{Resultant}(f(x), \det \mathbf{W}_n(x)).$$

Second we show that $\det \mathbf{W}_n(x) = c\Phi_n(x)$ for some scalar c. Third we compute c. Since the coefficients of $\det \mathbf{W}_n$ are v-integral if the coefficients of f are, this will provide enough information to prove Theorem 1.3.

Aside from the usual binomial coefficients $\binom{a}{m}$, the following quantity will appear frequently, so we give it a special symbol:

$$\begin{bmatrix} a \\ m \end{bmatrix} = a \cdot (a+2) \cdot (a+4) \cdots (a+2m-2) = 2^m m! \binom{\frac{1}{2}a + m - 1}{m}.$$

Here $m \geq 0$ is an integer, and a is arbitrary. By convention, we set $\begin{bmatrix} a \\ 0 \end{bmatrix} = 1$.

In order to compute the resultant of $f(x)$ and $\det \mathbf{W}_n(x)$, we will use the following determinant calculation. The elementary, but heavily computational proof will be postponed until later.

Proposition 5.1. *Let $I \geq J > 0$ be integers, let $L = I + J$, and let A, B, T be arbitrary quantities. Then*

$$\det \left(\binom{\ell}{i} \begin{bmatrix} A \\ \ell - i \end{bmatrix} T^{\ell - i} \;\middle|\; \binom{\ell}{j} \begin{bmatrix} B \\ \ell - j \end{bmatrix} T^{\ell - j} \right) = T^{IJ} \prod_{j=0}^{J-1} \begin{bmatrix} B - A - 2j \\ I \end{bmatrix}.$$

Here the rows of the matrix are indexed by $0 \leq \ell < L$, the columns on the left-hand-side are indexed by $0 \leq i < I$, and the columns on the right-hand-side are indexed by $0 \leq j < J$.

Proof. See Section 6.

Proposition 5.2.

$$\text{Resultant}(f(x), \det \mathbf{W}_n(x)) = \pm \Big(\prod_{j=0}^{J-1} \begin{bmatrix} -1 - 2j \\ I \end{bmatrix} \Big)^{\deg f} \cdot \text{Disc}(f)^{IJ}.$$

Proof. Recall the polynomials L_ℓ^m of Theorem 3.3 are defined recursively by the rule

$$L_{\ell+1}^m = 2 \cdot f \cdot DL_\ell^m - (m + 2\ell) \cdot Df \cdot L_\ell^m.$$

Since $L_0^m = 1$, an easy induction on ℓ shows that

$$L_\ell^m \equiv \begin{bmatrix} m \\ \ell \end{bmatrix} (-Df)^\ell \pmod{f\mathbb{C}[x]}. \tag{20}$$

Substituting (20) into the definition of $\mathbf{W}_n(x)$, we see that

$$\det \mathbf{W}_n(x) \equiv \det \left(\binom{\ell}{i} \begin{bmatrix} n \\ \ell - i \end{bmatrix} (-Df)^{\ell - i} \;\middle|\; \binom{\ell}{j} \begin{bmatrix} n-1 \\ \ell - j \end{bmatrix} (-Df)^{\ell - j} \right) \pmod{f\mathbb{C}[x]}.$$

This last determinant is of exactly the right form to apply Proposition 5.1 with

$$A = n, \qquad B = n - 1, \qquad T = -Df.$$

We obtain the formula

$$\det \mathbf{W}_n(x) \equiv (-Df)^{IJ} \prod_{j=0}^{J-1} \begin{bmatrix} -1 - 2j \\ I \end{bmatrix} \pmod{f\mathbb{C}[x]}. \tag{21}$$

For any polynomials $P(x), Q(x), R(x) \in \mathbb{C}[x]$ and any constant c, the following elementary properties of resultants and determinants are well known [[13], Sections 5.8,5.9]:

$$\text{Resultant}(P, Q + RP) = \text{Resultant}(P, Q),$$

$$\text{Resultant}(P, cQ) = c^{\deg P} \text{Resultant}(P, Q),$$

$$\text{Resultant}(P, Q^n) = \text{Resultant}(P, Q)^n,$$

$$\text{Resultant}(P, DP) = \lambda(P) \text{Disc}(P),$$

where $\lambda(P)$ is the leading coefficient of P.

Using these properties and (21), we let c denote the product on the right-hand-side of (21) and compute

$$\begin{aligned} \text{Resultant}(f, \det \mathbf{W}_n) &= \text{Resultant}(f, c(-Df)^{IJ}) \\ &= \pm c^{\deg f} \text{Resultant}(f, Df)^{IJ} \\ &= \pm c^{\deg f} (\text{Disc } f)^{IJ}. \end{aligned}$$

(Note we have taken f to be monic.) This completes the proof of Proposition 5.2.□

We are now going to make the assumption that

$$\deg f = 2g+1 \text{ is odd.}$$

Equivalently, this means that there is only one "point at infinity" on $y^2 = f(x)$, and necessarily that point is in $C[1]$.

Notice that the polynomial Φ_n is characterized up by the fact that it is monic and satisfies

$$\operatorname{ord}_P \Phi_n = \begin{cases} \operatorname{wt}_n(P) & \text{if } P \notin C[1] \\ 0 & \text{if } P \in C[1],\ P \neq \infty \end{cases}$$

Now we observe that Theorem 3.3 implies $\operatorname{ord}_P \det \mathbf{W}_n = \operatorname{ord}_P \Phi_n$ for all $P \notin C[1]$. On the other hand, Proposition 5.2 shows that $\det \mathbf{W}_n$ and f have no common zeros, so $\operatorname{ord}_P \det \mathbf{W}_n = 0$ for $P \in C[1]$, $P \neq \infty$. This proves that $\det \mathbf{W}_n = c\Phi_n$ for some scalar c. We now begin an alternative proof of this fact which has the advantage of determining the constant c.

Proposition 5.3. *As above, we assume that* $\deg f = 2g+1$ *is odd.*

(a)

$$\deg \det \mathbf{W}_n = 2gIJ.$$

(b) *The leading coefficient of* $\det \mathbf{W}_n(x)$ *is*

$$\prod_{j=0}^{J-1} \begin{bmatrix} -1-2j \\ & I \end{bmatrix}$$

(c)

$$\det \mathbf{W}_n(x) = \left(\prod_{j=0}^{J-1} \begin{bmatrix} -1-2j \\ & I \end{bmatrix} \right) \Phi_n(x).$$

Proof. For any polynomial $F(x)$, we let $\lambda(F)$ denote the leading coefficient of F. Proposition 4.1 tells us that

$$\deg L_\ell^m = 2g\ell \quad \text{and} \quad \lambda(L_\ell^m) = (-1)^\ell \begin{bmatrix} m(2g+1) \\ \ell \end{bmatrix}, \quad \text{so}$$

$$L_\ell^m(x) = \begin{bmatrix} m(2g+1) \\ \ell \end{bmatrix} (-x^{2g})^\ell + \cdots .$$

Then the definition of $\mathbf{W}_n$ (Theorem 3.3) shows that

$$\det \mathbf{W}_n = \det \left(\binom{\ell}{i} \begin{bmatrix} n(2g+1) \\ \ell - i \end{bmatrix} (-x^{2g})^{\ell-i} \,\middle|\, \binom{\ell}{j} \begin{bmatrix} (n-1)(2g+1) \\ \ell - j \end{bmatrix} (-x^{2g})^{\ell-j} \right) + \text{lower order terms.} \tag{22}$$

We now take the determinant of the matrix in (22). If that determinant turns out to be non-zero, then its value will give us both the degree and the leading coefficient of $\det \mathbf{W}_n$.

The matrix in (22) is of exactly the right form to apply Proposition 5.1 with

$$A = n(2g+1), \qquad B = (n-1)(2g+1), \qquad T = -x^{2g}.$$

The result is

$$\det \mathbf{W}_n = (-1)^{IJ} \left(\prod_{j=0}^{J-1} \begin{bmatrix} -2g-1-2j \\ & I \end{bmatrix} \right) x^{2gIJ} + \text{lower order terms.}$$

The product is clearly non-zero, which prove (a).

To prove (b), we rearrange the product as follows:

$$\begin{aligned} (-1)^{IJ} \prod_{j=0}^{J-1} \begin{bmatrix} -2g-1-2j \\ & I \end{bmatrix} &= \prod_{j=0}^{J-1} \prod_{i=0}^{I-1} (2g+1+2j-2i) \\ &= \prod_{k=I-J-g-1}^{I-g-2} \prod_{i=0}^{I-1} (-1-2k+2(I-1-i)) \\ &\qquad \text{where } k = I-g-2-j, \\ &= \prod_{k=I-J-g-1}^{I-g-2} \begin{bmatrix} -1-2k \\ & I \end{bmatrix}. \end{aligned}$$

Since one easily checks that

$$I-J-g-1 = 0 \qquad \text{and} \qquad I-g-2 = J-1,$$

This proves (b).

(c) As observed above, Theorem 3.3 and the definition of Φ_n imply that Φ_n divides $\det \mathbf{W}_n$ in $\bar{k}[x]$. Further, Proposition 3.1(c) tells us the

$$\deg \Phi_n(x) = \frac{1}{2} \sum_{\substack{P \in C[n] \\ P \notin C[1]}} \operatorname{wt}_n(P) = 2gIJ.$$

But we have just shown in (a) that $\deg \det \mathbf{W}_n(x) = 2gIJ$, from which we conclude that $\det \mathbf{W}_n(x) = c\Phi_n(x)$ for some $c \in \bar{k}$. Finally, we note that since Φ_n is monic, the constant c equals the leading coefficient of $\det \mathbf{W}_n$, which we determined in (b). □

Remark. If we had assumed instead that $\deg f = 2g+2$, then the determinant in (22) turns out to vanish. So for even values of $\deg f$, it is somewhat more difficult to compute the degree and leading coefficient of $\det \mathbf{W}_n$.

We now have all the tools needed to prove Theorem 1.3.

Proof (of Theorem 1.3). Let

$$c = \prod_{j=0}^{J-1} \begin{bmatrix} -1-2j & \\ & I \end{bmatrix}$$

be the constant appearing in Propositions 5.2 and 5.3.

(a) From Proposition 5.2 and 5.3(c) we have

$$\pm c^{\deg f} \operatorname{Disc}(f)^{IJ} = \operatorname{Resultant}(f, \det \mathbf{W}_n) = \operatorname{Resultant}(f, c\Phi_n).$$

By a standard property of the resultant, we can pull $c^{\deg f}$ out of the right-hand-side, which gives the desired result.

(b) Let $\mathbb{Z}_f$ denote the ring generated over $\mathbb{Z}$ by the coefficients of f. By assumption, every element of $\mathbb{Z}_f$ is v-integral. From the recursive definition of the L_ℓ^m's, it is then clear that every $L_\ell^m(x) \in \mathbb{Z}_f[x]$. Then the definition of $\mathbf{W}_n$ in terms of the L_ℓ^m's shows that $\det \mathbf{W}_n(x) \in \mathbb{Z}_f[x]$. So the coefficients of $\det \mathbf{W}_n(x)$ are v-integral.

From Proposition 5.3(c) we have $\Phi_n = c^{-1} \det \mathbf{W}_n$, so it remains to show that if $p > 2n(g-1)$, then p does not divide the integer c. We expand c as a double product

$$c = \prod_{j=0}^{J-1} \begin{bmatrix} -1-2j & \\ & I \end{bmatrix} = \prod_{j=0}^{J-1} \prod_{i=0}^{I-1} (-1-2j+2i).$$

By inspection, the smallest (most negative) integer appearing in the double product is $-2J+1 = -2(n-1)(g-1)+3$, and the largest integer appearing in the double product is $2I-3 = 2n(g-1)-1$ Hence if $p > 2n(g-1)$, then $\operatorname{ord}_p(c) = 0$, which completes the proof of (b). □

Proof (of Corollary 1.4.). Taking a finite extension of k if necessary, we may assume that $C[1]$ is contained in $C(k)$. Let $P_0 \in C[1]$, and let $\{1, x\}$ be a basis for $\Gamma(2(P_0))$, so we have a model for C of the form $y^2 = f(x)$. (Note $p \ge 3$.) $f(x)$ will have degree $2g+1$. Replacing (x,y) by $(a^2x, a^{2g+1}y)$, we may assume that $f(x)$ has v-integral coordinates and $v(\operatorname{Disc} f) = 0$. This is possible since we have assumed that C has good reduction at v.

Now we can apply Theorem 1.3. Let $P \in C[n] \setminus C[1]$. Then $\Phi_n(x(P)) = 0$. From Theorem 1.3(b) the coefficients of $\Phi_n(x)$ are v-integral, so

$$\Phi_n(\widetilde{x(P)}) \equiv 0 \pmod{v}.$$

On the other hand, $\operatorname{Disc} f \not\equiv 0 \pmod{v}$, so from Theorem 1.3(a) we see that

$$\operatorname{Resultant}(\Phi_n, f) \not\equiv 0 \pmod{v}.$$

Therefore $f(\widetilde{x(P)}) \not\equiv 0 \pmod{v}$.

The roots of $f(x)$ are the x-coordinates of the points in $C[1]$ (other than P_0.) Hence for any $P_1 \in C[1]$, $P_1 \ne P_0$, we have shown that

$$\widetilde{x(P)} \not\equiv \widetilde{x(P_1)} \pmod{v},$$

so $\tilde{P} \not\equiv \tilde{P}_1 \pmod{v}$.

Finally we observe that the choice of $P_0 \in C[1]$ was arbitrary, so

$$\tilde{P} \pmod{v} \notin \widetilde{C[1]} \pmod{v} \qquad \text{for all } P \in C[n] \setminus C[1].$$

□

6. A Messy Determinant

In this section we give the proof of Proposition 5.1, which we restate for the convenience of the reader.

Proposition 5.1. *Let $I \ge J > 0$ be integers, let $L = I+J$, and let A, B, T be arbitrary quantities. Then*

$$\det\left(\binom{\ell}{i} \begin{bmatrix} A \\ \ell - i \end{bmatrix} T^{\ell-i} \;\middle|\; \binom{\ell}{j} \begin{bmatrix} B \\ \ell - j \end{bmatrix} T^{\ell-j} \right) = T^{IJ} \prod_{j=0}^{J-1} \begin{bmatrix} B-A-2j & \\ & I \end{bmatrix}.$$

Note the rows of the matrix are indexed by $0 \le \ell < L$, the columns on the left-hand-side are indexed by $0 \le i < I$, and the columns on the right-hand-side are indexed by $0 \le j < J$.

Remark. Our proof of Proposition 5.1 is a straightforward, but lengthy computation. Ira Gessel has indicated an alternative approach which involves reinterpreting the determinant as the number of non-intersecting paths in a certain diagram. For details of similar computations, see the paper of Gessel and Viennot [2].

Proof. We can pull T^ℓ out of the ℓ^{th} row of the matrix, T^{-i} out of the i^{th} column on the left-hand-side, and T^{-j} out of the j^{th} column on the right-hand-side. This contributes

$$T^{\frac{1}{2}L(L-1)-\frac{1}{2}I(I-1)-\frac{1}{2}J(J-1)} = T^{IJ}$$

to the determinant. So it suffices to prove Proposition 5.1 for $T = 1$.

We begin by proving the elementary formula

$$\sum_{i=0}^{\ell} \begin{bmatrix} B-A \\ i \end{bmatrix} \binom{\ell}{i} \begin{bmatrix} A \\ \ell-i \end{bmatrix} = \begin{bmatrix} B \\ \ell \end{bmatrix}. \tag{23}$$

To prove (23), we observe that

$$\begin{bmatrix} 2C \\ m \end{bmatrix} = x^{C+m}(-2D)^m(x^{-C}),$$

where as usual $D^m = d^m/dx^m$. Hence

$$\begin{aligned}
&\sum_{i=0}^{\ell} \begin{bmatrix} 2B-2A \\ i \end{bmatrix} \binom{\ell}{i} \begin{bmatrix} 2A \\ \ell-i \end{bmatrix} \\
&\quad = \sum_{i=0}^{\ell} \binom{\ell}{i} \left\{x^{B-A+i}(-2D)^i(x^{A-B})\right\} \cdot \left\{x^{A+\ell-i}(-2D)^{\ell-i}(x^{-A})\right\} \\
&\quad = x^{B+\ell}(-2)^\ell \sum_{i=0}^{\ell} \binom{\ell}{i} D^i(x^{A-B}) \cdot D^{\ell-i}(x^{-A}) \\
&\quad = x^{B+\ell}(-2)^\ell D^\ell(x^{-B}) \\
&\quad = \begin{bmatrix} 2B \\ \ell \end{bmatrix}.
\end{aligned}$$

Replacing A and B by $\frac{1}{2}A$ and $\frac{1}{2}B$ gives (23).

We are going to perform column operations to simplify the matrix in Proposition 5.1. Let us write

$$V = \left(\vec{a}_i \;\middle|\; \vec{b}_j \right), \qquad 0 \le i < I, 0 \le j < J,$$

for the $L \times L$ matrix whose column vectors $\vec{a}_i$ and $\vec{b}_j$ are given by

$$(\vec{a}_i)_\ell = \binom{\ell}{i} \begin{bmatrix} A \\ \ell-i \end{bmatrix}, \qquad (\vec{b}_j)_\ell = \binom{\ell}{j} \begin{bmatrix} B \\ \ell-j \end{bmatrix}.$$

Taking successively $i = 0, 1, \ldots, I-2$, we replace the i^{th} column of V by

$$\vec{a}'_i = \vec{a}_i + \sum_{k=1}^{I-1-i} \binom{i+k}{i} \begin{bmatrix} B-A \\ k \end{bmatrix} \vec{a}_{i+k}.$$

This gives a new matrix V' whose i^{th} column on the left has as its ℓ^{th} entry

$$\begin{aligned}
(\vec{a}'_i)_\ell &= \sum_{k=0}^{I-1-i} \binom{i+k}{i} \begin{bmatrix} B-A \\ k \end{bmatrix} (\vec{a}_{i+k})_\ell \\
&= \sum_{k=0}^{I-1-i} \binom{i+k}{i} \begin{bmatrix} B-A \\ k \end{bmatrix} \binom{\ell}{i+k} \begin{bmatrix} A \\ \ell-i-k \end{bmatrix} \\
&= \binom{\ell}{i} \sum_{k=0}^{I-1-i} \binom{\ell-i}{k} \begin{bmatrix} B-A \\ k \end{bmatrix} \begin{bmatrix} A \\ \ell-i-k \end{bmatrix}.
\end{aligned} \tag{24}$$

Notice that all terms in the last sum with $k > \ell - i$ are zero. So if $\ell < I$, then the sum runs from $k = 0$ to $k = \ell - i$, and we can directly apply (23) to obtain

$$(\vec{a}'_i)_\ell = \binom{\ell}{i} \begin{bmatrix} B \\ \ell-i \end{bmatrix}, \qquad \text{for } 0 \le i < I \text{ and } 0 \le \ell < I.$$

Notice that $(\vec{a}'_i)_\ell = (\vec{b}'_i)_\ell$ for $0 \le i < J$ and $0 \le \ell < I$.

When $\ell \ge I$, the sum in (24) does not have enough terms to apply (23) directly. But we can add and subtract the necessary terms, leading to the formula

$$(\vec{a}'_i)_\ell = \binom{\ell}{i} \begin{bmatrix} B \\ \ell-i \end{bmatrix} - \binom{\ell}{i} \sum_{k=0}^{\ell-I} \begin{bmatrix} B-A \\ I-i+k \end{bmatrix} \binom{\ell-i}{I-i+k} \begin{bmatrix} A \\ \ell-I-k \end{bmatrix}, \tag{25}$$

for $0 \le i < I$ and $I \le \ell < L$.

So the new matrix V' looks like

$$\begin{array}{c} \\ \uparrow \\ I \\ \downarrow \\ \uparrow \\ J \\ \downarrow \end{array} \begin{array}{c} \begin{array}{cccccccccccc} \longleftarrow & J & \longrightarrow & \longleftarrow & I-J & \longrightarrow & \longleftarrow & J & \longrightarrow \end{array} \\ \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ & 1 & & \vdots & \vdots & & \vdots & & 1 & & \vdots \\ & & \ddots & 0 & & & & & & \ddots & 0 \\ & & & 1 & 0 & & & & & & 1 \\ & & (\vec{a}'_i)_\ell & & 1 & & & & (\vec{b}'_j)_\ell & & \\ & & & & & \ddots & 0 & & & & \\ & & & & & & 1 & & & & \\ * & \cdots & & * & * & \cdots & * & * & \cdots & & * \\ \vdots & \ddots & & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & & \vdots \\ * & \cdots & & * & * & \cdots & * & * & \cdots & & * \end{pmatrix} \end{array}.$$

In particular, since $(\vec{a}'_i)_\ell = (\vec{b}'_i)_\ell$ for $0 \le i < J$ and $0 \le \ell < I$, the $J \times I$ block in the upper right-hand corner is identical to the $J \times I$ block in the upper left-hand corner. So if we subtract the first J columns from the corresponding last J columns, we get a matrix that looks like

$$\begin{array}{c} \\ \uparrow \\ I \\ \downarrow \\ \uparrow \\ J \\ \downarrow \end{array} \begin{array}{c} \begin{array}{cccccccc} \longleftarrow & I & \longrightarrow & \longleftarrow & J & \longrightarrow \end{array} \\ \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ & 1 & & \vdots & \vdots & & \vdots \\ & & \ddots & 0 & & & \\ & & & 1 & 0 & \cdots & 0 \\ * & \cdots & & * & & & \\ \vdots & \ddots & & \vdots & & (\vec{b}'_j)_\ell & \\ * & \cdots & & * & & & \end{pmatrix} \end{array}.$$

Hence $\det V'$ is equal to the determinant of the matrix $\left(\vec{b}'_j\right)$. Using (25), we find that

$$(\vec{b}'_j)_\ell = (\vec{b}_j)_\ell - (\vec{a}'_j)_\ell = \binom{\ell}{j} \sum_{k=0}^{\ell-I} \begin{bmatrix} B-A \\ I-j+k \end{bmatrix} \binom{\ell-j}{I-j+k} \begin{bmatrix} A \\ \ell-I-k \end{bmatrix}.$$

We relabel by setting $\ell = m + I$. Then $\det V'$ is equal to the determinant of the $J \times J$ matrix $X = (x_{mj})$ whose $(mj)^{th}$ entry is

$$\begin{aligned} x_{mj} &= \binom{m+I}{j} \sum_{k=0}^{m} \begin{bmatrix} B-A \\ I-j+k \end{bmatrix} \binom{I-j+m}{I-j+k} \begin{bmatrix} A \\ m-k \end{bmatrix} \\ &= \sum_{k=0}^{m} \binom{I+m-k}{j} \begin{bmatrix} B-A \\ I+m-k-j \end{bmatrix} \binom{I+m}{k} \begin{bmatrix} A \\ k \end{bmatrix}. \end{aligned}$$

We are going to simplify $\det X$ by row operations, so we let $\vec{x}_m$ be the row vector $\vec{x}_m = (x_{m0}, x_{m1}, \ldots, x_{m,J-1})$. We form a new matrix X' by setting

$$\vec{x}'_0 = \vec{x}_0, \qquad \text{and} \qquad \vec{x}'_m = \vec{x}_m - \sum_{k=1}^{m} \binom{I+m}{k} \begin{bmatrix} A \\ k \end{bmatrix} \vec{x}'_{m-k}.$$

Then we claim that the j^{th} entry of the row vector $\vec{x}'_m$ is

$$x'_{mj} = (\vec{x}'_m)_j = \binom{I+m}{j} \begin{bmatrix} B-A \\ I+m-j \end{bmatrix}. \tag{26}$$

We prove (26) by induction on m. For $m = 0$ we have

$$x_{0j} = \binom{I}{j} \begin{bmatrix} B-A \\ I-j \end{bmatrix} \binom{I}{0} \begin{bmatrix} A \\ 0 \end{bmatrix}.$$

Next, assume (26) is true for $\vec{x}'_0, \ldots, \vec{x}'_{m-1}$. Using this and the definition of x_{mj}, we find

$$\begin{aligned} x'_{mj} &= x_{mj} - \sum_{k=1}^{m} \binom{I+m}{k} \begin{bmatrix} A \\ k \end{bmatrix} x'_{m-k,j} \\ &= \sum_{k=0}^{m} \binom{I+m-k}{j} \begin{bmatrix} B-A \\ I+m-k-j \end{bmatrix} \binom{I+m}{k} \begin{bmatrix} A \\ k \end{bmatrix} \\ &\quad - \sum_{k=1}^{m} \binom{I+m}{k} \begin{bmatrix} A \\ k \end{bmatrix} \binom{I+m-k}{j} \begin{bmatrix} B-A \\ I+m-k-j \end{bmatrix} \\ &= \binom{I+m}{j} \begin{bmatrix} B-A \\ I+m-j \end{bmatrix}. \end{aligned}$$

This proves (26).

We now know that $\det V = \det V' = \det X = \det X'$, where X' is the matrix

$$X' = \left(\binom{I+m}{j} \begin{bmatrix} B-A \\ I+m-j \end{bmatrix} \right)_{0 \le m,j < J}.$$

Since

$$\begin{bmatrix} B-A \\ I+m-j \end{bmatrix} = \begin{bmatrix} B-A \\ I-j \end{bmatrix} \begin{bmatrix} B-A+2(I-j) \\ m \end{bmatrix},$$

we can factor $\begin{bmatrix} B-A \\ I-j \end{bmatrix}$ out of the j^{th} column of the matrix X'. So if we define a new matrix $Y = Y_E$ by

$$Y_E = \left(\binom{I+m}{j} \begin{bmatrix} 2E-2j \\ m \end{bmatrix} \right)_{0 \le m,j < J},$$

then

$$\det X' = \left(\prod_{j=0}^{J-1} \begin{bmatrix} B-A \\ I-j \end{bmatrix} \right) \det Y_{\frac{1}{2}(B-A)+I}. \tag{27}$$

We claim that

$$\det Y_E = \prod_{j=0}^{J-1} \begin{bmatrix} 2(E-I-j) \\ j \end{bmatrix}. \tag{28}$$

Assuming for the moment that (28) is true, we can complete the proof of Proposition 5.1. Thus using (27) and (28) (with $E = \frac{1}{2}(B-A)+I$), we compute

$$\det V = \det X' = \prod_{j=0}^{J-1} \begin{bmatrix} B-A \\ I-j \end{bmatrix} \begin{bmatrix} B-A-2j \\ j \end{bmatrix} = \prod_{j=0}^{J-1} \begin{bmatrix} B-A-2j \\ I \end{bmatrix}.$$

This completes the proof of Proposition 5.1, subject to our proving (28), which we now restate and prove.

Lemma 6.1.

$$\det \left(\binom{I+m}{j} \begin{bmatrix} 2E-2j \\ m \end{bmatrix} \right)_{0 \le m,j < J} = \prod_{j=0}^{J-1} \begin{bmatrix} 2(E-I-j) \\ j \end{bmatrix}.$$

Proof. As above, we let Y_E denote the matrix whose determinant we are trying to compute. Notice that

$$(2D)^m (x^{E+m-1}) = \begin{bmatrix} 2E \\ m \end{bmatrix} x^{E-1}.$$

We are going to introduce variables $x_0, \ldots, x_{J-1}$, so we will write D_i to denote

differentiation with respect to x_i. Then the matrix

$$\begin{aligned} Y_E(\vec{x}) &= \left(\binom{I+m}{j} \begin{bmatrix} 2E-2j \\ m \end{bmatrix} x_m^{E-j-1} \right) \\ &= \left(\binom{I+m}{j} (2D_m)^m (x_m^{E-j+m-1}) \right) \end{aligned}$$

has the property that $Y_E(\vec{1}) = Y_E$. (Here we write $\vec{1}$ for the vector $(1,1,\ldots,1)$.) We can rewrite $\det Y_E(\vec{x})$ as

$$\det Y_E(\vec{x}) =$$

$$\left(\prod_{m=0}^{J-1} (2D_m)^m \right) \left\{ \left(\prod_{m=0}^{J-1} x_m^{E-I-1} \right) \det \left(\binom{I+m}{j} x_m^{I+m-j} \right) \right\}. \tag{29}$$

We now claim that the rightmost determinant in (29) has a Taylor expansion around $\vec{1}$ that looks like

$$\begin{aligned} \det \left(\binom{I+m}{j} x_m^{I+m-j} \right) &= 1 + (x_0-1)P_0(\vec{x}) \\ &\quad + (x_1-1)^2 P_1(\vec{x}) + \cdots + (x_{J-1}-1)^J P_{J-1}(\vec{x}). \end{aligned} \tag{30}$$

Here $P_0, \ldots, P_{J-1}$ are polynomials in $\mathbb{Z}[\vec{x}]$. So when we apply the differential operator $\prod (D_m)^m$ on the right-hand-side of (29) and evaluate at $\vec{x} = \vec{1}$, the only term which is not zero arises when the entire operator is applied to $\prod x_m^{E-I-1}$ and none of it is applied to the determinant. So assuming the expansion (30), we find the desired result,

$$\begin{aligned} \det Y_E &= \det Y_E(\vec{1}) \\ &= \prod_{m=0}^{J-1} (2D_m)^m (x_m^{E-I-1}) \Big|_{\vec{x}=\vec{1}} \\ &= \prod_{m=0}^{J-1} \begin{bmatrix} 2(E-I-m) \\ m \end{bmatrix}. \end{aligned}$$

It remains to verify (30). Since we are interested in the Taylor series around $\vec{1}$, we will make a change of variables and define

$$F(\vec{z}) = F(z_0, \ldots, z_{J-1}) = \det \left(\binom{I+m}{j} (1+z_m)^{I+m-j} \right)_{0 \le m,j < J}$$

We begin by showing that $F(\vec{0}) = 1$. (Thanks to Ira Gessel for showing me this quick proof.)

Let $p_j(T)$ be any polynomial of degree j with leading coefficient a_j. Then it is easy to see using elementary column operations that

$$\det(p_j(t_m)) = \det(a_j t_m^j).$$

This is a Vandermonde determinant after we pull out the a_j's. Now let $p_j(T) = \binom{T}{j} \in \mathbb{Q}[T]$, so $p_j(T)$ has degree j and leading coefficient $1/j!$. Then we can compute

$$F(\vec{0}) = \det\left(\binom{I+m}{j}\right) = \det(p_j(I+m)) = \det\left(\frac{1}{j!}(I+m)^j\right)$$
$$= \left(\prod_{j=0}^{J-1}\frac{1}{j!}\right)\cdot \prod_{0\le m<k\le J-1}(k-m) = 1.$$

Next, for each $0 \le t < J$ we let

$$F_t = F(0,0,\dots,0,z_t,\dots,z_{J-1}).$$

Since $F_{J-1} = F(\vec{0})$, we can write $F(\vec{z})$ as a telescoping sum

$$F(\vec{z}) = 1 + \sum_{t=0}^{J-2}(F_t - F_{t+1}).$$

So the proof of (30) will be complete if we can show that

$$F_t - F_{t+1} \in z_t^{t+1}\mathbb{Z}[\vec{z}] \qquad \text{for all } 0 \le t \le J-2. \tag{31}$$

Fix some $0 \le t \le J-2$, and to ease notation write

$$G(z_t) = F_t - F_{t+1} = F(0,\dots,0,z_t,z_{t+1},\dots) - F(0,\dots,0,z_{t+1},\dots).$$

Here we think of $G(z_t)$ as a polynomial in z_t with coefficients in $\mathbb{Z}[z_{t+1},\dots,z_{J-1}]$. Clearly $G(0) = 0$, so $G(z_t)$ is in the ideal generated by z_t. We want to show it is actually in the ideal $(z_t)^{t+1}$. To verify this, we will now show that $(D^rG)(0) = 0$ for every integer $1 \le r \le t$. (Here $D = d/dz_t$.)

The point is that z_t appears only in the t^{th} row of the matrix defining F_t, and does not appear at all in F_{t+1}. So for $r \ge 1$, D^rG is the determinant of the

matrix whose m^{th} row is given by the rules

$$m^{th}\text{ row} = \begin{cases} \left(\binom{I+m}{j}\right)_{0\le j<J} & \text{if } m < t \\ \left(\binom{I+t}{j}D^r(1+z_t)^{I+t-j}\right)_{0\le j<J} & \text{if } m = t \\ \left(\binom{I+m}{j}(1+z_m)^{I+m-j}\right)_{0\le j<J} & \text{if } m > t \end{cases}$$

In particular, evaluating the derivative in the t^{th} row and substituting $z_t = 0$, we find

$$\left\{\begin{matrix} t^{th}\text{ row} \\ \text{for } z_t=0 \end{matrix}\right\} = \left(\binom{I+t}{j}r!\binom{I+t-j}{r}(1+z_t)^{I+t-j-r}\Big|_{z_t=0}\right)$$
$$= \left(r!\binom{I+t}{r}\binom{I+t-r}{j}\right)$$
$$= r!\binom{I+t}{r}\cdot\{(t-r)^{th}\text{ row}\}.$$

Thus as long as $1 \le r \le t$, we see that $(D^rG)(0)$ is the determinant of a matrix whose t^{th} row is a multiple of its $(t-r)^{th}$ row, so $(D^rG)(0) = 0$. This proves that $G(z_t)$ is in the ideal $(z_t)^{t+1}$, which completes the proof of Lemma 6.1. □

Acknowledgements

I would like to thank Masato Kuwata for his assistance in the computations of Section 2, and Bob Accola, Joe Harris, Bob Lax and David Rohrlich for their encouragement and helpful suggestions. I would also like to thank Ira Gessel for his suggestions concerning the determinant computation in Section 6. The computations described in Section 2 were done using *Mathematica*.

References

1. Accola, R., *On generalized Weierstrass points on Riemann surfaces*, "Modular Functions in Analysis and Number Theory, ed. by T.A. Metzger, Lecture Notes in Math. and Stat.," Univ. of Pittsburgh, Pittsburgh, PA, 1978.

2. Gessel, I., Viennot, G., *Binomial determinants, paths, and hook length formulae*, Advances in Math. **58** (1985), 300–320.

3. Griffiths, P., Harris, J., "Principles of Algebraic Geometry," John Wiley & Sons, New York, 1978.

4. Hartshorne, R., "Algebraic Geometry," Springer, New York, 1977.

5. Lang, S., "Fundamentals of Diophantine Geometry," Springer, New York, 1983.

6. Laskov, D., *Weierstrass points on curves*, Astérisque **87–88** (1981), 221–248.

7. Mumford, D., Fogarty, J., "Geometric Invariant Theory," 2^{nd} edition, Springer, Berlin, 1982.

8. Mumford, D., "Curves and Their Jacobians," Univ. of Mich. Press, Ann Arbor.

9. Neeman, A., *The distribution of Weierstrass points on a compact Riemann surface*, Annals of Math. **120**, 317–328.

10. _____, *Weierstrass points in characteristic p*, Invent. Math. **75**, 359–376.

11. Rohrlich, D., *Some remarks on Weierstrass points*, Number Theory Related to Fermat's Last Theorem, N. Koblitz, ed., Boston, Birkhäuser.

12. Stöhr, K.-O., Voloch, J.F., *Weierstrass points and curves over finite fields*, Proc. London Math. Soc. **52**, 1–19.

13. van der Waerden, B.L., "Algebra," Vol. 1, Fred. Ungar. Publ. Co., New York.

Joseph H. Silverman
Mathematics Department
Brown University
Providence, RI 02912 USA