

WEIERSTRASS POINTS ON CYCLIC COVERS OF THE PROJECTIVE LINE

CHRISTOPHER TOWSE

ABSTRACT. We are interested in cyclic covers of the projective line which are totally ramified at all of their branch points. We begin with curves given by an equation of the form $y^n = f(x)$, where f is a polynomial of degree d . Under a mild hypothesis, it is easy to see that all of the branch points must be Weierstrass points. Our main problem is to find the total Weierstrass weight of these points, BW . We obtain a lower bound for BW , which we show is exact if n and d are relatively prime. As a fraction of the total Weierstrass weight of all points on the curve, we get the following particularly nice asymptotic formula (as well as an interesting exact formula):

$$\lim_{d \rightarrow \infty} \frac{BW}{g^3 - g} = \frac{n + 1}{3(n - 1)^2},$$

where g is the genus of the curve.

In the case that $n = 3$ (cyclic trigonal curves), we are able to show in most cases that for sufficiently large primes p , the branch points and the non-branch Weierstrass points remain distinct modulo p .

INTRODUCTION

A Weierstrass point is a point on a curve such that there exists a non-constant function which has a low order pole at the point and no other poles. By low order we mean that the pole has order less than or equal to the genus g of the curve. The Riemann-Roch theorem shows that every point on a curve has a (non-constant) function associated to it which has a pole of order less than or equal to $g + 1$ and no other poles, so Weierstrass points are somewhat special. In fact, it is a classical result due to Hurwitz [H] that all Weierstrass points on a given curve are zeros of a certain (higher order) differential form. The Weierstrass weight of a point is the order of the zero of this form at the point. Since this differential form has degree $g^3 - g$, there are only *finitely many* Weierstrass points on any curve—exactly $g^3 - g$ of them counted with weight. Additionally, these points are all algebraic over the field of definition of the curve.

A fundamental problem is, given a curve, to find some (eventually all) of the Weierstrass points on that curve. Once found, we would like to find the weight, or an estimate for the weight, of each Weierstrass point. Lewittes [L] showed that if the curve has an automorphism with five or more fixed points, then all of the fixed points must be Weierstrass points. For example, a hyperelliptic curve comes with an involution which fixes at least five points. It is well known that these branch

Received by the editors September 27, 1994 and, in revised form, October 16, 1995.

1991 *Mathematics Subject Classification*. Primary 14H55, 11G30.

Key words and phrases. Weierstrass points, branch points.

points are the only Weierstrass points on a hyperelliptic curve. Further, a simple formula is known for the weights of these points. Motivated by the hyperelliptic case, we generalize in the following way. Any hyperelliptic curve can be written $y^2 = f(x)$; we look at curves of the form $y^n = f(x)$, where f is a polynomial which, for simplicity, we assume has distinct roots. We let d denote the degree of f . These curves have an obvious automorphism: $(x, y) \mapsto (x, \zeta_n y)$ where ζ_n is a (primitive) n th root of unity. The fixed points are just the branch points of the curve. A natural question is whether or not there are any other Weierstrass points on these curves. This is equivalent to asking for the total weight of all of the branch points, which we will call BW . Our main result is an asymptotic lower bound on $BW/(g^3 - g)$, as d goes to infinity.

In section 1, we begin by using elementary methods to find that the genus of a curve $C : y^n = f(x)$ is given by

$$2g - 2 = nd - n - d - \gcd(n, d).$$

In particular, if n and d are relatively prime, $g = \frac{1}{2}(n - 1)(d - 1)$.

More importantly, we get a basis for the space of holomorphic differentials on C . This basis is used in section 2 to construct a certain Wronskian matrix. Locally, Hurwitz's differential form is given by the determinant of this Wronskian. After some manipulation and row reduction, we are left with a new "reduced" matrix denoted $\mathbb{W}(x)$. The entries of \mathbb{W} are polynomials (as opposed to rational functions). We get a bound on the degrees of these polynomials and thus a bound on the degree of $\det \mathbb{W}$. This follows the work of Silverman, who looked at the $n = 2$ (hyperelliptic) case in [S].

In section 3 we show that this bound on $\deg \det \mathbb{W}$ in turn gives us an explicit lower bound on BW . As an approximation, we are able to show that

$$\frac{BW}{g^3 - g} \geq \frac{n + 1}{3(n - 1)^2} + O\left(\frac{1}{nd}\right).$$

Further, if we assume n is relatively prime to d , then we show that our lower bound on BW is an equality if and only if a pair of determinants which we denote c and Λ do not vanish. Both c and Λ are determinants of integer matrices, potentially interesting in their own right since the entries of the matrices are products of binomial coefficients and "generalized binomial coefficients."

Unfortunately, except in the simplest cases, the constants c and Λ appear extremely difficult to compute. On the other hand, in [A2], Accola gives a formula for the so-called G -Weierstrass weights of points fixed by an automorphism of a curve. Here, G is a non-special divisor. The ordinary Weierstrass weight of a point is its K_C -Weierstrass weight, where K_C is the canonical divisor of the curve C . Although K_C is special, we are able to make slight modifications so in the case that n and d are relatively prime, his formula yields a value for BW which agrees with our lower bound. We conclude that whenever n and d are relatively prime,

$$\lim_{d \rightarrow \infty} \frac{BW}{g^3 - g} = \frac{n + 1}{3(n - 1)^2}$$

and the determinants c and Λ are non-zero.

Another interesting aspect of the quantity BW is that the exact formula, which we calculate, not only contains rational functions (of n , d , and g), but also has a term which involves Dedekind sums. A computation of BW will depend on the

value of d modulo n (or, via Dedekind reciprocity, n modulo d). We discuss this briefly and give some examples at the end of section 3.

In section 4, we look specifically at the case $n = 3$. Here, we are able to explicitly compute c and Λ . It is easy to see that neither is zero for any d . Our formula says that

$$\frac{BW}{g^3 - g} \sim \frac{1}{3},$$

which agrees with several previous papers (by Kato [Ka] and Coppens [C] for example).

Furthermore, we get some arithmetic information regarding the curve C . Up until now, C could have been defined over the complex numbers, since all of our work has been geometric. If we assume that C is defined over some number field K , then it is interesting to talk about the reduction of C modulo some prime v of K . Using our explicit computations of c and Λ , we show that if C has good reduction modulo v , and v does not divide any $p \leq 2g - 3$, then under some conditions, the non-branch-point Weierstrass points cannot “coalesce” with any branch point modulo v .

It is our hope that this may eventually lead to some information about the field generated by the non-branch-point Weierstrass points over the base field K :

$$K \left[\bigcup x(P) \right],$$

where the union is over the set $N(C)$ of non-branch Weierstrass points. Mumford [M] has remarked on some similarities between Weierstrass points (respectively, the so-called higher-order Weierstrass points) on curves of genus $g \geq 2$ and points of order 2 (respectively, points of finite order) on elliptic curves. Since the fields generated by points of finite order have been found to be interesting, Silverman [S] and Burnol [B] began looking at the fields generated by higher-order Weierstrass points on algebraic curves. In fact, many of the ideas presented in this paper stem from the paper of Silverman, [S]. In particular, the computations for $n = 3$ in section 4 follow those done by Silverman in the $n = 2$ case.

1. PRELIMINARIES

Throughout, C will be our curve of interest, defined over either \mathbf{C} or a number field K . We let g denote the genus of C . We assume C is given by

$$C : y^n = f(x),$$

where $f(x)$ is a polynomial of degree d with non-zero discriminant and n is some integer greater than 1. For simplicity, we assume $d > n$, although the proof of Theorem 8 (the main formula) will only require $d \geq 4$.

We will need to assume that f is monic. Certainly, over \mathbf{C} , we can assume this without loss of generality. In addition, if we assume n and d are relatively prime, we can make a simple change of coordinates which does not change the field of definition of the curve. A slight variation of this change of coordinates is given explicitly in the proof of Theorem 12.

Let $x \in K(C)$ be a non-constant function on C . If $\omega_1, \dots, \omega_g$ is a basis of holomorphic differentials on C , then we can write $\omega_i = \phi_i(x)dx$, where the ϕ_i 's are

regular functions on the Zariski open set

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

It is well known that a point $P \in U_x$ is a Weierstrass point if and only if it is a zero of the determinant of the following Wronskian matrix:

$$W(x) = \begin{pmatrix} \phi_1 & \phi_2 & \cdots & \phi_g \\ \frac{d\phi_1}{dx} & \frac{d\phi_2}{dx} & \cdots & \frac{d\phi_g}{dx} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{d^{g-1}\phi_1}{dx^{g-1}} & \frac{d^{g-1}\phi_2}{dx^{g-1}} & \cdots & \frac{d^{g-1}\phi_g}{dx^{g-1}} \end{pmatrix}.$$

Further, we call the order of vanishing of the function $\det W(x)$ at $P \in U_x$ the *Weierstrass weight* of the point. This is independent of the choice of basis and of the choice of parameter x (subject to $P \in U_x$). Since every point is in some U_x , we can look at $\sum_{P \in C} \text{wt}(P)$. It is a classical theorem due to Hurwitz ([H], or see [F-K]) that this sum is always $g^3 - g$.

In this section we will write down a basis of holomorphic differentials in terms of the given coordinate function x so that we will be able to manipulate the Wronskian in the next section. Note that in [Ko], Koo gives different holomorphic differentials, along with a genus formula, for an arbitrary curve of the form $y^n = f(x)$ (i.e. f is allowed to have multiple roots).

We begin by writing a smooth model for C which we will use to obtain a formula for the genus of C in terms of n and d .

Say $f(x) = \prod_{m=1}^d (x - \alpha_m)$. If we homogenize, C is given by:

$$Z^{d-n}Y^n = \prod_{m=1}^d (X - Z\alpha_m)$$

where $x = X/Z$ and $y = Y/Z$. There is only one point with $Z = 0$ (that is, at infinity) and it is the point $P_\infty = [0, 1, 0]$. One can easily check that this point is singular when $d > n + 1$. We are interested in a desingularization of our curve at P_∞ . Let $G = \text{gcd}(n, d)$, $L = nd/G$, $H = n/G$, $T = d/G$, and a and b be positive integers such that $bd - an = G$. Consider the birational transformation:

$$x = \frac{1}{r^b t^H}, \quad y = \frac{1}{r^a t^T};$$

then

$$C : r^G = \prod_{m=1}^d (1 - \alpha_m r^b t^H).$$

There are no points when $r = 0$, and of course when neither r nor t is zero, we are on the original affine piece of the curve given in terms of x and y . When $t = 0$ we get the G points $P_{\infty,s} = (r, t) = (\zeta_G^s, 0)$ where $s = 1, \dots, G$ and ζ_G is a primitive G th root of unity. It is easy to show that C , given by these two affine pieces, is smooth.

We conclude that there are $G = \text{gcd}(n, d)$ points at infinity, and that C is given by the two equations:

$$y^n = f(x) = \prod_{m=1}^d (x - \alpha_m)$$

and

$$r^G = \prod_{m=1}^d (1 - \alpha_m r^b t^H)$$

where $x = 1/(r^b t^H), y = 1/(r^a t^T)$, and L, H, T, a , and b are given as above.

Now that we have a smooth complete model for C , we compute the canonical divisor (dx) . Since the degree of a canonical divisor is $2g - 2$, we will be able to show the following special case of a genus formula noted in [Ko] (equation (4)):

Proposition 1. *Let $C : y^n = f(x)$ with $\text{Disc } f \neq 0$ and $\deg f = d > n$. Then C has genus g given by $2g - 2 = nd - n - d - \gcd(n, d)$. Thus, in particular, if n and d are relatively prime, $g = \frac{(n-1)(d-1)}{2}$.*

Proof. We leave it to the reader to show that

$$(1.1) \quad (dx) = (n - 1) \sum_{m=1}^d P_m - (H + 1) \sum_{s=1}^G P_{\infty,s}$$

where P_m is the point with $(x, y) = (\alpha_m, 0)$ and $P_{\infty,s}$ is the point with $(r, t) = (\zeta_G^s, 0)$. Thus, $\deg(dx) = d(n - 1) - G(H + 1) = nd - d - n - G$. Of course, $\deg(dx)$ is the degree of a canonical divisor, which is equal to $2g - 2$, where $g = g(C)$ is the genus of the curve. We conclude that

$$2g - 2 = nd - n - d - \gcd(n, d). \quad \square$$

Next, we want to write down a basis for the space of holomorphic differentials on C .

We would like to simplify the situation by assuming that $n \nmid d$. We can make this assumption without loss of generality by the following change of coordinates. If $C : y^n = a_0 \prod_{m=1}^d (x - \alpha_m)$ and $d/n \in \mathbf{Z}$, then let

$$u = \frac{1}{x - \alpha_d} \quad \text{and} \quad v = yu^{d/n}.$$

Then

$$v^n = A \prod_{m=1}^{d-1} (u - \beta_m),$$

where A is the constant $a_0 \prod_{m=1}^{d-1} (\alpha_d - \alpha_m)$ and $\beta_m = \frac{1}{\alpha_m - \alpha_d}$ for $m = 1, \dots, d - 1$. Now, since n and $d - 1$ are relatively prime, we can write C in the form:

$$v^n = f(u), \quad \text{where } \deg f \equiv -1 \pmod{n},$$

without changing the field of definition.

Henceforth, we assume $n \nmid d$.

Proposition 2. *Let $C : y^n = f(x)$ be as in the previous proposition. Write $\deg f = d = kn - e$, where $0 < e < n$. Then the following is a basis for the holomorphic differentials on C (that is, a basis for $H^0(C, \Omega_C^1)$):*

$$\left\{ \frac{x^i dx}{y^j} : 1 \leq j < n, 0 \leq i < \hat{I}_j \right\}$$

where $\hat{I}_j = kj - 1 - \lfloor \frac{e}{n} j \rfloor$ and the symbol $\lfloor z \rfloor$ stands for the greatest integer less than or equal to z .

Proof. We consider differentials of the form $x^i dx/y^j$. Clearly, this set of differentials will be linearly independent, so we only need to find g such holomorphic differentials. Therefore, we look at the divisors (x) and (y) .

Let $Q_\ell = (x, y) = (0, \zeta_n^\ell \sqrt[d]{a_d})$ for $\ell = 1, \dots, n$, where $f(x) = a_0 x^d + \dots + a_d$ (so $a_d = \prod_{m=1}^d \alpha_m$). (Note that a linear change of variables guarantees that $a_d \neq 0$.) We leave it to the reader to show that

$$(y) = \sum_{m=1}^d P_m - T \sum_{s=1}^G P_{\infty,s}$$

and that

$$(x) = \sum_{\ell=1}^n Q_\ell - H \sum_{s=1}^G P_{\infty,s}.$$

Putting this together with (1.1), we see that

$$\left(\frac{x^i dx}{y^j}\right) = i \sum Q_\ell + (n - 1 - j) \sum P_m + (jT - (H + 1) - iH) \sum P_{\infty,s}$$

is holomorphic if and only if

$$\begin{aligned} i &\geq 0, \\ j &\leq n - 1, \\ \text{and } i &\leq \frac{d}{n}j - 1 - \frac{G}{n}. \end{aligned}$$

It is easy to show that this last requirement is the same as

$$i < kj - 1 - \left\lfloor \frac{e}{n}j \right\rfloor,$$

where we write $d = kn - e$ with $0 < e < n$. Note that this also implies $j > 0$.

Thus, every differential in the set:

$$\left\{ \frac{x^i dx}{y^j} : 1 \leq j < n, 0 \leq i < \hat{I}_j \right\}$$

is holomorphic, where

$$\hat{I}_j = kj - 1 - \left\lfloor \frac{e}{n}j \right\rfloor.$$

Later on we will renumber the differentials and we will let $\hat{I}_j = I_{n-1-j}$ for $j = 1, \dots, n - 1$. This gives us $\sum_{j=1}^{n-1} (kj - 1 - \lfloor \frac{e}{n}j \rfloor)$ differentials. Using the fact that

$$\sum_{j=1}^{n-1} \left\lfloor \frac{e}{n}j \right\rfloor = \frac{(n-1)(e-1)}{2} + \frac{G-1}{2}$$

(see [U-H], page 97, exercise 6) we find that we do indeed have g holomorphic differentials in our set. □

2. THE WRONSKIAN

We assume that we begin with a curve C which can be written as

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

with $\text{Disc } f \neq 0$ and $d = \deg f$. To simplify the notation, we will write $N = n - 2$ when it makes things more readable.

Let us denote by $[\phi_1 \ \phi_2 \ \cdots \ \phi_g]$ the Wronskian matrix whose first row is the one given. It is well known that

$$\det [f\phi_1 \ \cdots \ f\phi_g] = f^g \det [\phi_1 \ \cdots \ \phi_g]$$

for any function f .

We are interested in the Wronskian matrix $W(x)$ whose zeros on U_x match the Weierstrass points on U_x . Recall that our basis of differentials is

$$\left\{ \frac{x^i dx}{y^j} : 1 \leq j < n, 0 \leq i < \hat{I}_j \right\}.$$

Thus,

$$\begin{aligned} \det W(x) &= \det \begin{bmatrix} \frac{1}{y^{n-1}} & \frac{x}{y^{n-1}} & \cdots & \frac{x^{\hat{I}_{n-1}-1}}{y^{n-1}} & \cdots & \frac{x}{y} & \frac{x}{y} & \cdots & \frac{x^{\hat{I}_1-1}}{y} \end{bmatrix} \\ &= \left(\frac{1}{y^{n-1}} \right)^g \det [1 \ x \ \cdots \ x^{I_0-1} \ \cdots \ y^N \ y^N x \ \cdots \ y^N x^{I_N-1}], \end{aligned}$$

where $I_{n-1-j} = \hat{I}_j$.

Proposition 3. *There is a matrix of recursively defined polynomials, $\mathbb{W}(x)$, such that $\text{wt } P$ is the order of vanishing of $\det \mathbb{W}(x)$ at P for any point in U_x .*

In the following we will construct such a matrix.

We notice that the first few columns of $W(x)$, after the above reduction, are

$$\begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} x \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} x^2 \\ 2x \\ 2 \\ \vdots \\ 0 \end{pmatrix}, \dots, \begin{pmatrix} x^{I_0-1} \\ \vdots \\ (I_0 - 1)! \\ \vdots \\ 0 \end{pmatrix},$$

so if we let D denote differentiation with respect to x (that is, $D = \frac{d}{dx}$), then we can simplify and get

$$\det W(x) = \left(\frac{1}{y^{n-1}} \right)^g \left(\prod_{i_0=0}^{I_0-1} i_0! \right) \det (D^l(yx^{i_1}) \mid \cdots \mid D^l(y^N x^{i_N})),$$

where $l = I_0, \dots, g - 1$ denotes the row and $i_j = 0, \dots, I_j - 1$ denotes the column within the j th block, as $j = 1, \dots, N = n - 2$.

It is this “reduced” matrix that we are interested in. We will now follow Silverman [S] and rewrite the derivatives in terms of derivatives of $f(x)$. Specifically, let us recursively define a set of polynomials in the following way:

$$P_0^m = 1; \quad P_{l+1}^m = nf(x)D(P_l^m) + (m - nl)D(f(x))P_l^m.$$

Notice that P_l^m depends on n (as well as f). Throughout, we will consider n fixed.

Proposition 4. $D^l(y^m) = \frac{y^m}{(ny^n)^l} P_l^m$.

We will use induction on l . We need the following lemma.

Lemma 5. *Let $F(x)$ be any polynomial in $K[x]$. Then*

$$D(y^k F(x)) = \frac{y^{k-n}}{n} (nf(x)F'(x) + kf'(x)F(x)).$$

Proof. Since $y^n = f(x)$, we know $ny^{n-1}y' = f'(x)$. Thus,

$$\begin{aligned} D(y^k F(x)) &= y^k F'(x) + ky^{k-1}y'F(x) \\ &= \frac{1}{y^{n-k}}(y^n F'(x) + ky^{n-1}y'F(x)) \\ &= y^{k-n}(f(x)F'(x) + \frac{k}{n}f'(x)F(x)) \\ &= \frac{y^{k-n}}{n}(nf(x)F'(x) + kf'(x)F(x)), \end{aligned}$$

which is what we wanted to show. □

Proof (of Proposition 4). When $l = 0$, the claim is clear. Assume the proposition is true for l . We use the lemma to compute

$$\begin{aligned} D^{l+1}(y^m) &= D(D^l(y^m)) = D\left(\frac{y^m}{(ny^n)^l} P_l^m\right) \\ &= \frac{1}{n^l} D(y^{m-nl} P_l^m) \\ &= \frac{1}{n^l} \frac{y^{m-n(l+1)}}{n} (nf(x)D(P_l^m) + (m-nl)f'(x)P_l^m) \\ &= \frac{y^m}{(ny^n)^{l+1}} P_{l+1}^m. \end{aligned}$$

This is the desired result. □

Additionally, Silverman showed that for any rational function $z = z(x)$,

$$(2.1) \quad D^l(x^i z) = i! \binom{l}{i} D^{l-i}(z) + (-1)^{i+1} \sum_{k=0}^{i-1} (-1)^k \binom{i}{k} x^{i-k} D^l(x^k z).$$

We use this column to reduce our matrix. All of this reduction will only be done within each “block:”

$$(D^l(x^{i_j} y^j))_{\substack{0 \leq l \leq g \\ 0 \leq i_j \leq l_j}}.$$

Let \vec{m}_{i_j} be the i_j th column in the j th block. We replace \vec{m}_{i_j} by

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

Using formula (2.1) with y^j in place of z , we see that the new (i_j, l) entry of the j th block is just

$$i_j! \binom{l}{i_j} D^{l-i_j}(y^j).$$

Thus, we have

$$\det W(x) = \left(\frac{1}{y^{n-1}}\right)^g \left(\prod_{i_0=0}^{I_0-1} i_0!\right) \cdot \det \left(i_1! \binom{l}{i_1} D^{l-i_1}(y) \mid \cdots \mid i_N! \binom{l}{i_N} D^{l-i_N}(y^N) \right),$$

where as always, $l = I_0, \dots, g - 1$ and each $i_j = 0, \dots, I_j - 1$. Now we use our formula for $D^{l-i_j}(y^j)$ in terms of the P_l^m 's. We get:

$$\det W(x) = \left(\frac{1}{y^{n-1}}\right)^g \left(\prod_{i_0=0}^{I_0-1} i_0!\right) \cdot \det \left(\frac{i_1! \binom{l}{i_1} y}{(ny^n)^{l-i_1}} P_{l-i_1}^1 \mid \cdots \mid \frac{i_N! \binom{l}{i_N} y^N}{(ny^n)^{l-i_N}} P_{l-i_N}^N \right).$$

We can pull $i_j!$ out of the i_j th column in the j th block and get:

$$\det W(x) = \left(\frac{1}{y^{n-1}}\right)^g \left[\prod_{j=0}^{n-2} \left(\prod_{i_j=0}^{I_j-1} i_j! \right) \right] \cdot \det \left(\frac{\binom{l}{i_1} y}{(ny^n)^{l-i_1}} P_{l-i_1}^1 \mid \cdots \mid \frac{\binom{l}{i_N} y^N}{(ny^n)^{l-i_N}} P_{l-i_N}^N \right).$$

We can also pull out $(ny^n)^{-l}$ from the l th row and $y^j (ny^n)^{i_j}$ from the i_j th column of the j th block. We get:

$$\det W(x) = \left[\prod_{j=0}^{n-2} \left(\prod_{i_j=0}^{I_j-1} i_j! \right) \right] \left(\frac{1}{y^{n-1}}\right)^g (ny^n)^{-\kappa} \prod_{j=0}^{n-2} y^{jI_j} \cdot \det \left(\binom{l}{i_1} P_{l-i_1}^1 \mid \cdots \mid \binom{l}{i_N} P_{l-i_N}^N \right)$$

where

$$\kappa = \frac{1}{2} \left(g(g-1) - \sum_{j=0}^{n-2} I_j(I_j-1) \right).$$

This quantity κ will continue to appear throughout a large portion of the following. Notice that in the case $n = 3$, since $g = I_0 + I_1$, we get the particularly simple form: $\kappa = I_0 I_1$.

We now turn our attention to the reduced matrix of *polynomials*

$$\mathbb{W}(x) = \left(\binom{l}{i_1} P_{l-i_1}^1 \mid \cdots \mid \binom{l}{i_N} P_{l-i_N}^N \right),$$

where $I_0 \leq l \leq g - 1$ and $0 \leq i_j \leq I_j - 1$. This is the matrix of Proposition 3. If P is a point on our curve and P is not a branch point with respect to the map x (so $y(P) \neq 0, \infty$), then the Weierstrass weight of P , $\text{wt } P$, is equal to the order of vanishing of $\det \mathbb{W}$ at P . Note that the points at infinity are all branch points since $\text{gcd}(n, d) \neq n$.

Now, we would like to look at the polynomials P_l^m in order to find out more about $\det \mathbb{W}(x)$. Let $\lambda(F)$ denote the leading coefficient of a polynomial F .

Definition. Let

$$\begin{bmatrix} \alpha \\ \beta \end{bmatrix}_n = \prod_{b=0}^{\beta-1} (\alpha - nb) = \alpha(\alpha - n)(\alpha - 2n) \cdots (\alpha - (\beta - 1)n).$$

Proposition 6.

1. $\deg P_l^m \leq (d - 1)l$, and this is an equality if n and d are relatively prime.
2. If n and d are relatively prime, then

$$\lambda(P_l^m) = \begin{bmatrix} md \\ l \end{bmatrix}_n = \prod_{b=0}^{l-1} (md - nb).$$

Proof. For $l = 0$, we have $P_0^m = 1$, so $\deg P_0^m = 0$ and $\lambda(P_l^m) = 1 =$ the empty product. We continue by induction on l . We have:

$$\begin{aligned} \deg P_{l+1}^m &= \deg(nf(x)D(P_l^m) + (m - nl)D(f(x))P_l^m) \\ &\leq \max(\deg f(x)D(P_l^m), \deg D(f(x))P_l^m) \\ &= \max(\deg f + \deg P_l^m - 1, \deg f - 1 + \deg P_l^m) \\ &\leq (d - 1)(l + 1). \end{aligned}$$

If $\gcd(n, d) \neq 1$ then we are done. We assume $\gcd(n, d) = 1$. Then the last inequality is an equality by the induction hypothesis. Also, since the two terms have the same degree, we only have to show that the $(d - 1)(l + 1)$ -coefficient of P_{l+1}^m is not zero. That is, we must show that $\lambda(nfD(P_l^m))$ and $\lambda((m - nl)D(f)P_l^m)$ do not cancel. We know

$$\begin{aligned} \lambda(nfD(P_l^m)) &= n\lambda(D(P_l^m)) \quad \text{since } f \text{ is monic} \\ &= n(\deg P_l^m)\lambda(P_l^m) \\ &= n(d - 1)l \begin{bmatrix} md \\ l \end{bmatrix}_n \quad \text{by induction hypothesis.} \end{aligned}$$

Meanwhile,

$$\begin{aligned} \lambda((m - nl)D(f)P_l^m) &= (m - nl)d\lambda(P_l^m) \\ &= (m - nl)d \begin{bmatrix} md \\ l \end{bmatrix}_n. \end{aligned}$$

We need to show that the sum of these two leading coefficients is not zero. But their sum is:

$$\begin{aligned} n(d - 1)l \begin{bmatrix} md \\ l \end{bmatrix}_n + (m - nl)d \begin{bmatrix} md \\ l \end{bmatrix}_n &= (ndl - nl + md - ndl) \begin{bmatrix} md \\ l \end{bmatrix}_n \\ &= (md - nl) \prod_{b=0}^{l-1} (md - nb) \\ &= \prod_{b=0}^l (md - nb) = \begin{bmatrix} md \\ l + 1 \end{bmatrix}_n. \end{aligned}$$

This product cannot be zero, since none of the factors in the product is zero. To see this, recall that n and d are relatively prime and that m only ranges between 1

and $n - 2$. Thus, each factor is

$$md - nb \equiv md \not\equiv 0 \pmod n$$

and hence not equal to zero. We conclude that if n and d are relatively prime, $P_l^m(x) = \begin{bmatrix} md \\ l \end{bmatrix}_n x^{(d-1)l} + \text{lower order terms}$. □

Since

$$\mathbb{W}(x) = \left(\begin{pmatrix} l \\ i_1 \end{pmatrix} P_{l-i_1}^1 \mid \cdots \mid \begin{pmatrix} l \\ i_N \end{pmatrix} P_{l-i_N}^N \right),$$

we can use the proposition to find the degree and leading coefficient of $\det \mathbb{W}(x)$. If $\gcd(n, d) = 1$, the highest degree term of $\det \mathbb{W}(x)$ will be

$$\det \left(\begin{pmatrix} l \\ i_1 \end{pmatrix} \begin{bmatrix} d \\ l - i_1 \end{bmatrix}_n x^{(d-1)(l-i_1)} \mid \begin{pmatrix} l \\ i_2 \end{pmatrix} \begin{bmatrix} 2d \\ l - i_2 \end{bmatrix}_n x^{(d-1)(l-i_2)} \mid \cdots \right. \\ \left. \cdots \mid \begin{pmatrix} l \\ i_N \end{pmatrix} \begin{bmatrix} Nd \\ l - i_N \end{bmatrix}_n x^{(d-1)(l-i_N)} \right)$$

as long as this determinant does not vanish. We can pull out $x^{(d-1)l}$ from the l th row and $x^{-(d-1)i_j}$ from the i_j th column of the j th block. Thus, the highest degree term of $\det \mathbb{W}(x)$ is

$$x^{(d-1)\kappa} \det \left(\begin{pmatrix} l \\ i_1 \end{pmatrix} \begin{bmatrix} d \\ l - i_1 \end{bmatrix}_n \mid \cdots \mid \begin{pmatrix} l \\ i_N \end{pmatrix} \begin{bmatrix} Nd \\ l - i_N \end{bmatrix}_n \right)$$

provided it is not zero. Notice that when n and d are not relatively prime, the degree of $\det \mathbb{W}(x)$ is at most $(d - 1)\kappa$, since each P_l^m has $\deg P_l^m \leq (d - 1)l$.

We have shown that $\deg \det \mathbb{W}(x) \leq (d - 1)\kappa$.

For now we will not worry about the remaining determinant. However, we will optimistically call it Λ (for leading coefficient). That is, let

$$(2.2) \quad \Lambda = \det \left(\begin{pmatrix} l \\ i_1 \end{pmatrix} \begin{bmatrix} d \\ l - i_1 \end{bmatrix}_n \mid \cdots \mid \begin{pmatrix} l \\ i_N \end{pmatrix} \begin{bmatrix} Nd \\ l - i_N \end{bmatrix}_n \right).$$

In the next section we will need a similar determinant which we call c :

$$(2.3) \quad c = \det \left(\begin{pmatrix} l \\ i_1 \end{pmatrix} \begin{bmatrix} 1 \\ l - i_1 \end{bmatrix}_n \mid \cdots \mid \begin{pmatrix} l \\ i_N \end{pmatrix} \begin{bmatrix} N \\ l - i_N \end{bmatrix}_n \right).$$

3. BRANCH WEIGHT

Let BW denote the total weight of the branch points (with respect to the map x). We wish to use our matrix $\mathbb{W}(x)$ to calculate exactly, or at least bound, BW .

Theorem 7. *Given $C : y^n = f(x)$ with $\text{Disc } f \neq 0$ and $d = \deg f > n$, then the total Weierstrass weight of the branch points is*

$$BW \geq g^3 - g - n(d - 1)\kappa.$$

If we assume that n and d are relatively prime, then this is an equality if and only if both Λ and c , defined in equations (2.2) and (2.3), respectively, are non-zero.

Proof. Recall that in our transformation from the matrix $W(x)$ to the matrix $\mathbb{W}(x)$ we pulled out a number of powers of y (along with some constants). We know that at all non-branch-points P on our curve we have

$$\text{ord}_P \det W(x) = \text{ord}_P \det \mathbb{W}(x) = \text{wt}(P).$$

That is, the order of vanishing of $\det \mathbb{W}(x)$ is the weight of P as long as $\text{ord}_P y = 0$ (i.e. $y(P) \neq 0, \infty$). Thus, the total weight of the non-branch-point Weierstrass points is

$$g^3 - g - BW \leq n \deg \det \mathbb{W}(x) \leq n(d-1)\kappa.$$

The second inequality is an equality if $\gcd(n, d) = 1$ and if $\Lambda \neq 0$. We would like to show that $\det \mathbb{W}(x)$ does not vanish at any of the branch points. Since $\det \mathbb{W}(x)$ is a polynomial, we do not have to worry about any points at infinity. We look at $\text{Res}(f, \det \mathbb{W})$, the resultant of f and $\det \mathbb{W}$. If we can show that

$$\text{Res}(f, \det \mathbb{W}) \neq 0,$$

then the total weight of the non-branch-points will be

$$n \deg \det \mathbb{W}.$$

Note that for each root x_0 of $\det \mathbb{W}$ which is not the x -coordinate of a branch point there are n Weierstrass points with this x -coordinate, $(x_0, \zeta_n^m \sqrt[n]{f(x_0)})$, where $m = 1, \dots, n$.

If $\text{Res}(f, \det \mathbb{W})$ is zero, this just says that some of the roots of $\det \mathbb{W}$ correspond to branch points, so the total weight of the non-branch-points is strictly less than $n(d-1)\kappa$. This weight is exactly $n(d-1)\kappa$ if and only if both Λ and $\text{Res}(f, \det \mathbb{W})$ are non-zero.

Let us look at $\text{Res}(f, \det \mathbb{W})$. Since $\text{Res}(P, Q + RP) = \text{Res}(P, Q)$, we look at $P_l^m \pmod{f\mathbf{C}[x]}$. We know $P_{l+1}^m = n f D(P_l^m) + (m - nl) D(f) P_l^m$, so a simple induction shows that

$$P_l^m(x) \equiv \begin{bmatrix} m \\ l \end{bmatrix}_n (Df)^l \pmod{f\mathbf{C}[x]}.$$

Thus

$$\begin{aligned} \det \mathbb{W} &= \det \left(\begin{pmatrix} l \\ i_1 \end{pmatrix} P_{l-i_1}^1 \mid \cdots \mid \begin{pmatrix} l \\ i_N \end{pmatrix} P_{l-i_N}^N \right) \\ &\equiv \det \left(\begin{pmatrix} l \\ i_1 \end{pmatrix} \begin{bmatrix} 1 \\ l-i_1 \end{bmatrix}_n (Df)^{l-i_1} \mid \cdots \mid \begin{pmatrix} l \\ i_N \end{pmatrix} \begin{bmatrix} N \\ l-i_N \end{bmatrix}_n (Df)^{l-i_N} \right) \\ &\hspace{15em} \text{modulo } f\mathbf{C}[x] \\ &= (Df)^\kappa \det \left(\begin{pmatrix} l \\ i_1 \end{pmatrix} \begin{bmatrix} 1 \\ l-i_1 \end{bmatrix}_n \mid \cdots \mid \begin{pmatrix} l \\ i_N \end{pmatrix} \begin{bmatrix} N \\ l-i_N \end{bmatrix}_n \right). \end{aligned}$$

But this last determinant is just our definition of c (equation (2.3)).

Thus,

$$\begin{aligned} \text{Res}(f, \det \mathbb{W}) &= \text{Res}(f, (Df)^\kappa c) \\ &= c^d \text{Res}(f, (Df)^\kappa) \\ &= c^d (\text{Res}(f, Df))^\kappa \\ &= c^d (\text{Disc } f)^\kappa. \end{aligned}$$

So this resultant is zero if and only if c is zero. \square

Consider any curve with an automorphism of degree n . Lewittes showed in [L] that as long as the automorphism has at least five fixed points, all fixed points are Weierstrass points. In [A2] (Theorem 6.19), Accola fixes n . Then for each fixed point P he defines two sets of integers: $t_j = t_{j,P}$ and $D_j = D_{j,P}$, where $j = 0, \dots, n-1$. Further he defines a third set of integers d_j (again $j = 0, \dots, n-1$) which depend on the curve and its automorphism. Given these integers, Accola shows that for any fixed point P ,

$$(3.1) \quad \text{wt}(P) \geq \frac{1}{2} \sum_{j=0}^{n-1} (2t_j + n(D_j + d_j))(d_j - D_j)$$

and equality holds as long as the image of P on the quotient curve is not a Weierstrass point of that curve. In our case, the quotient curve is \mathbf{P}^1 , the projective line. Since there are no Weierstrass points on \mathbf{P}^1 , the images of our fixed points cannot be Weierstrass points on the quotient curve. Thus, we can use Accola’s formula to find BW and compare it with the bound we have from Theorem 7. Note that although Accola assumes n is prime, it is clear that since we have total ramification at every branch point and $\gcd(n, d) = 1$, this assumption is not needed. Further, note that Accola deals with so-called G -Weierstrass points, where he assumes G is a non-special divisor. Ordinary Weierstrass points are K_C -Weierstrass points where K_C is a canonical divisor, which is actually special. However, the assumption that G is non-special is only used when Accola writes a formula for the d_j ’s which we have avoided.

Theorem 8. *If n and d are relatively prime, then the branch weight BW equals*

$$\frac{n+1}{3(n-1)^2}g^3 - \frac{2(n+1)}{3(n-1)}g^2 - \frac{n+4}{3}g + d(d-1) \sum_{j=0}^{n-1} j \left\{ \frac{e}{n} j \right\}.$$

Here $\{x\}$ is the fractional part of x (i.e., $\{x\} := x - [x]$). This formula may also be written:

$$BW = \frac{n+1}{3(n-1)^2}g^3 - \frac{g(d(n+1)+3)}{3} + d(d-1) \sum_{j=0}^{n-1} j \left\{ \frac{e}{n} j \right\}.$$

Proof. First, let us look at the t_j ’s. These are integers between 0 and $n-1$, and in the cases that P is not the point at infinity, Accola defines t_j to be the smallest non-negative residue of $j(\text{ord}_P y) + \text{ord}_P(dx)$ modulo n . Since $\text{ord}_P y = 1$ and $\text{ord}_P(dx) = n-1$, we get

$$t_j = \begin{cases} n-1, & \text{when } j = 0, \\ j-1, & \text{otherwise.} \end{cases}$$

Next, D_j is defined to be the number of integers in the set $\{0, 1, \dots, g-1\}$ which are congruent to t_j modulo n . Recall that $g = (n-1)(d-1)/2$. If we write $d = kn - e$, with $0 < e < n$, then we have the following three cases.

Case 1. e is odd.

We write

$$g = n \left(\frac{n-1}{2}k - \frac{e+1}{2} \right) + \frac{e+1}{2}.$$

Then

$$D_j = \begin{cases} \frac{n-1}{2}k - \frac{e+1}{2}, & \text{for } \frac{e+1}{2} \leq t_j \leq n-1, \\ \frac{n-1}{2}k - \frac{e-1}{2}, & \text{for } 0 \leq t_j \leq \frac{e-1}{2}. \end{cases}$$

Case 2. e is even, but $e \neq n-1$.

We write

$$g = n \left(\frac{n-1}{2}k - \frac{e+2}{2} \right) + \frac{n+e+1}{2}.$$

Then

$$D_j = \begin{cases} \frac{n-1}{2}k - \frac{e+2}{2}, & \text{for } \frac{n+e+1}{2} \leq t_j \leq n-1, \\ \frac{n-1}{2}k - \frac{e}{2}, & \text{for } 0 \leq t_j \leq \frac{n+e-1}{2}. \end{cases}$$

Case 3. e is even and $e = n-1$.

We write

$$g = n \left(\frac{n-1}{2}k - \frac{e}{2} \right).$$

Then

$$D_j = \frac{n-1}{2}k - \frac{e}{2} \quad \text{for all } j.$$

We only need to figure out the d_j 's. Accola decomposes the space of holomorphic differentials into subspaces depending on the effect of the automorphism on the differentials. The d_j 's are defined as the dimensions of these subspaces, and it is easy to check that in our language, we have

$$d_j = \begin{cases} 0, & j = 0, \\ I_{j-1}, & j \neq 0. \end{cases}$$

Now, we can compute the weights of the branch points. We rewrite equation (3.1) as follows:

$$\text{wt}(P) = \frac{1}{2} \sum_{j=0}^{n-1} (2t_{j,P}d_j + nd_j^2) - \frac{1}{2} \sum_{j=0}^{n-1} (2t_{j,P}D_{j,P} + nD_{j,P}^2).$$

Computing the first sum is straightforward once we see that I_j , which is defined to be

$$I_j = \hat{I}_{n-1-j} = k(n-1-j) - 1 - \left\lfloor \frac{e}{n}(n-1-j) \right\rfloor,$$

is the same as $d - k(j+1) + \left\lfloor \frac{e}{n}(j+1) \right\rfloor$. This is easy to show using the facts that $\gcd(n, e) = \gcd(n, d) = 1$ and j only ranges from 0 to $n-2$.

To compute the second sum, we permute the order of summation:

$$\sum_{j=0}^{n-1} (2t_{j,P}D_j + nD_j^2) = \sum_{t_j=0}^{n-1} (2t_{j,P}D_j + nD_j^2).$$

Now, for each of our three cases: e odd; e even, but not $n - 1$; $e = n - 1$, this sum is easy to compute, whether P is the point at infinity or one of the other branch points. We get, for any $e = 1, \dots, n - 1$ with $\gcd(n, e) = 1$,

$$\text{wt}(P) = \frac{g}{12}(d - 7)(n + 1) + (d - 1) \sum_{j=0}^{n-1} j \left\{ \frac{e}{n} j \right\} \quad (P \neq P_\infty)$$

and

$$\text{wt}(P_\infty) = \frac{g}{12}(d(n + 1) + n - 11).$$

Now, the true total branch weight BW is simply $d \text{wt}(P) + \text{wt}(P_\infty)$, and this is equal to

$$BW = \frac{n + 1}{3(n - 1)^2} g^3 - \frac{2(n + 1)}{3(n - 1)} g^2 - \frac{n + 4}{3} g + d(d - 1) \sum_{j=0}^{n-1} j \left\{ \frac{e}{n} j \right\}.$$

This proves Theorem 8. □

In fact, we will see that this is the same formula as in Theorem 7, once we examine κ .

Corollary 9. *If n and d are relatively prime, then the branch weight BW equals $g^3 - g - n(d - 1)\kappa$. Furthermore, given any two relatively prime positive integers $n < d$, write $d = kn - e$ with $1 \leq e \leq n - 1$ and let*

$$I_j = d - k(j + 1) + \left\lfloor \frac{e}{n}(j + 1) \right\rfloor$$

for $j = 0, \dots, n - 2$. Also, let $N = n - 2$ and $g = \sum_j I_j = (n - 1)(d - 1)/2$. Then the $(g - I_0) \times (g - I_0)$ determinants

$$\Lambda = \det \left(\begin{pmatrix} l \\ i_1 \end{pmatrix} \begin{bmatrix} d \\ l - i_1 \end{bmatrix}_n \mid \cdots \mid \begin{pmatrix} l \\ i_N \end{pmatrix} \begin{bmatrix} Nd \\ l - i_N \end{bmatrix}_n \right)$$

and

$$c = \det \left(\begin{pmatrix} l \\ i_1 \end{pmatrix} \begin{bmatrix} 1 \\ l - i_1 \end{bmatrix}_n \mid \cdots \mid \begin{pmatrix} l \\ i_N \end{pmatrix} \begin{bmatrix} N \\ l - i_N \end{bmatrix}_n \right)$$

are non-zero, where $l = I_0, \dots, g - 1$ and $i_j = 0, \dots, I_j - 1$.

Proof. We only need to show that this formula is the same as that in Theorem 8. The difficult part of the expression is the κ . Recall the definition of κ .

$$\begin{aligned} \kappa &= \sum_{l=I_0}^{g-1} l - \sum_{j=1}^{n-2} \left(\sum_{i_j=0}^{I_j-1} i_j \right) \\ &= \sum_{l=1}^{g-1} l - \sum_{j=0}^{n-2} \frac{I_j(I_j - 1)}{2} \\ &= \frac{1}{2} \left(g(g - 1) - \sum_{j=0}^{n-2} I_j(I_j - 1) \right). \end{aligned}$$

Also, recall that $I_j = \hat{I}_{n-1-j} = kj - 1 - \lfloor \frac{e}{n}j \rfloor$, where $d = kn - e$. So

$$\begin{aligned}
 (3.2) \quad \sum_{j=0}^{n-2} I_j(I_j - 1) &= \sum_{j=1}^{n-1} \hat{I}_j(\hat{I}_j - 1) \\
 &= \sum_{j=1}^{n-1} \left(kj - 1 - \lfloor \frac{e}{n}j \rfloor \right) \left(kj - 2 - \lfloor \frac{e}{n}j \rfloor \right).
 \end{aligned}$$

Now, we write $\lfloor \frac{e}{n}j \rfloor = \frac{e}{n}j - \{ \frac{e}{n}j \}$. Then

$$\begin{aligned}
 (3.2) &= \sum_{j=1}^{n-1} \left(\frac{kn - e}{n}j - 1 + \left\{ \frac{e}{n}j \right\} \right) \left(\frac{kn - e}{n}j - 2 + \left\{ \frac{e}{n}j \right\} \right) \\
 &= \frac{d^2(n-1)(2n-1)}{6n} - \frac{3d(n-1)}{2} + 2(n-1) - 3 \sum \left\{ \frac{e}{n}j \right\} \\
 &\quad + \sum \left\{ \frac{e}{n}j \right\}^2 + \frac{2d}{n} \sum j \left\{ \frac{e}{n}j \right\}.
 \end{aligned}$$

As j ranges between 1 and $n - 1$, it is easy to see that $\{ \frac{e}{n}j \}$ ranges between $1/n$ and $(n - 1)/n$, since $\gcd(n, e) = \gcd(n, d)$ is assumed to be 1. Thus,

$$\sum_{j=1}^{n-1} \left\{ \frac{e}{n}j \right\} = \sum \left(\frac{j}{n} \right) = \frac{n-1}{2}$$

and

$$\sum_{j=1}^{n-1} \left\{ \frac{e}{n}j \right\}^2 = \sum \left(\frac{j}{n} \right)^2 = \frac{(n-1)(2n-1)}{6n}.$$

Thus,

$$\kappa = \frac{1}{2} \left(g^2 - g - \frac{(d^2 + 1)(n-1)(2n-1)}{6n} + \frac{(3d-1)(n-1)}{2} - \frac{2d}{n} \sum j \left\{ \frac{e}{n}j \right\} \right).$$

The last term is a Dedekind sum and we won't get a rational expression out of it. Plugging the formula for κ into the equation for BW , we get (after some manipulation):

$$g^3 - g - n(d-1)\kappa = \frac{(n+1)}{3(n-1)^2}g^3 - \frac{2(n+1)}{3(n-1)}g^2 - \frac{n+4}{3}g + d(d-1) \sum_{j=1}^{n-1} j \left\{ \frac{e}{n}j \right\}.$$

The rest of the corollary follows directly from Theorem 7. □

The Dedekind sum is certainly no bigger than $\sum j = n(n-1)/2$, and g is on the order of $nd/2$, so the last terms are on the order of g^2 . Therefore, the first term is the dominant term. We can restate this as follows:

Corollary 10. *With the definitions as above, and n fixed,*

$$\liminf_{d \rightarrow \infty} \frac{BW}{g^3 - g} \geq \frac{n+1}{3(n-1)^2}.$$

Further, if we restrict to values of d that are relatively prime to n , we obtain

$$\lim_{\substack{d \rightarrow \infty \\ (d,n)=1}} \frac{BW}{g^3 - g} = \frac{n + 1}{3(n - 1)^2}.$$

Here we have written $g^3 - g$ for emphasis, since this is the total weight of all points on the curve, but as d goes to infinity, the $-g$ part has no effect.

We can look at some special cases of BW , such as $d \equiv \pm 1, \pm 2 \pmod n$. Of course, this will amount to looking at special cases of the Dedekind sum. It will be slightly more efficient, as well as possibly more enlightening, to use Dedekind reciprocity (see [R-G]). Let

$$S(e, n) = \sum_{j=1}^{n-1} j \left\{ \frac{e}{n} j \right\}.$$

Then Dedekind reciprocity states that

$$S(e, n) = \frac{(3e + 1)n^2 + (3e^2 - 9e)n + e^2 + 1}{12e} - \frac{n}{e} S(n, e).$$

Clearly, $S(n, e)$ depends only on the value of n modulo e .

Example. If $n \equiv 1 \pmod e$, then $\left\{ \frac{n}{e} i \right\} = \frac{i}{e}$, so

$$S(n, e) = \frac{(e - 1)(2e - 1)}{6}.$$

Thus,

$$S(e, n) = \frac{(3e + 1)n^2 - (e^2 + 3e + 2)n + e^2 + 1}{12e}$$

and

$$BW = \frac{n + 1}{3(n - 1)^2} g^3 + \frac{(e + 1)(n - (e + 1))}{3(n - 1)e} g^2 + \frac{en + n - e^2 - 8e - 1}{6e} g.$$

In particular, if $e = 1$ we get

$$BW = \frac{n + 1}{3(n - 1)^2} g^3 + \frac{2(n - 2)}{3(n - 1)} g^2 + \frac{n - 5}{3} g$$

and if $e = n - 1$ we get

$$BW = \frac{n + 1}{3(n - 1)^2} g^3 - g.$$

Further, if we assume n is odd and take $e = 2$, we have

$$BW = \frac{n + 1}{3(n - 1)^2} g^3 + \frac{(n - 3)}{2(n - 1)} g^2 + \frac{n - 7}{4} g.$$

Additionally, it is easy to show that if $e + b = n$ then

$$S(e, n) = \frac{n(n - 1)}{2} - S(b, n)$$

and we can use this fact to find BW when $e = n - 2$, for instance. Along with Dedekind reciprocity and some easy examples like the ones above, we can use this formula to find BW in a variety of different situations. However, the resulting formulae are not particularly illuminating, so we won't pursue the matter.

We can also take a quick look at some curves of this form and see what the remaining Weierstrass points look like. The simplest case is the Catalan curve:

$$y^n = x^d + 1.$$

In order to use our formula to get the weights of the $y = 0$ branch points, as well as the $x = 0$ branch points, we would need n and d relatively prime, and for both to be greater than 3. However, in any case, the matrix that we constructed in Section 2 allows us to look at the polynomial (in x) whose roots are (the x -coordinates of) the non-branch Weierstrass points. For small n and d , it is reasonable to calculate this. The author has done this for all cases with $g \leq 15$. Since there is an automorphism of the curve $x \mapsto \zeta_d x$, the polynomial is in fact a polynomial in x^d .

Interestingly, of these examples, in every case but one, this polynomial is irreducible (over \mathbf{Q}). In the case $n = 4, d = 5$, it factors into two irreducible pieces, one of which is linear in x^5 . This leads to the (rational) integer Weierstrass point (54000, 810000) on the curve

$$y^4 = x^5 - 2(30)^{15}.$$

Another similar example is the curve

$$y^3 = x^4 - 729,$$

which, aside from the points with $y = 0$ (weight 1 each), the $x = 0$ points (weight 2 each), and the point at infinity (weight 2), has 12 more weight 1 Weierstrass points consisting of (9, 18) and its conjugates.

4. CYCLIC TRIGONAL CURVES

It should be noted that many of the computations we have done for general n can be done directly (and more completely) for the case $n = 3$. In this section we restrict our attention to these curves.

Since \mathbb{W} is particularly simple (it only has one “block”), we are able to find simple formulae for Λ and c . We will use these to begin a look at what happens to the branch and non-branch points of our curves upon reduction modulo a prime.

Proposition 11.

$$\det \left(\binom{m+B}{i} \begin{bmatrix} M \\ m+B-i \end{bmatrix}_n \right)_{0 \leq i, m < I} = \prod_{i=0}^{I-1} \begin{bmatrix} M+ni \\ B \end{bmatrix}_n.$$

Proof. We begin by noting the easy identity

$$\begin{bmatrix} \alpha \\ \beta \end{bmatrix}_n \begin{bmatrix} \alpha - n\beta \\ \gamma \end{bmatrix}_n = \begin{bmatrix} \alpha \\ \gamma + \beta \end{bmatrix}_n.$$

So we write

$$\begin{aligned} & \det \left(\binom{m+B}{i} \begin{bmatrix} M \\ m+B-i \end{bmatrix}_n \right) \\ (4.1) \quad &= \det \left(\binom{m+B}{i} \begin{bmatrix} M \\ B-i \end{bmatrix}_n \begin{bmatrix} M+n(i-B) \\ m \end{bmatrix}_n \right) \\ &= \prod_{i=0}^{I-1} \begin{bmatrix} M \\ B-i \end{bmatrix}_n \det \left(\binom{m+B}{i} \begin{bmatrix} M+n(i-B) \\ m \end{bmatrix}_n \right). \end{aligned}$$

Now we need the following fact:

$$(4.2) \quad (-n)^m D^m(x^{\epsilon-1}) = \begin{bmatrix} n(m-\epsilon) \\ m \end{bmatrix}_n x^{\epsilon-m-1}$$

where, as usual, D stands for differentiation with respect to x . This fact is a variation on a standard identity and is left to the reader to verify.

We need to show that

$$\det \left(\begin{pmatrix} m+B \\ i \end{pmatrix} \begin{bmatrix} M+n(i-B) \\ m \end{bmatrix}_n \right) = \prod_{i=0}^{I-1} \begin{bmatrix} M+ni \\ i \end{bmatrix}_n,$$

so that (4.1) becomes

$$\prod_{i=0}^{I-1} \begin{bmatrix} M \\ B-i \end{bmatrix}_n \prod_{i=0}^{I-1} \begin{bmatrix} M+ni \\ i \end{bmatrix}_n = \prod_{i=0}^{I-1} \begin{bmatrix} M+ni \\ B \end{bmatrix}_n,$$

as required.

Now, let $E = B - \frac{1}{n}M$, so $M = n(B - E)$. We need to show that

$$\det \left(\begin{pmatrix} m+B \\ i \end{pmatrix} \begin{bmatrix} n(i-E) \\ m \end{bmatrix}_n \right) = \prod_{i=0}^{I-1} \begin{bmatrix} n(B+i-E) \\ i \end{bmatrix}_n.$$

Consider the following matrix:

$$Y(\vec{x}) = \left(\begin{pmatrix} m+B \\ i \end{pmatrix} \begin{bmatrix} n(i-E) \\ m \end{bmatrix}_n x_m^{E-i-1} \right),$$

where \vec{x} is the vector $(x_m) = (x_0, \dots, x_{I-1})$ and D_m denotes differentiation with respect to x_m . Clearly, $Y(\vec{1})$ is just the matrix we are interested in. Applying (4.2) with $\epsilon = E + m - i$, we see that

$$Y(\vec{x}) = \left(\begin{pmatrix} m+B \\ i \end{pmatrix} (-n)^m D_m^m(x_m^{E-i+m-1}) \right).$$

Since x_m only shows up in the m th row, we get the following:

$$\begin{aligned} \det Y(\vec{x}) &= \prod_{m=0}^{I-1} (-nD_m)^m \left\{ \det \left(\begin{pmatrix} m+B \\ i \end{pmatrix} x_m^{E-i+m-1} \right) \right\} \\ &= \prod_{m=0}^{I-1} (-nD_m)^m \left\{ \prod_{m=0}^{I-1} x_m^{E-B-1} \det \left(\begin{pmatrix} m+B \\ i \end{pmatrix} x_m^{B-i+m} \right) \right\}. \end{aligned}$$

We have a differential operator acting on a product. The second factor in that product is of exactly the form that Silverman looked at. He showed that

$$\begin{aligned} \det \left(\begin{pmatrix} m+B \\ i \end{pmatrix} x_m^{B-i+m} \right) &= 1 + (x_0 - 1)L_0(\vec{x}) \\ &\quad + (x_1 - 1)^2 L_1(\vec{x}) + \dots + (x_{B-1} - 1)^B L_{B-1}(\vec{x}), \end{aligned}$$

where L_0, L_1, \dots, L_{B-1} are polynomials in $\mathbf{Z}[\vec{x}]$ (see [S], formula (30)). Thus, when we apply the operator $\prod D_m^m$ and evaluate at $\vec{x} = \vec{1}$ the only non-zero term arises

when the entire operator is applied to the $\prod x_m^{E-B-1}$ and none of it is applied to the determinant. Then

$$\begin{aligned} \det Y(\vec{x})|_{\vec{1}} &= \prod_{m=0}^{I-1} (-n)^m D_m^m \left\{ \prod_{m=0}^{I-1} x_m^{E-B-1} \right\} \Big|_{\vec{x}=\vec{1}} \cdot 1 \\ &= \prod_{m=0}^{I-1} \left[\begin{matrix} n(m+B-E) \\ m \end{matrix} \right]_n x^{E-m-B-1} \Big|_{\vec{x}=\vec{1}} \quad \text{by (4.2) with } \epsilon = E - B \\ &= \prod_{m=0}^{I-1} \left[\begin{matrix} n(m+B-E) \\ m \end{matrix} \right]_n. \end{aligned}$$

Of course this is the same as $\prod_i \left[\begin{matrix} n(B+i-E) \\ i \end{matrix} \right]_n$, which is what we wanted. □

So in particular, we can let $M = d$, $B = I_0$, $I = I_1$, and $n = 3$, and get that

$$\begin{aligned} \Lambda &= \lambda(\det \mathbb{W}(x)) = \prod_{i_1=0}^{I_1-1} \left[\begin{matrix} d+3i_1 \\ I_0 \end{matrix} \right]_3 \\ &= \prod_{i_1=0}^{I_1-1} \prod_{i_0=0}^{I_0-1} (d+3(i_1-i_0)), \end{aligned}$$

and if we change M to 1, we get that

$$\begin{aligned} c &= \det \left(\binom{m+I_0}{i_1} \left[\begin{matrix} 1 \\ m+I_0-i_1 \end{matrix} \right]_3 \right)_{0 \leq i, m < I_1} \\ &= \prod_{i_1=0}^{I_1-1} \left[\begin{matrix} 1+3i_1 \\ I_0 \end{matrix} \right]_3 = \prod_{i_1=0}^{I_1-1} \prod_{i_0=0}^{I_0-1} (1+3i_1-3i_0). \end{aligned}$$

Recall that since we have assumed that $\text{Disc } f \neq 0$, we may conclude that the total weight of the non-branch-point Weierstrass points is $\text{deg det } \mathbb{W} = 3(d-1)\kappa = 3gI_0I_1$. We have

$$BW = g^3 - g - 3gI_0I_1.$$

In particular, we get

$$BW = \begin{cases} \frac{1}{3}(g^3 - g) + \frac{1}{3}(g^2 - g), & \text{when } d \equiv -1 \pmod{3}, \\ \frac{1}{3}g^3 - g, & \text{when } d \equiv 1 \pmod{3}. \end{cases}$$

(Note that in the $d \equiv 1$ case, there are no parentheses.)

It is well known that on a hyperelliptic curve (i.e., a curve of the form $y^2 = f(x)$) BW is equal to $g^3 - g$, the total weight of all Weierstrass points. That is, the branch points account for all of the Weierstrass weight on a hyperelliptic curve and therefore every Weierstrass point is a branch point. Now, we see that on curves of the form $y^3 = f(x)$ with $\text{Disc } f \neq 0$, the branch points account for approximately one-third of the Weierstrass weight. Corollary 10 gives BW as an approximate fraction of the total weight for any n .

We can use the previous result to bound the primes p for which a non-branch-point Weierstrass point and a (finite) branch point “coalesce” modulo p . Consider

$$\Phi(x) = \prod_{P \text{ not a branch pt}} (x - x(P))^{\text{wt}(P)/3}.$$

Notice that for each point in this product there are two other points with the same x -coordinate. It is easy to see that the Weierstrass weight of each of these points is the same since there is an automorphism of the curve which sends each one to each of the others. Thus, every factor in the product appears three times. It is for this reason that we only take powers of $\frac{1}{3} \text{wt}(P)$. This polynomial is characterized by being monic and by:

$$\text{ord}_P \Phi(x) = \begin{cases} 0, & \text{for } y(P) = 0, \infty, \\ \text{wt}(P), & \text{otherwise.} \end{cases}$$

Notice, however that $\det \mathbb{W}(x)$ is a polynomial characterized by:

$$\text{ord}_P \det \mathbb{W}(x) = \begin{cases} 0, & \text{for } y(P) = 0, \infty, \\ \text{wt}(P), & \text{otherwise,} \end{cases}$$

and

$$\lambda(\det \mathbb{W}(x)) = \Lambda = \prod_{i_1=0}^{I_1-1} \left[\begin{matrix} d + 3i_1 \\ I_0 \end{matrix} \right]_3.$$

Thus, $\det \mathbb{W}(x) = \Lambda \Phi(x)$.

Theorem 12. *Let $C : y^3 = f(x)$ be a curve defined over a number field K , with $\text{Disc } f \neq 0$ and $\text{deg } f = d \not\equiv 0 \pmod{3}$. Let $g = d - 1$ be the genus of C , and let Λ and c be the constants defined in equations (2.2) and (2.3), respectively. Then*

1. $\text{Res}(f, \Phi) = (\frac{c}{\Lambda})^d (\text{Disc } f)^\kappa$. Furthermore, if $d \equiv -1 \pmod{3}$, then $c = \pm \Lambda$, so $\text{Res}(f, \Phi) = \pm (\text{Disc } f)^\kappa$.
2. Let v be a place of \bar{K} of characteristic $p > 2g - 3$. If f has v -integral coefficients, then the coefficients of Φ are v -integral, as well.
3. Assume that $p > 2g - 3$ and that C has good reduction modulo v . Further, assume that $d \equiv -1 \pmod{3}$. If P is congruent to a branch point modulo v , then P is a branch point itself. That is, no non-branch-point Weierstrass points may coalesce with branch points modulo p for $p > 2g - 3$.

In (3), if we exclude the (branch) point at infinity, our proof will still work for $d \equiv 1 \pmod{3}$. That is, a non-branch point P will not coalesce with any of the finite branch points.

Proof. (1) We know

$$c^d (\text{Disc } f)^\kappa = \text{Res}(f, \det \mathbb{W}) = \text{Res}(f, \Lambda \Phi) = \Lambda^d \text{Res}(f, \Phi),$$

so $\text{Res}(f, \Phi) = (\frac{c}{\Lambda})^d (\text{Disc } f)^\kappa$. If we assume $d \equiv -1 \pmod{3}$, then we write $d = 3k - 1$ and recall that $I_0 = 2k - 1$ and $I_1 = k - 1$. Then

$$\Lambda = \prod_{i_0=0}^{I_0-1} \prod_{i_1=0}^{I_1-1} (d + 3(i_1 - i_0)).$$

Let $j_m = I_m - 1 - i_m$ for $m = 0, 1$. Then

$$\begin{aligned} \Lambda &= \prod_{j_0=0}^{I_0-1} \prod_{j_1=0}^{I_1-1} (d + 3(I_1 - I_0) - 3(j_1 - j_0)) \\ &= \prod_{j_0=0}^{I_0-1} \prod_{j_1=0}^{I_1-1} (3k - 1 + 3(k - 1) - 3(2k - 1) - 3(j_1 - j_0)) \\ &= \prod_{j_0=0}^{I_0-1} \prod_{j_1=0}^{I_1-1} (-1 - 3(j_1 - j_0)) \\ &= (-1)^{I_0 I_1} c = \pm c. \end{aligned}$$

(2) If the coefficients of f are v -integral, then clearly so are the coefficients of the P_l 's. And if the coefficients of all the P_l 's are v -integral, then clearly so are the coefficients of $\det \mathbb{W}$. Now, $\Phi = \frac{1}{\Lambda} \det \mathbb{W}$, and

$$\Lambda = \prod_{i_0=0}^{I_0-1} \prod_{i_1=0}^{I_1-1} (d + 3(i_1 - i_0)).$$

By observation, the largest (most positive) factor in the product is

$$d + 3(I_1 - 1) = \begin{cases} 6k - 7, & d \equiv -1 \pmod{3} \\ 6k - 8, & d \equiv 1 \pmod{3} \end{cases} = \begin{cases} 2g - 3, & d \equiv -1 \pmod{3} \\ 2g - 2, & d \equiv 1 \pmod{3} \end{cases}$$

The smallest (most negative) factor in the product is

$$d - 3(I_0 - 1) = \begin{cases} -(3k - 5), & d \equiv -1 \pmod{3} \\ -(3k - 7), & d \equiv 1 \pmod{3} \end{cases}$$

Thus, if $p > 2g - 3$, then $\frac{1}{\Lambda}$ is v -integral, so the coefficients of Φ are v -integral.

(3) Clearly, we can multiply through by an appropriate multiple so that f has v -integral coefficients and so that $v(\text{Disc } f) = 0$, since C has good reduction at v . Unfortunately, the equation for C is now in the form

$$ay^3 = bx^d + \text{lower order terms.}$$

Since n and d are relatively prime, we can find an integer m_1 so that $3 \mid m_1$ and $m_1 \equiv -1 \pmod{d}$. We also need an integer m_2 with $d \mid m_2$ and $m_2 \equiv -1 \pmod{3}$. Then we can multiply the equation for C by $a^{m_2} b^{m_1}$ and get

$$a^{m_2+1} b^{m_1} y^3 = a^{m_2} b^{m_1+1} x^d + \text{lower order terms.}$$

Thus, the change of coordinates

$$(x, y) \mapsto \left(x \left(a^{(m_2)} b^{(m_1+1)} \right)^{-1/d}, y \left(a^{(m_2+1)} b^{(m_1)} \right)^{-1/3} \right)$$

gives us an equation for C in the correct form and with v -integral coefficients and with $v(\text{Disc } f) = 0$. Now, let P be a non-branch-point Weierstrass point. So $\det \mathbb{W}(x(P)) = 0$. We know that the coefficients of $\det \mathbb{W}(x)$ are v -integral, so $\det \mathbb{W}(x(P)) \equiv 0 \pmod{v}$. Now, $p > 2g - 3$, so $p \nmid \Lambda$. When $d \equiv -1 \pmod{3}$, we

know that $\Lambda = \pm c$, so $p \nmid c$. Now, we note that even if $d \equiv 1 \pmod{3}$, p does not divide c . For

$$c = \prod_{i_0=0}^{I_0-1} \prod_{i_1=0}^{I_1-1} (1 + 3(i_1 - i_0)).$$

So the most positive factor in this product is $1 + 3(I_1 - 1) = 3k - 5$, while the most negative factor is $1 - 3(I_0 - 1) = -(6k - 10)$. We have assumed that $p > 2g - 3$, which is $6k - 9$ when $d \equiv 1 \pmod{3}$, so p cannot divide c . Now we know $\text{Disc } f \not\equiv 0 \pmod{v}$ and $p \nmid c$, so

$$c^d (\text{Disc } f)^\kappa = \text{Res}(f, \det \mathbb{W}) \not\equiv 0 \pmod{v}.$$

The roots of f are the x -coordinates of the branch points (other than the one at infinity), hence for any branch point P_1 we have shown that

$$x(\widetilde{P}) \not\equiv x(\widetilde{P}_1) \pmod{v}.$$

So $\widetilde{P} \not\equiv \widetilde{P}_1 \pmod{v}$.

In the $d \equiv -1 \pmod{3}$ case, we can change coordinates so that a different branch point is at infinity. The transformation discussed before Proposition 2 is a birational transformation. We can change coordinates so that none of the branch points are at infinity using the inverse transformation. Then C will be given by an equation like

$$w^3 = g(x),$$

where $\deg g$ is divisible by 3. Now, we simply change coordinates again, placing a different branch point at infinity.

Thus, we see that it is arbitrary which branch point is at infinity, so it is clear that modulo v none of the branch points on the curve come together with any of the non-branch-point Weierstrass points (for $p > 2g - 3$). \square

Notice that our proof has also dealt with the case when we are given a curve with $d \equiv 0 \pmod{3}$. In that case no non-branch Weierstrass point can coalesce with any branch point modulo v if $p > 2g - 3$.

REFERENCES

- [A1] Accola, R., *On generalized Weierstrass points on Riemann surfaces*, Modular Functions in Analysis and Number Theory (T. A. Metzger, ed.), Lecture Notes in Math. and Stat., Univ. of Pittsburgh, Pittsburgh, PA, 1978, pp. 1–19. MR **85f**:14016
- [A2] Accola, R., *Topics in the Theory of Riemann Surfaces*, Lecture Notes in Mathematics, vol. 1595, Springer-Verlag, Berlin, Heidelberg, 1994. CMP 1995:12
- [B] Burnol, J.-F., *Weierstrass points on arithmetic surfaces*, Diss. Harvard, Invent. Math. **107** (1992), 421–432. MR **93b**:14040
- [C] Coppens, M., *The Weierstrass gap sequences of the total ramification points of trigonal coverings of \mathbf{P}^1* , Indag. Math. **47** (1985), 245–270. MR **87c**:14028
- [F-K] Farkas, H. and Kra, I., *Riemann Surfaces*, Springer-Verlag, New York, 1980. MR **82c**:30067
- [H] Hurwitz, A., *Über algebraische Gebilde mit eindeutigen Transformationen in sich*, Mathematische Annalen **41** (1893), 403–442.
- [Ka] Kato, T., *On Weierstrass points whose first non-gaps are three*, J. Reine Angew. Math. **316** (1980), 99–109. MR **82f**:30046
- [Ko] Koo, J. K., *On Holomorphic Differentials of Some Algebraic Function Field of One Variable Over \mathbf{C}* , Bull. Austral. Math. Soc. **43** (1991), 399–405. MR **92e**:14019
- [L] Lewittes, J., *Automorphisms of compact Riemann surfaces*, American Journal of Mathematics **85** (1963), 734–752. MR **28**:4102

- [M] Mumford, D., *Curves and their Jacobians*, Univ. of Mich. Press, Ann Arbor, 1975. MR **54**:7451
- [R-G] Rademacher, H. and Grosswald, E., *Dedekind Sums*, Carus Mathematical Monographs, vol. 16, Math. Assoc. America, Washington, DC, 1972. MR **50**:9767
- [S] Silverman, J.H., *Some arithmetic properties of Weierstrass points: hyperelliptic curves*, Boletim de Sociedade Brasileira de Matematica **21** (1) (1990), 11–50. MR **92k**:11066
- [U-H] Uspensky, J. and Heaslet, M., *Elementary Number Theory*, McGraw-Hill, New York, 1939. MR **1**:38d

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF MICHIGAN, ANN ARBOR, MICHIGAN 48109-1003

E-mail address: `towse@math.lsa.umich.edu`