A Powerful Determinant

Alfred J. van der Poorten

CONTENTS

Supported in part by a grant from the Australian Research Council.

AMS Subject Classification: Primary, 11J25, 11J68, 41A28; Secondary, 11C20, 11J82, 41A21

Keywords: Diophantine approximation, Roth's Theorem, Simultaneous Padé approximation

We study the construction of auxiliary functions likely to aid in obtaining improved irrationality measures for cubic irrationalities and thence for arbitrary algebraic numbers. Specifically, we note that the construction of curves with singularities appropriately prescribed for our purpose leads to a simultaneous Padé approximation problem. The first step towards an explicit construction appears to be the evaluation of certain determinants. Our main task here is the computation of an example determinant, which turns out indeed to be a product of a small number of factors each to high multiplicity — whence the adjective 'powerful'. Our evaluation confirms a computational conjecture of Bombieri, Hunt and van der Poorten.

1. INTRODUCTION

Our object is to construct curves with prescribed singularities. To that end, we evaluate several determinants. As expected on the basis of computations reported in [Bombieri et al. 1995], we find that those determinants are products of a small number of distinct factors, each to high multiplicity; thus the adjective 'powerful' of the title.

Specifically, consider the easy problem of constructing a one variable polynomial with just three distinct zeros, say at α_1, α_2 and α_3, each of multiplicity k. Plainly, multiples of $(x-\alpha_1)^k(x-\alpha_2)^k(x-\alpha_3)^k$ will do. Suppose that the α_i are the three conjugates of some cubic irrational α, and recall that the binomial coefficients of order k have logarithmic height $O(k)$. Then we see that the 'easy problem' has a solution of degree $3k$ with coefficient vector of logarithmic height $O\big(kh(\alpha)\big)$, where $h(\alpha)$ is the logarithmic height of α.

We might more clumsily have proved the existence of such a solution by linear algebra. The required zeros constitute $3k$ linear conditions on the $3k+1$ coefficients of a polynomial of degree $3k$. Thus, solving by Cramer's rule provides the coefficients as $3k+1$ determinants each of dimension $3k$ by $3k$. It is now not immediate just what the height of that vector of coefficients is, given that it is appropriate to view the vector as a point in projective space. Clearly, each determinant is of logarithmic height $O\big(k^2h(\alpha)\big)$. However, the determinants are generalised Vandermonde and are readily seen to have a substantial common factor made up from powers of the three differences $\alpha_i-\alpha_j$. Removing that common factor makes it plain that the vector of determinants indeed is of logarithmic height $O\big(kh(\alpha)\big)$, as already seen by more straightforward considerations above.

The analogous problems for polynomials in two variables are more complicated. The clutter caused by studying the question at three arbitrary points (α_i,β_i) is already intolerable and quickly leads one to recall that there is no loss in dealing with the three points $(0,0)$, $(1,1)$, and (∞,∞), which are more amenable. At the end, linear fractional transformations in each variable retrieve generality. By the way, dealing with the 'more amenable' points trivialises the clumsy determinantal solution in the one variable case. In two variables, however, the analogues of the binomial coefficients remain difficult to tame.

Suppose one is to construct a polynomial $F(x,y)$ of bidegree (m,n) with prescribed vanishing at three points. First, we need a generalisation of 'multiplicity'. A correct one is to specify a *triangle* $\{k_0\geq k_1\geq \cdots\geq k_n=0\}$ of integers and to require vanishing of all the partial derivatives

$$D_x^uD_y^vF(x,y)$$

with $0\leq u<k_v$ and $v=0,\,1,\,\ldots,\,n$ at each of the points. The notation here is $D_x^u=(d^u/dx^u)/u!$ and $D_y^v=(d^v/dy^v)/v!$.

Dividing by the factorial is not just a frill. It is essential to use only divided derivatives if our arguments are to remain valid in positive characteristic.

There appears to be no alternative to solving for the polynomial's coefficients by some generalisation of Cramer's rule. We should therefore take

$$3\sum_{b=0}^{n}k_b<(m+1)(n+1)$$

and study the height of the resulting vector of $3\sum k_b$ by $3\sum k_b$ determinants. At a glance, each determinant has logarithmic height $O(m^2n+mn^2)$, which is far too large — we hope for $O(mn)$. Thus we need to find that the determinants share a very substantial common factor.

In the present paper we do no more than suggest the likelihood of such a common factor. We do that by studying the 'degenerate' case

$$3\sum_{b=0}^{n}k_b=(m+1)(n+1).$$

There is now just one determinant; we call it $\Delta=\Delta(k_0,k_1,\ldots,k_n)$.

Below, we evaluate several such determinants, and show in particular that each is a product of a small number of distinct small factors, each factor therefore with high multiplicity. That property of a 'master' determinant Δ is arguably necessary for its minors to share a very substantial common factor. We expect that our evaluation method will readily generalise to the nondegenerate case and will eventually show that our present techniques do suffice to prove the existence of the required common factor.

The context of the present work is the study of effective diophantine approximation of algebraic numbers; see [Bombieri et al. 1995] for an introduction, [Bombieri 1997] for an elegant summary and, for a detailed instance, [Bombieri et al. 1996].

2. PADÉ APPROXIMATION

Consider first the problem of constructing polynomials $Q(x)$ of degree at most l and $P(x)$ of degree at most $l-1$ so that the power series

$$R(x,y)=(1-x)^{-\alpha}Q(x)-yP(x) \qquad (2\text{–}1)$$

satisfies $R(x,1)=(1-x)^{-\alpha}Q(x)-P(x)=O(x^{2l})$.

For 'general' α, the requirement that $R(x,1)$ have a zero of order at least $2l$ at $x=0$ is $2l$ independent linear conditions on the $2l+1$ coefficients of Q and P and has a unique solution up to normalisation. Set $D^i=(d/dx)^i/i!$. Specifically, using Cramer's Rule, we may display a solution for which the leading coefficient of Q is the $2l$ by $2l$ determinant whose first l columns, with $1\leq j\leq l$, have the entries $D^{i-1}(1-x)^{-\alpha}x^{j-1}|_{x=0}$, and whose remaining l columns, with $l+1\leq j\leq 2l$, have entries $D^{i-1}x^{j-1-l}|_{x=0}$; here the rows are indexed by $1\leq i\leq 2l$.

The vanishing of the last l rows of the second set of columns suggests that in place of the approximation problem (2–1) we study the problem of finding $Q(x)$ of degree at most l so that

$$D^l R(x,1) = D^l (1-x)^{-\alpha} Q(x) = O(x^l). \qquad (2\text{–}2)$$

This 'new' problem constitutes l conditions on the $l+1$ coefficients of Q. As above, by Cramer's Rule there is a solution for which the leading coefficient of Q is the l by l determinant with entries

$$D^{l+i-1}(1-x)^{-\alpha}x^{j-1}|_{x=0} = (-1)^{l+i-j}\binom{-\alpha}{l+i-j} = \binom{\alpha+l+i-j-1}{l+i-j}. \qquad (2\text{–}3)$$

Clearly, having determined Q by Cramer's rule, as just begun, one determines P trivially, simply by requiring that it coincides with the first l terms of the power series $(1-x)^{-\alpha}Q(x)$. Thus we have solved the original problem. Because we are using *divided* derivatives it follows that the latter determinant can differ from the one we commenced to study at most by multiplication by ± 1. Specifically,

$$\left|\begin{array}{c:c} \binom{\alpha+i-j-1}{i-j} & -I \\ \hdashline \binom{\alpha+l+i-j-1}{l+i-j} & 0 \end{array}\right|$$

is the original determinant belonging to the system (2–1). Note also the remarks on evaluating the determinant at Useful Remark 5.8 below.

In the particular case $\alpha = -2l$ we name the determinant $\Delta(2l,0) = \Delta$. The value of Δ is well known. For example, recently Zeilberger [1996] recalls in a quite delightful way that

$$\left|\binom{a+i}{b+j}\right|_{1\le i,j\le l} = \frac{(a+l)!!\,(l-1)!!\,(a-b-1)!!\,(b)!!}{(a)!!\,(a-b+l-1)!!\,(b+l)!!}, \qquad (2\text{–}4)$$

where, $l!! := 1!2!3!\cdots l!$. We leave it as an exercise for the bemused reader to verify that this includes the evaluation of Δ mentioned in [Bombieri et al. 1995], namely, with $[k] := (k-1)!!$,

$$\pm\Delta = [l]^3[3l]/[2l]^3.$$

Actually, the construction discussed in [Bombieri et al. 1995] is not at all obviously equivalent to a Padé approximation problem. That construction is, given a 'triangle' $m \ge k_0 > k_1 > \cdots > k_s = \cdots = k_n = 0$, to find a polynomial $F(x,y)$ of bidegree (m,n) so that, with $D_x^u = (d^u/dx^u)/u!$ and $D_y^v = (d^v/dy^v)/v!$,

$$D_x^u D_y^v F(x,y)$$

vanishes at $(x,y) = (0,0)$, $(1,1)$, and (∞,∞) for $v = 0, 1, \ldots, n$ and $u = 0, 1, \ldots, k_v - 1$. This is $3\sum_{j=0}^n k_j$ conditions on $(m+1)(n+1)$ unknown coefficients.

Set

$$F(x,y) = \sum_{j=0}^n A_j(x)y^j.$$

The conditions at $(0,0)$ entail that for each j we must have $A_j(x) = x^{k_j}B_j(x)$, for some polynomial B_j of degree no more than $m-k_j$; and those at (∞,∞) mean that in fact B_{n-j} has degree k_j less than its apparent $m-k_{n-j}$. That, incidentally, leaves us with a construction where the $\sum_{j=0}^n k_j$ linear conditions at $(1,1)$ on

$$\sum_{j=0}^n x^{k_j}B_j(x)y^j, \qquad (2\text{–}5)$$

with $\deg B_j \le m - k_j - k_{n-j}$, apply to a function with $(m+1)(n+1) - 2\sum_{j=0}^n k_j$ unknown coefficients.

Put $C_j(x) = B_j(1-x)$. Then these conditions are the particular case $k_j = -\alpha_j$ and

$$F(x,y) = \sum_{j=0}^n (1-x)^{-\alpha_j}C_j(x)y^j$$

of the simultaneous approximation system

$$D_y^v F(x,y)\big|_{y=1} = \sum\nolimits_{j=v}^n (1-x)^{-\alpha_j}\binom{j}{v}C_j(x) = O(x^{k_v}) \qquad (2\text{–}6)$$

for $v = 0, 1, \ldots, n$. Set $T = \sum_{v=0}^n k_v$, $R = (m+1)\times(n+1)$, and suppose that $R \ge 3T$. The matrix of the preceding system is

$$\left(\binom{\alpha_j+u-k-1}{u-k}\binom{j}{v}\right)_{\substack{0\le u<k_v;\ 0\le v\le n \\ 0\le k\le m-k_j-k_{n-j};\ 0\le j\le n}} \qquad (2\text{–}7)$$

with its T rows given by the triangle of derivatives, while its $R-2T$ columns correspond to the coefficients of the polynomials C_j. The set of maximal, thus $T\times T$, minors of this matrix are the Grassmann co-ordinates of the linear manifold determined by the solutions to (2–6) and the goal is to estimate the height of that manifold; compare [Bombieri and Vaaler 1983]. That seems to remain very difficult to do, so a first step is to evaluate the one determinant remaining in the degenerate case $R = 3T$. We do that below in a very special case for general n, and in some generality for $n = 2$.

Elsewhere [van der Poorten 1991], I allude to the present Padé approximation viewpoint without at all suggesting that it might assist in evaluating determinants. In [Bombieri et al. 1995], the coefficients to be determined are those of the B_j in (2–5). Thus, there the matrix is

$$\left(\binom{j_1}{i_1}\binom{j_2}{i_2}\right)_{\substack{0\le i_1<k_{i_2};\ 0\le i_2\le n\\ k_{j_2}\le j_1\le m-k_{n-j_2};\ 0\le j_2\le n}} \tag{2–8}$$

with its T rows supported by the triangle of derivatives ($u = i_1, v = i_2$), and its $R-2T$ columns supported by the lozenge of points (j_1, j_2) remaining in the rectangle of coefficients after removal of its left lower and right upper triangles, and of course with the α_j left at their original $-k_j$, compare [Krattenthaler 1999, Theorem 49].

The two matrices (2–7) and (2–8) are not the same. One is obtained from the other by column operations induced by the transformation $x \mapsto 1-x$, and the substitutions $\alpha_j = -k_j$.

3. A SURPRISING EVALUATION

The computation in [Bombieri et al. 1995] to be explained here is the simplest case of the construction sketched above, namely the notional construction of the sequence of polynomials $F_n(x,y)$ of bidegree (m,n), where $m = 3nl-1$, and where the vanishing is respectively defined by the 'triangles' $\{2nl,\ 2(n-1)l,\ \dots,\ 2l,\ 0\}$. Then the shape of the triangle yields $k_j+k_{n-j} = 2nl$ for each $j = 0, 1, \dots,$ n so, as just explained, the construction is a matter of finding polynomials $C_{j,nl}$ each of degree at most $nl-1$ so that with

$$F_n(x,y) = \sum_{j=0}^{n}(1-x)^{-\alpha(n-j)}C_{j,nl}(x)y^j,$$

$$D_y^vF_n(x,y)\big|_{y=1} = \sum\nolimits_{j=v}^{n}(1-x)^{-\alpha(n-j)}\binom{j}{v}C_{j,nl}(x) = O\big(x^{2l(n-v)}\big) \tag{3–1}$$

for $v = 0, 1, \dots, n$; here $\alpha = -2l$. It is now easy to notice that the $C_{j,nl}$ provide just $nl(n+1)$ coefficients which are to satisfy $2l(1+2+\dots+n) = ln(n+1)$ conditions; thus the construction is indeed only notional. In the sequel we denote the determinant $\Delta(\alpha; 2nl, 2(n-1)l, \dots, 2l, 0) =$

$$\left|(-1)^{u-k}\binom{\alpha(n-j)+u-k-1}{u-k}\binom{j}{v}\right|_{\substack{0\le u<2l(n-v);\ 0\le v\le n\\ 0\le k<nl;\ 0\le j\le n}} \tag{3–2}$$

of this system by $\Delta_n(\alpha)$. In the special case $\alpha = -2l$, where $\alpha_j = -\alpha(n-j)$ becomes $k_j = 2l(n-j)$, we write $\Delta_n = \Delta_n(-2l)$.

If $n = 1$, $m = 3l-1$ and the triangle defining the conditions is $\{2l, 0\}$ we have $F_1(x,y) = (1-x)^{-\alpha}\times C_{0,l}(x)+C_{1,l}(x)$ with both $\deg C_{0,l} < l$ and $\deg C_{1,l} < l$; as suggested by the notation. The resulting 'system'

$$(1-x)^{-\alpha}C_{0,l}(x)+C_{1,l}(x) = O(x^{2l}) \tag{3–3}$$

of course has no solution in general — there are $2l$ conditions and just $2l$ unknown coefficients. The fact that there is indeed no nontrivial solution in general is given precisely by the nonvanishing of the determinant $\Delta(\alpha) = \Delta(\alpha; 2l, 0)$ discussed at Section 2 above. Indeed we might 'pretend that' $\deg C_{0,l} = l$, that is, study the problem

$$(1-x)^{-\alpha}D_{0,l+1}(x)+D_{1,l}(x) = O(x^{2l}), \tag{3–4}$$

with $\deg D_{0,l+1} < l+1$, and $\deg D_{1,l} < l$, so returning to the Padé approximation problem (2–1). We saw that in general this has a unique solution up to normalisation, moreover with $\Delta(\alpha)$ the coefficient of x^l in $D_{0,l+1}$. We might now choose to notice that $\Delta(\alpha)$ is a polynomial in α of degree l^2 and determine those 'non-general' α for which $\Delta(\alpha)$ vanishes so that there is a nontrivial solution to (3–3) after all. Whatever, we refer to $\Delta(\alpha)$ as *the determinant* of the system (3–3).

In [Bombieri et al. 1995] we numerically evaluated for $n = 1, 2, \dots$ the determinant $\Delta_n = \Delta(-2l; 2nl,$ $2(n-1)l, \dots, 0)$ of the system (3–1) with $\alpha = -2l$ and discovered to our surprise that

$$\Delta_n = \Delta^{\binom{n+2}{3}},$$

with $\Delta = \Delta_1$ as above.

At the time, we neither understood why the Δ_n should be powers of Δ, nor why — if the Δ_n are such powers — the power should be $\binom{n+2}{3}$.

4. SOME PARTIAL EXPLANATIONS

Observation 4.1. *If Δ_n is a power of Δ it's no surprise that that power is $\binom{n+2}{3}$.*

Idea of argument. Consider both $\Delta(\alpha)$ and $\Delta_n(\alpha)$ as polynomials in α. Recall that the entries of $\Delta_n(\alpha)$ are of the shape

$$\frac{d^{i-1}}{dx^{i-1}}(1-x)^{-(n-b)\alpha}\binom{b}{v}x^{j-1}\Big|_{x=0} = (-1)^{i-j}\binom{b}{v}\binom{-(n-b)\alpha}{i-j}$$

and thus of respective degree $i-j$ in α. One might therefore guess that $\Delta_n(\alpha)$ is of degree $\sum(i-j)$, where the sum runs over all relevant pairs i, j. More specifically,

$$\begin{aligned}\sum j = (n+1)\sum_{j=1}^{ln} j &= \tfrac12(n+1)ln(ln+1)\\ &= \tfrac12 n^2(n+1)l^2 + \tfrac12 n(n+1)l\end{aligned}$$

and

$$\begin{aligned}\sum i = \sum_{v=0}^{n}\sum_{i=1}^{2l(n-v)} i &= \sum_{v=0}^{n}(2(n-v)^2l^2 + (n-v)l)\\ &= \tfrac13 n(n+1)(2n+1)l^2 + \tfrac12 n(n+1)l.\end{aligned}$$

So, presumably,

$$\begin{aligned}\deg\Delta_n(\alpha) &= \sum(i-j) = \tfrac16 n(n+1)(4n+2-3n)l^2\\ &= \binom{n+2}{3}l^2.\end{aligned}$$

Specifically, for $n=1$ one guesses that $\deg\Delta(\alpha)=l^2$. This is confirmed by (2–4), whence—comparing presumed degrees — the inappropriateness of any surprise. $\square$

Observation 4.2. *$\Delta_n(a)$ vanishes if $\Delta(a)$ vanishes.*

Proof. If $\Delta(a)$ vanishes, there is a nontrivial solution to the approximation system (3–3), say corresponding to a solution $F_1(x,y)$ to (3–1) with $n=1$, $\alpha=a$. But then, for every n, $(F_1(x,y))^n$ is a solution for the system (3–1) with $\alpha=a$. There being a solution, $\Delta_n(a)$ must vanish. $\square$

Observation 4.3. *It is reasonable to believe that $\Delta_n(a)$ vanishes only if $\Delta(a)$ vanishes.*

Idea of argument. We suppose, for convenience, that $C_{0,nl}(0)$ does not vanish. Given a solution $F_n(x,y)$ to (3–1) with $\alpha=a$, observe that the conditions

$$\begin{aligned}D_y^vF_n(x,y)\big|_{y=1} &= \sum\nolimits_{j=v}^{n}(1-x)^{-a(n-j)}\binom{j}{v}C_{j,nl}(x)\\ &= O\big(x^{2l(n-v)}\big),\end{aligned} \tag{4–1}$$

for $v=0, 1, \ldots, n$, may be viewed as reporting that, in particular, $F_n(x,y)$ has an n-tuple zero mod x^{2l} at $y=1$. It follows there are power series $\psi_d(x)$ so that

$$F_n(x,y) = C_{0,nl}(x)\prod_{d=1}^{n}\big((1-x)^{-a}-\psi_d(x)y\big)$$

where $(1-x)^{-a}-\psi_d(x) = O(x^{2l})$ respectively for each d. We may then replace the $\psi_d(x)$ by rational approximants $-P_d(x)/Q_d(x)$ where,

$$x^{2l} \quad \text{divides} \quad (1-x)^{-a}Q_d(x)+P_d(x).$$

Indeed, we obtain

$$\begin{aligned}F_n(x,y) = C_{0,nl}(x)\prod_{d=1}^{n}\big((1-x)^{-a}\\ +y\big(P_d(x)/Q_d(x)+O(x^{2l}/Q_d(x))\big)\big),\end{aligned}$$

suggesting that we could choose the Padé approximants so that the product of the $Q_d(x)$s divides $C_{0,nl}(x)$. But, adding degrees, and recalling that $\deg C_{j,nl}(x) < nl$ for all j, we then find that for some d we must have both $\deg P_d(x) < l$ and $\deg Q_d(x) < l$, entailing that $\Delta(a)=0$. $\square$

5. AN EXPLANATION

Main Theorem. $\Delta(\alpha)$ is some constant multiple of

$$\prod_{a=-(l-1)}^{l-1}(\alpha-a)^{l-|a|}$$

so that $\Delta(-l)=\pm1$, and

$$\pm\Delta_n(\alpha) = \big(\Delta(\alpha)\big)^{n(n+1)(n+2)/6}.$$

Remark 5.1. It might seem preferable to set $l!! = 1!2!3!\cdots l!$ and $[k]=(k-1)!!$ and then to claim that

$$\pm\Delta_n(-\alpha) = \left(\frac{[\alpha-l][\alpha+l][l]^2}{[\alpha]^2[2l]}\right)^{n(n+1)(n+2)/6}.$$

This formulation of the result seems tidy and evocative, and it makes sense for $\alpha = 2l$. However, the meaning of the notation $[\beta]$, other than for integers β, is not obvious, and in any case the formulation hides the clearer statement we make in the Theorem. If that's no problem, a formulation that might be preferred is

$$\pm\Delta(\alpha) = \prod_{i=1}^{l}\binom{\alpha+i-1}{2i-1}\Big/\binom{l+i-1}{2i-1}$$

and

$$\pm\Delta_n(\alpha) = \big(\Delta(\alpha)\big)^{n(n+1)(n+2)/6}.$$

Proof of for the 'trivial' case. We deal first with the easy case $n = 1$. Consider the system (3–3) with $\alpha = -a$, some integer $a = 0, 1, \dots, l-1$. It is then plain that the $l-a$ conditions $D_x^{l+a+i-1}\big((1-x)^a \times C_{0,l} + C_{1,l}\big)\big|_{x=0} = 0$ with $i = 1, 2, \dots, l-a$ are redundant, seeing that $(1-x)^a C_{0,l} + C_{1,l}$ is of degree less than $l+a$, so those conditions are *automatically* satisfied and of course the rank of the determinant is at least $l-a$ less than its potential $2l$ when $\alpha = -a$ (perhaps in better phrasing, the 'co-rank' of the system is at least $l-a$). It follows that $(\alpha+a)^{l-a}$ divides $\Delta_1(\alpha)$. By symmetry — we may multiply by $(1-x)^\alpha$ and reindex the $C_{j,l}$ — also $(\alpha-a)^{l-a}$ divides $\Delta_1(\alpha)$. Given that the degree of $\Delta_1(\alpha)$ is bounded by l^2, as noticed above, it follows that $\Delta_1(\alpha)$ is a constant multiple of

$$\alpha^l \prod_{a=1}^{l-1}(\alpha^2 - a^2)^{l-a}.$$

Glancing at (2–3), we see immediately that $\pm\Delta_1(-l)$ equals 1, allowing us to nominate the leading coefficient of $\Delta_1(\alpha)$, as we had claimed. □

To deal with the general case we will need to introduce some convenient notation and to announce various useful facts. Accordingly, we set

$$f_b(x;\alpha) = (1-x)^{-b\alpha}\sum\nolimits_{j=b}^{n}\binom{j}{b}(1-x)^{j\alpha}C_{j,nl}(x)$$

In this notation, our given task is the evaluation of the determinant belonging to the approximation system

$$(1-x)^{-(n-b)\alpha} f_b(x;\alpha) = O(x^{2(n-b)l}), \qquad (5\text{–}1)$$

for $b = 0, 1, \dots, n$. In the sequel a always denotes one of the integers $0, 1 \dots, l-1$.

We follow the spirit of our argument for the case $n = 1$. Thus, we approach the task of evaluating $\Delta_n(\alpha)$ by endeavouring to discern the row rank of certain $n(n+1)l$ by $n(n+1)l$ matrices; and we achieve that by considering the rank of certain simultaneous approximation systems. However, we expect to discover that $\alpha - a$ divides $\Delta_n(\alpha)$ with multiplicity $\frac{1}{6}n(n+1)(n+2)(l-a)$. Plainly, we cannot do that by discovering that high a co-rank of the system at $\alpha = a$, given that the system consists of only $n(n+1)l$ conditions. It must therefore be that certain linear combinations of the conditions vanish to high order at $\alpha = a$.

We arrange matters so as to make the required vanishing as obvious as possible. To that end we begin with some guiding remarks.

Useful Remark 5.2. For arbitrary β, the set of conditions $f_b(x;\alpha) = O(x^m)$ is equivalent to the m conditions $(1-x)^\beta f_b(x;\alpha) = O(x^m)$.

Proof. Said in this manner, a mildly painful combinatorial argument — at least to write in detail — can be displaced by the remark 'obvious'. □

We therefore see immediately that we can replace study of the approximation system (5–1) by study of the system

$$f_b(x;\alpha) = O(x^{2(n-b)l}) \qquad b = 0, 1, \dots, n. \qquad (5\text{–}2)$$

That's very useful, because the $f_b(x;a)$ are polynomials in x, moreover of degree less than $n(l+a) - ba$. Indeed, as we see below at Useful Remark 5.4, given other conditions comprising our simultaneous approximation system their *implied degree* is less than $n(l+a) - 2ba$.

First, notice that the approximation system

$$f_b(x;\alpha) = O(x^{2(n-b)l}) \qquad b = 0, 1, \dots, n,$$

sequentially implies the following. The conditions on $f_0(x;a)$ entail that the a leading coefficients of $C_{n,nl}$ vanish. We see that immediately because those a leading coefficients of $f_0(x;a)$ are isolated. The said conditions also entail that a combination of each of the a leading coefficients of $C_{n-1,nl}$ and of the next a coefficients of $C_{n,nl}$ vanishes. But the conditions on $f_1(x;a)$ entail that an independent linear combination of those sets of coefficients vanishes. Thus, the sets of conditions on both $f_0(x;a)$ and on $f_1(x;a)$

entail that those two pairs of collections of a coefficients vanish. And so on. I report such entailments by saying that, in $f_1(x;a)$ the *implied degree* of $C_{n,nl}$ is less than $nl-a$; and then in $f_2(x;a)$ the implied degree of $C_{n-1,nl}$ is less than $nl-a$ whilst that of $C_{n,nl}$ is less than $nl-2a$; and so on.

Second, I speak below of conditions being satisfied *automatically*. For example, because $\deg_x f_b(x;a)$ has implied degree less than $n(l+a)-2ba$, it makes sense to say, for those b so that $n(l+a)-2ba$ is less than $2(n-b)l$, that $(n-2b)(l-a)$ of the conditions $f_b(x;a)=O(x^{2(n-b)l})$ are automatically satisfied.

Third, I speak below of *implied conditions*. In particular, if some of the system's conditions are implied by other conditions already taken into account then those implied conditions are redundant. That redundancy displays a reduction in the system's rank.

Useful Remark 5.3. Recall that D_x^l denotes $(d/dx)^l/l!$. Given any power series $h(x)$ in x, a set of conditions $D_x^l\big(h(x)f_b(x;a)\big)=O(x^m)$ is entailed by the set $f_b(x;a)=O(x^{m+l})$. This generalises the preceding Useful Remark. Moreover, if, say, all but r of the latter $m+l$ conditions are known to be satisfied then the former m conditions by implication comprise at most r independent conditions. That is, $m-r$ of those conditions are satisfied by implication at $\alpha=a$.

The point in saying this is to signal a sufficient condition which our argument can show to be satisfied. Supposing $m>r$, our 'moreover' is an example of *implied* vanishing of $m-r$ conditions at $\alpha=a$.

Useful Remark 5.4. It is easy to see that, once given the conditions $f_0(x;a)=O\big(x^{n(l+a)}\big)$, for example, we may amongst other things suppose in $f_1(x;a)$, $f_2(x;a)$, ... that $\deg C_{n,nl}<nl-a$. It follows that $f_1(x;a)$ has implied degree less than $n(l+a)-2a$. But there's more. The conditions

$$f_0(x;a)=O\big(x^{n(l+a)}\big)$$

do not just control the leading a coefficients of $C_{n,nl}$. Moreover, (5–1) entails $f_1(x;a)=O\big(x^{(n-1)(l+a)}\big)$, which together with the implied degrees just remarked upon, yields in $f_2(x;a)$, $f_3(x;a)$, ... both $\deg C_{n,nl}<nl-2a$ and $\deg C_{n-1,nl}<nl-a$. It follows that $f_2(x;a)$ has implied degree less that $n(l+a)-4a$. And so on.

In summary, we see sequentially that each set of conditions

$$f_b(x;a)=O\big(x^{(n-b)(l+a)}\big),\quad b=0,1,\dots,k-1,$$

allows us to speak as if $\deg\big(f_k(x;a)\big)<n(l+a)-2ka$.

Proof. Our remark here emphasises the opportunity to see a decrease in the implied degree of the polynomials we study.

We might have done better to speak of, say, the 'rank' of those polynomials — whereby a polynomial of rank r is, in the first instance, one of degree at most $r-1$. Generally, though, a polynomial of rank r is one with no more than r unspecified coefficients. We might then recognise that preceding conditions assist in implying that the implied *rank* of succeeding polynomials is decreased. In the present arguments it is happenstance that this decrease in rank manifests itself as a decrease in implied degree. $\square$

Useful Remark 5.5.

$$\binom{n+2}{3}=n^2+(n-2)^2+\cdots+\begin{cases}1^2 & \text{if } n \text{ is odd,}\\ 2^2 & \text{if } n \text{ is even.}\end{cases}$$

This is a core observation, because it provides the guide on what to aim for in the argument. It's an embarrassment that it is too easy to be misguided by the more seductive identity

$$\sum_{b=1}^{n}\sum_{j=1}^{b}\sum_{i=1}^{j}1=\tfrac{1}{6}n(n+1)(n+2),$$

which somehow seems 'more true'.

Useful Remark 5.6. Having discovered that some condition is satisfied, whether automatically, or *by implication* — thus that it plainly depends linearly on other conditions of the system — we must next endeavour to notice its vanishing to higher multiplicity. To that end, we divide vanishing conditions by $\alpha-a$ and study the resulting conditions. We obtain

$$\lim_{\alpha\to a}D_x^i\big(f_b(x;\alpha)-f_b(x;a)\big)/(\alpha-a)\big|_{x=0}$$
$$=(b+1)\cdot D_x^i\big((1-x)^a\log(1-x)f_{b+1}(x;a)\big)\big|_{x=0}.$$

The useful miracle is the appearance of $f_{b+1}(x;a)$, albeit multiplied by a power series.

Proof. We are to satisfy conditions

$$f_b(x;a)=O\big(x^{2(n-b)l}\big),$$

whilst we know by implication that the degree of $f_b(x;a)$ is less than $n(l+a)-2ba$. Note that

$$2(n-b)l-\bigl(n(l+a)-2ba\bigr)=(n-2b)(l-a).$$

It follows, for each b with $2b<n$, that $(n-2b)(l-a)$ conditions vanish automatically when $\alpha=a$, exhibiting the presence of $(n-2b)(l-a)$ factors $\alpha-a$ of $\Delta_n(\alpha)$.

This useful remark tells us that a set of conditions $f_b(x;\alpha)=O\bigl(x^{2(n-b)l}\bigr)$ yields, for each a, a total of $(n-2b)(l-a)$ factors $\alpha-a$ of the determinant *and* leaves the sets of conditions

$$f_b(x;a)=O\bigl(x^{n(l+a)-2ba}\bigr), \tag{5–3}$$

$$D_x^{n(l+a)-2ba}\bigl((1-x)^a\log(1-x)f_{b+1}(x;a)\bigr) = O\bigl(x^{(n-2b)(l-a)}\bigr). \tag{5–4}$$

The presence of the first set (5–3) of conditions explains their use in Useful Remark 5.4. Notice that, by Useful Remark 5.3, the second set 5–4 of $(n-2b)(l-a)$ conditions is implied by the conditions $f_{b+1}(x;a)=O(x^{2(n-b)l})$.

Iteration of our present remarks, and application of Useful Remark 5.3, will show that ultimately the conditions $f_b(x;a)=O\bigl(x^{2(n-b)l}\bigr)$ allow us to discover, for each a, and each b such that $2b<n$, a factor $\alpha-a$ to multiplicity $(n-2b)^2(l-a)$. Thus, the relevance of Useful Remark 5.5. □

Useful Remark 5.7. Our approximation system is symmetric; it is unchanged by the transformation $\alpha\to-\alpha$.

Proof. We deal with the vanishing of certain partial derivatives of the function

$$(1-x)^{n\alpha}F_n(x,y;\alpha)=\sum_{j=0}^{n}(1-x)^{j\alpha}C_{j,nl}(x)y^j$$

at $(0,1)$. It is then just a variant of Useful Remark 5.2 that it is no change to replace y by y^{-1}, nor, therefore, to deal with the same collection of partial derivatives at $(0,1)$ of

$$y^nF_n(x,y^{-1};-\alpha)=\sum_{j=0}^{n}(1-x)^{j\alpha}C_{n-j,nl}(x)y^j. \quad \square$$

Useful Remark 5.8. There's no need to study the determinant itself; one can see, say in the case $n=1$ that $\Delta(l)=\pm1$, by pure thought.

Proof. First, certainly $\Delta(\alpha)$ is not identically zero. Thus $\Delta(l)\neq0$, and so (3–4) with $\alpha=l$ has a unique solution, up to normalisation. But

$$(1-x)^{-l}\cdot(1-x)^l-1=O(x^{2l}) \tag{5–5}$$

displays that unique solution.

Now suppose that $\Delta(l)\equiv0 \pmod p$, for some prime p. Then the approximation problem $f_0(x;l)=O(x^{2l})$ has a nontrivial solution over the finite field $\mathbb{F}_p$ and that nontrivial solution will lift to a solution of

$$D_{0,l+1}+(1-x)^lD_{1,l}=O(x^{2l})$$

in characteristic zero. In that solution though, the leading coefficient of $D_{0,l+1}$ must vanish modulo p, contradicting the uniqueness of the solution (5–5).

In summary, our noticing a solution, in this example as at (5–5), allows us to conclude that the only primes dividing the determinant are those dividing the 'extra' coefficient in that solution. □

Proof of the Theorem. Our strategy is to fix a, and then to discover the multiplicity of the factor $\alpha-a$ of $\Delta_n(\alpha)$. We will find that the factor $\alpha-a$ appears as N groups of $l-a$ factors, with N independent of a. Given Observations 4.2 and 4.3, that would of itself suffice to show that $\Delta_n(\alpha)$ is an Nth power of $\Delta(\alpha)$, as we wish to show; and that $N=\frac16n(n+1)(n+2)$. However, those are just remarks, so as suggested by Useful Remark 5.5, we find that value of N directly.

We discover the groups of factors sequentially, as suggested by Useful Remark 5.5. Our primary tool is Useful Remark 5.4 whereby the implied degree of $f_b(x;a)$ — thus, in the presence of the rest of the system — is less than $n(l+a)-2ba$.

It is helpful to distinguish the cases $n=2m+1$, odd, and $n=2m$, even. We will deal with the case $n=2m+1$ odd in detail.

That leads us to view the system we are studying as the sets of conditions

$$f_{m-k}(x;\alpha)=O(x^{2(m+1+k)l}),\quad k=m,\,m-1,\,\ldots,\,0,$$

from which we discover the required factors of the determinant; and the sets of conditions

$$f_{m+k}(x;\alpha)=O(x^{2(m+1-k)l}),\qquad k=1,\,2,\,\ldots,\,m,$$

which assist in implying the required reductions in rank of the system at $\alpha=a$.

The reader should fasten her seatbelt before speeding through the following compacted proof. A test drive of the argument in the cases $m = 0, 1, 2, \ldots$ is recommended before the general journey.

We start an argument by induction on k by noticing that, because $f_m(x;a)$ has implied degree less than $(2m+1)(l+a)-2ma = 2(m+1)l-(l-a)$, it follows that $l-a$ of the conditions

$$f_m(x;a) = O(x^{2(m+1)l})$$

are automatically satisfied; the remaining conditions play a role in reducing the implied degree of the $f_{m+k}(x;a)$ in the sets of conditions 'below'. We suppose we have divided the automatically vanishing conditions by their factor $\alpha - a$ leaving us — see Useful Remark 5.6 — with $l-a$ conditions made up from the set of conditions $f_{m+1}(x;a) = O(x^{2(m+1)l})$; in brief, with '$l-a$ conditions implied by that set'.

Now suppose, as an induction assumption, that for $k = 0, 1, \ldots, s-1$ we have found that ultimately each of the conditions $f_{m-k}(x;a) = O(x^{2(m+1+k)l})$ yields $(2k+1)^2(l-a)$ factors $\alpha - a$ of the determinant $\Delta_n(\alpha)$, with the vanishing conditions leaving $(2k+1)(l-a)$ conditions implied by the set

$$f_{m+k+1}(x;a) = O(x^{2(m+1+k)l}).$$

We then show, given this context, that also the set of conditions $f_{m-s}(x;a) = O(x^{2(m+1+s)l})$ yields $(2s+1)^2(l-a)$ factors.

We have already done the case $s = 0$. In general, observe that the first $(s+1)(2s+1)(l-a)$ factors arise sequentially by noticing that, to begin, $(2s+1)(l-a)$ of the given conditions are satisfied automatically, and by Useful Remark 5.6 leave us with $(2s+1)(l-a)$ conditions implied by the set

$$f_{m-s+1}(x;a) = O(x^{2(m+1+s)l}).$$

However, for $t = 1, \ldots, s$, each set of conditions

$$f_{m-s+t}(x;a) = O(x^{2(m+1+s)l}) \tag{5–6}$$

is plainly satisfied by implication. Indeed, the implied degree of $f_{m-s+t}(x;a)$ must be less than the quantity $2(m+1+s-t)l$. So the conditions

$$f_{m-s+t}(x;a) = O(x^{2(m+1+s-t)l}),$$

elsewhere part of our system, in fact coincide with (5–6). Thus it suffices to recall that, now that we have noticed that the $(2s+1)(l-a)$ conditions implied by (5–6) *are* satisfied, we may extract the $(2s+1)(l-a)$ corresponding factors $\alpha - a$ and be led by Useful Remark 5.6 to consider the $(2s+1)(l-a)$ conditions implied by

$$f_{m-s+t+1}(x;a) = O(x^{2(m+1+s)l}).$$

Doing that for each t yields $(s+1)(2s+1)(l-a)$ factors in all, and also leaves us with $(2s+1)(l-a)$ conditions implied by the set

$$f_{m+1}(x;a) = O(x^{2(m+1+s)l}).$$

We find the remaining $s(2s+1)(l-a)$ factors in much the same manner, by verifying that each set

$$f_{m+t}(x;a) = O(x^{2(m+1+s)l}) \quad t = 1, 2, \ldots, s \tag{5–7}$$

is satisfied by implication. As above, we note that our system includes the set

$$f_{m+t}(x;a) = O(x^{2(m+1-t)l}),$$

and that the implied degree of $f_{m+t}(x;a)$ is less than $(2m+1)(l+a)-2(m+t)a$. However, that is just $(2m+1)l-(2t-1)a$, so not all, but all except

$$(2m+1)l-(2t-1)a-2(m+1-t)l = (2t-1)(l-a)$$

of the conditions

$$f_{m+t}(x;a) = O(x^{2(m+1+s)l})$$

can be immediately seen to vanish by implication. Fortunately — it's here that we use the induction assumption — the original conditions $f_{m-t+1}(x;a) = O(x^{2(m+t)l})$ have left us with $(2t-1)(l-a)$ conditions implied by the set

$$f_{m+t}(x;a) = O(x^{2(m+t)l}).$$

Thus the remaining $(2t-1)(l-a)$ conditions do also vanish by implication, verifying our claim to be able to extract an additional $s(2s+1)(l-a)$ factors $\alpha - a$ from the determinant of the approximation system.

In summary, the factor $(\alpha-a)^{l-a}$ divides the determinant $\Delta_n(\alpha)$ to multiplicity at least

$$1^2+3^2+\cdots+(2m+1)^2 = \tfrac{1}{6}n(n+1)(n+2).$$

The argument for $n = 2m$, even, is now an exercise almost fully effected by replacing $m+1$ — wherever it appears above as such — by m, and then undoing that 'correction' at the few places where it is inappropriate.

That argument shows that the factor $(\alpha-a)^{l-a}$ divides the determinant $\Delta_n(\alpha)$ to multiplicity at least

$$0^2+2^2+4^2+\cdots+(2m)^2 = \tfrac{1}{6}n(n+1)(n+2).$$

It remains to invoke Useful Remark 5.7 to see that the factors $(\alpha+a)^{l-a}$ also appear with that same multiplicity.

It follows, from the computation reported at Observation 4.1, that $\Delta_n(\alpha)$ differs from that power of the said product only by multiplication by a constant.

We may now apply Observations 4.2 and 4.3 to point out that $\pm\Delta_n(\alpha)$ is indeed the $\frac{1}{6}n(n+1)(n+2)$ th power of $\pm\Delta(\alpha)$. Those remarks hold regardless of field of definition, so that for all primes p we must have $\Delta_n(l)\equiv 0 \pmod p$ if and only if $\Delta(l)\equiv 0 \pmod p$.

However, we cannot rely on those Observations, and it is in any case more instructive to spell out some details and to prove $\pm\Delta_n(l)=1$ more directly, in the spirit of Useful Remark 5.8.

We first note, by Observation 4.1 on the degree of $\Delta_n(\alpha)$, and the main argument just completed, that $\Delta_n(l)=0$ if and only if $\Delta_n(\alpha)=0$ identically. But the latter is absurd; consider α transcendental.

Next, given that $\Delta_n(l)\neq 0$, if

$$g_b(x;\alpha)=(1-x)^{-b\alpha}\sum\nolimits_{j=i}^{n}\binom{j}{b}(1-x)^{j\alpha}D_{j,nl+\delta_{j,0}}(x),$$

then the system

$$(1-x)^{-(n-b)\alpha}g_b(x;\alpha)=O(x^{2(n-b)l}),$$

for $b=0,1,\ldots,n$, has a unique solution up to normalisation with $\Delta_n(l)$ the coefficient of x^{nl} in $D_{0,nl+\delta_{j,0}}(x)$. Moreover, a solution is given by

$$D_{j,nl+\delta_{j,0}}(x)=(-1)^j\binom{n}{n-j}(1-x)^{n-j},$$

for $j=0,1,\ldots,n$. Because the 'extra' coefficient — that of x^{nl} in $D_{0,nl+1}$ — is ±1 we may conclude that $\Delta_n(l)$ is not divisible by any prime. In other words, $\Delta_n(l)=\pm1$. $\square$

6. THE GENERAL TRIANGLE FOR $n = 2$

The preceding argument deals with a very special case of the general construction problem that inspired it, and moreover with a seemingly irrelevant case, in that the Observations at Section 4 provide a complete solution to the approximation problem, without any appeal to the evaluation of determinants. Nonetheless, we now show that precisely the ideas we employed — particularly the Useful Remarks of Section 5 — readily suffice to give a succinct evaluation of the determinant $\Delta(b,c,0)$ for the general triangle defined by $k_0=b$, $k_1=c$, $k_2=0$ in the case $n=2$. That is, we evaluate the determinant of the system

$$\begin{aligned}(1-x)^bC_{0,c}+(1-x)^cC_{1,b-c}+C_{2,c}&=O(x^b),\\ (1-x)^cC_{1,b-c}+2C_{2,c}&=O(x^c).\end{aligned}$$

Krattenthaler and Zeilberger [1997] provided a detailed evaluation of the determinant $\Delta(b,c,0)$, confirming a 'computational guess' reported in [Bombieri et al. 1995] to the following effect:

Principal Motivation. (i) $\Delta(b,c,0)=0$ if b is even and c is odd.

(ii) If one of these conditions does not hold, and $2c<b$, then,

$$\pm\Delta(b,c,0)=\frac{[\frac12(2b-c)]^2[b-2c][\frac12(b+c)]^2[\frac12(b-c)]^6[\frac12c]^6}{[b-c]^3[\frac12b]^6[\frac12(b-2c)]^2[c]^3},$$

where $[s]=\prod_{k=0}^{s-1}k!$ if s is an integer, and $[s]^2=(s+\frac12)!!\,(s-\frac12)!!$ if $s\in\mathbb{Z}+\frac12$.

(iii) Otherwise, if $2c>b$, then

$$\pm\Delta(b,c,0)=2^{b-2c}\Delta(b,b-c,0).$$

The present work is, as it were, a reaction to this result, my motive being to prove it, and much more, in considerably fewer pages and with very much less effort. To that end we set $\delta=\frac12(b-2c)$, and to emphasise the symmetry as at Useful Remark 5.7, we study the simultaneous approximation problem

$$\begin{aligned}(1-x)^{-\alpha}C_{0,c}+(1-x)^{-\delta}C_{1,c+2\delta}+(1-x)^{\alpha}C_{2,c}&=O(x^{2c+2\delta}),\\ (1-x)^{-\delta}C_{1,c+2\delta}+2(1-x)^{\alpha}C_{2,c}&=O(x^c).\end{aligned}$$

The determinant $\Delta(\alpha;b,c,0)$ of this system yields the $\Delta(b,c,0)$ required by the Principal Motivation on setting $\alpha=-\frac12b=-(c+\delta)$.

More precisely, much as here we actually evaluate $\Delta(\alpha;b,c,0)$ and get $\Delta(b,c,0)$ from it, Krattenthaler and Zeilberger [1997] also detail the evaluation of a more general determinant with one additional parameter. Indeed, without introduction of those respective extra parameters, the two evaluations could not be done at all. That issue is usefully discussed in the survey [Krattenthaler 1999]. However, the parameter x cleverly found in [Krattenthaler and Zeilberger 1997] is different from the rather more natural parameter α introduced here, and the values of the two determinants are not related in an obvious way, except that they coincide and yield the Principal Motivation for $x=0$ and respectively $\alpha=-\frac{1}{2}b$.

We show, as a corollary of the Useful Remarks leading to our Main Theorem:

Example Application. (i) $\Delta(\alpha;b,c,0)=0$ *if b is even and c is odd.*

(ii) *If one of these conditions does not hold, and $2c<b$, then $\Delta(\alpha;b,c,0)$ is a constant multiple of*

$$\prod_{|2a|\le b-2c}(\alpha-a)^c\prod_{0<|2a|<c}(\alpha-a)^{c-2a}\prod_{b-2c<|2a|<b-c}(\alpha-a)^{b-c-2a}.$$

(iii) *Otherwise, if $2c>b$, it is some constant multiple of $\Delta(\alpha;b,b-c,0)$. That is, $\Delta(\alpha;b,c,0)$ is a constant multiple of the polynomial above with c replaced by $b-c$, to wit of*

$$\prod_{|2a|\le 2c-b}(\alpha-a)^{b-c}\prod_{0<|2a|<b-c}(\alpha-a)^{b-c-2a}\prod_{2c-b<|2a|<c}(\alpha-a)^{c-2a}.$$

Here, thus in the two preceding expressions, the index a is such that $2a$ ranges over even integers when b, and thus also c, is even. If b is odd, the $2a$ are odd integers.

Proof. Set $\delta=\frac{1}{2}(b-2c)$. To emphasise the symmetry as at Useful Remark 5.7, it will be convenient to study the simultaneous approximation problem

$$(1-x)^{-\alpha}C_{0,c}+(1-x)^{-\delta}C_{1,c+2\delta}+(1-x)^{\alpha}C_{2,c}=O(x^{2c+2\delta}),$$
$$(1-x)^{-\delta}C_{1,c+2\delta}+2(1-x)^{\alpha}C_{2,c}=O(x^{c}),$$

where $\Delta(\alpha;b,c,0)$, the determinant of this system, yields the $\Delta(b,c,0)$ required by the Principal Motivation on setting $\alpha=-\frac{1}{2}b=-(c+\delta)$. It's not difficult to see — in the spirit of the explanations sketched at Section 4 — that the degree of the polynomial $\Delta(\alpha;b,c,0)$ is at most

$$\sum(i-j)=\tfrac{1}{2}b(b+1)+\tfrac{1}{2}c(c+1)-2\cdot\tfrac{1}{2}c(c+1)-\tfrac{1}{2}(b-c)(b-c+1)=(b-c)c.$$

We first suppose that δ is nonnegative. It will also be convenient to proceed as if b is even — so that δ is an integer. Given that, it is plain that for each integer a so that $0\le a\le\delta$ precisely c of the conditions in the set

$$(1-x)^{\delta-\alpha}C_{0,c}+C_{1,c+2\delta}+(1-x)^{\delta+\alpha}C_{2,c}=O(x^{2c+2\delta}) \tag{6–1}$$

are satisfied automatically. That reveals the factor $(\alpha-a)^c$ of $\Delta(\alpha;b,c,0)$ for each $0\le a\le\delta$. Then, on dividing the automatically satisfied conditions by $\alpha-a$, and again setting $\alpha=a$, we have

$$C_{1,c+2\delta}+2(1-x)^{\delta+a}C_{2,c}=O(x^c) \tag{6–2}$$

and, recall Useful Remark 5.3, a set of c conditions implied by the set

$$g_1(x;a)=-(1-x)^{\delta-a}C_{0,c}+(1-x)^{\delta+a}C_{2,c}=O(x^{2c+2\delta}).$$

It is plain by subtracting (6–2) from (6–1) that, in the context, (6–2) is equivalent to

$$g_1(x;a)=-(1-x)^{\delta-a}C_{0,c}+(1-x)^{\delta+a}C_{2,c}=O(x^c).$$

Since $(1-x)^{a-\delta}g_1(x;a)$ has degree less than $c+2a$, and given that the original system contains the set $g_1(x;a)=O(x^c)$, the conditions $g_1(x;a)=O(x^{2c+2\delta})$ clearly comprise at most $2a$ conditions neither vanishing automatically nor by implication. That entails, provided of course that $0\le 2a<c$, the implied vanishing of $c-2a$ conditions, thus revealing the additional factor $(\alpha-a)^{c-2a}$ of the determinant.

If $\delta<a<\delta+\frac{1}{2}c$, the set

$$f_0(x;\alpha)=C_{0,c}+(1-x)^{\alpha-\delta}C_{1,c+2\delta}+(1-x)^{2\alpha}C_{2,c}=O(x^{2c+2\delta})$$

displays $c+2\delta-2a$ automatically vanishing conditions, revealing for each of those a the factor

$$(\alpha-a)^{2\delta+c-2a}.$$

Moreover, if $0\le a<\delta+\frac{1}{2}c$ then the conditions $f_0(x;a)=O(x^b)$ entail that $f_0(x;a)$ vanishes identically. For $0\le a\le\delta$, above, that does not entail a reduction in the implied degree of $C_{2,c}$, but if $\delta<a<\delta+\frac{1}{2}c$ we see that the implied degree of $C_{2,c}(x)$ must be less than $c+a-\delta$.

Much as above, we are next led to study the $2\delta+c-2a$ derived conditions implied by the set

$$f_1(x;a)=C_{1,c+2\delta}+2(1-x)^{a+\delta}C_{2,c}=O(x^{2c+2\delta})$$

in the presence of the original set of conditions

$$f_1(x;a)=C_{1,c+2\delta}+2(1-x)^{a+\delta}C_{2,c}=O(x^c).$$

However, as just remarked, the implied degree of $f_1(x;a)$ is less than $c+2\delta$, and we have $f_1(x;a) = O(x^c)$. Thus the set $f_1(x;a) = O(x^{2c+2\delta})$ comprises at most 2δ conditions neither vanishing automatically nor by implication. That entails, provided that $\delta < a < \frac{1}{2}c$, the implied vanishing of $(2\delta+c-2a)-2\delta = c-2a$ conditions, revealing therefore the additional factor $(\alpha-a)^{c-2a}$ of the determinant. By the way, we note that if $c \le 2\delta$ then there were no additional factors to be found in the present manner for $a > \delta$.

Finally, we acknowledge that by symmetry it follows that for $a \neq 0$, each factor $\alpha-a$ we have noticed is partnered by a corresponding factor $\alpha+a$.

It is clear that $\alpha+\delta$ must be an integer in order for our arguments to make sense. Thus if b is not even the a must differ by half from integers; other than for that change our argument remains the same.

It is now convenient to count the number of factors of $\Delta(\alpha;b,c,0)$ thus far discovered.

We recall that on the first pass we found c factors corresponding to each a so that $0 \le a \le \delta$, and $2\delta+c-2a$ factors corresponding to each a so that $2\delta < 2a < 2\delta+c$. On the second pass we found $c-2a$ factors corresponding to each a so that $0 \le 2a < c$.

Suppose we first count those 'second pass' factors. If b is even there are c corresponding to $a=0$ and for $a = 1, 2, \ldots,$ a total of

$$(c-2)+(c-4)+\cdots+\begin{cases} 1 = \frac{1}{4}(c-1)^2 & \text{if } c \text{ is odd,} \\ 2 = \frac{1}{4}(c-2)c & \text{if } c \text{ is even.} \end{cases}$$

If b is odd, then for $a = \frac{1}{2}, 1\frac{1}{2}, \ldots,$ we have found a number of factors equal to

$$(c-1)+(c-3)+\cdots+\begin{cases} 1 = \frac{1}{4}c^2 & \text{if } c \text{ is even,} \\ 2 = \frac{1}{4}(c^2-1) & \text{if } c \text{ is odd.} \end{cases}$$

Similarly, we see that the number of 'first pass' factors corresponding to a so that $2\delta < 2a < 2\delta+c$ totals

$$(c-2)+(c-4)+\cdots+\begin{cases} 1 = \frac{1}{4}(c-1)^2 & \text{if } c \text{ is odd,} \\ 2 = \frac{1}{4}(c-2)c & \text{if } c \text{ is even.} \end{cases}$$

The point is that $2a-2\delta$ is even regardless of the parity of b.

We now also recall that for positive a each factor $\alpha-a$ is accompanied by a factor $\alpha+a$. Hence the total number of factors counted thus far is

- if b is odd and c is odd:
$$2\cdot\tfrac{1}{4}(c^2-1)+2\cdot\tfrac{1}{4}(c-1)^2 = c^2-c;$$
- if b is odd and c is even:
$$2\cdot\tfrac{1}{4}c^2+2\cdot\tfrac{1}{4}(c-2)c = c^2-c;$$
- if b is even and c is odd:
$$c+2\cdot\tfrac{1}{4}(c-1)^2+2\cdot\tfrac{1}{4}(c-1)^2 = c^2-c+1;$$
- if b is even and c is even:
$$c+2\cdot\tfrac{1}{4}(c-2)c+2\cdot\tfrac{1}{4}(c-2)c = c^2-c.$$

Finally, if b is even the number of 'first pass' factors for a in the range $0 \le a \le \delta$ is c corresponding to $a=0$ and for $a = 1, 2, \ldots,$ a total of $c\delta$. Recalling also the factors corresponding to negative a, we have here a total of $c+2c\delta$ factors. If b is odd the number of first pass factors for a in the range $-\delta \le a \le \delta$ corresponding to $a = \pm\frac{1}{2}, \pm1\frac{1}{2}, \ldots, \pm\delta$ is also $c+2c\delta$. That is, the number of additional factors in each case is $c+c(b-2c)$.

One now sees that, in summary, we have found a total of $(b-c)c$ factors unless b is even and c is odd, in which case we have discovered $(b-c)c+1$ factors! However, we commenced by noticing that the degree of the polynomial $\Delta(\alpha;b,c,0)$ is at most $(b-c)c$. Thus, if b is even and c is odd we must have $\Delta(\alpha;b,c,0) = 0$ identically.

In the other cases we have shown that $\Delta(\alpha;b,c,0)$ is some constant multiple of the polynomial

$$\prod_{|2a|\le b-2c}(\alpha-a)^c \prod_{0<|2a|<c}(\alpha-a)^{c-2a} \prod_{b-2c<|2a|<b-c}(\alpha-a)^{b-c-2a}. \tag{6–3}$$

The $2a$ are even integers if b is even and thus also c is even. If b is odd, the $2a$ are odd integers.

We now suppose that δ is negative. We've done less than half our work because, with $\eta = \frac{1}{2}(2c-b)$ positive, study of the determinant $\Delta(\alpha;b,c,0)$ of the system

$$\begin{aligned} f_0(x;\alpha) &= C_{0,c}+(1-x)^{\alpha+\eta}C_{1,c-2\eta}+(1-x)^{2\alpha}C_{2,c} \\ &= O(x^b), \\ f_1(x;\alpha) &= (1-x)^{\eta}C_{1,c-2\delta}+2(1-x)^{\alpha}C_{2,c} = O(x^c) \end{aligned}$$

seems somewhat different from our discussion above. It will again be convenient to conduct our principal discussion as if b is even. As before, when b is even a denotes a nonnegative integer.

Suppose first that $0 \le 2a < b-c$. Then $b-c-2a$ of the conditions $f_0(x;\alpha) = O(x^b)$ are automatically satisfied, revealing a factor $(\alpha-a)^{b-c-2a}$ of the determinant $\Delta(\alpha;b,c,0)$.

Therefore at the second pass we have for $0 \le 2a \le b-c$ a collection of $b-c-2a$ conditions implied by the set $f_1(x;a) = O(x^b)$, as well as the set $f_1(x;a) = O(x^c)$ of conditions.

Moreover, the conditions $f_0(x;a) = O(x^b)$ entail that $f_0(x;a)$ vanishes identically. It follows that, when $0 \le \alpha \le \eta$, the implied degree of $C_{2,c}$ is less than $c-2a$; and then also the implied degree of $(1-x)^{-a}f_1(x;a)$ is less than $c-2a$. However, if $\eta \le a$, the implied degree of $C_{2,c}$ is less than $c-\eta-a$, and the implied degree of $(1-x)^{-\eta}f_1(x;a)$ is less than $c-2\eta$.

Thus if $0 \le a \le \eta$, then because the implied degree of $(1-x)^{-a}f_1(x;a)$ is less than $c-2a$, the conditions $f_1(x;a) = O(x^c)$ themselves yield $2a$ factors $\alpha-a$. Further that implied degree also implies that the set $f_1(x;a) = O(x^b)$ is satisfied, yielding an additional $b-c-2a$ such factors. In all, we see an additional factor $(\alpha-a)^{b-c}$ of $\Delta(\alpha;b,c,0)$.

However, if $\eta \le a$ then, because the implied degree of $(1-x)^{-\eta}f_1(x;a)$ is less than $c-2\eta$, the conditions $f_1(x;a) = O(x^c)$ themselves yield 2η factors $\alpha-a$ and the set $f_1(x;a) = O(x^b)$ is satisfied. Since $(b-c-2a)+2\eta = c-2a$, this shows the additional factor $(\alpha-a)^{c-2a}$ of $\Delta(\alpha;b,c,0)$.

We now turn to the cases $b-c \le 2a$. Here we use the advice appended to Useful Remark 5.4. In this case none of the conditions $f_0(x;a) = O(x^b)$ is automatically satisfied. However, if $0 \le a \le \eta$, then $c \ge \eta+a+b-c$ so the set of conditions $f_0(x;a) = O(x^b)$ implies $b-c$ linear conditions on $(1-x)^{2a}C_{2,c}$, specifically, the set $D_x^c\big((1-x)^{2a}C_{2,c}\big) = O(x^{b-c})$. It follows that $b-c$ of the conditions $(1-x)^{-\alpha}f_1(x;a) = O(x^c)$ are satisfied by implication, revealing a factor $(\alpha-a)^{b-c}$ of the determinant $\Delta(\alpha;b,c,0)$.

Further, if $\eta \le a$ then $\eta+a+b-c \ge c$ and the set of conditions $f_0(x;a) = O(x^b)$ implies just $c-a-\eta$ linear conditions on $(1-x)^{2a}C_{2,c}$, namely, the set $D_x^{c+a-\eta}\big((1-x)^{2a}C_{2,c}\big) = O(x^{c-a-\eta})$. It follows that $(c-a-\eta)-(a-\eta) = c-2a$ of the conditions

$$(1-x)^{-\eta}f_1(x;a) = O(x^c)$$

are satisfied by implication, and this reveals a factor $(\alpha-a)^{c-2a}$ of $\Delta(\alpha;b,c,0)$.

We may now count the factors of $\Delta(\alpha;b,c,0)$ thus far obtained, once again remarking that $2a$ is odd if b is odd. Fortunately, there is little new work for us to do. We recall that if $0 \le 2a < b-c$ we have $(b-c-2a)+(b-c)$, or $(b-c-2a)+(c-2a)$, factors according as $0 \le a \le \eta$, or $\eta \le a$. Whereas if $b-c < 2a$, we have just $b-c$, or $c-2a$, factors.

The terms $b-c-2a$ yield $(b-c-2)+(b-c-4)+\cdots$ or $(b-c-1)+(b-c-3)+\cdots$ factors according as b is even or odd, and the $c-2a$, which have $a > \eta$, so $0 < c-2a < b-c$, provide $(c-2)+(c-4)+\cdots$ factors.

Again we recall that by symmetry it follows that for $a \ne 0$, each factor $\alpha-a$ we have noticed is partnered by a corresponding factor $\alpha+a$. In particular the multiplicities $b-c$ provide a total of $(b-c)+(2c-b)(b-c)$ factors.

Happily, it is now evident that we are about to repeat our calculations for the case $b \ge 2c$, other than that c is everywhere replaced by $b-c$. Thus, unless b is even and $b-c$ is odd, we see we have found

$$\big(b-(b-c)\big)(b-c) = c(b-c)$$

factors. If b is even and c is odd, we have found $c(b-c)+1$ factors, and it follows that $\Delta(\alpha;b,c,0)$ must vanish identically. In summary, we see that in the other cases the determinant $\Delta(\alpha;b,c,0)$ is some constant multiple of the polynomial

$$\prod_{|2a|\le 2c-b}(\alpha-a)^{b-c}\prod_{0<|2a|<b-c}(\alpha-a)^{b-c-2a}\prod_{2c-b<|2a|<c}(\alpha-a)^{c-2a} \tag{6–4}$$

with the $2a$ even integers if b is even, and odd integers if b is odd. ☐

7. REMARKS AND ACKNOWLEDGEMENTS

This work had its genesis in the early eighties; that story is recounted *in extenso* in [Bombieri et al. 1995]. Given that, it does seem some sort of achievement to finally be able to prove the Main Theorem. More to the point, it seems encouraging to have been able to do it by the use of ideas which are likely to be capable of application to the more general constructions alluded to in [Bombieri et al. 1995]. For an application of the implicit constructions see [Bombieri et al. 1996].

Nonetheless this draft of the argument is polluted by several disgraceful scandals. The worst of these surely is my inability to add an incidental remark to those in Section 4 making it immediately plain that $\Delta_n(\alpha)$ is a power of $\Delta(\alpha)$. That, however, is mitigated by the contribution of the 'Useful Remarks' of Section 5. I am therefore more outraged by the scandal whereby I can see only indirectly that there is a symmetry causing $\Delta(\alpha; b, c, 0)$ to be a constant multiple of $\Delta(\alpha; b, b-c, 0)$.

Moreover, that defect in my arguments helped to hold me back from sketching an elegant normalisation for $\Delta(\alpha; b, c, 0)$, as I had expected to be able to do. In the event, I am forced to rely on [Krattenthaler and Zeilberger 1997] by equating the cases $\alpha = -\frac{1}{2}b$ of this paper and $x = 0$ of [Krattenthaler and Zeilberger 1997]. I had expected to be able to effect the normalisation by evaluating the polynomials (6–3) and (6–4) at $\alpha = 0$, or $\alpha = \frac{1}{2}$, according as b is odd, or even.

I am very much indebted to Christian Krattenthaler and Doron Zeilberger for motivating me finally to produce the present efforts, and acknowledge the assistance of Paula Cohen in reading earlier drafts of these remarks. This paper has been a beneficiary of the unusually detailed and very helpful remarks of its referee.

REFERENCES

[Bombieri 1997] E. Bombieri, "The equivariant Thue–Siegel method", pp. 70–86 in *Arithmetic geometry* (Cortona, 1994), edited by F. Catanese, Symposia Math. **37**, Cambridge Univ. Press, Cambridge, 1997.

[Bombieri and Vaaler 1983] E. Bombieri and J. Vaaler, "On Siegel's lemma", *Invent. Math.* **73**:1 (1983), 11–32. Addendum in **75**:2 (1984), 377.

[Bombieri et al. 1995] E. Bombieri, D. C. Hunt, and A. J. van der Poorten, "Determinants in the study of Thue's method and curves with prescribed singularities", *Experiment. Math.* **4**:2 (1995), 87–96.

[Bombieri et al. 1996] E. Bombieri, A. J. Van der Poorten, and J. D. Vaaler, "Effective measures of irrationality for cubic extensions of number fields", *Ann. Scuola Norm. Sup. Pisa Cl. Sci.* (4) **23**:2 (1996), 211–248.

[Krattenthaler 1999] C. Krattenthaler, "Advanced determinant calculus", *Sém. Lothar. Combin.* **42** (1999), B42q, 67 pp. See http://www.mat.univie.ac.at/~slc.

[Krattenthaler and Zeilberger 1997] C. Krattenthaler and D. Zeilberger, "Proof of a determinant evaluation conjectured by Bombieri, Hunt and van der Poorten", *New York J. Math.* **3** (1997), 54–102. See http://nyjm.albany.edu:8000/nyjm.html.

[van der Poorten 1991] A. J. van der Poorten, "Generalised simultaneous approximation of functions", *J. Austral. Math. Soc. Ser. A* **51**:1 (1991), 50–61.

[Zeilberger 1996] D. Zeilberger, "Reverend Charles to the aid of Major Percy and Fields medalist Enrico", *Amer. Math. Monthly* **103**:6 (1996), 501–502.

Alfred J. van der Poorten, ceNTRe for Number Theory Research, Macquarie University, Sydney, NSW 2109, Australia (alf@math.mq.edu.au)

Received November 29, 1999; accepted in revised form January 24, 2001