ACTA ARITHMETICA
XIX (1971)

# On the Diophantine equation $f(y)-x=0$

by

MICHAEL FRIED (Stony Brook, N. Y.)

The author has made several investigations into problems related to the genus zero curve $f(y)-x=0$ for $f(y)\in K[y]$ where $K$ is a field. The progression of events can be best seen by a quick perusal of the sequence of papers [4], [5], [6], [7], [8], whereby the very particular problems of [4], [5], [6] have launched the general problems about the fields of definition of arbitrary models of Riemann surfaces in [7] and [8]. This latter work is just barely started, but it has already shown the need for the development of arithmetic tools that far transcend the simple techniques the author has so far mustered to attack the problems of [6] and [7]. Oftentimes it turns out, the most powerful tools are those of group theory. Especially when a precise formulation of the problem can be made in terms of the branching of certain Riemann surfaces. See Lemma 2 for example. However, the context of the problem is often much too difficult (or complicated) for the present state of group theory. See [5], Section 3, for a discussion of the general type of group theory question that needs to be answered. Also, even with complete knowledge of the group theory aspect of the problem, the arithmetic question may still remain unanswered for the very reason that we are often reduced to asking questions about the field of definition of curves which are apriori defined only over $C$. See [7], Section 1, for the precise formulation of general problems in this area.

In this paper, we will return to some of the particular problems related to the curve $f(y)-x=0$, and in so doing we will concentrate almost entirely on questions that can be answered by arithmetic; albeit simple arithmetic. A slight historical digression seems in order. We give a chronology of the author's work related to the problems to be discussed in this paper. Our comments on the other mathematicians who have made contributions to related problems is not meant to be complete. See [6] and [7] for references to these works.

We need some notation which will be used throughout this paper. Let $K^*$ be a fixed algebraic closure of a field $K$. A polynomial $f(y)\in K[y]$

is said to be *decomposable over* $K$ if we can write $f(y) = f_1(f_2)$, where $f_1, f_2 \in K[y]$ and degree of $f$ is not one for $i = 1, 2$. We call $f_1$ and $f_2$ *composition factors* of $f$.

LEMMA 1. (Theorem 3.5 of [2].) *If $f(y) \in K[y]$ is decomposable over $K^*$, then $f(y)$ is decomposable over $K$.*

Remark. The corresponding result for rational functions $f(y) \in K(y)$ does not hold. This is discussed in [4], Section 1.

Let $x$ be an indeterminate over $K^*$, so that the zeros $y_1, \ldots, y_n$ of $f(y)-x$ are also indeterminate over $K^*$, where $n =$ degree $f$. Let $\Omega_{f-x} = K(y_1, \ldots, y_n, x) = K(y_1, \ldots, y_n)$, and let $G(\Omega_{f-x}/K(x))$ denote the Galois group of $\Omega_{f-x}$ over $K(x)$. Much of our discussion is related to the case where $K$ is a number field. If $f(y) \in K(y)$, then we may reduce $f(y)$ modulo any prime $p$, of the ring of integers of $K$, not dividing the denominators of the coefficients of $f(y)$. Let $V_p(f)$ denote the values assumed by $f(y)$ modulo $p$ (we count $\infty$ as a coset modulo $p$).

**I. General Schur problem.** *For $K$ a number field, $f(y) \in K(y)$, when is it possible for $V_f(f)$ to consist of all cosets modulo $p$ for infinitely many primes $p$?* I. Schur conjectured that for $f(y) \in K[y]$ this could happen only if $f(y)$ was a composition of polynomials of two special types:

1) $ay^n + b$ (cyclic),

2) $T_n(y) = 2^{-n-1}\{(y+(y^2+4)^{1/2})^n + (y-(y^2+4)^{1/2})^n\}$ (Čebyšev polynomials).

For a simple proof of this see [6]; and for a discussion and partial results on the more complicated situation where $f(y)$ is not assumed to be a polynomial, see [5], Section 1.

MacCluer showed that if $f(y) \in L[y]$ where $L$ is a finite field and if

3) $\dfrac{f(y)-f(z)}{y-z}$ has no absolutely irreducible factors,

and

4) the function field $L(x, y)$ is tamely ramified over $L(x)$;

then $f$ is one-to-one over $L$. For a slight generalization of this to rational functions see [5], Section 1. It is known that polynomials satisfying 3) and 4) must be compositions of polynomials of type 1) and 2) (see [5], Theorem 3).

**II. Values of polynomials.** In [4] the author gave an analogue of Schur's conjecture. Let $f_1, \ldots, f_l \in K[y]$, and suppose $\bigcup_{i=1}^{l} V_p(f_i)$ consists of all cosets modulo $p$ for all but a finite number of primes $p$ (a.a.$p$). Then one of the polynomials $f_1, \ldots, f_l$ must be linear. If we only assume $f_1, \ldots$

$\ldots, f_l \in K(y)$, this result does not hold ([5], Section 2). Related situations have been studied in [4], [7]. One of the simplest of these is:

5) suppose $f, g \in K(y)$,

and

6) $V_p(f) = V_p(g)$ for a.a.$p$.

A reduction process appears in [4], p. 101; to show that if 5) and 6) hold, with no loss we may replace $f$ and $g$ by composition factors of $f$ and $g$, so that we may assume

7) both $f$ and $g$ are indecomposable.

In addition 5) and 6) imply:

8) $\Omega_{f-x} = \Omega_{g-x}$,

9) $\bigcup_{y_i} G\left(\Omega_{f-x}/K(y_i)\right) = \bigcup_{z_i} G\left(\Omega_{g-x}/K(z_j)\right)$ where $z_1, \ldots, z_m$ are the zeros of $g(z)-x$,

10) $f(y)-g(z)$ is reducible in the sense that it is a product of two elements of $K(y, z)$ of lower degree.

Note that 10) is a consequence of 9) by simple group theory.

The author has shown ([7], Section 4) that 7) and 10) together cannot happen if $f(y)$, $g(y) \in \boldsymbol{Q}[y]$ unless $f(ay+b) = g(y)$ for some $a, b \in \boldsymbol{Q}$. In particular, if $f, g \in \boldsymbol{Q}[y]$, and $V_p(f) = V_p(g)$ for a.a.$p$., then $f(ay+b) = g(y)$ for some $a, b \in \boldsymbol{Q}$. See [7], Section 3, for an example of polynomials $f, g \in \boldsymbol{Q}(\sqrt{-7})[y]$ for which $V_p(f) = V_p(g)$ for a.a.$p$; even though this simple relation does not exist between $f$ and $g$. It is not known if there are examples of polynomials $f$, $g$ of arbitrary large degree satisfying 5), 6), and 7), unless $f(ay+b) = g(y)$ for some $a$, $b \in K^*$. Examples of degree 7, 11, 13, 15 and 21 are now known.

**III. A global Diophantine problem.** Again let $K$ be a number field and $f \in K[y]$. We denote by $\mathfrak{o}_k$ the ring of integers of $K$. Let:

11) $W(f) = \{x_0 \in \mathfrak{o}_k \mid f(y)-x_0$ is reducible in $K[y]\}$,

12) $V(f) = \{x_0 \in \mathfrak{o}_k \mid f(y)-x_0$ has a linear factor in $K[y]\}$,

$V'(f) = \{x_0 \in \mathfrak{o}_k \mid f(y)-x_0$ has two linear factors in $K[y]\}$.

In [7], Section 4, the author has shown that: excluding a finite set of $x_0$, $W(f) \subset \bigcup_1^l V(g_i)$ where

13) $g_1, \ldots, g_l \in K[y]$,

14) $f(y)-g_i(z)$ is reducible over $K$, $i = 1, \ldots, l$.

The author has shown that if $K = \boldsymbol{Q}$, and either:

15) $f$ is indecomposable, or

16) degree $f = p^r$ for some rational prime $p$, $p \neq 2$,

then $W(f)$ is $V(f)$ plus a finite set. This holds when 15) is satisfied by use of a result already mentioned. Under very general circumstances $V'(f)$ is finite ([7], Section 4).

Schinzel and the author have separately shown that if $f, g \in \boldsymbol{Q}[y]$ and $f(y)-g(z)$ is reducible where degree $f=p$ ($p$ a rational prime), then there exists a polynomial $h(y) \in \boldsymbol{Q}^*[y]$ such that $f(h(y)) = g(y)$. In the next section we give a generalization of this to the case where 16) is satisfied.

**1. Reducibility of polynomials of form $f(y)-g(z)$ over $\boldsymbol{Q}$.** We first give a lemma that demonstrates how questions about the existence of polynomials $f(y)$, $g(y) \in \boldsymbol{C}[y]$ satisfying 10) can be reduced to group theory. As already has been mentioned, we may without loss assume that $\Omega_{f-x} = \Omega_{g-x}$ (as in Lemma 4). Let $G^*$ be a finite permutation group on the letters $\{w_1, \ldots, w_n\}$. For $\sigma \in G^*$, write $\sigma$ as a product of disjoint cycles $\gamma_1 \cdots \gamma_s$.

We define $\operatorname{ind}\sigma$ (read index of $\sigma$) as

$$\sum_{i=1}^{s} [\operatorname{ord}(\gamma_i)-1].$$

LEMMA 2. *Let $G^*$ be a finite group with two permutation representations on the letters $\{y_1^*, \ldots, y_n^*\}$ and $\{z_1^*, \ldots, z_n^*\}$ respectively. If $\{w_1, \ldots, w_n\}$ are any set of letters on which it makes sense to represent $G^*$, for $\sigma^* \in G^*$ let $\sigma_w^*$ be the permutation of $\{w_1, \ldots, w_n\}$ corresponding to $\sigma^*$. Then, there exist polynomials $f, g \in \boldsymbol{C}[y]$ such that:*

17) $\Omega_{f-x} = \Omega_{g-x}$,

18) $G^* = G\big(\Omega_{f-x}/C(x)\big) \overset{\text{def}}{=} G$, *and*

19) $f(y)-g(z)$ *is reducible,*

*if and only if there exist elements $\sigma^*(1), \ldots, \sigma^*(r)$ in $G$, such that:*

20) $\sigma^*(1), \ldots, \sigma^*(r)$ *generate $G^*$,*

21) *if we let $\sigma^*(\infty) = \sigma^*(1) \ldots \sigma^*(r)$, then $\sigma_{x^*}^*(\infty) = (x_1^*, \ldots, x_n^*)$ (an $n$-cycle) and $\sigma_{z^*}^*(\infty) = (z_1^*, \ldots, z_n^*)$,*

22) $\sum_{j=1}^{r} \operatorname{ind}\sigma_{x^*}^*(j) = \sum_{j=1}^{r} \operatorname{ind}\sigma_{z^*}^*(j) = n-1$.

Remark. Lemma 2 is a particular case of Proposition 5 of [7]. According to Schinzel, Cassels has formulated problems about the reducibility of polynomials of type $f(y)-g(z)$ in terms similar to those expressed by Lemma 2 (see [1]). Riemann's existence theorem is the main tool used in the proof of Lemma 2. When group theory can be used to show the existence of a Riemann surface of certain type, there still remains a question

as to the field of definition of its function field. This is the general problem dealt with in [7] and [8].

In what follows we will use the Puiseux expansion about $\infty$ of $f(y)-x$. The notation we use is that of [4], p. 101, which contains a very down-to-earth explanation of these expansions.

THEOREM. *Let $K$ be a field such that*

23) $[K(\zeta_{p^r}): K] = (p-1)p^{r-1}$ *(where $p \neq 2$, and $\zeta_{p^r}$ is a primitive $p^r$-th root of 1).*

*Suppose that $f,\ g \in K[y]$ where:*

24) *degree $f = p^r$,*

25) *$f(y)-g(z)$ is reducible in $K[y,z]$, but $f_1(y)-g(z)$ is not reducible for any polynomial $f_1(y)$ such that $f_1\ (f_2) = f$ for some polynomial $f_2$.*

*Then,*

26) *there exists $h(y) \in K^*[y]$ such that $f(h(y)) = g(y)$.*

The next two lemmas are needed for the proof of the theorem. They have been useful to the author in computations unrelated to the theorem above. For these lemmas we introduce the field $F$ of any characteristic.

LEMMA 3. *Let $f,\ g \in F[y]$ be polynomials of degree $n$ with the same leading coefficient. Assume that $(2 \cdot \operatorname{char} F, n) = 1$. Then, either*

27) *$F(w_1, x) = F(y_1, z_1)$ where $w_1 = y_1 - z_1$, or*

28) *$w_1$ is a constant, where $f(y_1) = x$, $g(z_1) = x$ and the leading terms of the expansions for $y_1$ and $z_1$ over $\infty$ are the same.*

Note. If $w_1$ is a constant, then $f(y) = g(y+b)$ for some $b \in F$.

Proof. Let $\zeta_n$ be a primitive $n$th root of one. The Puiseux expansions over $\infty$ for $y_{i+1}$ are of the form

29) $y_{i+1} = a_{-1}\zeta_n^i x^{1/n} + a_0 + a_1\zeta_n^{-i}x^{-1/n} + \ldots$ for $i = 0, 1, \ldots, n-1$.

30) $z_{i+1} = b_{-1}\zeta_n^i x^{1/n} + b_0 + b_1\zeta_n^{-i}x^{-1/n} + \ldots$ for $i = 0, 1, \ldots, n-1$,

where we may assume $a_{-1} = b_{-1}$ because of the assumption of equality of the leading coefficients of $f$ and $g$. Thus, $w_1$ has no $x^{1/n}$ term in its expansion over $\infty$. Let $\Omega_x = \Omega_{g-x} \cdot \Omega_{f-x}$.

The quantity $w_1 + z_1 = y_1$ has exactly $n$ conjugates, all obtained by the substitutions $x^{1/n} \to \zeta_n^j x^{1/n}$ for $j = 0, \ldots, n-1$. If $w_1$ has no conjugates over $F(z_1)$, then by the fundamental theorem of Galois theory, $w_1 \in F(z_1)$. Thus, $y_1 \in F(z_1)$. This implies $\dfrac{ay_1+b}{cy_1+d} = z_1$ for some $a, b, c, d \in F$. Therefore, $f(y_1) = g\left(\dfrac{ay_1+b}{cy_1+d}\right)$, and since $g$ is a polyno-

mial we easily deduce that we may take $c = 0$, $d = 1$. From the assumption that leading coefficients of the expansions for $y_1$ and $z_1$ over $\infty$ are the same we deduce $a = 0$. Thus $w_1$ is a constant.

Now, assume $w_1$ is not a constant, and let $w_2$ be a conjugate of $w_1$ over $F(z_1)$. If $w_2$ has no $x^{1/n}$ term in its Puiseux expansion over $\infty$, then $w_2+z_1$ would be a conjugate of $y_1$ with leading term $a_{-1}x^{1/n}$. Thus, $w_2+z_1 = y_1$, and $w_2 = w_1$. Therefore, $w_2$ has leading term $ax^{1/n}$ where $a \neq 0$. We have

31) $y_r = z_1+w_2$ for some integer $r$.

Look at conjugates of $z_1$ over $F(w_2, x)$. If $z_s$ is conjugate to $z_1$ over $F(w_2, x)$, then for some integer $t$,

32) $y_t = z_s+w_2$.

An examination of the $x^{1/n}$ terms in 31) and 32) yields $\zeta_n^{r-1} = 1+ +aa_{-1}^{-1}$ and $\zeta_n^{t-1} = \zeta_n^{s-1}+aa_{-1}^{-1}$.

We obtain:

33) $\zeta_n^{r-1}-1 = \zeta_n^{t-1}-\zeta_n^{s-1}$ with $s \neq 1$.

If we arrange the $n$th roots of one in a regular polygon about the origin in the complex plane, we see that 33) implies that two pairs of vertices of this polygon are separated by parallel line segments of the same length. This can only happen when $n$ is even, contrary to $(n, 2) = 1$.

We conclude from the above argument that $z_1$ has no conjugates over $F(w_2, x)$, or $F(z_1, y_r) = F(w_2, x)$. Since $w_2$ was obtained as a conjugate of $w_1$ over $F(z_1)$, for some $\sigma \in G(\Omega_x/F(x))$, $\sigma w_2 = w_1$, $\sigma y_r = y_1$, $\sigma z_1 = z_1$. From this we obtain $F(z_1, y_1) = F(w_1, x)$. ■

If Lemma 3 were true without the assumption $(n, 2) = 1$, the proof of our theorem would apply to the case degree $f = 2^r$. However, there exist polynomials $f, g \in \boldsymbol{Q}[y]$ such that $f(ay+b) \neq g(y)$ for $a, b \in \boldsymbol{Q}$, degree $f$ = degree $g = 4$, and 25) holds.

Consider:

34) $(y^2+yz+\frac{1}{2}z^2+1)(y^2-yz+\frac{1}{2}z^2+1) = y^4+2y^2+\frac{1}{2}+z^4/4+z^2+\frac{1}{2}$ where $f(y) = y^4+2y^2+\frac{1}{2}$ and $g(y) = y^4/4+y^2+\frac{1}{2}$.

LEMMA 4. (Proposition 7 of [7].) *Let $f, g \in K[y]$. Assume:*

35) *$f(y)-g(z)$ is reducible, but*

36) *$f_1(y)-g_1(z)$ is not reducible if degree $f_1 <$ degree $f$ or degree $g_1 <$ degree $g$, and $f_1$, $g_1$ are composition factors (respectively) of $f$ and $g$.*

*Then*

37) *degree $f$ = degree $g$, and $\Omega_{f-x} = \Omega_{g-x}$.*

*As usual, let $y_1, \ldots, y_n$ be the zeros of $f(y)-x$, and $z_1$ a zero of $g(z)-x$. Let $y_1, y_{a_2}, \ldots, y_{a_s}$ be the conjugates of $y_1$ over $K(z_1)$.*

*Then,*

38) $K(z_1)$ = *field obtained by adjoining to* $K$ *the symmetric polynomials in* $y_1, y_{a_2}, \ldots, y_{a_s}$.

*In addition, if there exists* $x_0 \in K^* \cup \{\infty\}$ *such that*

39) $f(y)-x_0$ *has a zero of multiplicity* $p^u$ (*where* $p$ *is some rational prime*) *and* $p^u$ *does not divide the multiplicity of any other zero of* $f(y)-x_0$, *then*

40) $az_1+b = y_1+y_{a_2}+\ldots+y_{a_s}$ *for some constants* $a \neq 0$, $b \in K^*$.

Proof of Theorem. If we replace $g(y)$ by some composition factor (say $g_1(y)$, where $g(y) = g_1(g_2(y))$), then the hypotheses 35) and 36) of Lemma 4 are satisfied. By Lemma 1 we may assume that both $g_1$ and $g_2$ are in $K[y]$.

With these assumptions we proceed to show that there exist constants $a, b$ such that $f(ay+b) = g(y)$. Since degree $f = p^u$, the hypothesis 39) is satisfied for $x_0 = \infty$. Thus, there exist constants $a, b \in K^*$ such that 40) holds. With no loss we may assume that $g(y)$ (but not necessarily $f(y)$) is monic. Let $d$ be the coefficient of the leading term in $f(y)$.

Factor $f(y)-g(z)$ into absolutely irreducible factors in $K[y, z]$ to obtain

$$41)\quad f(y)-g(z) = \prod_{i=1}^{t} h_i(y, z).$$

Let $H_i(y, z)$ be the highest degree term of $h_i(y, z)$, so that

$$42)\quad dy^n-z^n = \prod_{1}^{t} H_i(y, z).$$

This expression was introduced by Schinzel and earlier by MacCluer.

Since $n = p^r$ where $p \neq 2$, an $n$th root of $d$ is not in $K(\zeta_n)$ unless $d$ is an $n$th power from $K$. If $d$ is an $n$th power from $K$ by a change of variable ($y \to \sqrt[n]{d}\, y$) we could then assume $d$ is 1.

Let $\sqrt[n]{d}$ be any one of the zeros of $x^n-d$. Suppose $\sqrt[n]{d}\, y-z \mid H_i(y_i z)$. Then $\dfrac{H_i(y, z)}{\sqrt[n]{d}\, y-z}$ has coefficients in $K(\sqrt[n]{d})$. Otherwise, the coefficients of $h_i(y, z)$ would not be in $K(\sqrt[n]{d})$ and $h_i(y, z)$ would have a conjugate $h_j(y, z)$ over $K(\sqrt[n]{d})$. This would imply that $\sqrt[n]{d}\, y-z \mid H_i$ and $H_j$, which is contrary to the fact that all factors of $dy^n-z^n$ are simple.

Assumption 23) for arbitrary $n$ may be phrased as $[K(\zeta_n): K]$ equals $[Q(\zeta_n): Q]$. This implies for $n = p^r$, $p \neq 2$, that the only polynomials dividing $dy^n - z^n$ having coefficients in $K(\sqrt[n]{d})$, are polynomials of form

43) $\Psi(\sqrt[n]{d}\, y, z) = \prod_{j=1}^{l} \varphi_{r_j}(\sqrt[n]{d}\, y, z)$ where $\prod_{1}^{r} \varphi_i(y, z) = y^n - z^n$ and $\varphi_i(y, 1)$ has as its zeros the primitive $p^i$th roots of 1.

The expansions for $y_i, z_j$ over $\infty$ are of the form 29) and 30) where $a_{-1} = \sqrt[n]{d}$ and $b_{-1} = 1$. If, in the expression $H_1(y_{a_j}, z_1)$ we let the variable $x^{1/n} \to \infty$, then this expression approaches $h_1(y_{a_j}, z) = 0$. Therefore, $y = \sqrt[n]{d}\,\zeta^{a_i - 1} z$ is one of the zeros of $H_1(y, z)$ (as a polynomial in $z$).

Since $\dfrac{H_1(y, z)}{\sqrt[n]{d}\, y - z}$ is of form 43),

44) $\{\zeta^{a_2 - 1}, \ldots, \zeta^{a_s - 1}\}$ run over the union of all primitive $p^{rj}$ th roots of 1, for $j = 1, \ldots, l$.

Again, remember $n = p^r$. In the equation 40), consider the Puiseux expansion over $\infty$ of the right side of 40). The coefficient of $x^{-m/n}$ for $(m, p) = 1$, is

45) $a_m(1 + \zeta^{(a_2 - 1)m} + \ldots + \zeta^{(a_s - 1)m}) = a_m c_m$.

From 44), $c_m$ is independent of $m$ if we restrict $m$ to integers relatively prime to $p$. With this restriction we let $c_m = c$. Thus, $\dfrac{a}{c} z_1$ and $y_1$ have Puiseux expansions at $x = \infty$ whose coefficients differ only at terms of form $x^{-u/n}$ where $(u, n) \neq 1$.

Let $z_1^* = \dfrac{a}{c} z_1$. From Lemma 3, if $w_1$ is not a constant we must have

46) $K(w_1, x) = K(z_1^*, y_1)$ where $w_1 = z_1^* - y_1$.

However, $w_1$ only has terms in its expansion over $\infty$ of form $x^{-u/n}$ where $p \mid u$. Therefore $y_1$ cannot be a rational function of $x$ and $w_1$. Thus $w_1$ must be a constant, and $a' z_1 + b' = y_1$ for some constants $a', b'$. ■

## References


[1] J. W. S. Cassels, *Factorization of polynomials in several variables*, Proc. 15th Scandinavian Congress, Oslo 1968 (Lecture Notes in Mathematics 118), pp. 1–17.

[2] M. Fried and R. E. MacRae, *On the invariance of chains of fields*, Illinois Journ. Math. 13 (1969), pp. 165–171.

[3] — — *On curves with separated variables*, Math. Ann. 180 (1969), pp. 220–226.

[4] — *Arithmetical properties of value sets of polynomials*, Acta Arith. 15 (1969), pp. 91–125.

[5] M. Fried, *Arithmetical properties of value sets of polynomials, II*, Acta Arith. to appear.

[6] — *On a conjecture of Schur*, Michigan Math. Journ. 17 (1970), pp. 41–55.

[7] — *Families of Riemann surfaces with application to Diophantine problems, I*, Number Theory Journal, to appear.

[8] — *Families of Riemann surfaces with application to Diophantine problems, II*, in preparation.

INSTITUTE FOR ADVANCED STUDY
Princeton, New Jersey
STATE UNIVERSITY OF NEW YORK
Stony Brook, L.I., New York

*Received on 11. 3. 1970* (58)