ANNALS OF MATHEMATICS
Vol. 72, No. 3. November, 1960
Printed in Japan

ARITHMETIC VARIETY OF MODULI FOR GENUS TWO

BY JUN-ICHI IGUSA*

(Received March 14, 1960)

1. Introduction

One of the most interesting problems in algebraic geometry is to obtain a generalization of the theory of elliptic modular functions to the case of higher genus. We know that in the elliptic case the so-called *Weierstrass absolute invariant* j plays a fundamental role. We recall that, as a function on the upper-half plane, it is holomorphic and invariant under the modular group. Moreover, if τ is the variable in the upper-half plane, the Fourier expansion of j, i.e., the expansion of j with respect to the variable $q = \exp(2\pi i\tau)$ starts as follows

$$j = q^{-1} + 744 + \cdots .$$

We observe that the function j is uniquely determined by these properties, and a remarkable fact, which is crucial in some arithmetic applications, is that *the coefficients of powers of q are all rational integers*. Now, it is quite natural to ask whether such a function or a system of functions exists in the case of higher genus. The problem is classical and there are many important papers in this direction. Nevertheless, the problem has been unsettled even for genus two in which case we claim that we have a complete solution. In order to obtain a better understanding by the reader, we shall explain how the problem can be interpreted in terms of *arithmetic geometry*. We recall that j takes the same value at τ_1 and τ_2 if and only if they are equivalent under the modular group, i.e., if and only if the complex tori determined by 1, τ_1 and 1, τ_2 are complex-analytically isomorphic. Since every complex torus (of dimension one) can be obtained in this way, we can consider j as a function on the space of complex tori. We observe that complex tori are objects of algebraic geometry, i.e., they are just elliptic curves. Moreover, elliptic curves exist not only in characteristic zero but also in characteristic p for $p = 2, 3, 5, \cdots$. We observe also that the function j is not quite intrinsic. In fact, any function of the form $\pm j +$ integer and only such function equally serves the purpose. This suggests that the ring $\mathbf{Z}[j]$ is intrinsic. Here $\mathbf{Z}$ is the ring of rational integers. We shall show that $\mathbf{Z}[j]$ is indeed intrinsic. Let $\mathbf{Q}$ be the field of fractions of $\mathbf{Z}$ and let $\mathrm{GF}(p)$ be the prime field of characteristic p for $p = 2, 3, 5, \cdots$. We fix one universal domain

[*] Alfred P. Sloan Research Fellow.

for each characteristic and we take their union as an *arithmetic universal domain.* We fix also one universal projective space for each characteristic and, unless we add "arithmetic" or "over $\mathbf{Z}$," we assume that every variety is contained in some universal projective space. Consider the set $\mathcal{A}$ of all elliptic curves. We say that a function f, defined on $\mathcal{A}$ with values in the arithmetic universal domain, is "characteristic-preserving" if, at any elliptic curve A of characteristic p, the value $f(A)$ of f at A is an element of the universal domain of characteristic p for every p. Such function f is called "continuous" if, over an arbitrary specialization $A \rightarrow A'$ within the set $\mathcal{A}$, the quantity $f(A)$ specializes uniquely to the quantity $f(A')$. If we use this terminology, we can say that the ring $\mathbf{Z}[j]$ can be identified with the ring of characteristic-preserving continuous functions defined on $\mathcal{A}$ with values in the arithmetic universal domain and which depend only on birational classes of elliptic curves. This view-point was originated by Deuring [4] and we gave an explicit formulation in our previous paper [9-III, Section 2]. At any rate, we observe that the above interpretation can be generalized to the case of genus two. We call a non-singular curve of genus two simply a *hyperelliptic curve* and we consider the set C of all hyperelliptic curves. A characteristic-preserving continuous function defined on C with values in the arithmetic universal domain is called an *absolute invariant* if it depends only on birational classes of hyperelliptic curves. The set of absolute invariants forms a ring $\mathfrak{R}$ and we shall determine its structure. Let y_1, y_2, y_3 be independent variables over $\mathbf{Q}$ and define y_4 as

$$y_4 = 1/4 \cdot (y_1 y_3 - y_2^2) .$$

Then, if ζ is a primitive fifth root of unity, the ring of invariant elements of $\mathbf{Z}[y] = \mathbf{Z}[y_1, y_2, y_3, y_4]$ under the transformation $y_i \rightarrow \zeta^i y_i$ for $i = 1, 2, 3, 4$ can be identified with $\mathfrak{R}$. In particular, it is an integrally closed noetherian domain over $\mathbf{Z}$ with ten generators. Moreover, if we associate an arithmetic affine variety $\mathfrak{M}$ to $\mathfrak{R}$ consisting of points which are homomorphisms of $\mathfrak{R}$ in universal domains, it will be shown as Theorem 2 that *the "space of birational classes of hyperelliptic curves" and* $\mathfrak{M}$ *are "homeomorphic."* Therefore, we call $\mathfrak{M}$ the *arithmetic variety of moduli.* The geometric structure of $\mathfrak{M}$ can be deduced quite easily from the algebraic structure of $\mathfrak{R}$. Moreover, the variety of moduli with reference to a field k can be obtained by just tensorizing $\mathfrak{M}$ with k.

Finally, we shall explain more technical points. It is known that Bolza [1] applied the invariant theory of binary sextics (cf. [3]) to the investigation of moduli of hyperelliptic curves. However, he did not construct the variety of moduli. It was straightforward to construct an affine variety, which coincides with our $\mathfrak{M} \otimes \mathbf{Q}$, out of invariants, but *the classical theory gave no guarantee as to the fact that every point of the variety represents a birational class of hyperelliptic curves.* Therefore, even in the complex case it was unknown that the variety of moduli is precisely the quotient variety of an affine space with complex co-ordinates z_1, z_2, z_3 with respect to a rotation group defined by $z_i \rightarrow \zeta^i z_i$ for $i = 1, 2, 3$ in which $\zeta^5 = 1$. This was the first difficulty when we started working on the subject. In overcoming this difficulty, we found a new way of constructing an invariant theory of binary sextics, which is valid for any characteristic different from 2. The second difficulty was, then, how to include the characteristic 2 case. We trusted in the God who created integers and introduced what we call *arithmetic invariants.* These invariants worked perfectly in characteristic 2 and the existence of the arithmetic variety of moduli came out together with its structure. One serious drawback of our method is that it depends too heavily on calculations of various kinds. It is, therefore, greatly desired to re-establish our structure theorem from a new conceptual view-point.

2. Normal forms of hyperelliptic curves

We shall assume that the reader is familiar with linear systems, the Riemann-Roch theorem and coverings, all for curves. In general, if P is a point of a non-singular curve C of genus g, the complete linear system determined by gP has a positive dimension if and only if it is "partially contained" in the canonical system. This is a simple consequence of the Riemann-Roch theorem. A point P having this property is called a *Weierstrass point.* Now, suppose that C is a hyperelliptic curve, i.e., a non-singular curve of genus two. Then, we observe, first of all, that the canonical system is characterized uniquely as a linear system of degree two and of dimension one. Again, this is a consequence of the Riemann-Roch theorem. Therefore, if D is the so-called projective straight-line, up to automorphisms of D, *there exists one and only one way to convert C into a two-sheeted covering of D.* In particular, the different $\mathfrak{d}$ of the covering has an intrinsic meaning and a point P of C is a Weierstrass point if and only if P is a component of $\mathfrak{d}$. Also, the Riemann-Hurwitz relation implies that $\mathfrak{d}$ is a member of the triple canonical system, hence it is of degree six. Therefore, if the characteristic is different from 2, we get six Weierstrass points. Let P be one of them and take a function x having $2P$ as its polar divisor. We then normalize x so that it takes the value 0, 1 at two other Weierstrass points. Let $\lambda_1, \lambda_2, \lambda_3$ be the values of x at the remaining three Weierstrass points. Then the lambdas are different from each other and from 0, 1, ∞. Moreover, the divisor of the function $x(x-1)(x-\lambda_1)(x-\lambda_2)(x-\lambda_3)$ is $2(\mathfrak{d}-6P)$. Since $\mathfrak{d}$ and $6P$ are both members of the triple canonical system, there exists a function y with $\mathfrak{d}-6P$ as its divisor. If we normalize the constant factor of y suitably, the pair (x, y) uniformizes a plane curve defined inhomogeneously by

$$Y^2 = X(X-1)(X-\lambda_1)(X-\lambda_2)(X-\lambda_3)\,,$$

which we call a *Rosenhain normal form* of C, cf. [11]. This curve is, by its construction, birationally equivalent to C. We observe that the *birational equivalence* of C is characterized by the *projective equivalence* of the six image points on D of the Weierstrass points of C. We observe also that we have $6\cdot5\cdot4\cdot2=240$ possibilities for (x, y) and 120 possibilities for x. Conversely, if lambdas are different from each other and from 0, 1, ∞, the corresponding Rosenhain normal form can be considered as a "normal form" of its non-singular model which is indeed a hyperelliptic curve.

In the case of characteristic 2, the situation is entirely different because of the wildness of ramification. First of all, instead of a Kummer generation, we must use an Artin-Schreier generation $y^2-y=R(x)$. The generator y is unique up to transformations of the form $y\rightarrow y+B(x)$. Here $R(x)$ and $B(x)$ are rational functions of x. Let $\sum n_a(a)$ be the polar divisor of the function $R(x)$ on D. If we take y suitably, we can make n_a odd. Then C is ramified at each a and, if P_a is the unique point of C over a, the different $\mathfrak{d}$ is given by $\sum(n_a+1)P_a$. This is a consequence of Hasse's theory [7]. Thus we get the following relation $\sum(n_a+1)=6$. This equation in the unknown odd positive integers n_a has three types of solutions, and they are (1, 1, 1), (3, 1), (5). Accordingly C can be transformed birationally into plane curves defined inhomogeneously by

$$Y^2 - Y = \begin{cases} \alpha X + \beta X^{-1} + \gamma(X-1)^{-1} \\ X^3 + \alpha X + \beta X^{-1} \\ X^5 + \alpha X^3\,, \end{cases}$$

and these are the *normal forms in characteristic* 2. Here α, β, γ are finite quantities and α, β, γ in the first case and β in the second case are different from 0. The verification is straightforward except possibly for the last case. In this case, we first transform C birationally into $Y^2-Y=X^5+\alpha X^3+\beta X$ and, then, we apply a linear transformation to X and, at the same time, we modify Y by a quadratic polynomial in X to get $Y^2-Y=X^5+\alpha X^3$. Since the calculation is elementary, we leave it to the reader. Conversely, if three constants α, β, γ satisfy the above mentioned conditions, non-singular models of the corresponding normal forms are hyperelliptic curves of the respective type.

We can also discuss conditions under which hyperelliptic curves in characteristic 2 are birationally equivalent. If α, β, γ are the constants in type (1, 1, 1), since x is unique up to projective transformations which permute $0, 1, \infty$ among themselves, i.e., up to six projective transformations $x \to x$, $1-x$, $x(x-1)^{-1}$, x^{-1}, $(1-x)^{-1}$, $(x-1)x^{-1}$, we see that the elementary symmetric functions $\alpha+\beta+\gamma$, $\alpha\beta+\beta\gamma+\gamma\alpha$, $\alpha\beta\gamma$ of α, β, γ depend only on the birational class of the curve and characterize this class uniquely. Similarly, if α, β are the constants in type (3, 1), since x is unique up to multiplications by third roots of unity, we see that α^3, $\alpha\beta$, β^3 characterize the birational class uniquely. Finally, if α is the constant in type (5), we see that α^5 characterizes the birational class uniquely. In fact, suppose that x, y satisfying $Y^2 - Y = X^5 + \alpha' X^3$ and x', y' satisfying $Y^2 - Y = X^5 + \alpha X^3$ generate the same function-field over the universal domain. Since x and x' generate the same simple transcendental extension of the universal domain over which the function-field is ramified only at $x = x' = \infty$, we see that x' is linear in x. If we replace x' by a linear expression of x and y' by $y + B(x)$ in $(y')^2 - y' = (x')^5 + \alpha(x')^3$ and require that the new relation should be $y^2 - y = x^5 + \alpha' x^3$, by a simple calculation, we get

$$\alpha' = \zeta\alpha\ , \quad x' = \zeta^2(x + c_1^2)\ ,$$
$$y' = y + c_1x^2 + c_2x + c_3$$

in which

$$\zeta^5 = 1\ , \quad c_1^{16} + (\alpha')^2c_1^8 + \alpha' c_1^2 + c_1 = 0\ ,$$
$$c_2^2 = \alpha' c_1^2 + c_1\ , \quad c_3^2 - c_3 = c_1^2c_2\ .$$

Conversely, if x, y satisfy $Y^2 - Y = X^5 + \alpha' X^3$ and if we define α, x', y' as above, clearly x', y' satisfy $Y^2 - Y = X^5 + \alpha X^3$. This proves the assertion. We see also that the above thirty-two transformations, which are certainly distinct, exhaust all possible ways of transforming $Y^2 - Y = X^5 + \alpha X^3$ birationally into $Y^2 - Y = X^5 + \alpha' X^3$. In case α, hence also α', are 0, we can vary ζ, hence the number of transformations is 160.

Now, in order to define arithmetic invariants in the next section, we shall introduce one *universal normal form* which is valid "universally" for all characteristics. Let C be a hyperelliptic curve and let P be one of its Weierstrass points. Let Q be a non-Weierstrass point and let $Q + Q'$ be a member of the canonical system. Then the complete linear system $|3P + Q| = |P + 2Q + Q'|$ is of degree four and of dimension two. Moreover, if M is a point of C, the residual system $|3P + Q| - M$ has a fixed point if and only if the complete linear system $|P + Q - M|$ exists. Since the complete linear system $|P + Q|$ is of dimension zero, this is the case if and only if M coincides with either P or Q. All these are consequences of the Riemann-Roch theorem. In particular, the linear system $|3P + Q|$ gives rise to a birational transformation of C to a plane quartic C'. The transformation is biregular except at P and Q, which are mapped to the same, necessarily multiple, point of C'. Moreover, since we have $|3P + Q| - 2P = P + Q$ and $|3P + Q| - 2Q = P + Q'$, the multiple point is an ordinary double point and the corresponding two tangents cut C' at the branches $3P + Q$ and $P + 2Q + Q'$. On the other hand, we observe that the curve C' is determined only up to projective transformations. We take three co-ordinate axes so that the double point has co-ordinates $(0, 1, 0)$. We can assume thereby that $Z = 0$ and $X = 0$ are tangent respectively to the branches P and Q of C'. Also, we can assume that $Y = 0$ is tangent to C' at the image point of Q'. If we adjust, finally, the proportionality factors of X, Y, Z suitably, the equation for C' can be written inhomogeneously as follows

$$XY^2 + (1 + aX + bX^2)Y + X^2(c + dX + X^2) = 0 \ .$$

We observe that, once P and Q are fixed, the above normalization is unique up to transformations of the form $(a, b, c, d) \rightarrow (\zeta a, \zeta^2 b, \zeta^3 c, \zeta^4 d)$ with $\zeta^5 = 1$. Conversely, if a curve C' defined by the above equation has no other singularity except at $(0, 1, 0)$, its non-singular model is a hyperelliptic curve.

We now possess three kinds of normal forms. In the following we shall discuss their relations. If the characteristic is different from 2, the universal normal form can be transformed into the Rosenhain normal form. Therefore a, b, c, d and the corresponding lambdas must be related in some may. This relation can be obtained as follows. If x is different from 0 and ∞, the two roots of the equation

$$xY^2 + (1 + ax + bx^2)Y + x^2(c + dx + x^2) = 0$$

determine a member of the canonical system of C other than $2P$ and $Q + Q'$. These canonical divisors themselves correspond respectively to $x = \infty$ and $x = 0$. Therefore, the five missing Weierstrass points correspond to those values of x at which the above quadratic equation in Y has double roots, i.e., to the roots of the equation

$$(1 + aX + bX^2)^2 - 4X^3(c + dX + X^2) = 0 \ .$$

In other words, the six roots of this equation and the six roots of the equation $X(X-1)(X-\lambda_1)(X-\lambda_2)(X-\lambda_3)=0$, both considered as inhomogeneous sextic equations, are projectively equivalent. This is the relation we are looking for. If the characteristic is 2, the universal normal form can be transformed, according to its type, into one of the three normal forms we introduced before. The classification of these types in terms of a, b, c, d is immediate. In fact, the previous argument shows that the Weierstrass points of C other than P, if any, correspond to the roots of the equation $1+aX+bX^2=0$. Therefore, if ab is different from 0, we get three Weierstrass points, i.e., we get the type (1, 1, 1). In the same way, we see that the type (3, 1) corresponds to $ab=0$, but not $a=b=0$, and the type (5) corresponds to $a=b=0$. We shall determine the relation between a, b, c, d and the corresponding α, β, γ explicitly. Suppose first that ab is different from 0. If we multiply $x(1+ax+bx^2)^{-2}$ to $xy^2+(1+ax+bx^2)y+x^2(c+dx+x^2)=0$ and replace $x(1+ax+bx^2)^{-1}y$ by y, we get $y^2-y=R(x)$ with

$$R(x)=x^3(1+ax+bx^2)^{-2}(c+dx+x^2).$$

Then, we decompose $R(x)$ into partial fractions and apply the transformation $y\to y+B(x)$ with a suitable $B(x)$ and, finally, we apply a linear transformation to x. In this way, we get

$$y^2-y=\alpha x+\beta x^{-1}+\gamma(x-1)^{-1}$$

with

$$\begin{aligned}\alpha&=ab^{-3}\,,\\ \beta&=a^{-3}b\xi^{-2}\bigl(c+\xi^{-2}+a(c\xi+d+\xi^{-1})^{1/2}\bigr)\,,\\ \gamma&=a^{-3}b\eta^{-2}\bigl(c+\eta^{-2}+a(c\eta+d+\eta^{-1})^{1/2}\bigr)\end{aligned}$$

in which $\xi+\eta=a$, $\xi\eta=b$. The reader will find no difficulty in carrying through the indicated calculation and verifying the statement. Similarly, if a is different from 0 and $b=0$, we can transform

$$xy^2+(1+ax)y+x^2(c+dx+x^2)=0$$

into

$$y^2-y=x^3+\alpha x+\beta x^{-1}$$

with

$$\begin{aligned}\alpha&=a^{5/3}\bigl(a^{-3}c+(a^{-5}+a^{-4}d)^{1/2}\bigr)\,,\\ \beta&=a^{-5/3}\bigl(a^{-5}+a^{-3}c+(a^{-5}+a^{-4}d+a^{-3}c)^{1/2}\bigr)\,.\end{aligned}$$

If b is different from 0 and $a=0$, we get similar but more complicated expressions for α, β in terms of b, c, d. Finally, if we have $a=b=0$, then $xy^2+y+x^2(c+dx+x^2)=0$ can be transformed into $y^2-y=x^5+\alpha x^3$ with $\alpha=c$. The verification is left to the reader.

3. Construction of arithmetic invariants

The discussion in the previous section shows that "moduli" of hyperelliptic curves in characteristic different from 2 are closely connected with projective invariants of binary sextics. *A surprising fact is that this connection subsists also in characteristic* 2. At any rate, we shall first introduce the notion of integral invariants of binary sextics. Consider an affine space over **Z** with a generic point u. The co-ordinates of u are, therefore, independent variables over **Q**. If J is a characteristic-preserving continuous function on this space with values in the arithmetic universal domain, the continuity being defined again by specializations, the value $J(u)$ of J at u is a polynomial in the co-ordinates of u with coefficients in **Z**. In fact, since J is characteristic-preserving, we see that $J(u)$ is an element of the universal domain over **Q**. Since J is continuous, we see that $J(u)$ has no other specialization than itself over $\mathbf{Q}(u)$. Hence $J(u)$ is an element of $\mathbf{Q}(u)$. Moreover $J(u)$ must be a polynomial in the co-ordinates of u for, otherwise, the specialization of u to a general zero of the denominator of $J(u)$ will map $J(u)$ to ∞. Also $J(u)$ has integral coefficients for, otherwise, a generic specialization of u modulo a prime factor of the denominator of some coefficient of $J(u)$ will map $J(u)$ again to ∞. We note that the above proof subsists even if J admits removable discontinuities, i.e., apparent discontinuities of co-dimension two. Conversely, if $J(u)$ is a polynomial in the co-ordinates of u with coefficients in **Z**, we can define $J(u')$ for any point u' of the affine space by just replacing u by u' and, in this way, we get a characteristic-preserving continuous function on the affine space. We next observe that the set of binary sextics is an affine space over **Z** of dimension seven minus the arithmetic line consisting of origins. Consider a projective transformation $X \rightarrow X'$ of a projective straight-line and an arbitrary sextic of the same characteristic written inhomogeneously as $u_0X^6 + u_1X^5 + \cdots + u_6$. Then, we get a new sextic of the same characteristic by

$$(u'_0X^6 + u'_1X^5 + \cdots + u'_6)^{-1}(dX)^3 = \big(u_0(X')^6 + u_1(X')^5 + \cdots + u_6\big)^{-1}(dX')^3 .$$

In this way, the group of projective transformations operates on the space of sextics of the same characteristic, hence we can talk about its orbits. A characteristic-preserving continuous function J on the space of all sextics with values in the arithmetic universal domain is called an *integral invariant* if it depends only on orbits. If a sextic

$$u_0X^6 + u_1X^5 + \cdots + u_6$$

corresponds to a generic point, the previous observation shows that the value $J(u)$ of J at this sextic is a polynomial in $u_0, u_1, \cdots, u_6$ with coefficients in $\mathbf{Z}$. Therefore, we can decompose J into homogeneous parts and, then, each one of them will itself be an integral invariant. In other words, the set of integral invariants forms a graded integral domain over $\mathbf{Z}$. If we tensorize this ring with $\mathbf{Q}$ over $\mathbf{Z}$, we get the graded integral domain of *rational invariants*. On the other hand, if two sextics are equivalent up to constant factors under the group of projective transformations, we say that they are projectively equivalent. This is the case if and only if the six roots of one sextic and the six roots of another sextic are projectively equivalent. We observe that the quotient of two homogeneous integral invariants of the same degree takes the same value at projectively equivalent sextics. Now, we shall use the following well-known elementary lemma for an actual construction of integral invariants:

LEMMA 1. *Let $x_1, x_2, \cdots, x_6$ be roots of a sextic $u_0X^6 + u_1X^5 + \cdots + u_6$. Then, an expression of the form*

$$u_0^m \sum (x_i - x_j)(x_k - x_l) \cdots ,$$

in which every x_i appears m times in each product and which is symmetric in $x_1, x_2, \cdots, x_6$, can be considered as an irrational form of the value at the sextic of a homogeneous integral invariant J of degree m.

PROOF. We can obviously assume that $u_0, u_1, \cdots, u_6$ are independent variables over $\mathbf{Q}$. Then, by the so-called fundamental theorem of symmetric functions (cf. [12-I, Section 29]), we can rewrite the above expression as a homogeneous polynomial $J(u)$ in $u_0, u_1, \cdots, u_6$ of degree m with coefficients in $\mathbf{Z}$. The rest, i.e., the fact that J depends only on orbits, is a straightforward verification.

A well-known example of integral invariants is the *discriminant* D which is defined by the following expression

$$D(u) = u_0^{10} \prod_{i<j} (ij)^2 .$$

Here (ij) is an abridged notation for $x_i - x_j$. The above lemma shows that D is indeed a homogenous integral invariant of degree ten. A rational form of $D(u)$ is also known (cf. [12-I, Section 31]). In fact, if $R(u)$ is the *Sylvester resultant* of the sextic and its derivative with respect to X, we have $R(u) = - u_0D(u)$. The lemma shows also that the following expressions

$$A(u) = u_0^2 \sum_{\text{fifteen}} (12)^2(34)^2(56)^2$$

$$B(u) = u_0^4 \sum_{\text{ten}} (12)^2(23)^2(31)^2(45)^2(56)^2(64)^2$$

$$C(u) = u_0^6 \sum_{\text{sixty}} (12)^2(23)^2(31)^2(45)^2(56)^2(64)^2(14)^2(25)^2(36)^2$$

define homogeneous integral invariants A, B, C of degree two, four, six. Rational forms of these invariants can be obtained following the proof of the fundamental theorem of symmetric functions. The calculations are elementary but long and tedious. If we evaluate A at a sextic of the form $v_0X^5 - v_1X^4 + \cdots - v_5$, we get $2(2^2 5v_0v_4 - 2^3v_1v_3 + 3v_2^2)$. Similarly, if we evaluate B and C at this sextic, we get rational forms for B and C. The one for B consists of nine terms and the greatest common divisor of the coefficients is 2^2; the one for C consists of twenty-seven terms and the greatest common divisor of the coefficients is 2. We recall here that the reason why we introduced invariants at all was to discuss moduli of hyperelliptic curves in terms of associated sextics of Weierstrass points. However, *if we restrict to integral invariants, the discussion will break down in characteristic* 2 *simply because Weierstrass points behave badly under reduction modulo* 2. Therefore, we must introduce a certain type of rational invariants. In the previous section, we introduced a normal form depending on four, instead of three, variable coefficients a, b, c, d such that its associated sextic of Weierstrass points is

$$(1 + aX + bX^2)^2 - 4X^3(c + dX + X^2) .$$

Since this normal form is valid universally, we define an *arithmetic invariant* to be a rational invariant whose value at the above sextic has integral coefficients as a polynomial in a, b, c, d. According to this definition, every integral invariant is an arithmetic invariant. We shall, now, construct five basic arithmetic invariants out of A, B, C, D.

If we express A as a polynomial in a, b, c, d with coefficients in $\mathbf{Z}$, i.e., if we evaluate A at the sextic $(1 + aX + bX^2)^2 - 4X^3(c + dX + X^2)$ in which a, b, c, d are independent variables over $\mathbf{Q}$, we can see that the greatest common divisor of the coefficients is 2^3. Moreover, if we define J_2 as $2^{-3}A$, we get the following congruence modulo 2

$$J_2 \equiv a^2b^2 .$$

In the same way, the greatest common divisor of the coefficients of D as a polynomial in a, b, c, d is shown to be 2^{12}. Moreover, if we define J_{10} as $2^{-12}D$, we get the following congruence modulo 2

$$\begin{aligned} J_{10} \equiv a^7c &+ a^6(b^3c + d) + a^5(b^4cd + b^3d + bc + 1) \\ &+ a^4(b^5c^2 + b^4d^2 + b^3 + c^2) \\ &+ a^3b(bc + 1)^3 + (bc + 1)^4 . \end{aligned}$$

The construction of the three remaining basic invariants is a little bit more complicated. Let m be an integer and consider mJ_2^2-B. Then, the greatest common divisor of the coefficients of $mJ_2^2 - B$ as a polynomial in a, b, c, d attains its maximal value $2^5 3 = 96$ when m satisfies the congruence $m \equiv 4 \bmod 96$. Moreover, if we define J_4 as $2^{-5}3^{-1}(4J_2^2 - B)$, we get the following congruence modulo 2

$$J_4 \equiv a^5b + a^4b^2d + a^3b^2(bc + 1) + a^2b^2c^2 + ab^3(bc + 1) .$$

In the same way, the greatest common divisor of the coefficients of $mJ_2^3 + nJ_2J_4 - C$ as a polynomial in a, b, c, d attains its maximal value $2^6 3^2 = 576$ when m, n satisfy the congruences $m \equiv 8$, $n \equiv -160 \bmod 576$. Moreover, if we define J_6 as $2^{-6}3^{-2}(8J_2^3 - 160J_2J_4 - C)$, we get the following congruence modulo 2

$$J_6 \equiv a^8 + a^6b^2d^2 + a^4b^2(bc + 1)^2 + a^2b^2c^4 + b^4(bc + 1)^2 .$$

The above expressions for J_2, J_4, J_6 show that J_2J_6 and J_4^2 are congruent modulo 2. If we compute the greatest common divisor of the coefficients of $J_2J_6 - J_4^2$ as a polynomial in a, b, c, d, we get 2^2. Moreover, if we define J_8 as $2^{-2}(J_2J_6 - J_4^2)$, we get the following congruence modulo 2

$$\begin{aligned} J_8 \equiv a^9bd^2 &+ a^8d^3(b^2 + d) + a^7b(bc + 1)(bc + bd^2 + 1) \\ &+ a^6b^2d(b^2c^2 + c^2d + 1) \\ &+ a^5b(b^4c^3 + b^3c^2 + b^3cd^2 + b^2c + b^2d^2 + b + c^4) \\ &+ a^4(b^2c^4d + b^2c^2 + 1) + a^3b^2c^3(b^4 + bc^2 + c) \\ &+ a^2b^2c(b^5 + b^4cd + c^5) \\ &+ ab^3(b^5cd + b^4c^3 + b^4d + b^3c^2 + bc^5 + c^4) \\ &+ (b^9c^2 + b^8d^2 + b^7 + b^6c^4 + c^8) . \end{aligned}$$

We have not yet explained the motivation of the above construction. This comes from a part of the main theorem of classical invariant theory of binary sextics which asserts that every homogeneous rational invariant of even degree is a polynomial in A, B, C, D with rational coefficients (cf. [3]). However, we shall not make any use of this theorem except for the motivation. At any rate, it is clear from the construction that each J_{2i} is a homogeneous arithmetic invariant of degree $2i$ for $i = 1, 2, \cdots, 5$. We shall examine whether these *basic arithmetic invariants* are actually useful to discuss moduli of hyperelliptic curves in characteristic 2.

We know that birational classes of hyperelliptic curves in characteristic 2 can be characterized in terms of the constants α, β, γ. We also know the relations between a, b, c, d and the corresponding α, β, γ explicitly. Therefore, we can discuss relations between α, β, γ and the basic arithmetic invariants and see whether they work well or not. We first observe that the type (1, 1, 1) is characterized by the non-vanishing of J_2; the type (3, 1) is characterized by $J_2 = 0$ and the non-vanishing of J_6, and the type (5) by $J_2 = J_6 = 0$. Moreover, if α, β, γ are the constants for the type (1, 1, 1), we have the following relations

$$\alpha^2 + \beta^2 + \gamma^2 = J_2^{-2}J_4 , \qquad \alpha^2\beta^2\gamma^2 = J_2^{-5}J_{10} ,$$
$$\alpha^2\beta^2 + \beta^2\gamma^2 + \gamma^2\alpha^2 = J_2^{-4}J_8 + (J_2^{-2}J_4)^3 + (J_2^{-2}J_4)^4 .$$

Also, if α, β are the constants for the type (3, 1), we have the following relations

$$\alpha^6 = J_6^{-1}J_8^{3/4} , \quad \beta^6 = J_6^{-5}J_{10}^3 , \quad \alpha^2\beta^2 = J_6^{-2}J_8^{1/4}J_{10} .$$

Finally, if α is the constant for the type (5), we have

$$\alpha^{10} = J_{10}^{-1}J_8^{5/4} .$$

Therefore, the basic arithmetic invariants work perfectly well in characteristic 2.

4. Some lemmas

We shall devote this section to proving some lemmas which depend on explicit rational forms of the basic arithmetic invariants. If we take a sextic of the form $v_0X^5 - v_1X^4 + \cdots - v_5$ in characteristic different from 2 and if we evaluate J_2, J_4, J_6 at this sextic, we get

$$\begin{aligned}
J_2 &= 5v_0v_4 - 2v_1v_3 + 2^{-2}3v_2^2 ,\\
J_4 &= -2^{-3}[5^2v_0^2v_3v_5 - 3\cdot5v_0^2v_4^2 - 3\cdot5v_0v_1v_2v_5 + 7v_0v_1v_3v_4\\
&\quad + 2^{-1}v_0v_2^2v_4 - v_0v_2v_3^2 + 2^2v_1^3v_5 - v_1^2v_2v_4 - v_1^2v_3^2\\
&\quad + v_1v_2^2v_3 - 2^{-4}3v_2^4] ,\\
J_6 &= -2^{-4}[2^{-1}5^3v_0^3v_2v_5^2 - 5^2v_0^3v_3v_4v_5 + 5v_0^3v_4^3 - 5^2v_0^2v_1^2v_5^2\\
&\quad - 2\cdot5v_0^2v_1v_2v_4v_5 + 2\cdot5v_0^2v_1v_3^2v_5 - v_0^2v_1v_3v_4^2\\
&\quad - 2^{-2}5v_0^2v_2^2v_3v_5 - 2^{-2}11v_0^2v_2^2v_4^2 + 2^{-1}7v_0^2v_2v_3^2v_4\\
&\quad - v_0^2v_3^4 + 2\cdot3v_0v_1^3v_4v_5 - 3v_0v_1^2v_2v_3v_5 + 2^{-1}7v_0v_1^2v_2v_4^2\\
&\quad - 2v_0v_1^2v_3^2v_4 + 2^{-2}3v_0v_1v_2^3v_5 - 2^{-2}7v_0v_1v_2^2v_3v_4\\
&\quad + v_0v_1v_2v_3^3 + 2^{-4}7v_0v_2^4v_4 - 2^{-2}v_0v_2^3v_3^2 - v_1^4v_4^2\\
&\quad + v_1^3v_2v_3v_4 - 2^{-2}v_1^2v_2^3v_4 - 2^{-2}v_1^2v_2^2v_3^2 + 2^{-3}v_1v_2^4v_3\\
&\quad - 2^{-6}v_2^6] .
\end{aligned}$$

It is not necessary to mention the rational form for J_8, because it is simply $2^{-2}(J_2J_6 - J_4^2)$. Also, we know the rational form of J_{10}, because we know it for the discriminant D. Now, we shall first obtain an exact information about the dependence of the basic arithmetic invariants.

Let $\mathbf{Z}[X]$ be a graded ring of polynomials in five letters X_i of degree $2i$ for $i = 1, 2, \cdots, 5$ with coefficients in $\mathbf{Z}$. Let k be a field of arbitrary characteristic and consider the tensor product $k[X] = \mathbf{Z}[X] \otimes k$ taken over $\mathbf{Z}$. If $F(X)$ is an arbitrary element of the new graded ring $k[X]$, we can consider $F(J) = F(J_2, J_4, \cdots, J_{10})$ as a polynomial in a, b, c, d with coefficients in k.

LEMMA 2. *Suppose that $F(X)$ is a homogeneous element of $k[X]$. Then, we have $F(J) = 0$ if and only if $F(X)$ is a multiple of $X_1X_3 - X_2^2 - 4X_4$.*

PROOF. Suppose first that the characteristic of k is 2. Then $F(X)$ can be written uniquely in the form $F_0(X_1, X_3, X_4, X_5) + F_1(X_1, X_3, X_4, X_5)X_2$ modulo $X_1X_3 - X_2^2$. We shall show that $F(J) = 0$ implies

$$F_0(X) + F_1(X)X_2 = 0 .$$

We introduce a lexicographic ordering in the set of monomials in a, b, c, d. Then two distinct monomials which might appear in $F_0(X) + F_1(X)X_2$ have distinct highest terms. We recall that $J_2, J_4, J_6, J_8, J_{10}$ have a^2b^2, a^5b, a^8, a^9bd^2, a^7c as their highest terms. We shall show that two distinct monomials which might appear both in $F_0(X)$ or both in $F_1(X)$ have distinct highest terms. Let e_1, e_3, e_4, e_5 and e'_1, e'_3, e'_4, e'_5 be the exponents of X_1, X_3, X_4, X_5 in two monomials of the same degree which have the same highest term. Then, if we put $e_i - e'_i = n_i$ for $i = 1, 3, 4, 5$, we have

$$\begin{aligned} n_1 + 3n_3 + 4n_4 + 5n_5 &= 0 \\ 2n_1 + 8n_3 + 9n_4 + 7n_5 &= 0 \\ 2n_1 \qquad + \ n_4 \qquad &= 0 \\ n_5 &= 0 \\ 2n_4 \qquad &= 0 . \end{aligned}$$

This is a contradiction, because the system has only the trivial solution $(0, 0, 0, 0)$. We shall show also that two monomials one in $F_0(X)$ and another in $F_1(X)X_2$ have distinct highest terms. Otherwise, a similar system of linear equations in which the right hand sides are 2, 5, 1, 0, 0 has an integer solution. Again, this is a contradiction, because the only solution of the system is $(1/2, 1/2, 0, 0)$. This proves the statement. Therefore, if we have $F(J) = F_0(J) + F_1(J)J_4 = 0$, since there is no cancellation, we get $F_0(X) + F_1(X)X_2 = 0$. Suppose next that the characteristic of k is different from 2. Then $F(X)$ can be written uniquely in the form $F_0(X_1, X_2, X_3, X_5)$ modulo $X_1X_3 - X_2^2 - 4X_4$. Hence, we have only to show that $F_0(J) = 0$ implies $F_0(X) = 0$. Choose $\sigma_1, \sigma_2, \sigma_3$ so that $X(X-1)(X^3 - \sigma_1X^2 + \sigma_2X - \sigma_3)$ is projectively equivalent to the sextic $(1 + aX + bX^2)^2 - 4X^3(c + dX + X^2)$. Then, firstly the choice is possible and secondly $\sigma_1, \sigma_2, \sigma_3$ are independent variables over k. Moreover, the vanishing of $F_0(J)$ does not depend on whether we consider J_2, J_4, J_6, J_{10} as polynomials in a, b, c, d or as polynomials in $\sigma_1, \sigma_2, \sigma_3$ with coefficients in k. Now, we shall see later in this section that the values of $J_2^5J_{10}^{-1}$, $J_4^5J_{10}^{-1}$, $J_6^5J_{10}^{-1}$ at the sextic $X(X-1)(X^3-\sigma_1X^2+\sigma_2X-\sigma_3)$ are independent variables over k. If we grant this fact, certainly $F_0(J)=0$ will imply $F_0(X)=0$.

We shall assume, in the remaining lemmas that the characteristic is different from 2. Then, we can simply disregard J_8, i.e., we have only to consider J_2, J_4, J_6, J_{10}.

LEMMA 3. *A sextic has a triple root if and only if the basic arithmetic invariants take the following form*

$$J_2 = 2^2 3r^2\,, \qquad J_4 = 2\cdot 3r^4\,, \qquad J_6 = 2^2r^6\,, \qquad J_{10} = 0$$

with some r different from 0.

PROOF. Since the only-if part is straightforward, we shall prove the if-part. Since J_{10} vanishes, the sextic has a multiple root and at least one more root for, otherwise, all basic arithmetic invariants will vanish. We bring the multiple root to ∞ and another root to 0 by a suitable projective transformation. Then, the sextic takes the form

$$-v_1X^4 + v_2X^3 - v_3X^2 + v_4X\,.$$

This being done, we eliminate v_3 from the three equations $2^4J_2 = 2^2 3r^2$, $2^8J_4 = 2\cdot 3r^4$, $2^{12}J_6 = 2^2r^6$. Then, by a simple calculation, we get $2^6v_1^2v_2v_4 = 3(v_2^2-r^2)^2$, $2^9v_1^4v_4^2 = (v_2^2-r^2)^3$. We, then, eliminate v_1, v_4 from these two equations and we get

$$(v_2^2-r^2)^3(v_2^2-(3r)^2) = 0\,.$$

A remarkable fact is that in the process of elimination, we multiply powers of 2 but we need not cancel any integer at all. Indeed, since the characteristic is only assumed to be different from 2, no cancellation except a power of 2 is allowed. At any rate, if we have $v_2^2 = r^2$, we get $v_1v_3 = v_1v_4 = 0$, i.e., either $v_3 = v_4 = 0$ or $v_1 = 0$. Therefore, either 0 or ∞ is a triple root. On the other hand, if we have $v_2^2 = (3r)^2$, we get $v_1v_3 = 3r^2$ and $v_1^2v_4 = r^3$ or $-r^3$. Therefore, either rv_1^{-1} or $-rv_1^{-1}$ is a triple root. This completes the proof.

The following lemma can be proved in the same way:

LEMMA 4. *A sextic has a root of multiplicity at least equal to four if and only if the basic arithmetic invariants vanish simultaneously.*

As a consequence of this lemma, all arithmetic invariants vanish at a sextic of the form $X(X-1)(s_0X^3-s_1X^2+s_2X-s_3)$ if and only if, up to a constant factor, it coincides with $X^4(X-1)$, $X(X-1)^4$ or $X(X-1)$. Now, before we pass to the next lemma, we shall introduce some notations. Let s be a projective point with co-ordinates s_0, s_1, s_2, s_3. Then, we shall denote by $J_{2i}(s)$ the value of J_{2i} at the sextic

$$X(X-1)(s_0X^3 - s_1X^2 + s_2X - s_3)$$

for $i = 1, 2, 3, 5$. We note that each $J_{2i}(s)$ is a polynomial in s_0, s_1, s_2, s_3. In case s_0 is different from 0, we can introduce an affine point σ with co-ordinates $\sigma_i = s_is_0^{-1}$ for $i = 1, 2, 3$. We shall denote by $J_{2i}(\sigma)$ the value of J_{2i} at the sextic $X(X-1)(X^3 - \sigma_1X^2 + \sigma_2X - \sigma_3)$ for $i = 1, 2, 3, 5$. These notations will be used throughout this paper and, in particular, in the following important lemma:

LEMMA 5. *Let $\sigma_1, \sigma_2, \sigma_3$ be independent variables over a field k. Then, no matter how we extend the specialization $(\sigma_1, \sigma_2, \sigma_3) \to (0, 0, 0)$ with reference to k to a specialization including three quotients*

$$J_2(\sigma)^5J_{10}(\sigma)^{-1}, \qquad J_4(\sigma)^5J_{10}(\sigma)^{-2}, \qquad J_6(\sigma)^5J_{10}(\sigma)^{-3},$$

at least one of them will be specialized to ∞.

PROOF. Suppose that there exists an extension of the specialization under which all three quotients are specialized to finite quantities. Then, by a well-known argument (see e.g., [13, pp. 276-277]), we can find power-series expansions of $\sigma_1, \sigma_2, \sigma_3$ without constant terms which convert the three quotients into integral power-series. We shall show that this will bring a contradiction. We observe that the set of monomials in $\sigma_1, \sigma_2, \sigma_3$ forms a partially ordered multiplicative monoid. Therefore, we can talk about minimal elements in the set of monomials which appear in $J_{2i}(\sigma)$ with non-vanishing coefficients. We observe that, if we replace $\sigma_1, \sigma_2, \sigma_3$ by their power-series expansions, unless there is a cancellation, the lowest term of the power-series expansion for $J_{2i}(\sigma)$ comes from the minimal elements of $J_{2i}(\sigma)$. In this way, if $at^\alpha, bt^\beta, ct^\gamma$ are the lowest terms of $\sigma_1, \sigma_2, \sigma_3$ and if the characteristic of k is different from 3, we get the following power-series expansions:

$$\begin{aligned}
J_2(\sigma) &= 2^{-2}3a^2t^{2\alpha} - 2bt^\beta + 3ct^\gamma + \cdots, \\
J_4(\sigma) &= 2^{-3}[2^{-4}3a^4t^{4\alpha} - a^2bt^{2\alpha+\beta} + act^{\alpha+\gamma} + b^2t^{2\beta} \\
&\qquad - 2^2bct^{\beta+\gamma} + 3^2c^2t^{2\gamma}] + \cdots, \\
J_6(\sigma) &= 2^{-4}[2^{-6}a^6t^{6\alpha} - 2^{-3}a^4bt^{4\alpha+\beta} + 2^{-2}a^3ct^{3\alpha+\gamma} \\
&\qquad + 2^{-2}a^2b^2t^{2\alpha+2\beta} - 2^{-1}ab^3t^{\alpha+3\beta} - abct^{\alpha+\beta+\gamma} \\
&\qquad + 2^{-2}b^4t^{4\beta} + b^2ct^{2\beta+\gamma} + c^2t^{2\gamma}] + \cdots,
\end{aligned}$$

$$J_{10}(\sigma) = -2^{-12}[2^2a^3c^3t^{3\alpha+3\gamma} - a^2b^2c^2t^{2\alpha+2\beta+2\gamma} - 2\cdot 3^2abc^3t^{\alpha+\beta+3\gamma} + 2^2b^3c^2t^{3\beta+2\gamma} + 3^3c^4t^{4\gamma}] + \cdots .$$

The reader must keep in mind that the next smallest power of t in the above table might not be the lowest term in case the coefficients of the smallest power of t cancel each other. This being remarked, we separate four cases. Suppose first that 2α is smaller than β and γ. If 3α is greater than γ, the order of $J_4(\sigma)$ is $\alpha + \gamma$ and the order of $J_{10}(\sigma)$ is 4γ, hence the order of $J_4(\sigma)^5J_{10}(\sigma)^{-2}$ is $5\alpha - 3\gamma$, which is less than $-\alpha$. If 3α is at most equal to γ, the order of $J_{10}(\sigma)$ is at least equal to the minimum of $3\alpha+3\gamma$ and $2\alpha + 2\beta + 2\gamma$. Therefore, the order of $J_2(\sigma)^5J_{10}(\sigma)^{-1}$ is at most equal to the maximum of $7\alpha - 3\gamma$ and $8\alpha - 2\beta - 2\gamma$, which is -2α at most. Suppose next that β is smaller than 2α and γ. Then, the order of $J_{10}(\sigma)$ is at least equal to the minimum of $\alpha+\beta+3\gamma$, $3\beta + 2\gamma$ and 4γ, hence the order of $J_2(\sigma)^5J_{10}(\sigma)^{-1}$ is at most equal to the maximum of $-\alpha + 4\beta - 3\gamma$, $2\beta - 2\gamma$ and $5\beta - 4\gamma$, which is negative if 5β is smaller than 4γ. If 5β is at least equal to 4γ, the order of $J_6(\sigma)$ is 2γ and the order of $J_{10}(\sigma)$ is 4γ. Therefore, the order of $J_6(\sigma)^5J_{10}(\sigma)^{-3}$ is -2γ. Thirdly, suppose that γ is at most equal to 2α and β. Then, we have the same situation, i.e., the order of $J_6(\sigma)$ is 2γ and the order of $J_{10}(\sigma)$ is 4γ. Finally, suppose that 2α and β are equal and that they are smaller than γ. If 3α is greater than γ, we see again that the order of $J_6(\sigma)$ is 2γ and the order of $J_{10}(\sigma)$ is 4γ.. If 3α is at most equal to γ, the order of $J_{10}(\sigma)$ is at least equal to $6\alpha + 2\gamma$. On the other hand, since we have

$$J_2(\sigma) = 2^{-2}(3a^2 - 2^3b)t^{2\alpha} + \cdots ,$$

in case $3a^2$ is different from 2^3b, the order of $J_2(\sigma)^5J_{10}(\sigma)^{-1}$ is at most equal to $4\alpha - 2\gamma$, which is -2α at most. In case $3a^2$ is equal to 2^3b, we consider two subcases. If 3α is smaller than γ, since we have

$$J_4(\sigma) = -2^{-9}3a^4t^{4\alpha} + \cdots ,$$

the order of $J_4(\sigma)^5J_{10}(\sigma)^{-2}$ is at most equal to $8\alpha - 4\gamma$, which is less than -4α. If 3α is equal to γ, since we have

$$J_4(\sigma) = -2^{-9}a(3a^3 - 2^6c)t^{4\alpha} + \cdots ,$$

in case $3a^3$ is different from 2^6c, the order of $J_4(\sigma)^5J_{10}(\sigma)^{-2}$ is again less than -4α. In case $3a^3$ is equal to 2^6c, however, we fortunately have

$$J_6(\sigma) = 2^{-16}a^6t^{6\alpha} + \cdots .$$

Therefore, the order of $J_6(\sigma)^5J_{10}(\sigma)^{-3}$ is -6α at most. Hence, in all possible cases, at least one of the three quotients has a negative order. This is a contradiction. We have been assuming so far that the characteristic of k is different from 3. If the characteristic of k is 3, we have only to know the orders of minimal elements in $J_2(\sigma)$, $J_4(\sigma)$, $J_6(\sigma)$, $J_{10}(\sigma)$ and they are as follows:

$$(\alpha+\gamma, \beta),$$
$$(2\alpha+\beta, \alpha+\gamma, 2\beta, \beta+\gamma),$$
$$(6\alpha, 4\alpha+\beta, 3\alpha+\gamma, 2\alpha+2\beta, \alpha+3\beta, \alpha+\beta+\gamma, 4\beta, 2\beta+\gamma, 2\gamma),$$
$$(3\alpha+3\gamma, 2\alpha+2\beta+2\gamma, 3\beta+2\gamma).$$

Since $J_6(\sigma)^5J_{10}(\sigma)^{-3}$ has a non-negative order, we see that the minimum of α and β has to be smaller than γ. Moreover, since $J_4(\sigma)^5J_{10}(\sigma)^{-2}$ also has a non-negative order, we see that α has to be smaller than β. Hence α is smaller than β and γ. This being remarked, we separate two cases. Suppose first that β is at least equal to γ. Then, the order of $J_4(\sigma)$ is $\alpha+\gamma$ and the order of $J_{10}(\sigma)$ is at least equal to the minimum of $3\alpha+3\gamma$ and $3\beta+2\gamma$. Therefore, the order of $J_4(\sigma)^5J_{10}(\sigma)^{-2}$ is at most equal to the maximum of $-\alpha-\gamma$ and $5\alpha-6\beta+\gamma$, which is negative. Suppose next that β is smaller than γ. Then, the order of $J_2(\sigma)$ is β, hence the order of $J_2(\sigma)^5J_{10}(\sigma)^{-1}$ is at most equal to the maximum of $-3\alpha+5\beta-3\gamma$, $-2\alpha+3\beta-2\gamma$ and $2\beta-2\gamma$, and this is negative if 5β is smaller than $3\alpha+3\gamma$. On the other hand, if 5β is at least equal to $3\alpha+3\gamma$, we observe that the order of $J_{10}(\sigma)$ is $3\alpha+3\gamma$. Moreover, in case $\alpha+\beta$ is different from γ, the order of $J_4(\sigma)$ is equal to the minimum of $2\alpha+\beta$ and $\alpha+\gamma$. Therefore, the order of $J_4(\sigma)^5J_{10}(\sigma)^{-2}$ is $-\alpha-\gamma$ at most. In case $\alpha+\beta$ is equal to γ, the order of $J_6(\sigma)$ is equal to 6α. Therefore, the order of $J_6(\sigma)^5J_{10}(\sigma)^{-3}$ is equal to $12\alpha-9\beta$, which is -15α at most. Again, in all possible cases, at least one of the three quotients has a negative order. This is a contradiction. The lemma is thereby proved.

We shall discuss some consequences of this lemma. First of all, we shall show that $\sigma_1, \sigma_2, \sigma_3$ are integral over the ring generated over k by the three quotients in the lemma. We have only to show that $\sigma_1, \sigma_2, \sigma_3$ are finite over every finite specialization of the ring in question with reference to k. Suppose, therefore, that there exists a finite specialization of the ring with reference to k which extends to a non-finite specialization of $\sigma_1, \sigma_2, \sigma_3$. If we introduce a projective point s whose co-ordinates s_0, s_1, s_2, s_3 are proportional to $1, \sigma_1, \sigma_2, \sigma_3$ and if s' with co-ordinates s'_0, s'_1, s'_2, s'_3 is a specialization of s over this specialization, we have $s'_0=0$, hence $J_{10}(s')=0$. Since the three quotients $J_{2i}(\sigma)^5J_{10}(\sigma)^{-i}=J_{2i}(s)^5J_{10}(s)^{-i}$ for $i=1, 2, 3$ are specialized to finite quantities, we also get $J_2(s')=J_4(s')=J_6(s')=0$. Then, by a remark following Lemma 4, we get $s'_1=s'_2=0$.

On the other hand, the sextic $X(X-1)(s_0X^3 - s_1X^2 + s_2X - s_3)$ is projectively equivalent to $X(X-1)(s_3X^3 - s_2X^2 + s_1X - s_0)$. Therefore $J_{2i}^5J_{10}^{-i}$ for $i = 1, 2, 3$ take the same values at these two sextics. Moreover, if we define τ_1, τ_2, τ_3 by $s_2s_3^{-1}, s_1s_3^{-1}, s_0s_3^{-1}$, they are independent variables over k and we get an extension of the specialization $(\tau_1, \tau_2, \tau_3) \to (0, 0, 0)$ with reference to k under which $J_{2i}(\sigma)^5J_{10}(\sigma)^{-i} = J_{2i}(\tau)^5J_{10}(\tau)^{-i}$ for $i = 1$, 2, 3 are specialized to finite quantities. This contradicts Lemma 5. The statement about the integrality of $\sigma_1, \sigma_2, \sigma_3$ is thereby proved.

One consequence of this observation is that the field $k(\sigma)$ is a finite algebraic extension of the field generated over k by

$$J_2(\sigma)^5J_{10}(\sigma)^{-1}, \quad J_4(\sigma)^5J_{10}(\sigma)^{-2}, \quad J_6(\sigma)^5J_{10}(\sigma)^{-3}.$$

Therefore, these three quotients must be independent variables over k.

5. Integral dependence theorem

We shall denote by $\mathbf{Z}[J]$ the graded subring of the integral domain of rational invariants generated over $\mathbf{Z}$ by the basic arithmetic invariants J_{2i} for $i = 1, 2, \cdots, 5$. Since the powers of J_{10} are different from zero and form a multiplicatively stable subset of $\mathbf{Z}[J]$, we can construct a ring of fractions of $\mathbf{Z}[J]$ with respect to this set and, in this way, we get a new graded ring with elements of negative degrees. The set of homogeneous elements of degree zero in this graded ring forms its subring, which we shall denote by $\mathfrak{R}$. We observe that $\mathfrak{R}$ is an integral domain and it is finitely generated over $\mathbf{Z}$. We shall find later a minimal set of generators of $\mathfrak{R}$. Let k be a field of arbitrary characteristic and consider the tensor product $\mathfrak{R} \otimes k$ taken over $\mathbf{Z}$. Let a, b, c, d be independent variables over k and consider the sextic $(1+aX+bX^2)^2-4X^3(c+dX+X^2)$. Then, the evaluation of the basic arithmetic invariants at this sextic will give rise to a k-homomorphism of $\mathfrak{R} \otimes k$ in the field $k(a, b, c, d)$. Lemma 2 shows that this homomorphism is a monomorphism. If the characteristic of k is different from 2, we have the same situation, for instance, for the sextic $X(X-1)(X^3 - \sigma_1X^2 + \sigma_2X - \sigma_3)$ in which $\sigma_1, \sigma_2, \sigma_3$ are independent variables over k. At any rate, we see that $\mathfrak{R} \otimes k$ is also an integral domain.

PROPOSITION 1. *The ring $\mathfrak{R} \otimes k$ is normal,* i.e., *it is integrally closed in its field of fractions.*

PROOF. Let x_1, x_2, x_3, x_5 be independent variables over k and let x_2 be a root of $X^2 - x_1x_3 + 4x_4 = 0$. We shall show that $k[x]$ is normal. Suppose first that the characteristic of k is different from 2. Then, we have $k[x] = k[x_1, x_2, x_3, x_5]$ and x_1, x_2, x_3, x_5 are independent variables over k. Hence $k[x]$ is normal. Suppose next that the characteristic of k is 2. Then x_4, x_5 are independent variables over $k(x_1, x_2, x_3)$. Therefore, we have only to show that $k[x_1, x_2, x_3]$ is normal. This is a consequence of a well-known general criterion in algebraic geometry. However, we can proceed also as follows. Let t_1, t_2 be independent variables over k. Then, it is the same thing to say that $k[t_1^2, t_1t_2, t_2^2]$ is normal. Let D be a derivation of $k[t_1, t_2]$ which is trivial on k and such that $Dt_i = t_i$ for $i = 1, 2$. Then $k[t_1^2, t_1t_2, t_2^2]$ is precisely the subring of $k[t_1, t_2]$ on which D is trivial. Therefore $k[t_1^2, t_1t_2, t_2^2]$ is normal. Once we know that $k[x]$ is normal, we can derive the normality of $\mathfrak{R} \otimes k$ in the following way. Lemma 2 implies that $\mathfrak{R} \otimes k$ is k-isomorphic to the subring of $k(x)$ generated over k by elements of the form $x_1^{e_1}x_2^{e_2}x_3^{e_3}x_4^{e_4}x_5^{-e_5}$ in which e_i are non-negative integers satisfying $e_1 + 2e_2 + 3e_3 + 4e_4 = 5e_5$. Therefore, if an element θ of the field of fractions of $\mathfrak{R} \otimes k$ is integral over $\mathfrak{R} \otimes k$, we can find a positive integer e such that $x_5^e \cdot \theta$ is integral over $k[x]$. Since $k[x]$ is normal, this implies that $x_5^e \cdot \theta$ is contained in $k[x]$, hence θ is contained in $\mathfrak{R} \otimes k$.

We can show in the same way that the ring $\mathfrak{R}$ is normal. Actually, we can derive the normality of $\mathfrak{R}$ as a corollary of Proposition 1. Now, we have shown in Section 3 that our basic arithmetic invariants work well in characteristic 2. We shall show that they also work in characteristic different from 2. Let k be, therefore, a field of characteristic different from 2 and let σ_1, σ_2, σ_3 be independent variables over k. Then, the evaluation of J_{2i} at the sextic $X(X-1)(X^3 - \sigma_1X^2 + \sigma_2X - \sigma_3)$, i.e., the correspondence $J_{2i} \rightarrow J_{2i}(\sigma)$, converts $\mathfrak{R} \otimes k$ into a subring of the field $k(\sigma_1, \sigma_2, \sigma_3)$. We shall denote by s the projective point whose co-ordinates s_0, s_1, s_2, s_3 are proportional to 1, σ_1, σ_2, σ_3 and by σ the affine point with co-ordinates σ_1, σ_2, σ_3. Then, by a standard agreement, we have $k(s) = k(\sigma)$. We know that there exist precisely 120 sextics of the form $X(X-1)(X^3 - \sigma_1X^2 + \sigma_2X - \sigma_3)$ which are projectively equivalent to $X(X-1)(X^3 - \sigma_1X^2 + \sigma_2X - \sigma_3)$. Consequently, we get 120 points like s. Let $c(s)$ be the corresponding "Chow point." Then, we have the following situation:

PROPOSITION 2. (Integral dependence theorem). *The field of fractions of $\mathfrak{R} \otimes k$ is generated over k by*

$$J_4(s)J_2(s)^{-2}, \qquad J_6(s)J_2(s)^{-3}, \qquad J_{10}(s)J_2(s)^{-5}$$

and it coincides with $k(c(s))$. Moreover, the point σ is integral over $\mathfrak{R} \otimes k$.

PROOF. Let λ_1, λ_2, λ_3 be the roots of the equation $X^3 - \sigma_1X^2 + \sigma_2X - \sigma_3 = 0$.

Then $k(\lambda) = k(\lambda_1, \lambda_2, \lambda_3)$ contains $k(s)$ and also $k(c(s))$. Moreover, it is clear that $k(\lambda)$ is a Galois extension of $k(c(s))$ with the symmetric group of permutations of six letters as the corresponding Galois group. Therefore $k(s)$ is a separable extension of $k(c(s))$ of degree 120. This being remarked, we put $t_1 = J_4(s)J_2(s)^{-2}$, $t_2 = J_6(s)J_2(s)^{-3}$, $t_3 = J_{10}(s)J_2(s)^{-5}$. The invariance property of J_{2i} implies, then, that they generate a subfield $k(t)$ of $k(c(s))$. Moreover, it is clear that $k(t)$ coincides with the field of fractions of $\mathfrak{R} \otimes k$. We shall show that $k(t)$ also coincides with $k(c(s))$. We first observe that t_1, t_2, t_3 are independent variables over k by Lemma 2. Therefore, the extension degree, say N, of $k(s)$ over $k(t)$ is a positive integer and we have $N = 120 \cdot [k(c(s)) : k(t)]$. Let S_0, S_1, S_2, S_3 be letters to write equations and consider three surfaces of degree 4, 6, 10 defined respectively by the following equations

$$J_4(S) - t_1J_2(S)^2 = 0\ , \quad J_6(S) - t_2J_2(S)^3 = 0\ , \quad J_{10}(S) - t_3J_2(S)^5 = 0\ .$$

These equations have their coefficients in $k(t)$ and certainly s and its conjugates over $k(t)$ are in their intersection. Let s' be an arbitrary point of the intersection. If we have $J_2(s')=0$, we get $J_4(s')=J_6(s')=J_{10}(s')=0$. Hence, by the remark following Lemma 4 we see that s' must be one of the three points (1, 0, 0, 0), (1, 3, 3, 1) and (0, 0, 0, 1). On the other hand, if $J_2(s')$ is different from 0, clearly $k(s')$ contains $k(t)$. Therefore s' is a generic specialization of s over k and, in fact, over $k(t)$. Hence s' is a conjugate of s over $k(t)$. Therefore, the intersection-product of the three surfaces is, possibly a multiple of, the prime rational cycle over $k(t)$ determined by s plus a linear combination of the three points (1, 0, 0, 0), (1, 3, 3, 1), (0, 0, 0, 1) with positive integer coefficients. If we apply the Bezout theorem (cf. [12-II, Section 89 or 13]) to this situation, we get $N + 3 \leqq 4 \cdot 6 \cdot 10 = 240$. Since we also know that N is a multiple of 120, the only possibility is $N = 120$, hence $k(t) = k(c(s))$. This proves the first part. As for the second part, it is a consequence of an observation made after Lemma 5.

We shall discuss some consequences of this proposition. Let Σ be the complete set of conjugates of σ over the field $k(c(s))$ and let Σ' be a specialization of Σ over an arbitrary finite specialization, say φ, of $\mathfrak{R} \otimes k$ with reference to k. We know by Proposition 2 that all points in Σ' are finite. We also know, because of the normality of $\mathfrak{R} \otimes k$, that the set Σ' is uniquely determined by φ (cf. [15 or 13]). We shall show that J_{10} can not vanish at any one of the corresponding sextics. Suppose that J_{10} vanishes at a sextic which corresponds to a point σ' in Σ'. Then J_2, J_4, J_6 must also vanish at the sextic, because φ is finite. It then follows from Lemma 4 and Lemma 5 that the co-ordinates of σ' are 3, 3, 1. On the other hand, the sextic $X(X-1)(X^3-\sigma_1X^2+\sigma_2X-\sigma_3)$ is projectively equivalent to $X(X-1)\big((1-X)^3-\sigma_1(1-X)^2+\sigma_2(1-X)-\sigma_3\big)$. Therefore, the point with co-ordinates $3-\sigma_1$, $3-2\sigma_1+\sigma_2$, $1-\sigma_1+\sigma_2-\sigma_3$ is a member of Σ, hence the point with co-ordinates 0, 0, 0 must be contained in Σ'. This contradicts Lemma 5. Therefore, the 120 sextics which correspond to points in Σ' are all free from multiple roots. Then, these sextics, which are not necessarily distinct, are projectively equivalent to each other on account of the following general lemma:

LEMMA 6. *Let* $\mathfrak{a}$ *and* $\mathfrak{b}$ *be two positive divisors of degree* d *on a projective straight-line and let* $(\mathfrak{a}', \mathfrak{b}')$ *be a specialization of* $(\mathfrak{a}, \mathfrak{b})$; *let* m *and* n *be two positive integers satisfying* $m+n=d-1$. *Then, if* $\mathfrak{a}'$ *and* $\mathfrak{b}'$ *have no components of multiplicity greater than* m *and* n *respectively, the projective equivalence of* $\mathfrak{a}$ *and* $\mathfrak{b}$ *implies the projective equivalence of* $\mathfrak{a}'$ *and* $\mathfrak{b}'$.

PROOF. Let D and D' be the projective straight-lines carrying $\mathfrak{a}$, $\mathfrak{b}$ and $\mathfrak{a}'$, $\mathfrak{b}'$. Also, let T be the graph in the product $D \times D$ of the projective transformation sending $\mathfrak{a}$ to $\mathfrak{b}$ and let T' be its specialization over the given specialization. Then T' is a positive divisor of $D' \times D'$ with D' as its algebraic projection to each factor. Hence, if T' is irreducible, it is certainly a graph of a projective transformation and it sends $\mathfrak{a}'$ to $\mathfrak{b}'$. Therefore, we have only to show that T' is irreducible. If T' is reducible, there exist two points a' and b' on D' such that $T' = a' \times D' + D' \times b'$. Since $\mathfrak{b}'$ has no component of multiplicity greater than n, it contains $m+1$ points b'_j for $j=1, 2, \cdots, m+1$, which are in general not distinct but certainly distinct from b'. Consequently T' and $D' \times \mathfrak{b}'$ intersect properly at $a' \times b'_j$. If we apply the so-called "principle of conservation of number" to $T, D \times \mathfrak{b}$ and $T', D' \times \mathfrak{b}'$, we see that a' is a component of $\mathfrak{a}'$ of multiplicity at least equal to $m+1$. This contradicts the assumption. Hence T' can not be reducible. The lemma is thereby proved.

As a consequence of the previous consideration, we can state the following corollary:

COROLLARY. *Two sextics without multiple roots are projectively equivalent if and only if the corresponding basic arithmetic invariants differ only by a transformation of the form* $J_{2i} \rightarrow r^{2i}J_{2i}$ *for* $i = 1, 2, 3, 5$ *with some* r *different from* 0.

The reader must observe that our integral dependence theorem is not by any means a consequence of this corollary. At any rate, we are now sure that our basic arithmetic invariants work perfectly well also in characteristic different from 2.

6. Absolute invariants and the arithmetic variety of moduli

The concept of absolute invariants was already explained in the Introduction: An *absolute invariant* is a characteristic-preserving continuous function on the set C of hyperelliptic curves with values in the arithmetic universal domain taking same value on birationally equivalent hyperelliptic curves. The totality of absolute invariants forms a ring in an obvious way, which we call the ring of absolute invariants. We observe that the ring of absolute invariants is a natural, intrinsic object. Therefore, the following proposition will show that our ring $\mathfrak{R}$ is also intrinsic:

PROPOSITION 3. *The ring of absolute invariants can be identified with $\mathfrak{R}$ in a natural way.*

In the course of the proof of this proposition, we shall use a technical lemma which we proved elsewhere in a slightly different form [10]. Since we shall also need this lemma later, we restate it for the convenience of the reader:

LEMMA 7. *Let C_0 and C_0' be absolutely irreducible curves of the same genus and suppose that C_0' is a specialization of C_0. Then, we can construct a non-singular model C of C_0 and an extension of the specialization so that C specializes to a non-singular model C' of C_0' under this extended specialization.*

Now, we start proving Proposition 3. We shall first define a monomorphism from the ring of absolute invariants into $\mathfrak{R}$. Let f be, therefore, an arbitrary absolute invariant. Let s with co-ordinates s_0, s_1, s_2, s_3 be a generic point of a projective space over $\mathbf{Q}$ and consider a plane curve C_0 defined inhomogeneously by $Y^2 = X(X-1)(s_0X^3 - s_1X^2 + s_2X - s_3)$. Then, the normalization C of C_0 over $\mathbf{Q}(s)$ is a hyperelliptic curve. We recall that C does not determine s uniquely. In fact, we know that s has 120 possibilities and the corresponding Chow point $c(s)$ is uniquely determined by C. Let $f(C)'$ be an arbitrary specialization of $f(C)$ with reference to $\mathbf{Q}(c(s))$ and let s' be a specialization of s over this specialization. Then s' is just one of the conjugates of s over $\mathbf{Q}(c(s))$, hence C is specialized to the corresponding conjugate, say C', over the above specialization. Moreover C' is also a hyperelliptic curve and it is birationally equivalent to C. This implies $f(C)' = f(C') = f(C)$, hence $f(C)$ has no other specialization than itself over $\mathbf{Q}(c(s))$. Therefore $f(C)$ must be an element of $\mathbf{Q}(c(s))$, which coincides with the field of fractions of $\mathfrak{R} \otimes \mathbf{Q}$ by Proposition 2. We shall show that $f(C)$ is an element of $\mathfrak{R} \otimes \mathbf{Q}$. Let $f(C)'$ be a specialization of $f(C)$ over an arbitrary finite specialization of $\mathfrak{R} \otimes \mathbf{Q}$ with reference to $\mathbf{Q}$ and let s' be a specialization of s over this specialization.

Then, if s'_0, s'_1, s'_2, s'_3 are the co-ordinates of s', an observation following Proposition 2 shows that the equation $Y^2 = X(X-1)(s'_0X^3 - s'_1X^2 + s'_2X - s'_3)$ defines a plane curve C'_0 of genus two. Therefore, we can apply Lemma 7 to get a non-singular model C_1 of C_0 which specializes to a non-singular model C'_1 of C'_0 over the so-extended specialization. This implies $f(C)' = f(C'_1)$, hence $f(C)'$ is a finite quantity. We have shown, therefore, that $f(C)$ is finite over every finite specialization of $\mathfrak{R} \otimes \mathbf{Q}$ with reference to $\mathbf{Q}$. Since $\mathfrak{R} \otimes \mathbf{Q}$ is normal by Proposition 1, we see that $f(C)$ is contained in $\mathfrak{R} \otimes \mathbf{Q}$. This implies that $f(C)$ can be written in the form $(F_0(J(s)) + F_1(J(s))J_4(s))J_{10}(s)^{-e}$ in which $F_0(J)$, $F_1(J)$ are polynomials in J_2, J_6, J_8, J_{10} with coefficients in $\mathbf{Q}$ and e is a positive integer. We shall show that the coefficients are actually contained in $\mathbf{Z}$, i.e., that the least common multiple, say n, of the denominators of the coefficients is 1. Otherwise, let p be a prime factor of n and let a', b', c', d' be independent variables over GF(p). Let C'_0 be the plane curve defined inhomogeneously by $XY^2 + (1 + a'X + b'X^2)Y + X^2(c' + d'X + X^2) = 0$; let a, b, c, d be independent variables over $\mathbf{Q}$ and let C_0 be the plane curve defined similarly. Then C'_0 is the unique specialization of C_0 over the specialization $(a, b, c, d) \rightarrow (a', b', c', d')$ and by Lemma 7 we can find a compatible specialization $C_1 \rightarrow C'_1$ of their non-singular models. We can assume that C_1 is birationally equivalent to C. Then $f(C_1) = f(C)$ will specialize uniquely to the finite quantity $f(C'_1)$ over this specialization. Therefore, if we consider J_{2i} to be evaluated at the sextic $(1 + aX + bX^2)^2 - 4X^3(c + dX + X^2)$, the specialization of $nJ_{10}^e \cdot f(C) = nF_0(J) + nF_1(J)J_4$ over the above specialization will produce a non-trivial relation of the form $G_0(J) + G_1(J)J_4 = 0$ over GF(p). This contradicts Lemma 2, hence we have $n = 1$. Therefore, we can identify f with the element $(F_0(J) + F_1(J)J_4)J_{10}^{-e}$ of $\mathfrak{R}$. It is clear that this identification is compatible with ring structures. We shall show that the monomorphism so obtained is an epimorphism.

Let f be an element of $\mathfrak{R}$ and let C be an arbitrary hyperelliptic curve. Let a, b, c, d be the constants which appear in the universal normal form of C. If we express J_{2i} as polynomials in a, b, c, d, we get a quantity $f(C)$ in the universal domain for C. Moreover, this $f(C)$ depends only on the birational class of C. In fact, if the characteristic is different from 2, it is a consequence of the invariance property of J_{2i}. If the characteristic is 2, it is a consequence of the explicit relations we calculated at the end of Section 3. For instance, if C has three Weierstrass points, i.e., if C is of type (1, 1, 1), the relations show that $J_4J_2^{-2}$, $J_8J_2^{-4}$, $J_{10}J_2^{-5}$, hence also $J_6J_2^{-3} = (J_4J_2^{-2})^2$, depend only on the birational class of C. Therefore $f(C)$ depends only on the birational class of C. The proof is similar for the other two cases. In order to show that f is an absolute invariant, therefore, we have only to prove its continuity. Suppose that a hyperelliptic curve C' is a specialization of a hyperelliptic curve C. Call this specialization φ. Let $f(C)'$ be a specialization of $f(C)$ over φ. Let P be a Weierstrass point of C and let P' be a specialization of P over this specialization. Then P' is a Weierstrass point of C', because of the upper semi-continuity $\dim |2P'| \geqq \dim |2P| = 1$. Let Q be a point of C such that its specialization Q' over the above specialization is not a Weierstrass point of C'. We can take as Q a generic point of C over its field of definition. Call the so-extended specialization φ^*. Then the complete linear system $|3P' + Q'|$ on C' is the unique specialization of the complete linear system $|3P + Q|$ on C over φ^*. In this case, we can choose compatible regular birational transformations of C and C' to plane quartics which correspond to $|3P+Q|$ and $|3P'+Q'|$. This can be proved, *for instance*, by representing these complete linear systems as residual systems of complete linear systems determined by hypersurface sections (cf. [9-I, Section 2]). Therefore, if

$$XY^2 + (1 + aX + bX^2)Y + X^2(c + dX + X^2) = 0$$

and

$$XY^2 + (1 + a'X + b'X^2)Y + X^2(c' + d'X + X^2) = 0$$

are the normalized equations for the image quartics and if we replace (a', b', c', d') by $(\zeta a', \zeta^2 b', \xi^3 c', \zeta^4 d')$ for some ζ satisfying $\zeta^5 = 1$, then (a', b', c', d') so modified will become a specialization of (a, b, c, d) over φ^*. This implies that $f(C')$ is a specialization of $f(C)$ over φ^*, hence $f(C)' = f(C')$. In other words $f(C')$ is the unique specialization of $f(C)$ over the specialization φ. This completes the proof.

Suppose that k is an arbitrary field of characteristic p. Then, we can consider the ring of continuous functions on the set of hyperelliptic curves in characteristic p with values in the universal domain over k depending only on birational classes of hyperelliptic curves. It is clear that $\mathfrak{R} \otimes k$ can be considered as its subring. A similar consideration as in the proof of Proposition 3 shows, moreover, that this ring consists of purely inseparable elements over $\mathfrak{R} \otimes k$. Therefore, it coincides with $\mathfrak{R} \otimes k$ only in the case $p = 0$.

Now, we consider the set $\mathfrak{M}$ of homomorphisms of $\mathfrak{R}$ in universal domains of various characteristics. We then introduce the "Zariski topology" in the set $\mathfrak{M}$. The topological space $\mathfrak{M}$ so obtained will be called an *arithmetic variety of moduli*. We observe that specializations of points of $\mathfrak{M}$ can be defined in an obvious way. In the same way $\mathfrak{R} \otimes k$ defines a variety which we call a *variety of moduli* with reference to k. We shall denote this variety by $\mathfrak{M} \otimes k$. The following theorem is a consequence of Proposition 1 and a remark after that:

THEOREM 1. *The arithmetic variety of moduli $\mathfrak{M}$ and the variety of moduli $\mathfrak{M} \otimes k$ are normal.*

The following theorem will justify our terminology:

THEOREM 2. *There exists a natural mapping j from the set $\mathcal{C}$ onto the variety $\mathfrak{M}$ with the following three properties*:

(i) *We have $j(C_1) = j(C_2)$ if and only if C_1 and C_2 are birationally equivalent*;

(ii) *j is continuous*, i.e., *if $C' \rightarrow C''$ is a specialization in $\mathcal{C}$, then $j(C'')$ is the unique specialization of $j(C')$ over this specialization*;

(iii) *If $j' \rightarrow j''$ is a specialization of points of $\mathfrak{M}$, we can find a compatible specialization $C' \rightarrow C''$ in $\mathcal{C}$ satisfying $j(C') = j'$ and $j(C'') = j''$.*

PROOF. We must first define the mapping j. Pick an element C of $\mathcal{C}$ and an element f of $\mathfrak{R}$. Then, the correspondence $f \rightarrow f(C)$ is a homomorphism of $\mathfrak{R}$ in the universal domain for C. Denote it by $j(C)$:

$$j(C)f = f(C).$$

Then j is certainly a naturally defined mapping of $\mathcal{C}$ in $\mathfrak{M}$. Moreover, the if-part of (i) and the property (ii) can be verified trivially. We shall show that j is a surjection and also the only if-part of (i). Let j' be an arbitrary point of $\mathfrak{M}$. Suppose that j' is a homomorphism of $\mathfrak{R}$ in the universal domain of characteristic 2. Then, according as j' does not annihilate $J_2^5 J_{10}^{-1}$ or annihilate $J_2^5 J_{10}^{-1}$ but not $J_6^5 J_{10}^{-3}$ or annihilate $J_2^5 J_{10}^{-1}$ and $J_6^5 J_{10}^{-3}$, we can determine α, β, γ or α, β or α by the formulas at the end of Section 3. In this way, we get a hyperelliptic curve C' satisfying $j(C') = j'$. Moreover, the same formulas show that the birational class of C' is uniquely determined by j'. Suppose next that j' is a homomorphism of $\mathfrak{R}$ in a universal domain of characteristic different from 2. If k is the corresponding prime field, we get a finite specialization of $\mathfrak{R} \otimes k$ with reference to k. Then, it follows from an observation following Proposition 2 that there exists a hyperelliptic curve C' satisfying $j(C') = j'$. Moreover, the same observation shows that the birational class of C' is uniquely determined by j'. Therefore j is a surjection and it has the property (i). We shall show, finally, that j has the property (iii). Consider the set of quadruples (a, b, c, d) which appear as coefficients in the universal normal forms of hyperelliptic curves. This set, say U, is precisely an affine space over $\mathbf{Z}$ of dimension four minus the hypersurface defined by $J_{10} = 0$. Let (a, b, c, d) be a point of U and let $\pi(a, b, c, d)$ be the image under j of the corresponding hyperelliptic curve. Then π is an everywhere regular rational mapping of U onto $\mathfrak{M}$. Moreover, the set-theoretic inverse image under π of every point of $\mathfrak{M}$ is of dimension one and, moreover, every point in the inverse image is contained in a component of dimension one. This follows from the fact that, if C is an arbitrary hyperelliptic curve, the inverse image of $j(C)$ under π corresponds in a very simply way to the set of pairs of Weierstrass and non-Weierstrass points of C. Therefore, the inverse image is a relatively closed subset of U of which every component is of dimension one. This being said, let j^* be a generic point of $\mathfrak{M}$ and define a positive U-cycle whose support coincides with the inverse image of j^* and which is rational over $\mathbf{Q}(j^*)$. Since j^* is certainly a simple point of $\mathfrak{M} \otimes \mathbf{Q}$, we can get such a cycle by using an intersection-theory in characteristic zero. At any rate, we denote this cycle by $\pi^{-1}(j^*)$. Let j' and j'' be as in (iii). Then, we define $\pi^{-1}(j')$ as a specialization of $\pi^{-1}(j^*)$ over the specialization $j^* \to j'$. More precisely, we embed U into a projective space over $\mathbf{Z}$ of dimension four and, at the same time, we complete every component of $\pi^{-1}(j^*)$. Then, we specialize the completion of $\pi^{-1}(j^*)$ over the specialization $j^* \to j'$. Finally, we restrict the specialized cycle to U to get $\pi^{-1}(j')$. This description shows that, *a priori*, there is no guarantee as to the fact that $\pi^{-1}(j')$ is unique. However, since j' is a normal point of $\mathfrak{M}$, we can apply the so-called "principle of connectedness" (cf. [2, 6]) to see that the set of all specializations of the completion of $\pi^{-1}(j^*)$ over the specialization $j^* \to j'$ is "connected." This implies that $\pi^{-1}(j')$ is unique. Moreover, since the support of $\pi^{-1}(j^*)$ contains a generic point of U which can, therefore, be specialized to every point of U, every component of the inverse image of j' appears with a positive multiplicity in $\pi^{-1}(j')$. In other words, the support of $\pi^{-1}(j')$ is precisely the inverse image of j'. We also have the same situation at the point j''. Therefore $\pi^{-1}(j'')$ is the unique specialization of $\pi^{-1}(j')$ over the specialization $j' \to j''$. In particular, every point of the inverse image of j'' is a specialization of some point in the inverse image of j'. If we apply Lemma 7 to this situation, we can find a compatible specialization $C' \to C''$ in $\mathcal{C}$ satisfying $j(C') = j'$ and $j(C'') = j''$. This completes the proof.

On account of this theorem, the "distribution of birational classes of hyperelliptic curves" is completely determined by the geometric structure of the variety $\mathfrak{M}$ and it simply reflects the algebraic structure of the ring $\mathfrak{R}$.

7. Structure theorems

We shall first explain the structure of the ring $\mathfrak{R}$. Let y_1, y_2, y_3 be independent variables over $\mathbf{Q}$ and define y_4 as follows:

$$y_4 = 1/4\cdot(y_1y_3 - y_2^2) .$$

Then, Lemma 2 shows that the correspondence

$$J_2^{e_1}J_4^{e_2}J_6^{e_3}J_8^{e_4}J_{10}^{-e_5} \longrightarrow y_1^{e_1}y_2^{e_2}y_3^{e_3}y_4^{e_4} ,$$

in which e_i are non-negative integers satisfying $e_1 + 2e_2 + 3e_3 + 4e_4 = 5e_5$, gives rise to a monomorphism of the ring $\mathfrak{R}$ in $\mathbf{Z}[y] = \mathbf{Z}[y_1, y_2, y_3, y_4]$. We observe that, *if ζ is a primitive fifth root of unity, the image ring is,* i.e., *the ring $\mathfrak{R}$ can be considered as, the ring of invariant elements of* $\mathbf{Z}[y]$ *under the transformation $y_i \to \zeta^i y_i$ for* $i = 1, 2, 3, 4$. Therefore, if k is a field of characteristic different from 2 and if y_1, y_2, y_3 are independent variables over k, then $\mathfrak{R} \otimes k$ can be considered as the ring of invariant elements of $k[y_1, y_2, y_3]$ under the transformation $y_i \to \zeta^i y_i$ for $i = 1, 2, 3$. In case the characteristic happens to be 5, we must use the corresponding infinitesimal transformation, i.e., a derivation of $k[y_1, y_2, y_3]$ over k defined by $Dy_i = iy_i$ for $i = 1, 2, 3$. If k is a field of characteristic 2 and if t_1, t_2, t_3 are independent variables over k, then $\mathfrak{R} \otimes k$ can be considered as the ring of invariant elements of $k[t_1, t_2, t_3]$ under the transformation $t_1 \to \zeta^3 t_1$, $t_2 \to \zeta^4 t_2$, $t_3 \to \zeta^4 t_3$ as well as under the infinitesimal transformation $Dt_1 = t_1$, $Dt_2 = t_2$, $Dt_3 = 0$. In fact, if we put $y_1 = t_1^2$, $y_2 = t_1t_2$, $y_3 = t_2^2$, $y_4 = t_3$, then $\mathfrak{R} \otimes k$ is the ring of invariant elements of $k[y]$ under the transformation $y_i \to \zeta^i y_i$ for $i = 1, 2, 3, 4$.

The entire statement made so far in this section should be called the "structure theorem," because the theorems we are going to state, e.g., the following one, are its simple corollaries:

THEOREM 3. *The variety of moduli $\mathfrak{M} \otimes k$ admits an affine space over k of dimension three as a finite covering.*

We can determine the singular locus of the variety of moduli. The result can be stated as follows:

THEOREM 4. *If the characteristic of k is different from* 2, *the variety $\mathfrak{M} \otimes k$ has one and only one singular point, and it corresponds to $J_2 = J_6 = J_8 = 0$. If the characteristic of k is* 2, *the singular locus of $\mathfrak{M} \otimes k$ is a rational curve, and it corresponds to $J_2 = J_6 = 0$,* i.e., *to the birational classes of hyperelliptic curves each with only one Weierstrass point.*

PROOF. We can prove this theorem using some well-known general properties of coverings. However, we can prove it also quite elementally in the following way. Suppose first that the characteristic of k is different from 2. Take a point of the variety $\mathfrak{M}\otimes k$ and consider the corresponding local ring, say $\mathfrak{o}$. Then, if y_1^5 is a unit in $\mathfrak{o}$, we have the following inclusion relation $\mathfrak{R}\otimes k\subset k[y_1^5, y_1^{-5}, y_1^3y_2, y_1^2y_3]\subset\mathfrak{o}$. Therefore, we can consider $\mathfrak{o}$ as a local ring of an affine space over k with $k[y_1^5, y_1^3y_2, y_1^2y_3]$ as its co-ordinate ring. Hence the center of $\mathfrak{o}$ is simple. If either y_2^5 or y_3^5 is a unit in $\mathfrak{o}$, a similar argument shows that $\mathfrak{o}$ can be considered as a local ring of an affine space over k with either $k[y_1y_2^2, y_2^5, y_2y_3]$ or $k[y_1y_3^3, y_2y_3, y_3^5]$ as its co-ordinate ring. Hence the center of $\mathfrak{o}$ is again simple. Finally, if none of the y_1^5, y_2^5, y_3^5 is a unit in $\mathfrak{o}$, we can see quite easily that the Zariski tangent space (see [16]) of $\mathfrak{M}\otimes k$ at the center of $\mathfrak{o}$ is of dimension greater than three. Since we shall shortly return to a closer examination of this space, we do not go into detail. At any rate, we can conclude that the center of $\mathfrak{o}$ is singular. Thus, the variety $\mathfrak{M}\otimes k$ has one and only one singular point, and it corresponds to $J_2 = J_4 = J_6 = 0$, i.e., to $J_2 = J_6 = J_8 = 0$. Suppose next that the characteristic of k is 2. Take a point of the variety $\mathfrak{M}\otimes k$ and consider the corresponding local ring $\mathfrak{o}$. Then, if either y_1^5 or y_3^5 is a unit in $\mathfrak{o}$, we can consider $\mathfrak{o}$ as a local ring of an affine space over k with either $k[y_1^5, y_1^3y_2, y_1y_4]$ or $k[y_2y_3, y_3^5, y_3^2y_4]$ as its co-ordinate ring. Hence the center of $\mathfrak{o}$ is simple. If none of the y_1^5 and y_3^5 is a unit in $\mathfrak{o}$, the center of $\mathfrak{o}$ is singular. In fact, if the center of $\mathfrak{o}$ is general under the above condition in the sense y_4^5 is a unit in $\mathfrak{o}$, we have the following inclusion relation $\mathfrak{R}\otimes k\subset k[y_1y_4, y_2y_4^2, y_3y_4^3, y_4^5, y_4^{-5}]\subset\mathfrak{o}$. Therefore, the variety $\mathfrak{M}\otimes k$ is biregularly equivalent at the center of $\mathfrak{o}$ to a cylindrical hypersurface defined over k by $Y_1Y_3 - Y_2^2 = 0$ such that the image of the center of $\mathfrak{o}$ is on the fourth co-ordinate axis. Since this axis is the singular locus of the hypersurface, the center of $\mathfrak{o}$ is singular. Then, the same is true even if the center of $\mathfrak{o}$ is not general. Therefore, the singular locus of the variety $\mathfrak{M}\otimes k$ corresponds to $J_2 = J_6 = 0$.

We observe that the singular locus of the arithmetic variety of moduli $\mathfrak{M}$ consists of two "rational curves" of which one is arithmetic and another is geometric. These curves correspond, respectively, to $J_2 = J_6 = J_8 = 0$ and $2 = J_2 = J_6 = 0$, hence they intersect at one point which corresponds to $2 = J_2 = J_6 = J_8 = 0$. We know that points on the geometric curve represent birational classes of hyperelliptic curves which are birationally equivalent to plane curves defined inhomogeneously by $Y^2 - Y = X^5 + \alpha X^3$. Also, the point of the intersection corresponds to $Y^2 - Y = X^5$. On the other hand, we do not know as yet the meaning of the arithmetic curve except for this point. Actually, it is a straight-forward verification that points on this curve, not on $2 = 0$, represent birational classes of hyperelliptic curves which are birationally equivalent to plane curves defined inhomogeneously by[1]

$$Y^2 = X(X-1)(X-1-\zeta)(X-1-\zeta-\zeta^2)(X-1-\zeta-\zeta^2-\zeta^3)\,.$$

We shall discuss a more general class of hyperelliptic curves systematically in the next section. The following theorem, which we could have stated much earlier, is contained in the proof of Theorem 4:

THEOREM 5. *A generic point of the variety of moduli $\mathfrak{M}\otimes k$ with reference to k generates a purely transcendental extension of k of dimension three.*

We shall now discuss an affine embedding of the variety $\mathfrak{M}$, i.e., a selection of generators of the ring $\mathfrak{R}$. For this purpose, we consider the set of non-negative integer triples (e_3, e_4, e_5) satisfying an inequality of the form $3e_3 + 4e_4 \leqq 5e_5$. It is clear that this set forms a monoid M in the additive group of all integer triples. We shall show that M is generated by eight elements, which are (0, 0, 1), (0, 1, 1), (1, 0, 1), (2, 1, 2), (3, 0, 2), (1, 3, 3), (5, 0, 3), (0, 5, 4). Let (e_3, e_4, e_5) be an arbitrary element of M; let g_i be the least residue of e_i modulo 5 and put $e_i = 5f_i + g_i$ for $i = 3, 4$. Then (e_3, e_4, e_5) decomposes, within the monoid M, into $(g_3, g_4, e_5-3f_3-4f_4)$ and $(5f_3, 5f_4, 3f_3+4f_4)$. The decomposition is trivial if either $g_3 = g_4 = 0$ and $e_5 = 3f_3 + 4f_4$ or $f_3 = f_4 = 0$. In the first case, the given triple can be expressed by (5, 0, 3) and (0, 5, 4). In the second case, it is congruent modulo (0, 0, 1) to one of the $5^2 - 1 = 24$ triples. If we omit those triples among them which can be expressed by others, we get only five additional triples, and they are (0, 1, 1), (1, 0, 1), (1, 3, 3), (2, 1, 2), (3, 0, 2). This proves the statement. We consider next the submonoid N of M defined by a stricter inequality $2 + 3e_3 + 4e_4 \leqq 5e_5$. Then the monoid M operates on N in an obvious way, hence we can talk about M-generators of N. We shall show that (0, 0, 1), (1, 0, 1), (0, 2, 2) are M-generators of N. At any rate, every element of N is a sum of the previous eight elements of M. We observe also that the differences between $3e_3 + 4e_4$ and $5e_5$ for the eight elements are 5, 1, 2, 0, 1, 0, 0, 0. Therefore, in addition to (0, 0, 1) and (1, 0, 1), we have only to consider three sums of (0, 1, 1) and (3, 0, 2). In this way, we get one more namely (0, 2, 2). This proves the assertion.

[1] The fact that the variety of moduli does have a singular point was discovered, as we understand it, in the complex case by Gerstenhaber. Also, at the end of Fischer's dissertation [5], it is remarked that the variety of moduli for characteristic different from 2 and 5 is singular at the point which "represents the function field $K = k(x, y)$, where (x, y) satisfies the equation $y^2 = x^6 - x$." Fischer's method, however, "can not settle the question whether there exist any other singular points."

Now, we shall apply the above consideration to our problem. We know that the ring $\mathfrak{R}$ is generated by elements of the form $J_2^{e_1}J_4^{e_2}J_6^{e_3}J_8^{e_4}J_{10}^{-e_5}$ in which e_i are non-negative integers satisfying $e_1 + 2e_2 + 3e_3 + 4e_4 = 5e_5$. Since we have $J_4^2 = J_2J_6 - 4J_8$, we can assume that e_2 is either 0 or 1. In the first case, the corresponding triple (e_3, e_4, e_5) is an element of the monoid M and conversely. In the second case, the corresponding triple is an element of N and conversely. Therefore $\mathfrak{R}$ is generated by $8 + 3 = 11$ elements. However, because of the relation $J_2J_6 = J_4^2 + 4J_8$, one of them namely $J_2J_6^3J_{10}^{-2}$ becomes redundant. If we omit this element, therefore, we get the following ten generators:

$$J_2^5J_{10}^{-1}\ ,\quad J_2^3J_4J_{10}^{-1}\ ,\quad J_2^2J_6J_{10}^{-1}\ ,\quad J_2J_8J_{10}^{-1}\ ,\quad J_4J_6J_{10}^{-1}\ ,$$
$$J_4J_8^2J_{10}^{-2}\ ,\quad J_6^2J_8J_{10}^{-2}\ ,\quad J_6^5J_{10}^{-3}\ ,\quad J_6J_8^3J_{10}^{-3}\ ,\quad J_8^5J_{10}^{-4}\ .$$

If we fix these generators in this order, we get a definite embedding of the variety $\mathfrak{M}$ in an affine space over $\mathbf{Z}$ of dimension ten. We shall prove the following theorem:

THEOREM 6. *The variety $\mathfrak{M}$ can be considered as a subvariety of an affine space over $\mathbf{Z}$ of dimension ten. If k is a field of characteristic 2, the Zariski tangent space of the variety $\mathfrak{M} \otimes k$ at the singular point which corresponds to $J_2 = J_6 = J_8 = 0$ is of dimension ten. Therefore, the variety $\mathfrak{M}$ can not be embedded in an affine space over $\mathbf{Z}$ of dimension less than ten even locally at this point.*

PROOF. We have only to prove the second statement. Let $\mathfrak{o}$ be the local ring of the variety $\mathfrak{M} \otimes k$ at the singular point and let $\mathfrak{m}$ be the ideal of non-units in $\mathfrak{o}$. Then, the Zariski tangent space in question is the residue module $\mathfrak{m}/\mathfrak{m}^2$ considered as a vector space over the residue field $\mathfrak{o}/\mathfrak{m} = k$. If we denote the co-ordinate ring $\mathfrak{R} \otimes k$ by $\mathfrak{o}'$ and the restriction $\mathfrak{o}' \cap \mathfrak{m}$ of $\mathfrak{m}$ to $\mathfrak{o}'$ by $\mathfrak{m}'$, we can identify this vector space with the residue module $\mathfrak{m}'/\mathfrak{m}'^2$ considered as a vector space over the residue field $\mathfrak{o}'/\mathfrak{m}' = k$. We know that $\mathfrak{m}'$ is generated by ten elements, which are y_1^5, $y_1^3y_2$, $y_1^2y_3$, y_1y_4, y_2y_3, $y_2y_4^2$, $y_3^2y_4$, y_3^5, $y_3y_4^3$, y_4^5. We shall show that they are linearly independent over k modulo $\mathfrak{m}'^2$. We know that $y_1y_3 = y_2^2$ is the only relation between y_1, y_2, y_3, y_4. Therefore, every element of $k[y]$ can be expresesed uniquely as a linear combination with coefficients in k of monomials of the form $y_1^{e_1}y_2^{e_2}y_3^{e_3}y_4^{e_4}$ in which e_2 is 0 or 1. Since the ten elements in question are of this form, we have only to show that none of them is contained in $\mathfrak{m}'^2$. For this purpose, if $y_1^{e_1}y_2^{e_2}y_3^{e_3}y_4^{e_4}$ is contained in $\mathfrak{o}'$, we define its weight by $e = 1/5\cdot(e_1 + 2e_2 + 3e_3 + 4e_4)$. We note that the weight is a non-negative integer and it is at least equal to one if the monomial is contained in $\mathfrak{m}'$. Since the weight is additive, therefore, it is at least equal to two if the monomial is in $\mathfrak{m}'^2$. Now, since the first five elements are of weight one, they are not in $\mathfrak{m}'^2$. The next two elements namely $y_2y_4^2$ and $y_3^2y_4$ are of weight two. However, since they contain y_4 and since the only monomial in $\mathfrak{o}'$ of weight one containing y_4 is y_1y_4, certainly they are not in $\mathfrak{m}'^2$. The last three elements are y_3^5, $y_3y_4^3$ and y_4^5. Since y_3^5 is the smallest power of y_3 contained in $\mathfrak{m}'$, it is not in $\mathfrak{m}'^2$. Similarly y_4^5 is not in $\mathfrak{m}'^2$. As for $y_3y_4^3$, it is of weight three and contains y_4. Since it can not contain any one of the y_1y_4, $y_2y_4^2$, $y_3^2y_4$ as a factor, it is not in $\mathfrak{m}'^2$. This completes the proof.

If k is a field of characteristic different from 2, the ring $\mathfrak{R} \otimes k$ can be generated in a simpler way. In fact, if we make the substitution $J_8 = 1/4\cdot(J_2J_6 - J_4^2)$ in the ten elements we obtained before and if we omit redundant elements, we only get the following eight elements:

$$J_2^5J_{10}^{-1}\,, \quad J_2^3J_4J_{10}^{-1}\,, \quad J_2J_4^2J_{10}^{-1}\,, \quad J_2^2J_6J_{10}^{-1}\,,$$
$$J_4J_6J_{10}^{-1}\,, \quad J_2J_6^3J_{10}^{-2}\,, \quad J_4^5J_{10}^{-2}\,, \quad J_6^5J_{10}^{-3}\,.$$

Therefore, the variety $\mathfrak{M} \otimes k$ can be embedded in an affine space over k of dimension eight. On the other hand, we can show more or less in the same way as in the proof of Theorem 6 that the Zariski tangent space of $\mathfrak{M} \otimes k$ at its singular point is of dimension eight. Therefore, none of the above eight elements is redundant and the variety $\mathfrak{M} \otimes k$ can not be embedded in an affine space over k of dimension less than eight even locally at this point.[2] On the other hand, the Zariski tangent spaces of $\mathfrak{M} \otimes k$ with reference to a field k of characteristic 2 at singular points different from the worst point mentioned in Theorem 6 are all of dimension four. In fact, this is contained in the proof of Theorem 4.

8. Hyperelliptic curves with many automorphisms

We recall that every hyperelliptic curve can be considered in a unique way as a two-sheeted covering of a projective straight-line. Therefore, the group of automorphisms of the curve contains the Galois group of the covering, which is just a cyclic group of order 2, in the center. Moreover, the factor group can be considered as a group of automorphisms of the projective straight-line, which we call the *reduced group of automorphisms*. If the reduced group of automorphisms contains at least two elements, we say that the curve or rather its birational class has

[2] Hecke considered, in his dissertation [8], a ring generated over **Q** by six, instead of eight, invariants. We can show that the variety over **Q** with this as its co-ordinate ring is in a one-to-one correspondence with our $\mathfrak{M} \otimes$ **Q**. The correspondence is birational but not everywhere biregular, because the variety of Hecke is not normal.

many automorphisms. In this last section, we shall determine the distribution of birational classes of hyperelliptic curves with many automorphisms. The following lemma will simplify some of the arguments:

LEMMA 8. *Let C and C' be two non-singular curves of genus greater than one and assume that C' is a specialization of C. Then, the group of automorphisms of C' contains at least one subgroup which is isomorphic to the group of automorphisms of C.*

PROOF. Construct the Jacobian variety J of C by Chow's method. Then, as we know [9–I], it specializes uniquely to the Jacobian variety J' of C' over the specialization $C \rightarrow C'$. The reader must keep in mind that the group law of J specializes also uniquely to the group law of J'. Therefore (cf. [14]), the specialization $C \rightarrow C'$ gives rise to a homomorphism of the ring of endomorphisms of J in the ring of endomorphisms of J'. We shall show that this homomorphism is a monomorphism. Let L be a graph in the product $J \times J$ of an endomorphism of J different from 0 and let L' be a specialization of L over the specialization $J \rightarrow J'$. We take a generic hyperplane H in the ambient space of J and also a generic hyperplane H' in the ambient space of J' so that H' is a specialization of H over the above specialization. Then, the intersection $L \cap J \times H$ is certainly not empty, hence $L' \cap J' \times H'$ is not empty. Therefore, the point-set theoretic projection of L' to the second factor can not reduce to a point, hence L' is a graph of an endomorphism of J' different from 0. On the other hand, we observe that a linear extension (cf. [14, p. 77]) of an automorphism of C is an automorphism of J. In this way, we get a homomorphism of the group of automorphisms of C in the group of automorphisms of J. Since C is of genus greater than one, this homomorphism is a monomorphism. This is not a trivial statement, but we suppose that it is well known. In the same way, we get a monomorphism of the group of automorphisms of C' in the group of automorphisms of J'. Moreover, if X is a graph in the product $C \times C$ of an automorphism of C and if X' is a specialization of X over the specialization $C \rightarrow C'$, then X' is certainly irreducible, hence it is a graph of an automorphism of C'. Therefore, the monomorphism of the ring of endomorphisms of J in the ring of endomorphisms of J' induces a monomorphism of the group of automorphisms of C in the group of automorphisms of C'. This completes the proof.

We shall prove another lemma which gives a supplementary information about Jacobian varieties of hyperelliptic curves with many automorphisms. We note that a Jacobian variety of a hyperelliptic curve is either simple or isogeneous to a product of two elliptic curves (cf. [14, p. 94]).

LEMMA 9. *If a hyperelliptic curve has two automorphisms of order 2, its Jacobian variety is isogeneous to a product of elliptic curves.*

PROOF. Let C be the hyperelliptic curve. Since C has two automorphisms of order 2, one of them is not the covering automorphism of the projective straight-line. Call it τ. Then τ can be considered as a covering automorphism of a non-singular curve, say A. If g_0 is the genus of A, by the choice of τ, it is positive. Also, if we apply the Riemann-Hurwitz relation to this covering, we get

$$2 = 2(2g_0 - 2) + \text{deg. different.}$$

The only possibility is then $g_0 = 1$. Hence, we can consider A as an elliptic curve. If J is the Jacobian variety of C, the projection $C \to A$ gives rise to an epimorphism $J \to A$. In this case J is isogeneous to the product of A and the kernel of the epimorphism, which is an elliptic curve.

Now, suppose that C is a hyperelliptic curve in characteristic different from 2. Then, the group of automorphisms of C can be discussed in terms of its Rosenhain normal form, which is $Y^2 = X(X-1)(X-\lambda_1)(X-\lambda_2)(X-\lambda_3)$. In fact, the reduced group of automorphisms can be considered as the group of projective transformations which permute the six points of ramification with co-ordinates $0, 1, \infty, \lambda_1, \lambda_2, \lambda_3$ among themselves. Since every projective transformation is uniquely determined by its effect on three of these six points, the order of the group is at most 120. Not only that, all possible cases where the six points are permuted among themselves can be determined by an actual computation. This is a more naive way of determining hyperelliptic curves with many automorphisms than Bolza's method [1]. However, the advantage is that this naive method is not restricted to the case of characteristic zero. At any rate, if we denote by D_{2n} the dihedral group of order $2n$, the results can be summarized as follows:

(1) The lambdas are $\lambda, \mu, \lambda(1-\lambda)^{-1}(1-\mu)$. The reduced group of automorphisms is in general, i.e., unless this case reduces to cases below, cyclic of order 2;

(2) The lambdas are obtained by specializing μ in (1) to $(\lambda-1)\lambda^{-1}$, i.e., they are $\lambda, (\lambda-1)\lambda^{-1}, (1-\lambda)^{-1}$. The reduced group of automorphisms is in general D_6, which is the symmetric group of permutations of three letters;

(3) The lambdas are obtained by specializing μ in (1) to λ^{-1}, i.e., they are $\lambda, \lambda^{-1}, -1$. The reduced group of automorphisms is in general D_4, which is the Klein four group;

(4) The lambdas are obtained by specializing λ in (2) or in (3) to 2, i.e., they are $2, 2^{-1}, -1$. This case, unlike other cases, disappears in characteristic 3. The reduced group of automorphisms is D_{12};

(5) The lambdas are obtained by specializing λ in (3) to $i = (-1)^{1/2}$, i.e., they are $i, -i, -1$. The reduced group of automorphisms is the octahedral group, which is the symmetric group of permutations of four letters;

(6) The lambdas are $1 + \zeta, 1 + \zeta + \zeta^2, 1 + \zeta + \zeta^2 + \zeta^3$ in which ζ is a primitive fifth root of unity. The reduced group of automorphisms is cyclic of order 5.

There is *one modification* to be made in the above table. If the characteristic is 5, the last three cases reduce to a single case and the reduced group of automorphisms will swell up to PL(2, 5), i.e., to the group of projective transformations with coefficients in GF(5). The order of this group attains the upper bound 120. On the other hand, it must be understood that, in each case, we just mentioned one out of many lambda triples which correspond to the same birational class of hyperelliptic curves. For instance, if we specialize λ in (2) to i, the specialized triple also belongs to the same birational class as the triple in (5). At any rate, we can say that (2) and (3) are limiting cases of (1) and both (4) and (5) are limiting cases of (2) and (3). However, Lemma 8 shows that, except in characteristic 5, the case (6) is isolated in the sense it is not a limiting case of any other cases.

We shall also examine hyperelliptic curves with many automorphisms in characteristic 2. Suppose that $Y^2 - Y = \alpha X + \beta X^{-1} + \gamma(X - 1)^{-1}$ is a normal form for the type (1, 1, 1). Then, the reduced group of automorphisms is a subgroup of the group of projective transformations permuting the three points of ramification with co-ordinates $0, 1, \infty$ among themselves. Hence, the order of the group is at most 6. Moreover, this group is cyclic of order 2 for $\alpha = \beta$ or $\beta = \gamma$ or $\gamma = \alpha$ and it swells up to D_6 for $\alpha = \beta = \gamma$. These cases will be referred to respectively as case (1′) and case (2′). Suppose next that $Y^2 - Y = X^3 + \alpha X + \beta X^{-1}$ is a normal form for the type (3', 1). Then, as we can see, the reduced group of automorphisms reduces to the identity. Finally, suppose that $Y^2 - Y = X^5 + \alpha X^3$ is a normal form for the type (5). Then, the discussion in Section 2 shows that the reduced group of automorphisms consists of linear transformations of the form $x \rightarrow x + c_1^2$ in which $c_1^{16} + \alpha^2 c_1^8 + \alpha c_1^2 + c_1 = 0$. Therefore, it is an abelian group of type (2, 2, 2, 2). In the limiting case $\alpha = 0$, we can multiply fifth roots of unity to x. The discussion in Section 2 shows again that we get the full reduced group of automorphisms in this way. Therefore, the group is a meta-abelian group of order 80. These cases will be referred to as case (3′) and case (6′). We shall show that (3′) is a limiting case of (1′). Consider a plane quartic defined inhomogeneously by $XY^2 + (1 + aX + bX^2)Y + X^2(c + X^2) = 0$. Then, the calculation at the end of Section 2 shows that this curve is birationally equivalent to a hyperelliptic curve in case (1′) if it is of genus two and if a, b, c satisfy the following relation

$$ab^4c^2 + (a^2 + b)b^5c + (a^5 + a^4b^3 + a^2b^4 + b^5) = 0$$

with ab different from 0. In fact, we get this relation from the equation $\beta = \gamma$ by eliminating ξ and η. We shall show that the above polynomial is absolutely irreducible. At any rate, it is primitive as a polynomial in c. Therefore, if it is reducible, it splits into two polynomials both linear in c. However, for $b = a^2$ it becomes $a^5(a^4c^2 + a^5 + 1)$, and this brings a contradiction. Also, the polynomial does not divide J_{10} in which J_{10} is considered as a polynomial in a, b, c. In fact, if the polynomial divides J_{10}, the divisibility must subsist, e.g., for $b = 0$. However, for $b = 0$ they become a^5 and $a^7c + a^5 + a^4c^2 + 1$, and this is a contradiction. Therefore, if we take a generic solution (a, b, c) over GF(2), we get a plane quartic of genus two. Moreover, we can specialize (a, b, c) to $(0, 0, c)$ over GF(2) and we get a plane quartic defined inhomogeneously by

$$XY^2 + Y + X^2(c + X^2) = 0.$$

We remarked at the very end of Section 2 that this curve is birationally equivalent to a plane curve defined by the normal form $Y^2 - Y = X^5 + cX^3$. Since c is transcendental over GF(2), it can be specialized to any quantity. Therefore, Lemma 7 shows that (3′) is a limiting case of (1′). Now, the above labeling will be justified by the following theorem:

THEOREM 7. *If C is a hyperelliptic curve in case* (n) *in which $n = 1$, 2, 3, 6 and if a hyperelliptic curve C' in characteristic 2 is a specialization of C, then C' is in case* (n′) *or in its limiting case. Conversely, every hyperelliptic curve C' in case* (n′) *is birationally equivalent to a hyperelliptic curve which is a specialization of a hyperelliptic curve C in case* (n). *As for case* (4) *and case* (5), *they simply disappear in characteristic* 2.

PROOF. The first and the third parts follow from Lemma 8. Therefore, we shall prove only the second part. Consider $X(X - 1)(Y^2 - Y) = \alpha X^2(X - 1) + \beta(X - 1) + \gamma X$ as a quadratic equation with respect to Y in characteristic zero. Then, its discriminant is given by

$$X(X - 1)\big(4\alpha X^3 - (4\alpha - 1)X^2 + (4\beta + 4\gamma - 1)X - 4\beta\big) .$$

If we denote the roots of the third cubic factor by $\lambda_1, \lambda_2, \lambda_3$, therefore, the corresponding Rosenhain normal form is birationally equivalent to a plane curve defined inhomogeneously by $Y^2 - Y = \alpha X + \beta X^{-1} + \gamma(X-1)^{-1}$. This being remarked, we try to find conditions under which a projective transformation of the form $x \rightarrow \text{constant}\cdot x^{-1}$ keeps the above discriminant invariant up to a constant factor. In this way, we get $x \rightarrow (\alpha^{-1}\beta)^{1/2}\cdot x^{-1}$ and the condition $\alpha(8\beta + 4\gamma - 1)^2 - \beta(8\alpha - 1)^2 = 0$. We observe that this polynomial in α, β, γ is absolutely irreducible. Therefore, we can take a generic solution (α, β, γ) over $\mathbf{Q}$. Then, the plane curve defined by $Y^2 - Y = \alpha X + \beta X^{-1} + \gamma(X - 1)^{-1}$, more precisely its non-singular model, is in case (1) or in its limiting case. However, since the equation in α, β, γ reduces modulo 2 to $\alpha \equiv \beta \bmod 2$, the curve in characteristic zero is in case (1) and it specializes to a generic curve in case (1′). This is all we have to know for $n = 1$. The other cases, e.g., the case $n = 2$, can be treated similarly. We observe that roots of a cubic equation

$$s_0X^3 - s_1X^2 + s_2X - s_3 = 0$$

are of the form $\lambda, (\lambda - 1)\lambda^{-1}, (1 - \lambda)^{-1}$ if and only if the coefficients are related as $3s_0 - s_1 + s_2 = 0,\ s_0 + s_3 = 0$. Therefore, if we determine α, β, γ from $\alpha + \beta = \alpha + \gamma = 0$, the plane curve defined by

$$Y^2 - Y = \alpha X + \beta X^{-1} + \gamma(X - 1)^{-1}$$

is in case (2) or in its limiting case. We take a variable α over $\mathbf{Q}$ and consider a plane curve defined by $Y^2 - Y = \alpha\big(X - X^{-1} - (X - 1)^{-1}\big)$. Then, this curve is in case (2) and its reduction modulo 2 is a generic curve in case (2′). For other cases, we shall give a different treatment shortly.

We observe that the theorem can be proved more systematically by examining the corresponding points on the arithmetic variety of moduli. For instance, case (1) corresponds to zeros of the so-called ''skew-invariant.'' However, the above calculations are simpler for $n = 1. 2$. The second method is not too bad for $n \geqq 3$. If we introduce a new modulus ρ in case (3) by $\rho = (\lambda - \lambda^{-1})^2$, the corresponding basic arithmetic invariants will become

$$2^2(3 + 2\rho),\ 2(3 + \rho + \rho^2),\ 2^2(1 - \rho), 2^2(1 - \rho)(3 + 2\rho) - (3 + \rho + \rho^2)^2,\ \rho^3.$$

Hence, if we reduce them modulo 2, we get $0, 0, 0, (1 + \rho + \rho^2)^2, \rho^3$. This shows that case (3′) is a limiting case of case (3). Now, in case (4), i.e., if we specialize ρ to $2^{-2}3^2$, the ten co-ordinates we introduced in the previous section take the following values:

$$2^{11}3^{-1}5^{5}\,,\quad 2^{6}3^{-2}5^{4}11\,,\quad -2^{8}3^{-4}5^{3}\,,\quad -2^{-1}3^{-4}5^{3}491\,,\quad -2^{3}3^{-5}5^{2}11\,,$$
$$2^{-7}3^{-9}5^{5}11\cdot 491^{2}\,,\quad -2^{4}3^{-11}5^{4}491\,,\quad -2^{18}3^{-18}5^{5}\,,\quad 2^{-6}3^{-15}5^{7}491^{3}\,,\quad -2^{-16}3^{-19}5^{10}491^{5}\,.$$

Also, in case (5), i.e., if we specialize ρ to -2^2, the ten co-ordinates take the following values:

$$2^{4}5^{4}\,,\quad 2\cdot 3\cdot 5^{4}\,,\quad -5^{3}\,,\quad -2^{-4}5^{3}13\,,\quad -2^{-3}3\cdot 5^{2}\,,$$
$$2^{-11}3\cdot 5^{5}13^{2}\,,\quad -2^{-8}5^{4}13\,,\quad -2^{-8}5^{5}\,,\quad 2^{-16}5^{7}13^{3}\,,\quad -2^{-24}5^{10}13^{5}\,.$$

Since some of the co-ordinates are not integral at 2 and 3 in case (4) and at 2 in case (5), these cases must disappear at these characteristics. On the other hand, we get $J_2 = J_4 = J_6 = J_8 = 0$ in case (6) and in case (6′). Therefore, the ten co-ordinates are all zero in these cases.

Finally, we shall show that the Jacobian variety of case (n) and case (n′) is for $n \leqq 5$ isogeneous to a product of elliptic curves. We observe that in case (1) the group of automorphisms is a four group. Lemma 8 shows, therefore, that in case (n) or in case (n′) the group of automorphisms contains this group as a subgroup, whence the assertion by Lemma 9. Actually, the isogeny itself can be determined explicitly. Let A, B be two elliptic curves defined by two automorphisms in the four group. Then, the Jacobian variety J covers the product $A \times B$ four times. Moreover, if the group law in J is with respect to a Weierstrass point of C and if, for instance, the characteristic is different from 2, the kernel of the epimorphism $J \rightarrow A \times B$ is as follows. Let τ be the covering automorphism of $C \rightarrow A$. Then τ permutes the six Weierstrass points among themselves without fixed-point, hence they can be written as P_i, τP_i for $i=1, 2, 3$. With this notation, non-zero elements in the kernel of $J \rightarrow A \times B$ are $P_i + \tau P_i$ for $i = 1, 2, 3$. At any rate, we see that case (6) is the only case of hyperelliptic curves with many automorphisms which might give simple Jacobian varieties. We can show that if, for instance, the characteristic is zero, the Jacobian variety is indeed simple and the ring of endomorphisms is the principal order of the cyclotomic field of fifth roots of unity.

INSTITUTE FOR ADVANCED STUDY AND
JOHNS HOPKINS UNIVERSITY

REFERENCES

1. O. BOLZA, *On binary sextics with linear transformations into themselves,* Amer. J. Math., 10 (1888), 47–70.
2. W. L. CHOW, *On the connectedness theorem in algebraic geometry,* Amer. J. Math., 81 (1959), 1033–1074.

3. A. CLEBSCH, Theorie der binären algebraischen Formen, Leipzig, 1872.

4. M. DEURING, *Invarianten und Normalformen elliptischer Funktionenkörper*, Math. Z., 47 (1941), 47–56.

5. I. FISCHER, *The moduli of hyperelliptic curves*, Trans. Amer. Math. Soc., 82 (1956), 64–84.

6. A. GROTHENDIECK, Géométrie formelle et géométrie algébrique, Séminaire Bourbaki (1958–59), exposé 182.

7. H. HASSE, *Theorie der relativ-zyklischen algebraischen Funktionenkörper, insbesondere bei endlichem Konstantenkörper*, J. Reine Angew., Math. 172 (1934), 37–54.

8. E. HECKE, *Höhere Modulfunktionen und ihre Anwendung auf die Zahlentheorie*, Math. Ann., 71 (1912), 1–37.

9. J. IGUSA, *Fibre systems of Jacobian varieties*, I, Amer. J. Math., 78 (1956), 171–199, III, *ibid.*, 81 (1959), 453–476.

10. ———, *Kroneckerian model of fields of elliptic modular functions*, Amer. J. Math., 81 (1959), 561–577.

11. G. ROSENHAIN, Abhandlung über die Functionen zwier Variabler mit vier Perioden, 1851, Ostwald's Klassiker der Exacten Wissenschaften, no. 65 (1895).

12. B. L. VAN DER WAERDEN, Algebra, I, II, Springer, 1955.

13. A. WEIL, Foundations of Algebraic Geometry, New York, 1946.

14. ———, Variétés Abéliennes et Courbes Algébriques, Paris, 1948.

15. O. ZARISKI, *Foundations of a general theory of birational correspondences*, Trans. Amer. Math. Soc., 53 (1943), 490–542.

16. ———, *The concept of a simple point of an abstract algebraic variety*, Trans. Amer. Math. Soc., 62 (1947), 1–52.