

## COUNTING SOLUTIONS TO TRINOMIAL THUE EQUATIONS: A DIFFERENT APPROACH

EMERY THOMAS

ABSTRACT. We consider the problem of counting solutions to a trinomial Thue equation — that is, an equation

$$(*) \quad |F(x, y)| = 1,$$

where  $F$  is an irreducible form in  $\mathbb{Z}[x, y]$  with degree at least three and with three non-zero coefficients. In a 1987 paper J. Mueller and W. Schmidt gave effective bounds for this problem. Their work was based on a series of papers by Bombieri, Bombieri-Mueller and Bombieri-Schmidt, all concerned with the “Thue-Siegel principle” and its relation to  $(*)$ . In this paper we give specific numerical bounds for the number of solutions to  $(*)$  by a somewhat different approach, the difference lying in the initial step — solving a certain diophantine approximation problem. We regard this as a real variable extremal problem, which we then solve by elementary calculus.

### 1. INTRODUCTION

We consider the Diophantine equation

$$(*) \quad |F(x, y)| = 1,$$

where  $F$  is an irreducible form in  $\mathbb{Z}[x, y]$  with degree at least three.  $F$  is often called a Thue form and  $(*)$  the Thue equation. In 1909 Thue [23] showed that  $(*)$  has only a finite number of solutions. Davenport and Roth [6] — using material from the proof of Roth’s theorem [18] — gave an effective upper bound for the number of solutions to  $(*)$ . In a recent series of papers (Bombieri-Schmidt [5], Mueller-Schmidt [16], [17]) Bombieri, Mueller and Schmidt obtain an effective upper bound by quite a different method. This approach uses the “Thue-Siegel principle”, in the sense of Bombieri [3] and as further developed in Bombieri-Mueller [4] and Bombieri-Schmidt [5]. At the end of the section we give a brief summary of the Bombieri-Mueller-Schmidt method. (Note important earlier work by Siegel [21], Dyson [7], Mahler [13], Lewis-Mahler [12] and Evertse [8].)

In this paper we suggest a somewhat different approach to the above problem. We apply this approach to give explicit numerical bounds when  $F$  is a *trinomial* (= 3 non-zero coefficients) Thue form (cf. [16]). One can also treat Thue tetranomials (= 4 non-zero coefficients) in a similar way.

We regard a solution  $(p, q)$  and its negative  $(-p, -q)$  as *equivalent*, and so count the pair as a single solution. Also, we do not count *trivial* solutions — i.e., solutions

---

Received by the editors May 23, 1997 and, in revised form, July 29, 1998.

2000 *Mathematics Subject Classification*. Primary 11D41, 11J68; Secondary 11Y50.

*Key words and phrases*. Thue equation, Thue-Siegel principle, diophantine approximation, trinomial equation, counting solutions.

TABLE 1.

$n$	5	6	7	8	9	10-11	12-16	17-37	$\geq 38$
$w(n)$	27	16	13	11	9	8	7	6	5

with  $|pq| \leq 1$ . We will say that a solution  $(p, q)$  to  $(*)$  is *regular* if:

$$(1.1) \quad p \neq 0, \quad |p| \neq q, \quad q > 0.$$

Note that every non-trivial solution is equivalent to a unique regular solution. For any Thue form  $F$ , set

$$N_F = \text{number of regular solutions to } (*) \text{ for } F.$$

In order to state the main theorem — giving an upper bound for  $N_F$ , for Thue trinomials — we need two numerical functions,  $v$  and  $w$ . For  $n$  a positive integer, set

$$(1.2) \quad v(n) = \begin{cases} 3, & \text{if } n \text{ odd,} \\ 4, & \text{if } n \text{ even.} \end{cases}$$

Define  $w(n)$  by Table 1.

We will prove

**Theorem 1.1.** *Let  $F(x, y)$  be a trinomial Thue form of degree  $n$ , where  $n \geq 5$ . Then,*

$$N_F \leq v(n) \cdot w(n).$$

Thus if  $n = 6$ , then  $N_F \leq 64$ , while if  $n = 39$ , then  $N_F \leq 15$ .

Theorem 1.1 will follow from a second theorem, proved in section 2, which gives a more precise value for the upper bound.

*Remark 1.* Mueller–Schmidt [16] show that there is an effectively computable universal constant bounding  $N_F$  for all Thue trinomials, though no explicit numerical computations are given in [16]. (See also [17].) As noted below, the approach taken here differs from that used in [16] at several key points.

**Comparison of methods:** We describe briefly the steps in the Bombieri–Mueller–Schmidt approach. We then indicate where the approach taken here differs. The common starting point is a Thue form  $F$  together with a regular solution  $(p, q)$  to  $(*)$  for  $F$ ; set  $f(x) = F(x, 1)$ .

*Step 1.* With the solution  $(p, q)$  associate the (real or complex) root  $\omega$  of  $f(x)$  closest to  $p/q$ . The first step then consists in giving an effective answer to the question: how well does  $p/q$  approximate  $\omega$ ?

*Step 2.* Let  $Y$  be a (large) number. Using the data from the approximation result above together with the “gap principle”, one shows that there is an effective (and efficient) procedure for bounding the number of solutions  $(p, q)$  with  $q \leq Y$ . (In fact the  $q$ ’s increase very rapidly.)

*Step 3.* Using data from  $F$ , one constructs a large number  $Y_F$ . Suppose that  $(p, q)$  and  $(p', q')$  are two solutions to  $(*)$  with  $Y_F \leq q < q'$ . By the “Thue–Siegel principle”, one shows that  $q'$  is then bounded in terms of  $q$  and  $F$ . And thus one is able to bound (using the gap principle again) the number of solutions  $(p, q)$  with  $q > Y_F$ .

The approach taken here differs from the above mainly in Step 1. Here, given a form  $F$  and solution  $(p, q)$ , we associate to  $(p, q)$  either a *real root* of  $f(x)$  or a *real root* of  $f'(x)$ , i.e. a *critical point* of  $f(x)$ . (Only certain critical points—called *proper*—will be used.) We call the set of all these roots the *exceptional set* of  $F$ . If  $(p, q)$  is associated with an exceptional point  $\tau$ , we again have the question: how well does  $p/q$  approximate the (real) number  $\tau$ ? The method used here to solve this approximation problem is quite different from that used above. We work with polynomials with real coefficients, and we regard a given Thue trinomial as a member of a 1-parameter family of real trinomials. In this family we then select one or two trinomials called *maximal*. These trinomials have two key properties: first, the approximation problem is readily solved for these trinomials (and the solution is “almost” sharp); and second, the solution for these trinomials is “maximal” for all trinomials in the family. In this way we solve the approximation problem for an arbitrary Thue trinomial. With Step 1 completed, Steps 2 and 3 follow much as above.

*Remark 2.* There is an entirely different approach to studying  $(*)$ , initiated independently by Gelfond [10] and Schneider [20], and culminating in the fundamental work of Baker [1] and of Feldman [9] on linear forms in logarithms. One consequence of this work is the celebrated theorem of Baker [2]: *if  $(p, q)$  is a solution to  $(*)$ , then  $|p|$  and  $|q|$  are effectively bounded.* However, this method does not seem to work if one wishes to find a *universal* bound for  $N_F$  for an *infinite family* of Thue forms—e.g., all trinomials or all forms of a given degree. See Bombieri [3] for a discussion of pre-1980 work — also Schmidt [19] and Stewart [22] for more recent work.

*Remark 3.* Equation  $(*)$  is the “classical” form of the Thue equation. In recent years considerable research has been done on bounding the number of solutions to Thue *inequalities*—see, for example, Mueller–Schmidt [16],[17]. In a subsequent note we will indicate how the approach taken here can be modified to give explicit numerical bounds for the number of solutions to Thue inequalities for trinomial Thue forms.

We conclude the section by giving a brief description of the remaining material. In section 2 we define the exceptional set and use this to state a more precise version of Theorem 1.1. Section 3 contains several results on approximation. In section 4 we show how the gap principle enables one to bound the number of “small” solutions to  $(*)$ , while section 5 does the same thing for “large” solutions using the Thue–Siegel principle. In sections 6 and 7 we use these results to calculate the specific numbers occurring in Theorem 1.1. Section 8 is the heart of the new approach. We define the *maximal* polynomials and show how, using these, one proves the main approximation theorem in section 3. In section 9 we give conditions for finding upper and lower bounds for “small” roots of polynomials; we then use this result in section 10 to estimate the roots of the maximal polynomials. Finally, in section 11 we prove a different approximation theorem, involving the height  $H$  of the polynomial.

## 2. THE EXCEPTIONAL SET

We “localize” the problem of bounding  $N_F$  as follows. Consider polynomials with real coefficients, say  $g(x)$ , with  $\deg g(x) = n \geq 5$ . (From now on we assume that

TABLE 2.

$n$	5	6-7	8	$\geq 9$
$z(n)$	8	4	3	2

any such polynomial has a positive leading coefficient.) As usual, a real number  $\lambda$  is a *real root* of  $g(x)$  if  $g(\lambda) = 0$ , and a real number  $\mu$  is a *real critical point* if  $g'(\mu) = 0$ .

**Definition 2.1.** We will say that a real critical point  $\mu$  is *proper* (resp. *improper*) if  $g(\mu) \neq 0$  and if there is a deleted neighborhood  $U$  of  $\mu$  such that

$$g''(x)g(x) > 0 \quad (\text{resp. } < 0),$$

for all  $x$  in  $U$ .

We define the exceptional set of  $g(x)$ ,  $\mathcal{E}(g)$ , to be *the set of all real roots and all proper critical points*. (Note Lagrange [11].)

Suppose that  $F(x, y)$  is a Thue form with associated polynomial  $f(x) (= F(x, 1))$ . Define

$$(2.1) \quad \begin{aligned} R_F &= \text{number of real roots of } f(x), \\ C_F &= \text{number of non-zero proper critical points of } f(x). \end{aligned}$$

Recall the function  $w(n)$  defined in Table 1. Define a numerical function  $z(n)$  by Table 2.

Our main result is

**Theorem 2.1.** *Let  $F(x, y)$  be a trinomial Thue form with  $\deg F = n \geq 5$ . Then,*

$$N_F \leq w(n)R_F + z(n)C_F.$$

For example, suppose that  $F(x, y)$  is a trinomial Thue form such that  $f(x)$  has a non-zero proper critical point. If  $\deg F = 6$ , then  $N_F \leq 36$ ; while if  $\deg F = 39$ , then  $N_F \leq 7$ .

Let  $v(n)$  be the function defined in (1.2). Since  $F(x, y)$  is a *trinomial* form, it follows that  $R_F + 2C_F \leq v(n)$ . Also,  $z(n) < w(n)$ , and so Theorem 1.1 follows at once from Theorem 2.1.

In the remainder of this section we show how Theorem 2.1, in turn, follows from a result (Theorem 2.2) bounding the number of solutions that “belong to” a given exceptional point. The rest of the paper is then devoted to proving Theorem 2.2.

We continue to work with polynomials  $g(x)$  in  $\mathbb{R}[x]$ . Let  $\mathcal{E}(g)$  be the exceptional set of  $g$  as above, say

$$(2.2) \quad \mathcal{E}(g) = \{\tau_1, \dots, \tau_c\} \quad \text{with} \quad \tau_1 < \dots < \tau_c, \quad c \geq 1.$$

By calculus we then have

**Lemma 2.1.** *Let  $g(x)$  be a polynomial in  $\mathbb{R}[x]$ , with exceptional set  $\mathcal{E}(g)$ . If  $c > 1$ , there are improper critical points  $\eta_1, \dots, \eta_{c-1}$  such that*

$$\tau_1 < \eta_1 < \tau_2 < \dots < \eta_{c-1} < \tau_c.$$

For completeness, we set  $\eta_0 = -\infty$ ,  $\eta_c = +\infty$ , and we write  $J_i = [\eta_{i-1}, \eta_i]$ ,  $1 \leq i \leq c$ . Note that  $\tau_i$  is in  $J_i$ . (If  $c = 1$ ,  $J_1 = \mathbb{R}$ .)

**Definition 2.2.** Given any real number  $\rho$ , we say that  $\rho$  *belongs to*  $\tau_i$  (and  $\tau_i$  *belongs to*  $\rho$ ), if  $\rho$  is in  $J_i$ .

Let  $F(x, y)$  be a Thue form with  $f(x)$  the associated Thue polynomial. Suppose that  $(p, q)$  is a regular solution to  $(*)$  for  $F$ ; then,

$$F(p, q) = w, \quad w = \pm 1,$$

and so,  $f(\rho) = \alpha$ , where

$$\rho = p/q, \quad \alpha = w/q^n, \quad 0 < |\alpha| \leq 1, \quad \rho \neq 0.$$

We carry this over to polynomials  $g(x)$  in  $\mathbb{R}[x]$ .

**Definition 2.3.** A pair of non-zero real numbers  $(\rho, \alpha)$ , with  $0 < |\alpha| \leq 1$ , is a *solution* for  $g(x)$  if  $g(\rho) = \alpha$ .

By Definition 2.2 we then have

**Proposition 2.1.** Let  $g(x)$  be a polynomial in  $\mathbb{R}[x]$  and let  $(\rho, \alpha)$  be a solution for  $g(x)$ . Then there is a unique exceptional point of  $g$ , say  $\tau$ , such that  $\rho$  belongs to  $\tau$ .

*Remark 1.* Let  $(p, q)$  be a pair of co-prime integers. By analogy with (1.1), we will say that  $(p, q)$  is a *regular solution* for  $g(x)$  if (1.1) holds and  $g(p/q) = \pm 1/q^n$ . If  $p/q$  belongs to the exceptional point  $\tau$  we often will say that *the solution*  $(p, q)$  *belongs to*  $\tau$ .

Suppose now that  $F(x, y)$  is a Thue form and that  $\tau$  is an exceptional point for  $f(x)$  ( $= F(x, 1)$ ). We set

$N_F(\tau)$  = number of regular solutions  $(p, q)$  to  $(*)$  such that  $p/q$  belongs to  $\tau$ .

Note that if  $\tau = 0$ , then  $N_F(\tau) = 0$ ; thus, by Proposition 2.1, we obtain

**Corollary 2.1.** For any Thue form  $F$ ,

$$N_F = \sum_{\tau} N_F(\tau),$$

where the sum is over all non-zero exceptional points in  $\mathcal{E}(f)$ .

For any Thue form  $F$ , define

$S_F$  = set of all regular solutions to  $(*)$  for  $F$ .

Thus,  $N_F = |S_F|$ .

We proceed to partition  $S_F$ . Denote by  $\mathbb{R}^0$  the subset of  $\mathbb{R}$  obtained by deleting the three points  $\{-1, 0, 1\}$ . Partition  $\mathbb{R}^0$  into four disjoint subsets,  $\mathbb{R}(1), \dots, \mathbb{R}(4)$ , by setting:

$$\mathbb{R}(1) = (-\infty, -1), \quad \mathbb{R}(2) = (-1, 0), \quad \mathbb{R}(3) = (0, 1), \quad \mathbb{R}(4) = (1, +\infty).$$

Using this, we partition  $S_F$ . For  $1 \leq j \leq 4$ , define

$$S_F(j) = \{(p, q) \in S_F \mid p/q \text{ is in } \mathbb{R}(j)\}.$$

Thus,

$$(2.3) \quad S_F = \bigcup_{j=1}^4 S_F(j), \quad S_F(i) \cap S_F(j) = \emptyset, \quad i \neq j.$$

With each Thue form  $F(x, y)$  we associate three other Thue forms:

$$(2.4) \quad F_1(x, y) = \pm F(-x, y), \quad F_2(x, y) = \pm F(-y, x), \quad F_3(x, y) = \pm F(y, x),$$

where in each case the sign is chosen so that the new form has positive coefficient on  $x^n$ . We also set

$$(2.5) \quad F_0(x, y) = F(x, y).$$

We then have the following important fact.

**Lemma 2.2.** *Suppose that  $F$  is a fixed Thue form and that  $(p, q)$  is in  $S_F$ . Then,  $(-p, q)$  is a solution for  $F_1$ ,  $(-q, p)$  is a solution for  $F_2$ , and  $(q, p)$  is a solution for  $F_3$ .*

Consequently, if we define

$$S_i(j) = S_{F_i}(j), \quad 0 \leq i \leq 3, \quad 1 \leq j \leq 4,$$

then by Lemma 2.2, we obtain

**Corollary 2.2.** *For  $1 \leq i \leq 3$ ,  $|S_i(4)| = |S_0(i)|$ .*

Thus, by (2.3) we have

$$(2.6) \quad N_F = \sum_{i=0}^3 |S_i(4)|.$$

This motivates our next result.

**Theorem 2.2.** *Let  $F(x, y)$  be a trinomial Thue form, with degree  $F = n \geq 5$ . Suppose that  $\tau$  is an exceptional point of  $f(x)$  ( $= F(x, 1)$ ) and that  $\tau$  is in  $\mathbb{R}(4)$ , i.e.,  $\tau > 1$ . Then,*

$$N_F(\tau) \leq \begin{cases} w(n), & \text{if } \tau \text{ is a root } \lambda, \\ z(n), & \text{if } \tau \text{ is a proper critical point } \mu. \end{cases}$$

The remainder of the section is devoted to showing that Theorem 2.2 implies Theorem 2.1. We then prove Theorem 2.2 in sections 6 and 7.

Suppose then that  $F(x, y)$  is a given trinomial Thue form, and let  $F_i$ ,  $0 \leq i \leq 3$ , be the Thue forms defined in (2.4)–(2.5). Set

$$(2.7) \quad f_i(x) = F_i(x, 1).$$

In counting solutions, we need to distinguish between those that belong to a root and those that belong to a proper critical point. For  $0 \leq i \leq 3$ ,  $1 \leq j \leq 4$ , we set

$$\begin{aligned} R_i(j) &= \text{number of real roots of } f_i(x) \text{ that lie in } \mathbb{R}(j), \\ C_i(j) &= \text{number of proper critical points of } f_i(x) \text{ that lie in } \mathbb{R}(j), \\ N_i(j, R) &= \text{number of solutions in } S_i(j) \text{ that belong to a root,} \\ N_i(j, C) &= \text{number of solutions in } S_i(j) \text{ that belong to a proper critical point.} \end{aligned}$$

We need to establish some elementary facts before we will be able to show that Theorem 2.2 implies Theorem 2.1. Note that if  $g(x)$  is any trinomial, then  $g$  has at most two non-zero real critical points, and these—if they both occur—have opposite signs. Using this fact, one readily sketches the graph of  $y = g(x)$ , from which one reads off the following information. (We now take  $g(x) = f(x) = f_i(x)$ , as given in (2.7).)

**Lemma 2.3.** *Let  $F(x, y)$  be a trinomial Thue form, with  $F_i$  and  $f_i$ ,  $0 \leq i \leq 3$ , as defined above. Then, for  $1 \leq j \leq 4$ ,*

- (i)  $0 \leq R_i(j) \leq 2$ ,  $0 \leq C_i(j) \leq 1$ .
- (ii)  $R_i(j) \cdot C_i(j) = 0$ .

The proof is immediate, from the above remarks.

For any integer  $c$ ,  $|f(c)| \geq 1$ , since  $f(x)$  is irreducible and is in  $\mathbb{Z}[x]$ . As a consequence we have

**Lemma 2.4.** *Let  $(p, q)$  be a solution in  $S_F(j)$ ,  $1 \leq j \leq 4$ , and let  $\tau$  be the exceptional point that belongs to  $p/q$ . Then  $\tau$  is in  $\mathbb{R}(j)$ .*

Using these facts, we prove

**Lemma 2.5.** *Let  $F(x, y)$  be as above. Then, for  $1 \leq i \leq 3$ ,*

- (i)  $R_i(4) = R_0(i)$ .
- (ii)  $N_i(4, R) = N_0(i, R)$ .
- (iii)  $N_0(i, C) = N_i(4, C)$ .
- (iv) *If  $N_0(i, C) \neq 0$ , then  $C_0(i) = C_i(4)$ .*

*Proof.* Let  $\lambda$  be any real root of  $f(x)$ . Then,  $-\lambda$  is a root of  $f_1(x)$ ,  $-\lambda^{-1}$  is a root of  $f_2(x)$  and  $\lambda^{-1}$  is a root of  $f_3(x)$ . Using this, one easily deduces (i) above.

To prove Lemma 2.5(ii), suppose that  $N_0(i, R) > 0$ , for some  $i$ ,  $1 \leq i \leq 3$ . Then, by Lemma 2.4,  $R_0(i) > 0$ , and so by Lemma 2.3(ii),  $C_0(i) = 0$ . Consequently,  $|S_0(i)| = N_0(i, R)$ . But by Lemma 2.5(i),  $R_i(4) = R_0(i) > 0$ , and so by the same argument,

$$N_i(4, R) = |S_i(4)| = |S_0(i)| = N_0(i, R),$$

by Corollary 2.2. Since the argument is reversible,  $N_0(i, R) = 0$  implies  $N_i(4, R) = 0$ . This proves part (ii) above. The proof for parts (iii) and (iv) is similar and is left to the reader.

*Remark 2.* In general it is *not* true that  $C_0(3) = C_3(4)$ . See for example the form  $x^5 - xy^4 + y^5$ .

We now have:

*Proof of Theorem 2.1* (assuming Theorem 2.2). Suppose then that  $F(x, y)$  is a trinomial Thue form, with degree  $F = n \geq 5$ , and that  $F_i$  and  $f_i$  are the forms and polynomials defined above, for  $0 \leq i \leq 3$ . Then, by Corollary 2.1,

$$N_F = \sum_1^4 N_0(j, R) + \sum_1^4 N_0(j, C).$$

But by Lemma 2.5 and Theorem 2.2, for  $1 \leq i \leq 3$ ,

$$N_0(i, R) = N_i(4, R) \leq w(n)R_i(4) = w(n)R_0(i).$$

Therefore,

$$\sum_1^4 N_0(j, R) \leq w(n) \sum_1^4 R_0(j) = w(n)R_F.$$

By a similar argument,  $N_0(j, C) \leq z(n)C_0(j)$ , and so,

$$\sum_1^4 N_0(j, C) \leq z(n)C_F.$$

Thus,

$$N_F \leq w(n)R_F + z(n)C_F,$$

as required.

## 3. APPROXIMATION

We are left with the proof of Theorem 2.2. For this we will need a result from this section on Diophantine approximation — see Theorem 3.2. Combining this result with the “gap principle” (section 4) and the “Thue-Siegel principle” (section 5), we then prove Theorem 2.2 in sections 6 and 7. The remainder of the paper, sections 8 through 11, is devoted to the proof of Theorem 3.2.

We begin by studying the approximation problem in a broader setting. Suppose, as in section 2, that  $g(x)$  is a polynomial with *real* coefficients, where degree  $g = n \geq 5$ . Let  $(\rho, \alpha)$  be a solution for  $g(x)$ , see Definition 2.3. We say that  $(\rho, \alpha)$  is a *special solution* if

$$(3.1) \quad \begin{aligned} & \text{(i)} \quad \rho > 1, \\ & \text{(ii)} \quad \rho \geq p_0(n) \cdot |\alpha|^{1/n}, \end{aligned}$$

where

$$(3.2) \quad p_0(n) = \begin{cases} 3, & \text{if } 5 \leq n \leq 8, \\ 2, & \text{if } n \geq 9. \end{cases}$$

For a given special solution  $(\rho, \alpha)$ , if  $\tau$  is the unique exceptional point of  $g(x)$  that belongs to  $\rho$ , how well does  $\rho$  approximate  $\tau$ ?

Suppose that

$$(3.3) \quad g(x) = ax^n - sx^k + t, \quad n > k > 0, \quad n \geq 5$$

where  $a$ ,  $s$  and  $t$  are non-zero real numbers with  $a \geq 1$ . We set

$$(3.4) \quad \Delta(a, n, k) = \sqrt{\frac{2}{a(n-2)(n-k)}}.$$

Our main technical result is

**Theorem 3.1.** *Let  $g(x)$  be a trinomial in  $\mathbb{R}[x]$ , as above, and suppose that  $(\rho, \alpha)$  is a special solution for  $g(x)$ . Let  $\tau$  be the unique exceptional point of  $g(x)$  that belongs to  $\rho$ . Then,*

$$|\tau - \rho| < \Delta \sqrt{\frac{|\alpha|}{\rho^{n-2}}}, \quad \Delta = \Delta(a, n, k).$$

The proof is given in sections 8–10. We note several corollaries to the theorem. Let  $p$  and  $q$  be positive co-prime integers. Set

$$\rho = p/q, \quad \alpha = \pm 1/q^n.$$

We have at once

**Lemma 3.1.**  *$(\rho, \alpha)$  is a special solution for a polynomial  $g(x)$  in  $\mathbb{R}[x]$  if and only if*

$$g(\rho) = \alpha, \quad p > q, \quad \text{and} \quad p \geq p_0.$$

Combining Theorem 3.1 with the above result, we obtain

**Corollary 3.1.** *Suppose that  $g(x)$  is a trinomial in  $\mathbb{R}[x]$ , as in (3.3), and that  $(p/q, \pm 1/q^n)$  is a special solution for  $g(x)$ , as in Lemma 3.1. Let  $\tau$  be the unique exceptional point of  $g(x)$  that belongs to  $p/q$ . Set  $n^* = (n-2)/2$ . Then,*

$$|\tau - p/q| < \frac{\Delta}{\sqrt{\rho^{n-2}q^n}} = \frac{\Delta}{p^{n^*}q}.$$



From this result we have

**Corollary 3.2.** *Let  $g(x)$ ,  $p/q$ , and  $\tau$  be as above. Then,  $p/q$  is a principal convergent in the continued fraction expansion of the number  $\tau$ .*

*Proof.* By hypothesis,  $a \geq 1$  and  $p \geq 2$  (see (3.3) and (3.2)). Thus, by Corollary 3.1,

$$|\tau - p/q| < 1/2q^2,$$

which implies the desired result. For a related result (on polynomials with integer coefficients) see Lagrange [11], page 680.

Given any form  $F$  (or polynomial  $f$ ) with real coefficients, we denote by  $H(F)$  (or  $H(f)$ ) the maximum of the absolute values of the coefficients. ( $H$  is commonly called the *height*.) In order to prove Theorem 2.2 we need a result analogous to Theorem 3.1 that involves the height  $H$  of  $g(x)$ . We now state such a result.

Let  $n$  be an integer  $\geq 5$ ; for  $b$  a real number, with  $0 \leq b \leq n^*$ , define a number  $K_b(n)$  as follows: Let

(3.5)

$$m_n = 2\sqrt{\frac{2n}{(n-1)(n-2)}}, \quad r_n = (2.032)^{1/n}, \quad u_n = \sqrt{\frac{2}{(n-2)p_0^n}}.$$

Define

$$(3.6) \quad K_b(n) = m_n(r_n(1 + u_n))^b.$$

We will prove

**Theorem 3.2.** *Let  $g(x)$ ,  $(p/q, \pm 1/q^n)$  and  $\tau$  be as in Corollary 3.1. Choose a number  $b$  such that  $0 \leq b \leq n^*$ ; set  $c = n^* - b$ . Then,*

$$|\tau - p/q| < \frac{K_b(n)}{2H^{b/n}p^c q^{b+1}}, \quad H = H(g).$$

The proof uses Theorem 3.1 and is given in section 11. Note that for all  $n \geq 5$  and all  $b$  as above,  $K_b(n) \leq 2.44$ .

#### 4. COUNTING SMALL SOLUTIONS: THE GAP PRINCIPLE

In this section we see how the “gap principle” can give information about the following problem.

Suppose that  $g(x)$  is a trinomial in  $\mathbb{R}[x]$ , as in section 2, and that  $\lambda$  is a real root of  $g(x)$ , with  $\lambda > 1$ . Let  $Y$  be a large positive number.

**Problem.** How many special solutions  $(p, q)$  are there (for  $g(x)$ ) that belong to  $\lambda$  and have

$$1 \leq q \leq Y?$$

We will need the following result. Combining Theorem 3.2 with the gap principle (e.g., see [6], [15], [16], [19]), we obtain

**Theorem 4.1.** *Let  $(p, q)$ ,  $(p', q')$  be two special solutions that belong to the root  $\lambda$ . Suppose that  $q' > q$ . Then,*

$$q' > \frac{H^{b/n}p^c q^b}{K_b(n)},$$

where the notation is given in Corollary 3.1, (3.6), and Theorem 3.2.

Suppose that  $t$  is an integer ( $\geq -1$ ) such that there are precisely  $t + 2$  special solutions that belong to  $\lambda$ , say  $\{(p_0, q_0), \dots, (p_{t+1}, q_{t+1})\}$ , with

$$q_i \leq Y, \quad 0 \leq i \leq t + 1.$$

By Corollary 3.2 we can index these solutions so that

$$1 \leq q_0 < \dots < q_{t+1} \leq Y.$$

Since  $(p_0, q_0)$  is a special solution, we have  $p_0 \geq p_0(n)$ , see (3.2). Choose a number  $b_0$  so that  $0 \leq b_0 \leq n^* - 1.4$ , where  $n^* = (n - 2)/2$ . Set

$$(4.1) \quad c_0 = n^* - b_0, \quad K_0 = K_{b_0}(n), \quad Q_1 = p_0^{c_0}/K_0;$$

choose  $b_0$  so that  $Q_1 > 1$ . By Theorem 4.1

$$(4.2) \quad q_1 > H^{b_0/n} Q_1, \quad p_1 > Q_1 + 1.$$

In order to bound the  $q_i$ 's for  $i > 1$ , we choose a second number  $b$ , with  $1 < b \leq n^*$ , and set

$$(4.3) \quad c = n^* - b, \quad K = K_b(n)/(Q_1 + 1)^c, \quad B = (b_0 + 1)/n.$$

Choose  $b$  (and  $b_0$ ) so that  $Q_1 > K$ . By iteration of Theorem 4.1, applied to (4.2), we obtain

$$(4.4) \quad q_{t+1} > \frac{H^{Bb^t} Q_1^{b^t}}{K^{P_t(b)}},$$

where  $P_t(x) = (x^t - 1)/(x - 1)$ ,  $t \geq 1$ .

Suppose that  $t \geq 1$ . We consider 2 cases: (i)  $K > 1$ , (ii)  $K \leq 1$ .

**Case (i),**  $K > 1$ . Then, by (4.4), since  $P_t(b) < b^t$ ,

$$(4.5) \quad q_{t+1} > H^{Bb^t} Q_2^{b^t}, \quad \text{where } Q_2 = Q_1/K > 1.$$

**Case (ii),**  $K \leq 1$ . Then by (4.4), since  $b^{t-1} \leq P_t(b)$ ,

$$(4.6) \quad q_{t+1} > H^{Bb^t} Q_3^{b^{t-1}}, \quad Q_3 = Q_1^b/K.$$

Since  $q_{t+1} \leq Y$ , by (4.5) and (4.6) we obtain

**Lemma 4.1.** (a) *In Case (i),*

$$Bb^t \log H + b^t \log Q_2 < \log Y.$$

(b) *In Case (ii),*

$$Bb^t \log H + b^{t-1} \log Q_3 < \log Y.$$

Suppose we are given a (large) positive number  $Y_F$ , that depends upon  $F$  in the following way: Assume that we have two sequences of positive numbers  $\{\chi_n\}$  and  $\{\pi_n\}$ ,  $n \geq 5$ . Define  $Y_F$  by

$$(4.7) \quad \log(Y_F) = \chi_n \log(H) + \pi_n,$$

where degree  $F = n$  ( $\geq 5$ ),  $H(F) = H$ . Set

$$(4.8) \quad \begin{aligned} T_1 &= \left\lfloor \frac{\log(\chi_n) - \log(B)}{\log(b)} \right\rfloor, \\ T_2 &= \left\lfloor \frac{\log(\pi_n) - \log \log(Q_2)}{\log(b)} \right\rfloor, \quad \text{if } K > 1, \\ T_3 &= \left\lfloor \frac{\log(\pi_n) - \log \log(Q_3)}{\log(b)} \right\rfloor, \quad \text{if } K \leq 1, \\ T^* &= \begin{cases} \max(T_1 + 2, T_2 + 2), & \text{if } K > 1, \\ \max(T_1 + 2, T_3 + 3), & \text{if } K \leq 1. \end{cases} \end{aligned}$$

By Lemma 4.1 and (4.7) we have proved

**Theorem 4.2.** *Let  $g(x)$  be a trinomial in  $\mathbb{R}[x]$ , with degree  $g = n \geq 5$ , and let  $\lambda$  be a real root of  $g(x)$  with  $\lambda > 1$ . Suppose that either  $T_1 \geq 1$  or  $T_2 \geq 1$  or  $T_3 \geq 0$ . Then, there are at most  $T^*$  special solutions  $(p, q)$  for  $g(x)$  that belong to  $\lambda$  and have*

$$1 \leq q \leq Y_F.$$

Note that  $T^* \geq 3$ . We now give conditions that ensure that there are at most 2 “small” solutions, relative to  $Y_F$ , that belong to the root  $\lambda$ . By (4.4), taking  $t=1$ , we obtain

**Corollary 4.1.** *Let  $g(x)$  be a trinomial with real coefficients and  $\lambda$  a real root of  $g(x)$ , as above. If*

$$(i) \quad Bb > \chi_n,$$

and

$$(ii) \quad b \log(Q_1) - \log(K) > \pi_n,$$

*then there are at most 2 special solutions  $(p, q)$  for  $g(x)$  that belong to the root  $\lambda$  and have*

$$1 \leq q \leq Y_F.$$

## 5. COUNTING LARGE SOLUTIONS: THE THUE–SIEGEL PRINCIPLE

In the previous section we considered the problem of counting “small” special solutions for a trinomial  $g(x)$  in  $\mathbb{R}[x]$ . In this section we assume that  $g(x) = f(x) = F(x, 1)$ , where  $F(x, y)$  is a Thue form; we take up the problem of counting “large” special solutions that belong to a real root  $\lambda$  of  $f(x)$ , where  $\lambda > 1$ . We show that by using the recent striking results of Bombieri [3], Bombieri–Mueller [4] and Bombieri–Schmidt [5], we obtain a good bound for the number of such solutions that are *large*, relative to  $Y_F$ —i.e., solutions  $(p, q)$  with  $q > Y_F$ . Here  $Y_F$  is a number that depends *inter alia* on the height  $H$  of  $F$ .

Following Bombieri–Schmidt [5], we fix an integer  $n \geq 5$  and choose numbers  $u$  and  $v$  so that

$$(5.1) \quad 0 < u < v < 1 - (2w)^{1/2}, \quad w = (n + u^2)/n^2.$$

In terms of  $n$ ,  $u$  and  $v$ , define

$$(5.2) \quad L = (2(n + u^2))^{1/2}/(1 - v), \quad D = L/(n - L), \quad U = u^{-2}.$$

Given any trinomial Thue form  $F$ , with  $n = \deg F$  and  $H = H(F)$ , define a positive number  $Y_F$  by

$$(5.3) \quad \log(Y_F) = \chi_n \log(H) + \pi_n,$$

where

$$(5.4) \quad \chi_n = D(U+1) + 1, \quad \pi_n = (D(4+U) + 2) \log(2) + (D+1) \log(n)/2 + nUD/2.$$

To bound the number of large solutions, relative to  $Y_F$ , we use the Thue–Siegel principle in the sense of Bombieri [3], as further developed in Bombieri–Mueller [4] and Bombieri–Schmidt [5]. (With a change in notation we essentially follow [5].) Set

$$(5.5) \quad E = (2(v^2 - u^2))^{-1}, \quad Z = \left\lceil \frac{\log E + 2 \log(n) - \log(L-2)}{\log(n-1)} \right\rceil + 2.$$

By Lemma 2 in Bombieri–Schmidt [5] we obtain

**Proposition 5.1.** *Let  $F(x, y)$  be a trinomial Thue form with degree  $F$  at least 5. Suppose that  $\lambda$  is a real root of  $f(x) (= F(x, 1))$ , with  $\lambda > 1$ . Let  $Y_F$  and  $Z$  be the numbers defined above. Then there are at most  $Z$  special solutions  $(p, q)$  in  $S_F$  that belong to the root  $\lambda$  and have  $q > Y_F$ .*

*Remark.* In the notation of section 3 of [5],  $Y_0 < Y_F$ , and so it is appropriate to use Lemma 2 of [5] to prove Proposition 5.1.

## 6. PROOF OF THEOREM 2.2, $\tau$ A REAL ROOT

By combining the results given by Theorem 4.2, Corollary 4.1, and Proposition 5.1, we now can prove Theorem 2.2 for the case  $\tau = \lambda$ , a real root of  $f(x)$ .

For this section let  $F(x, y)$  be a trinomial Thue form, with  $\deg F = n \geq 5$ , and let  $\lambda$  be a real root of  $f(x) (= F(x, 1))$ . By (2.6) and Lemma 2.4 we may assume that  $\lambda > 1$ . Let  $Y_F$  be the number defined in (5.3), using the values of  $u$  and  $v$  given in Table 3 below, and let  $Z$  be the integer defined in (5.5). At the end of the section we prove

**Lemma 6.1.** *If  $n \geq 38$ , then  $F(x, y)$  has at most two special solutions  $(p, q)$  that belong to  $\lambda$  and have  $1 \leq q \leq Y_F$ . Moreover,  $Z = 3$ .*

Define functions  $\delta(n)$  and  $T(n)$  by

**Definition 6.1.** (a)  $\delta(n) = 1$ , if  $5 \leq n \leq 8$ ;  $\delta(n) = 0$ , if  $n \geq 9$ .  
(b)  $T(n) = T^*$ , if  $5 \leq n \leq 37$ ;  $T(n) = 2$ , if  $n \geq 38$ .

Here  $T^*$  is the integer defined in (4.8)—we show below that the hypotheses of Theorem 4.2 are satisfied for  $5 \leq n \leq 37$ .

The integers  $T$ ,  $Z$  and  $\delta$  have the following interpretation.  $T$  is an upper bound for the number of *small* special solutions, and  $Z$  an upper bound for the number of *large* special solutions, relative to  $Y_F$ , that belong to the root  $\lambda$ . The fact that  $\delta = 0$  for  $n \geq 9$  indicates that in this case every regular solution that belongs to  $\lambda$  is also special; whereas, for  $5 \leq n \leq 8$ , there is possibly one regular solution that is not special, namely  $(2, 1)$ . Hence, for this case we set  $\delta = 1$ .

Thus, we now set

$$(6.1) \quad w(n) = T + Z + \delta(n), \quad n \geq 5.$$

TABLE 3.

$n$	5	6	7	8	9	10-11	12-16	17-37	$\geq 38$
$b_0, b$	0, 1.5	0, 2	1, 2.5	1, 3	1, 3.5	1, $n^*$	2, $n^*$	3.5, $n^*$	.65 $n^*$ , $n^*$
$u, v$	.19, .25	.16, .28	.2, .3	.25, .45	.3, .45	.3, .44	.32, .44	.41, .48	.46, .50
$T, Z$	21, 5	11, 4	8, 4	7, 3	6, 3	5, 3	4, 3	3, 3	2, 3

By the above remarks, we have  $N_F(\lambda) \leq w(n)$ .

To complete the proof of Theorem 2.2, in this case, we are left with showing that  $w(n)$  has the values shown in Table 1 (in section 1). In Table 3 we give values for  $(b_0, b)$  and  $(u, v)$  to be used for each  $n$ . One may then check that, for  $5 \leq n \leq 37$ ,  $T$  and  $Z$  have the values shown. Combining this with Lemma 6.1, we see that  $w(n)$  takes on the values given in Table 1.

We are left with proving Lemma 6.1. Using the values given in Table 3 for  $b_0$ ,  $b$ ,  $u$  and  $v$ , one finds by direct calculation that conditions (i) and (ii) in Corollary 4.1 hold for  $38 \leq n \leq 43$ . Also, for  $n \geq 44$  we have  $D \leq 3/4$  and  $U \leq 5$ ; and so,

$$\chi_n \leq 5.5 \quad \text{and} \quad \pi_n \leq (2.1)n.$$

Thus, conditions (i) and (ii) in Corollary 4.1 are satisfied for all  $n \geq 44$ . Finally, it follows directly from (5.5) that  $Z = 3$ . This completes the proof of the lemma.

## 7. PROOF OF THEOREM 2.2, $\tau$ A PROPER CRITICAL POINT

We continue to assume that  $F$  is a trinomial Thue form with degree  $F = n \geq 5$ . Suppose that  $\mu$  is a proper critical point; as in Theorem 2.2, we may assume that  $\mu > 1$ . We wish to show that the function  $z(n)$ , see Table 2, is an upper bound for  $N_F(\mu)$ .

**Theorem 7.1.** *Suppose that  $(p, q)$  is a regular solution to  $(*)$  that belongs to the proper critical point  $\mu$ . Then,*

$$1 \leq q \leq 1.32nH^{(n-1)/n}, \quad H = H(F).$$

Assuming this for the moment, we complete the proof of Theorem 2.2. Let  $t$  be a non-negative integer, and suppose there are  $t + 2$  special solutions to  $(*)$  that belong to  $\mu$ , say  $(p_i, q_i)$ ,  $0 \leq i \leq t + 1$ . By Corollary 3.2 we assume that

$$1 \leq q_0 < q_1 < \cdots < q_{t+1},$$

and by Theorem 7.1 we have

$$(7.1) \quad \log(q_{t+1}) \leq ((n-1)/n) \log(H) + \log(1.32n).$$

Suppose that  $n \geq 8$ . As in section 4, set  $n^* = (n-2)/2$ , and choose  $b_0 = 1.5$  and  $c_0 = n^* - 1.5$ ,  $b = n^*$ ,  $c = 0$ . Using these numbers and (4.1), (4.3), define the numbers  $K_0$ ,  $Q_1$  and  $K$ . Note that  $Q_1 > K$ . One easily shows that, for  $n \geq 8$ ,

$$(b_0 + 1)b > (n-1), \quad \text{and} \quad b \log(Q_1) - \log(K) > \log(1.32n).$$

Thus, by (4.4) and (7.1), we must have  $t = 0$ , i.e., there are at most 2 special solutions to  $(*)$  that belong to  $\mu$ .

Recall that in section 6 we introduced the numerical function  $\delta(n)$  that measures how many solutions there could be in  $S_F(4)$  that are *not* special. Thus we have shown: *If  $n \geq 8$ , then*

$$N_F(\mu) \leq 2 + \delta(n) = z(n).$$

Thus we have completed the proof of Theorem 2.2, for  $n \geq 8$ .

TABLE 4.

$n$	5	6	7
$b_0, b$	0, 1.3	.3, 2	0, 2.5
$T$	6	2	2

Suppose now that  $5 \leq n \leq 7$ . In Table 4 we give the values of  $b_0$  and  $b$  to take for each such  $n$ . Using these, set  $Q_2 = Q_1/K$ , as in (4.5). We find that if we take  $t = T$  (see the table), then

$$(i) \quad (b_0 + 1)b^t > (n - 1)$$

and

$$(ii) \quad b^t \log(Q_2) > \log(1.32n), \quad Q_2 = Q_1/K.$$

Consequently, we must have  $t \leq T - 1$ , and so there are at most  $T + 1$  special solutions that belong to  $\mu$ . Since  $\delta(n) = 1$  for  $5 \leq n \leq 7$ , there are at most

$$T + 2 = z(n)$$

regular solutions that belong to  $\mu$ . This completes the proof of Theorem 2.2.

*Proof of Theorem 7.1.* Set  $f(x) = F(x, 1)$  and suppose that

$$f(x) = ax^n - sx^k + t.$$

Here  $a$  is a positive integer, as is  $s$  (since  $\mu > 1$ ). The integer  $t$  is also positive, as we show below. Since  $f(x)$  is a trinomial,  $\mu$  must be a local minimum, with  $f(\mu) > 0$ . Hence,

$$(7.2) \quad 1 \geq f(p/q) = 1/q^n \geq f(\mu) > 0.$$

Set

$$\Omega = f(\mu) > 0.$$

Then, by (7.2),

$$(7.3) \quad 1 \leq q \leq \Omega^{-1/n}.$$

Note that  $\mu^{n-k} = ks/an$ , since  $f'(\mu) = 0$ . Consequently,

$$\Omega = t - (a(n-k)/k) \left( \frac{ks}{an} \right)^{n/(n-k)}.$$

Set

$$(7.4) \quad T = \frac{kt}{a(n-k)}, \quad S = \left( \frac{ks}{an} \right)^{n/(n-k)}, \quad W = T - S, \quad \delta = (k/(n-k))^{n-k} (an)^{-n}.$$

One readily proves

**Lemma 7.1.**  $W = k\Omega/a(n-k)$  and  $T^{n-k} - S^{n-k} \geq \delta$ .

Note that

$$T^{n-k} - S^{n-k} = (T - S) \cdot \sum_{i+j=n-k-1} T^i S^j \leq W(n-k)T^{n-k-1}.$$

Therefore, by Lemma 7.1,

$$W \geq \frac{\delta}{(n-k)T^{n-k-1}}.$$

Hence, by (7.4) and Lemma 7.1,

$$\Omega \geq (a^k(n-k)t^{n-k-1}n^n)^{-1} \geq ((n-k)n^n H^{n-1})^{-1}.$$

Since  $4^{1/5} \leq 1.32$ , we obtain Theorem 7.1, by (7.3).

## 8. MAXIMAL POLYNOMIALS

With Theorem 2.2 now proved, we are left with proving Theorems 3.1 and 3.2. In sections 8–10 we prove Theorem 3.1; Theorem 3.2 is then proved (using Theorem 3.1) in section 11.

We now revert to the general situation of a trinomial  $g(x)$  with real coefficients, say

$$g(x) = ax^n - s_0x^k + t_0, \quad n > k > 0, \quad n \geq 5$$

where  $a, s_0, t_0$  are non-zero real numbers with  $a \geq 1$ . We proceed to embed  $g(x)$  in a 1-parameter family of real trinomials.

Suppose that  $(\rho, \alpha)$  is a solution for  $g(x)$ ; i.e., by Definition 2.3,

$$(8.1) \quad g(\rho) = \alpha.$$

Keeping  $a$  and  $s_0$  fixed, (8.1) holds if, and only if,

$$g(x) = a\omega_n(x) - s_0\omega_k(x) + \alpha,$$

where

$$(8.2) \quad \omega_j(x) = x^j - \rho^j, \quad j \geq 1, \quad \rho \text{ fixed, all } x.$$

For  $s$  in  $\mathbb{R}$  define

$$(8.3) \quad g_s(x) = a\omega_n(x) - s\omega_k(x) + \alpha.$$

We denote by  $\mathcal{F}$  the 1-parameter family  $\{g_s(x)\}$ ,  $s \in \mathbb{R}$ . Note that  $g_{s_0}(x) = g(x)$ ; also, for every  $s$  in  $\mathbb{R}$ ,  $g_s(\rho) = \alpha$ . Thus, we say that  $(\rho, \alpha)$  is a solution for the family  $\mathcal{F}$ .

A key role will be played by the following polynomial. Define

$$(8.4) \quad \phi(x) = (an/k)x^{n-k} \quad (= a\omega'_n(x))/\omega'_k(x).$$

Note the identity

$$g'_s(x) = \omega'_k(x)(\phi(x) - s).$$

Hence, we have

**Lemma 8.1.** *Let  $s$  and  $\mu$  be real numbers with  $\mu \neq 0$ . Then,*

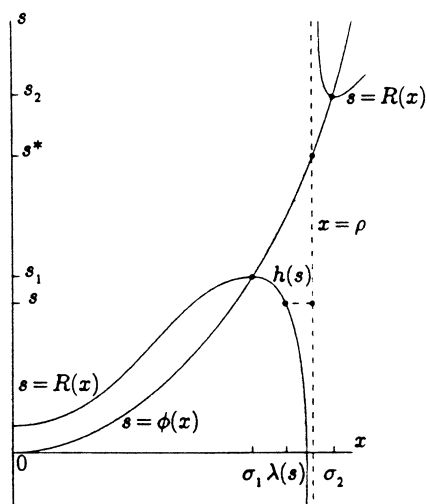
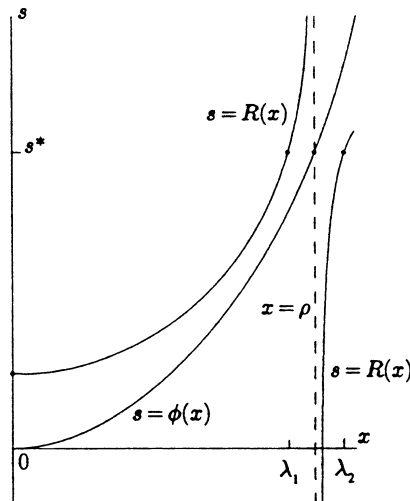
$$g'_s(\mu) = 0 \Leftrightarrow s = \phi(\mu).$$

In Lemma 8.1 we regard  $s$  as a function of the critical point  $\mu$ . In like fashion we may regard  $s$  as a function of a root  $\lambda$ . Define

$$(8.5) \quad R(x) = (a\omega_n(x) + \alpha)/\omega_k(x), \quad x \neq \rho.$$

Then,

$$g_s(\lambda) = 0 \Leftrightarrow s = R(\lambda).$$

FIGURE 1.  $\alpha > 0$ FIGURE 2.  $\alpha < 0$ 

Suppose now that  $(\rho, \alpha)$  is a *special solution* for  $\mathcal{F}$ —in particular, by (3.1),  $\rho > 1$  and  $\rho^n > |\alpha|$ . Thus,

$$(8.6) \quad R(0) = (\rho^n - \alpha)/\rho^k > 0.$$

Moreover, by (8.4) and (8.5), we obtain

$$(8.7) \quad R'(x) = \frac{\omega'_k(x)(\phi(x) - R(x))}{\omega_k(x)},$$

$$(8.8) \quad R''(x) = \frac{\omega''_k(x)(\phi(x) - R(x)) + \omega'_k(x)(\phi'(x) - 2R'(x))}{\omega_k(x)}.$$

In particular, we have

**Corollary 8.1.** *For  $\sigma > 0$ ,  $R'(\sigma) = 0 \Leftrightarrow R(\sigma) = \phi(\sigma)$ .*

We wish to graph the functions

$$s = \phi(x), \quad s = R(x).$$

Note that the second function depends in a crucial way on the *sign* of  $\alpha$ . Thus, we now consider two cases,  $\alpha > 0$  and  $\alpha < 0$ .

**Case I:  $\alpha > 0$ .** Suppose first that  $0 < x < \rho$ . For  $x$  small and positive, (8.6) and (8.7) imply that  $R'(x) > 0$ , and hence  $R(x)$  is an increasing function. But for  $x$  close to  $\rho$  and  $x < \rho$ ,  $R(x)$  is large and negative. Thus, there is a point  $\sigma_1$ , with  $0 < \sigma_1 < \rho$ , such that  $\sigma_1$  is a local maximum for  $R(x)$ . Moreover, by (8.8), we see that  $R''(x) < 0$ , for  $\sigma_1 \leq x < \rho$ . Thus, we obtain the graph shown in Figure 1 (the drawing is highly schematic!) for  $0 < x < \rho$ . A similar argument gives the graph in Figure 1 for  $\rho < x$ . Here, the point  $\sigma_2 > \rho$  is a local minimum.

How do we find the key numbers  $\sigma_1$  and  $\sigma_2$ ? For this we will use an *auxiliary polynomial*. Define

$$(8.9) \quad A(x) = -(a\omega_n(x) - \phi(x)\omega_k(x) + \alpha).$$

We will show that  $A(x)$  has the following properties.



**Proposition 8.1.** *Let  $A(x)$  be defined as above. Then,*

- (i)  $A'(\rho) = 0$ ,
- (ii) *there are unique numbers,  $\sigma_1$  and  $\sigma_2$ , such that*

$$A(\sigma_1) = A(\sigma_2) = 0, \quad 0 < \sigma_1 < \rho < \sigma_2,$$

$$(iii) |\sigma_i - \rho| < \sqrt{\frac{2|\alpha|}{a(n-2)(n-k)\rho^{n-2}}}, \quad i = 1, 2.$$

Assuming Proposition 8.1, we have at once

**Corollary 8.2.** *Let  $\sigma_1$  and  $\sigma_2$  be the two positive roots of the polynomial  $A(x)$ , as above. Then,*

- (a)  $R'(\sigma_i) = 0$ ,  $i = 1, 2$ , and
- (b) *if  $x \notin \{\sigma_1, \rho, \sigma_2\}$ , then  $R'(x) \neq 0$ .*

*Proof.* (a) By (8.9) and Proposition 8.1,

$$(8.10) \quad 0 = -A(\sigma_i) = a\omega_n(\sigma_i) - \phi(\sigma_i)\omega_k(\sigma_i) + \alpha.$$

Thus, by (8.5),  $R(\sigma_i) = \phi(\sigma_i)$ , and hence by Corollary 8.1,  $R'(\sigma_i) = 0$ . Part (b) follows from (8.7).

Now set

$$s_i = \phi(\sigma_i), \quad i = 1, 2; \quad s^* = \phi(\rho);$$

and define, for  $i = 1, 2$ ,

$$(8.11) \quad M_i(x) = a\omega_n(x) - s_i\omega_k(x) + \alpha \quad (= g_{s_i}(x)).$$

We call  $M_1(x)$  and  $M_2(x)$  the *maximal polynomials* in  $\mathcal{F}$ , for  $\alpha$  positive.

The following result justifies the terminology. Let  $g(x)$  be the polynomial we considered at the beginning of the section. Since  $g(\rho) = \alpha$ , there is a unique exceptional point, say  $\tau_0$ , that belongs to  $\rho$ . We will prove

**Theorem 8.1.** *Let  $M_i(x)$  be the polynomials and  $\sigma_i$  the numbers defined above. Then,*

- (i)  $M_i(\sigma_i) = M'_i(\sigma_i) = 0$ ,  $i = 1, 2$ .
- (ii) (a)  $|\tau_0 - \rho| < |\sigma_1 - \rho|$ , if  $\tau_0 < s^*$ ;  
(b)  $|\tau_0 - \rho| < |\sigma_2 - \rho|$ , if  $\tau_0 > s^*$ .

Combining Proposition 8.1 and Theorem 8.1, we have proved Theorem 3.1 for the case when  $\alpha$  is positive.

**Case II:**  $\alpha < 0$ . With  $\alpha$  changed in sign, the function  $R(x)$  behaves very differently; by (8.4), (8.5) and (8.7) we see that  $R'(x) > 0$ , for all  $x > 0$ ,  $x \neq \rho$ . Thus,  $R(x)$  is a strictly increasing function which has no point in common with  $\phi(x)$ . The line  $x = \rho$  is again a vertical asymptote, but with the directions switched from those in Case I. These properties are shown in Figure 2.

Set  $s^* = \phi(\rho)$ ; define

$$(8.12) \quad M(x) = a\omega_n(x) - s^*\omega_k(x) + \alpha \quad (= g_{s^*}(x)).$$

We call  $M(x)$  the *maximal polynomial* in the family  $\mathcal{F}$ ,  $\alpha$  negative. We will prove that  $M(x)$  has the following properties.

**Theorem 8.2.** *Let  $M(x)$  be the maximal polynomial in the family  $\mathcal{F}$ , when  $\alpha$  is negative. Then,*

- (i)  $M'(\rho) = 0$ ,
- (ii) *there are real numbers  $\lambda_1$  and  $\lambda_2$  such that  $M(\lambda_1) = M(\lambda_2) = 0$ , with  $1 < \lambda_1 < \rho < \lambda_2$ , and*
- (iii) *for  $i = 1, 2$*

$$|\lambda_i - \rho| < \sqrt{\frac{2|\alpha|}{a(n-2)(n-k)\rho^{n-2}}}.$$

(iv) *Let  $g_{s_0}(x)$  be the initial trinomial  $g(x)$ ; suppose that  $\tau_0$  is the exceptional point of  $g(x)$  belonging to  $\rho$ . Then*

- (a)  $|\tau_0 - \rho| \leq |\lambda_1 - \rho|$ , *if  $s_0 > s^*$ ,*
- (b)  $|\tau_0 - \rho| \leq |\lambda_2 - \rho|$ , *if  $s^* > s_0$ .*

By Theorem 8.2 (iii)–(iv) we have proved Theorem 3.1 for the case when  $\alpha$  is negative, and hence the theorem is now proved.

It remains to prove, Proposition 8.1 and Theorems 8.1 and 8.2. The proofs for Proposition 8.1(i) and Theorem 8.2(i) are routine calculations, while Theorem 8.1(i) follows from (8.11). In this section we prove Theorem 8.1(ii) and Theorem 8.2(iv), while in sections 9 and 10 we prove parts (ii) and (iii) of Proposition 8.1 and Theorem 8.2.

*Proof of Theorem 8.1(ii) and Theorem 8.2(iv).* The functions  $R(x)$  and  $\phi(x)$  have inverse functions:

$$\begin{aligned} x &= \lambda(s), \text{ inverse to } s = R(x), \\ x &= \mu(s), \text{ inverse to } s = \phi(x). \end{aligned}$$

By (8.4), (8.7) and the inverse function theorem, we obtain

**Proposition 8.2.** *Let  $\lambda(s), \mu(s)$  be inverse functions as above. Then,*

$$\begin{aligned} \text{(a)} \quad \frac{d\lambda}{ds} &= \frac{\omega_k(\lambda)}{\omega'_k(\lambda)(\phi(\lambda) - s)}, & \lambda &= \lambda(s), \\ \text{(b)} \quad \frac{d\mu}{ds} &= \frac{1}{\phi'(\mu)}, & \mu &= \mu(s). \end{aligned}$$

Recall that Theorems 8.1 and 8.2 correspond respectively to the case  $\alpha$  positive and  $\alpha$  negative. Thus to complete the proofs of these theorems we again consider these two cases.

**Case I:**  $\alpha > 0$ . Write  $\mathbb{R}$  as the union of three intervals:

$$J_1 = (-\infty, s_1], \quad J_2 = (s_1, s_2), \quad J_3 = [s_2, +\infty).$$

Note that by Corollary 8.2(b) and Figure 1 there is a unique inverse function  $\lambda_i(s)$  defined on all of  $J_i$ ,  $i = 1, 3$ , such that  $\lambda_i(s)$  is the unique exceptional point of  $g_s(x)$  that belongs to  $\rho$ . Moreover, there is a unique inverse function  $\mu(s)$  defined on all of  $J_2$ ; and again,  $\mu(s)$  is the exceptional point of  $g_s(x)$  that belongs to  $\rho$ .

For  $s$  in  $\mathbb{R}$ , define a function  $\tau(s)$  by

$$\tau(s) \mid J_i = \lambda_i(s), \quad i = 1, 3; \quad \tau(s) \mid J_2 = \mu(s).$$

Thus,  $\tau$  is continuous for all  $s$ , but is not differentiable for  $s$  in  $\{s_1, s^*, s_2\}$ . Also, as noted above,

**Lemma 8.2.** *For all  $s$  in  $\mathbb{R}$ ,  $\tau(s)$  is the unique exceptional point of  $g_s(x)$  that belongs to  $\rho$ .*

Note that if  $s = s^*$ , then  $\tau(s^*) = \rho$ , and so Theorem 3.1 is trivially true. Thus, from now on we assume that  $s \neq s^*$ . Define

$$(8.13) \quad h(s) = |\tau(s) - \rho| > 0.$$

Let  $v = \pm 1$  be defined by

$$(8.14) \quad \tau(s) = \rho + vh(s), \quad s \neq s^*.$$

For  $m$  a positive integer, define a polynomial in  $\mathbb{Z}[x, y]$  by

$$(8.15) \quad P_m(x, y) = \sum_{i+j=m-1} x^i y^j.$$

Using the fact that, for all  $j \geq 1$ ,

$$\omega_j(\tau) = \tau^j - \rho^j = vh(s)P_j(\tau, \rho), \quad \tau = \tau(s),$$

we obtain, from Proposition 8.2, (8.13), (8.14) and (8.15)

**Corollary 8.3.** (a) *Suppose that  $s \in J_1 \cup J_3$ ,  $s \notin \{s_1, s_2\}$ . Then,*

$$\frac{dh(s)}{ds} = \frac{h(s)P_k(\tau, \rho)}{\omega'_k(\tau)(\phi(\tau) - s)}.$$

(b) *Suppose that  $s \in J_2$ ,  $s \neq s^*$ . Then,*

$$\frac{dh(s)}{ds} = \frac{v}{\phi'(\tau)}.$$

By Figure 1 we have

**Lemma 8.3.** (i) *If  $s < s_1$ , then  $v = -1$  and  $\phi(\tau(s)) > s$ .*

(ii) *If  $s > s_2$ , then  $v = +1$  and  $\phi(\tau(s)) < s$ .*

Consequently we obtain

**Corollary 8.4.** (i) *If  $s < s_1$  or  $s^* < s < s_2$ , then  $dh/ds > 0$ .*

(ii) *If  $s_1 < s < s^*$  or  $s_2 < s$ , then  $dh/ds < 0$ .*

Recall the maximal polynomials  $M_i(x)$  and the positive roots  $\sigma_i$ ,  $i = 1, 2$ . By Corollary 8.3, Lemma 8.3 and Corollary 8.4 we have

$$h(s) < \begin{cases} |\sigma_1 - \rho|, & \text{if } s < s^*, \\ |\sigma_2 - \rho|, & \text{if } s^* < s. \end{cases}$$

This completes the proof of Theorem 8.1(ii).

**Case II:**  $\alpha < 0$ . This case is simpler (see Figure 2), so we only sketch the proof. Set

$$\hat{J}_1 = (s^*, +\infty), \quad \hat{J}_2 = (-\infty, s^*].$$

As before, for  $i = 1, 2$ , there is a unique inverse function  $\hat{l}_i(s)$ , defined on all of  $\hat{J}_i$ , such that  $\hat{l}_i(s)$  is the unique exceptional point of  $g_s(x)$  that belongs to  $\rho$ . For all  $s$  in  $\mathbb{R}$ , define a function  $\hat{\tau}(s)$  by

$$\hat{\tau}(s) \mid \hat{J}_i = \hat{l}_i(s).$$

Thus,  $\hat{\tau}(s)$  is continuous and differentiable for all  $s \neq s^*$ . As before, set

$$\hat{h}(s) = |\hat{\tau}(s) - \rho|, \quad \hat{\tau}(s) = \rho + \hat{v}\hat{h}(s), \quad \hat{v} = \pm 1.$$

We then have

**Lemma 8.4.** (i)  $d\hat{h}/ds < 0$ , if  $s \in \hat{J}_1$ ;  
(ii)  $d\hat{h}/ds > 0$ , if  $s \in \hat{J}_2$ ,  $s \neq s^*$ .

The proof is similar to the proof for Corollary 8.4—we leave the details to the reader. And the proof of Theorem 8.2(iv) then follows at once from Lemma 8.4.

It remains to prove (ii)–(iii) in Proposition 8.1 and in Theorem 8.2. This will be done in sections 9 and 10.

## 9. FINDING SMALL ROOTS

In this section we give sufficient conditions for a polynomial in  $\mathbb{R}[x]$  to have a “small” root. If these conditions are satisfied, we then give (fairly sharp) upper and lower bounds for such a root. In section 10 we apply this result to complete the proofs of Proposition 8.1 and Theorem 8.2.

Let  $n$  and  $d$  be fixed positive integers with  $n \geq d + 1$ . Suppose that  $B(x)$  is a polynomial with real coefficients, given by

$$(9.1) \quad B(x) = -\alpha + \beta x^d + \sum_{i=1}^{n-d} B_{d+i} x^{d+i},$$

where  $\alpha$  and  $\beta$  are positive numbers and  $B_n \neq 0$ .

Let  $\Phi$  be a positive number such that

$$(9.2) \quad |B_{d+i}|/\beta \leq \Phi^i, \quad 1 \leq i \leq n-d.$$

We prove

**Theorem 9.1.** *Let  $B(x)$  be a polynomial as given in (9.1), with  $\Phi$  as above. Choose a real number  $c \geq 1$ . If*

$$(9.3) \quad \beta/\alpha \geq (2(1+c)\Phi)^d(1+c^{-1}),$$

*then there is a positive number  $y_2$  such that  $B(y_2) = 0$ , and*

$$(9.4) \quad 0 \leq \{(1-\varepsilon)\alpha/\beta\}^{1/d} < y_2 < \{(1+\varepsilon)\alpha/\beta\}^{1/d} \quad (\varepsilon = 1/c).$$

*Moreover, if  $d$  is even there is a positive number  $y_1$  such that  $B(-y_1) = 0$ , where  $y_1$  also satisfies the inequalities in (9.4).*

*Proof.* Set  $D(x) = -\alpha + \beta x^d$ , and for any number  $\varepsilon$ , with  $0 < \varepsilon \leq 1$ , let

$$\gamma_\varepsilon = \{(1+\varepsilon)\alpha/\beta\}^{1/d}, \quad \bar{\gamma}_\varepsilon = \{(1-\varepsilon)\alpha/\beta\}^{1/d}.$$

Note that  $D(\gamma_\varepsilon) = \varepsilon\alpha$ ,  $D(\bar{\gamma}_\varepsilon) = -\varepsilon\alpha$ . Consequently, we have

**Lemma 9.1.** *Suppose that*

$$(9.5) \quad |B_{d+i}| \cdot \gamma_\varepsilon^{d+i} \leq \varepsilon\alpha/2^i, \quad 1 \leq i \leq n-d.$$

*Then there is a positive root  $y_2$  of  $B(x)$  such that (9.4) holds.*

The proof is routine.

If we divide both sides of (9.5) by  $\beta$ , use (9.2) and rearrange terms, we obtain

**Lemma 9.2.** *If*

$$(9.6) \quad (\beta/\alpha)^i \geq (2^i \Phi^i (1+\varepsilon^{-1}))^d (1+\varepsilon)^i, \quad 1 \leq i \leq n-d,$$

*then (9.5) holds.*

Note that (9.3) is simply the case  $i = 1$  in (9.6), taking  $\varepsilon = c^{-1}$ . Suppose then that (9.3) holds. Then,

$$\begin{aligned} (\beta/\alpha)^i &\geq (2(1 + \varepsilon^{-1})\Phi)^{di}(1 + \varepsilon)^i \\ &\geq (2^i \Phi^i (1 + \varepsilon^{-1}))^d (1 + \varepsilon)^i, \end{aligned}$$

since  $(1 + \varepsilon^{-1}) > 1$ . Thus, (9.6) holds, and hence so does (9.5).

This completes the proof of Theorem 9.1 for the root  $y_2$ . To show the existence of the root  $-y_1$ , simply note that if  $d$  is *even* then  $D(-x) = D(x)$  for all  $x$ .

We will use a special case of the above theorem in section 10 to complete the proofs in section 8. For this we will take  $d = 2$ ,  $n \geq 5$ , and  $\Phi (= \Phi(n)) = 2(n-2)/3$ . Moreover, for all  $n$  we will have

$$(9.7) \quad \beta/\alpha \geq np_0^n/2,$$

where  $p_0$  is defined in (3.2). Taking  $c = (n-2)/2$ , we find that condition (9.3) is always satisfied by this data. Thus, we have

**Corollary 9.1.** *Let  $B(y)$  be a polynomial as in (9.1) with  $d = 2$  and  $n \geq 5$ . Suppose that  $\Phi(n)$  satisfies (9.2) and that  $\beta/\alpha$  satisfies (9.7). If we take  $c = (n-2)/2$ , then (9.3) holds and hence (since  $d$  is even) there are two roots of  $B(y)$ ,  $(-1)^i y_i$ ,  $i = 1, 2$ , each  $y_i$  satisfying (9.4).*

## 10. PROOF OF PROPOSITION 8.1 AND THEOREM 8.2

We will use the material from section 9 to prove these results. We begin with a result needed for the two proofs. Let  $N$  and  $K$  be integers, with  $N \geq 4$ ,  $0 \leq K \leq N-1$ ; and let  $\gamma_0$  be a fixed positive number. For each  $j$ , with  $1 \leq j \leq N$ , set

$$(10.1) \quad C_j = (\gamma_0/(j+1)) \left\{ \binom{N}{j} - \binom{K}{j} \right\}.$$

We prove

**Lemma 10.1.** (i)  $C_2/C_1 = (N+K-1)/3$ .  
(ii)  $C_2/C_1 > C_3/C_2 > \cdots > C_N/C_{N-1}$ .

Part (i) follows at once from (10.1). The proof of (ii) is given at the end of the section.

By repeated use of part (ii) we have

**Corollary 10.1.** *For  $1 \leq j \leq N-1$ ,*

$$C_{j+1}/C_1 \leq ((N+K-1)/3)^j.$$

We apply this to the material in section 8, beginning with Proposition 8.1; this corresponds to the case  $\alpha > 0$ . Recall the polynomial  $A(x)$ , defined in (8.9). Using (8.4), we see that

$$A(x) = (a(n-k)/k)x^n - (an/k)\rho^k x^{n-k} + a\rho^n - \alpha, \quad n > k \geq 1, \quad n \geq 5.$$

By construction,  $A(\rho) = -\alpha < 0$ ; and by Proposition 8.1(i),  $A'(\rho) = 0$ . To find the roots  $\sigma_i$ , set

$$(10.2) \quad \sigma_i = \rho(1 + v_i y_i), \quad v_i = (-1)^i, \quad i = 1, 2,$$

and define

$$B(y) = A(\rho(1 + vy)), \quad v = \pm 1.$$

Therefore,

$$B(y) = \sum_{j=0}^n B_j (vy)^j,$$

where, for  $0 < j \leq n$ ,

$$B_j = A^{(j)}(\rho) \rho^j / j! = (an(n-k)/jk) \left\{ \binom{n-1}{j-1} - \binom{n-k-1}{j-1} \right\} \rho^n.$$

Thus,

$$(10.3) \quad B_0 = -\alpha, \quad B_1 = 0, \quad B_2 = \frac{1}{2}an(n-k)\rho^n.$$

If we set  $\beta = B_2$ , then we may take  $B(y)$  to be the polynomial  $B$ , given in (9.1), with  $d = 2$ . We apply (10.1), taking  $N = n - 1$ ,  $K = n - 1 - k$ ,  $\gamma_0 = an(n-k)/k$ . Then,  $B_{j+1} = C_j$ , as given in (10.1). Hence, by Corollary 10.1,

$$(10.4) \quad B_{2+j}/\beta \leq \Phi_1(n, k)^j, \quad 1 \leq j \leq n-2,$$

where  $\Phi_1(n, k) = (2n - k - 3)/3$ . Note that by (3.1) and (10.3),  $\beta/\alpha$  satisfies (9.7); also, for  $1 \leq k \leq n-1$ ,

$$\Phi_1(n, k) \leq \Phi(n) \quad (= 2(n-2)/3).$$

Thus, we may apply Corollary (9.1), taking  $c = (n-2)/2$ ; and so we obtain the two roots  $(-1)^i y_i$ ,  $i = 1, 2$ . By Theorem 9.1,

$$(10.5) \quad 0 < y_i < \Delta \sqrt{\frac{\alpha}{\rho^n}}, \quad \Delta = \Delta(a, n, k),$$

since  $(1 + c^{-1})/(an(n-k)) = 1/(a(n-2)(n-k))$ . Parts (ii) and (iii) of Proposition 8.1 now follow, since  $\sigma_i = \rho(1 + (-1)^i y_i)$ ,  $i = 1, 2$ .

The proof of Theorem 8.2(ii)–(iii) is similar, using the polynomial  $M(x)$ , given in (8.12), with  $M(\rho) = \alpha < 0$ . Proceeding as above, set

$$\tilde{B}(y) = M(\rho(1 + vy)) = \sum_0^n \tilde{B}_j (vy)^j, \quad v = \pm 1.$$

Thus,

$$\tilde{B}_j = (an/j) \left\{ \binom{n-1}{j-1} - \binom{k-1}{j-1} \right\} \rho^n,$$

and so,

$$\tilde{B}_1 = 0, \quad \tilde{B}_2 = \frac{1}{2}an(n-k)\rho^n, \quad \tilde{B}_{j+1} = C_j, \quad 1 \leq j \leq N,$$

where  $N = n - 1$ ,  $K = k - 1$ ,  $\gamma_0 = an$ . If we take  $\beta = \tilde{B}_2$ , then  $\tilde{B}(y)$  is a polynomial as in (9.1), with  $d = 2$ . By Corollary 10.1,

$$\tilde{B}_{2+j}/\tilde{B}_2 \leq \Phi_2(n, k)^j, \quad 1 \leq j \leq n-2,$$

where  $\Phi_2(n, k) = (n + k - 3)/3$ . Since

$$\Phi_2(n, k) \leq \Phi(n) = 2(n-2)/3, \quad 1 \leq k \leq n-1,$$

we obtain the proof of Theorem 8.2(ii)–(iii) by using Corollary 9.1. We leave the details to the reader.

We are left with proving Lemma 10.1(ii). Given integers  $b$  and  $c$ , with  $b > 0$  and  $c \geq 0$ , set

$$[c, b] = c(c-1) \cdots (c+1-b).$$

Note the key identity:

$$(10.6) \quad [c, b+1] = [c, b](c-b).$$

For  $1 \leq i \leq N$ , set

$$D_i = [N, i] - [K, i].$$

Thus by (10.1), for  $2 \leq j \leq N$ ,

$$C_j/C_{j-1} = (1/(j+1))D_j/D_{j-1}.$$

Set

$$(10.7) \quad E_j = D_j/D_{j-1} - D_{j+1}/D_j, \quad 2 \leq j \leq N-1.$$

Since  $1/(j+1)$  is a decreasing function, to prove part (ii) it suffices to show:

**Lemma 10.2.** *For all  $j$ , with  $2 \leq j \leq N-1$ ,  $E_j > 0$ .*

This follows at once from three simple identities, each proved by repeated use of (10.6).

**Lemma 10.3.** *For  $1 \leq j \leq N-1$ ,*

- (i)  $D_{j+1} - D_j(N-j) = [K, j](N-K)$ ,
- (ii)  $D_{j+1} - D_j(K-j) = [N, j](N-K)$ ,
- (iii)  $D_{j+1}[K, j] - D_j[K, j+1] = [N, j][K, j](N-K)$ .

Combining (10.7) with (i) and (ii), we obtain

$$E_{j+1} = 1 + \left\{ \frac{[K, j]}{D_j} - \frac{[K, j+1]}{D_{j+1}} \right\} (N-K), \quad 1 \leq j \leq N-2.$$

Thus, by (iii)

$$E_{j+1} = 1 + \left( \frac{[N, j] \cdot [K, j]}{D_{j+1}D_j} \right) (N-K)^2 > 0.$$

This completes the proof of Lemma 10.2 and hence of Lemma 10.1.

## 11. PROOF OF THEOREM 3.2

We return to the situation considered in section 3:  $g(x)$  is a polynomial in  $\mathbb{R}[x]$ ,  $(p/q, \pm 1/q^n)$  is a special solution for  $g(x)$  and  $\tau$  is the exceptional point that belongs to  $p/q$ . Let

$$(11.1) \quad g(x) = ax^n - sx^k + t, \quad n > k > 0, \quad n \geq 5,$$

where  $a, s, t$  are non-zero real numbers with  $a \geq 1$ .

We first state 3 lemmas, without proof. Using these, we prove Theorem 3.2. We then go back and prove the 3 lemmas. Set

$$(11.2) \quad V_k = |s|\rho^k, \quad V_n = a\rho^n, \quad W = \max\{V_k, V_n\}, \quad \rho = p/q.$$

We will prove

**Lemma 11.1.** *Let  $H = H(g)$ . Then  $H \leq (2.032)W$ .*

As in section 3, set  $\alpha = \pm 1/q^n$ , so that  $g(\rho) = \alpha$ . Since  $(\rho, \alpha)$  is special,  $\tau > 1$ . Recall the distance function,  $h(s) = |\tau - \rho|$ , of (8.13). We will prove

**Lemma 11.2.** *Choose a real number  $j$  with  $0 \leq j < \frac{1}{2}$ . Then,*

$$|s|^j h(s) \leq ((1+u_n)^n s^*)^j \Delta(\rho^{n^*} q^{n/2})^{-1}, \quad \Delta = \Delta(a, n, k).$$

Here  $s^* = \phi(\rho)$ , see (8.4),  $u_n$  is given in (3.5), and  $n^* = (n-2)/2$ .

The final result we need is

**Lemma 11.3.** *Let  $m_n$  be the number defined in (3.5). Then,*

$$2(an/k)^j \Delta \leq m_n, \quad \Delta = \Delta(a, n, k), \quad 0 \leq j < \frac{1}{2}.$$

Using these three results, we have

*Proof of Theorem 3.2.* We consider two cases:  $W = V_k$  or  $W = V_n$ . Suppose first that  $W = V_k$ . Take  $j = b/n$ , where  $0 \leq b \leq n^*$ . Then by Lemma 11.1, Lemma 11.2, (3.5), and Corollary 3.1,

$$\begin{aligned} H^{b/n} h(s) &\leq (2.032)^{b/n} W^{b/n} h(s) \leq r_n^b (|s| \rho^k)^{b/n} h(s) \\ &\leq r_n^b (1 + u_n)^b (s^* \rho^k)^{b/n} \Delta (\rho^{n^*} q^{n/2})^{-1}, \end{aligned}$$

where  $r_n$  is given in (3.5) and  $\Delta$  in (3.4). Thus by (8.4), Lemma 11.3 and (3.6), taking  $c = n^* - b$ ,

$$\begin{aligned} H^{b/n} h(s) &\leq (r_n(1 + u_n))^b (an\rho^n/k)^{b/n} \Delta (\rho^{n^*} q^{n/2})^{-1} \\ &\leq \frac{K_b(n)}{2\rho^{n^*-b} q^{n/2}} = \frac{K_b(n)}{2\rho^c q^{n/2}} = \frac{K_b(n)}{2p^c q^{b+1}}. \end{aligned}$$

Consequently,

$$h(s) \leq \frac{K_b(n)}{2H^{b/n} p^c q^{b+1}}.$$

This completes the proof for this case. The proof for  $W = V_n$  is quite similar and is left to the reader.

*Proof of Lemma 11.1.* Note that  $H \in \{a, |s|, |t|\}$ , by (11.1). If  $H = a$  or  $|s|$ , then Lemma 11.1 follows trivially, since  $\rho > 1$ . Suppose then that  $H = |t|$ . Since  $f(\rho) = \alpha$ , we have

$$\begin{aligned} H = |t| &\leq a\rho^n + |s|\rho^k + |\alpha| = V_n + V_k + |\alpha| \\ &\leq 2W + |\alpha| = (2 + |\alpha|/W)W. \end{aligned}$$

If  $q > 1$ , then  $|\alpha| \leq 1/32$  and so  $(2 + |\alpha|/W) < 2.032$ . If  $q = 1$ , then  $\rho \geq 2$  (since  $(p, q)$  is special), and so  $W \geq 32$ , which implies that  $(2 + |\alpha|/W) < 2.032$ . This completes the proof.  $\square$

*Proof of Lemma 11.3.* By assumption,  $j < \frac{1}{2}$ , so  $(an/k)^j < (an/k)^{\frac{1}{2}}$ . Thus, the proof follows at once from the fact that

$$(n-k)k \geq (n-1), \quad 1 \leq k \leq n-1.$$

*Proof of Lemma 11.2.* Recall that the proof of Theorem 3.1 hinged on a study of the function  $\tau(s)$  and its derivative. We prove Lemma 11.2 in a similar fashion, except that now we use a function that involves a coefficient of  $g(x)$ .

Let  $s$  be a real variable, with  $s \neq 0$ . Set

$$r = |s| = ws, \quad \text{where } w = \pm 1.$$

Thus,  $dr/ds = w$ . Given a number  $j$ , with  $0 \leq j < \frac{1}{2}$ , define

$$(11.3) \quad \xi_j(s) = r^j h(s), \quad s \neq 0,$$

where  $h(s)$  is defined in (8.13). Suppose first that  $\tau(s) = l(s)$ , a root of  $g_s(x)$ . Recall the function  $P_k(x, y)$ , defined in (8.15). Since  $\tau = \lambda$ , the term  $P_k(\lambda, \rho)$



occurs in the formula for  $dh/ds$ , given in Corollary 8.3(a). Define a non-negative number  $\varepsilon$  by

$$(11.4) \quad P_k(\lambda, \rho) = (1 - v\varepsilon)P_k(\lambda, \lambda)$$

where  $v = \pm 1$ , see (8.14). By (8.15),  $P_k(\lambda, \lambda) = \omega'_k(\lambda)$ , and so we may rewrite Corollary 8.3(a) as

$$\frac{dh}{ds} = \frac{h(1 - v\varepsilon)}{(\phi - s)}, \quad \phi = \phi(\lambda(s)).$$

By a routine calculation we then obtain

**Proposition 11.1.** *If  $\tau(s) = \lambda(s)$ , then*

$$\frac{d\xi_j(s)}{ds} = \frac{hr^{j-1}}{(\phi - s)} \{wj\phi + r(1 - j - v\varepsilon)\}.$$

At the end of the section we establish

**Claim 11.1.**  $0 \leq \varepsilon < \frac{1}{2}$ .

We break up the proof of Lemma 11.2 (assuming that  $\tau = \lambda$ ) into several cases, starting with: I,  $\alpha > 0$ ; II,  $\alpha < 0$ .

**Case I:**  $\alpha > 0$ . Recall the two roots,  $\sigma_1$  and  $\sigma_2$ ; as before, set

$$s_i = \phi(\sigma_i), \quad i = 1, 2, \quad s^* = \phi(\rho).$$

Thus,

$$0 < s_1 < s^* < s_2.$$

We will work with the root function  $l(s)$ , and so we assume that either  $s < s_1$  or  $s > s_2$ . In fact we need 3 cases;

$$\text{A) } s > s_2, \quad \text{B) } 0 < s < s_1, \quad \text{C) } s < 0.$$

**Case A.**  $s > s_2$ . Thus,  $r = s$ , so  $w = +1$ ; moreover,  $v = +1$  and  $\phi < s$ . Consequently, by Proposition 11.1,

$$d\xi_j(s)/ds < 0,$$

and so (see Figure 1)

$$\xi_j(s) < s_2^j |\sigma_2 - \rho|,$$

by (11.3) and Theorem 8.1. We will prove that

$$(11.5) \quad s_2 \leq (1 + u_n)^n s^*.$$

Thus,

$$\xi_j(s) < ((1 + u_n)^n s^*)^j |\sigma_2 - \rho|,$$

which proves Lemma 11.2 for this case.

**Case B.**  $0 < s < s_1$ . We now have  $r = s$ ,  $w = +1$ ,  $v = -1$ ,  $\phi > s$ ; and so,  $d\xi_j(s)/ds > 0$ , by Proposition 11.1. Since  $s_1 < s^*$ , the proof follows at once—we omit the details.

**Case C.**  $s < 0$ . Thus,  $r = -s$ ,  $w = -1$ ,  $v = -1$ ,  $\phi > s$ . Consequently,

$$\frac{d\xi_j(s)}{ds} = \frac{hr^{j-1}}{(\phi - s)}\{-j\phi + r(1 - j + \varepsilon)\},$$

and so,

$$(11.6) \quad d\xi_j(s)/ds > \frac{hr^{j-1}}{2(\phi - s)}(r - \phi).$$

Recall that as  $\lambda$  gets close to  $\rho$ , with  $\lambda < \rho$ , then  $|R(\lambda)|$  gets large. Since  $\phi(\rho) = s^*$ , we obtain

**Lemma 11.4.** *There is a unique number  $\lambda_3$  such that  $\sigma_1 < \lambda_3 < \rho$  and  $|R(\lambda_3)| = s^*$ .*

Let

$$s_3 = R(\lambda_3), \quad r_3 = |s_3|.$$

Suppose that  $s \leq s_3$ ; i.e.,  $r \geq r_3$ . Then,  $r - \phi \geq 0$ , and so, by (11.6),

$$(11.7) \quad d\xi_j(s)/ds \geq 0.$$

If  $s = s_3$ , then

$$r = r_3 = s^*,$$

and so, by Theorem 8.1,

$$(11.8) \quad r^j h(s) < s^{*j} |\sigma_1 - \rho| < ((1 + u_n)^n s^*)^j |\sigma_1 - \rho|.$$

Moreover, if  $0 > s > s_3$ , then  $r < r_3$ , and so (11.8) continues to hold. But by Proposition 8.1, (11.8) implies Lemma 11.2 for this case.

**Case II:**  $\alpha < 0$ . We now use Figure 2 in section 8. Suppose first that  $s > s^*$ . Then,

$$w = +1, \quad v = -1, \quad \phi < s.$$

Thus, by Proposition 11.1,

$$d\xi_j(s)/ds < 0,$$

and so,

$$s^j h(s) < s^{*j} |\lambda_1 - \rho|,$$

which proves Lemma 11.2 for this case. If  $0 < s < s^*$ , or if  $s < 0$ , the proof of Lemma 11.2 is similar and is left to the reader.

This completes the proof of Lemma 11.2 for the case  $\tau = \lambda$ . Suppose then that  $\tau = \mu$ , a proper critical point of  $g_s(x)$ . By Figures 1 and 2 we see that  $\alpha$  is positive and that  $s_1 \leq s \leq s_2$ . Thus Lemma 11.2, for this case, follows at once from (11.5) and Proposition 8.1(iii).

We are left with proving claim 11.1 and (11.5).

*Proof of Claim 11.1.* If  $k = 1$ , the claim is trivially true since then  $\varepsilon = 0$ . Suppose now that  $2 \leq k \leq n - 1$  and  $\varepsilon > 0$ .

We consider first the case when  $\alpha > 0$  and  $\lambda < \rho$ , so that  $v = -1$ . By the definition of  $\varepsilon$ ,

$$P_k(\lambda, \rho) = (1 + \varepsilon)P_k(\lambda, \lambda).$$

Note that

$$(11.9) \quad P_k(\rho, \rho) > P_k(\lambda, \rho) > P_k(\lambda, \lambda) > P_k(\sigma_1, \sigma_1),$$

where  $\sigma_1$  is defined in Proposition 8.1. If we set

$$P_k(\rho, \rho)/P_k(\sigma_1, \sigma_1) = 1 + \varepsilon_1,$$

then,  $0 < \varepsilon < \varepsilon_1$ , since  $\sigma_1 < \lambda < \rho$ , by Theorem 8.1(ii). Thus, we will prove Claim 11.1 by showing that

$$(11.10) \quad 0 < \varepsilon_1 < \frac{1}{2}.$$

By (8.15) and (10.2),

$$(11.11) \quad P_k(\rho, \rho)/P_k(\sigma_1, \sigma_1) = (\rho/\sigma_1)^{k-1} = (1 - y_1)^{-(k-1)}.$$

But by (10.5) and (3.4), since  $(p, q)$  is a special solution,

$$0 < y_1 < \Delta \sqrt{\frac{\alpha}{\rho^n}} \leq \sqrt{\frac{2}{(n-2)(n-k)p_0^n}}.$$

Thus,

$$(k-1)y_1 \leq (n-2)y_1 \leq \sqrt{\frac{2(n-2)}{p_0^n}} < 1/3.$$

Consequently,

$$(1 - y_1)^{-(k-1)} < (1 - (k-1)y_1)^{-1} < (2/3)^{-1} = 3/2.$$

Thus, by (11.11) we have proved (11.10), and so Claim 11.1 is proved for this case. The proofs for the remaining cases are similar; we omit the details.

*Proof of (11.5).* Recall that

$$s_2 = \phi(\sigma_2), \quad s^* = \phi(\rho).$$

Thus, by (8.4),

$$s_2/s^* = \phi(\sigma_2)/\phi(\rho) = (\sigma_2/\rho)^{n-k}.$$

But, by (10.2), (10.5), (3.4) and (3.5),

$$\sigma_2 = \rho(1 + y_2) \leq \rho(1 + u_n),$$

and so,

$$(\sigma_2/\rho)^{n-k} \leq (1 + u_n)^{n-k} < (1 + u_n)^n.$$

This completes the proof.

#### ACKNOWLEDGEMENTS

I would like to thank Alf van der Poorten and Wolfgang Schmidt for their useful comments, and Paul Brown for the computer graphics.

#### REFERENCES

- [1] Baker, A. (1964). Rational approximations to certain algebraic numbers. *Proc. London Math Soc.* (3) **14**, 385–398. MR **28**:5029
- [2] Baker, A. (1968). Contributions to the theory of diophantine equations. *Phil. Trans. Roy. Soc. London* **A263** 173–208. MR **37**:4005; MR **37**:4006
- [3] Bombieri, E. (1982). On the Thue-Siegel-Dyson Theorem. *Acta Math.* **148**, 255–296. MR **83m**:10052
- [4] Bombieri, E. and Mueller, J. (1983). On effective measures of irrationality for  $\sqrt[n]{a/b}$  and related numbers. *J. Reine Angew. Math.* **342**, 173–196. MR **84m**:10023
- [5] Bombieri, E. and Schmidt, W. M. (1987). On Thue's equation, *Invent. Math.* **88**, 69–81. MR **88d**:11026

- [6] Davenport, H. and Roth, K. F. (1955). Rational approximations to algebraic numbers. *Mathematika* **3**, 160–167. MR **17**:1060d
- [7] Dyson, F. J. (1947). The approximation to algebraic numbers by rationals. *Acta Math.* **79**, 225–240. MR **9**:412h
- [8] Evertse, J. H. (1983). Upper bounds for the numbers of solutions of Diophantine equations. Math Centrum tract 168, Amsterdam. MR **85k**:11015
- [9] Feldman, N. I. (1968). Estimation of a linear form of logarithms of algebraic numbers, *Mat. Sb.* **77** (119), 423–436; English transl., Math. USSR Sb. 6 (1968), 393–406. MR **38**:1059
- [10] Gelfond, A. O. (1952). *Transcendental and Algebraic Numbers*, Dover, New York, 1960. MR **22**:2598
- [11] Lagrange, J. L. (1867). Oeuvres, vol. 2, section 2, Gauthier-Villars, Paris, 675–693. MR **55**:12437 (reprint)
- [12] Lewis, D. and Mahler, K. (1961). On the representation of integers by binary forms. *Acta Arith.* **6**, 333–363. MR **22**:10952
- [13] Mahler, K. (1984). On Thue’s Theorem, *Math. Scand.* **55**, 188–200. MR **86h**:11030
- [14] Mignotte, M. (1992). *Mathematics for Computer Algebra*, Springer-Verlag, New York. MR **92i**:68071
- [15] Mueller, J. (1987). Counting solutions of  $|ax^r - by^r| \leq h$ . *Quarterly J. Math. Oxford Ser. (2)* **32**, 503–513. MR **89b**:11026
- [16] Mueller, J. and Schmidt, W. M. (1987). Trinomial Thue equations and inequalities. *J. Reine Angew. Math.* **379**, 76–99. MR **88j**:11011
- [17] Mueller, J. and Schmidt, W. M. (1988). Thue’s equation and a conjecture of Siegel. *Acta Math.* **160**, 207–247. MR **89g**:11029
- [18] Roth, K. F. (1955). Rational approximations to algebraic numbers, *Mathematika* **2**, 1–20. MR **17**:242d
- [19] Schmidt, W. M. (1991). *Diophantine Approximations and Diophantine Equations*, LNM **1467**, Springer-Verlag, New York. MR **94f**:11059
- [20] Schneider, T. (1957). *Einführung in die transzendenten Zahlen*, Springer-Verlag, Berlin. MR **19**:252f
- [21] Siegel, C. L. (1929). Über einige Anwendungen diophantischer Approximationen. *Abh. Preuss. Akad. Wiss., Math. Phys. Kl., Nr. 1=Ges. Abh. I.*, 209–266.
- [22] Stewart, C. L. (1991). On the number of solutions of polynomial congruences and Thue equations. *J. Amer. Math. Soc.* **4**, 793–835. MR **92j**:11032
- [23] Thue, A. (1909). Über Annäherungswerte algebraischer Zahlen. *J. Reine Angew. Math.* **135**, 284–305.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CALIFORNIA, BERKELEY, CALIFORNIA 94720–3840