JOURNAL OF COMBINATORIAL THEORY, Series A 43, 70–84 (1986)

Shanks' Convergence Acceleration Transform, Padé Approximants and Partitions

GEORGE E. ANDREWS*

Pennsylvania State University University Park, Pennsylvania 16802

IAN P. GOULDEN†

AND

DAVID M. JACKSON‡

The University of Waterloo, Waterloo, Ontario N2L 3G1

Communicated by the Managing Editors

Received November 10, 1985

Shanks developed a method for accelerating the convergence of sequences. When applied to classical sequences in number theory, Shanks' transform yields some famous identities of Euler and Gauss. It is shown here that the Padé approximants for the little q-Jacobi polynomials can be used to explain and extend Shanks' observations. The combinatorial significance of these results is also discussed. © 1986 Academic Press, Inc.

1. INTRODUCTION

In the early 1950's, Shanks [10] studied the following convergence acceleration method: Let A_n be a sequence converging to the limit L. Define for $n \geqslant k$,

* Partially supported by National Science Foundation Grant DMS-8503324

† Partially supported by a grant from the Natural Sciences and Engineering Research Council of Canada.

‡ Partially supported by a grant from the Natural Sciences and Engineering Research Council of Canada.

0097-3165/86 $3.00

$$B_{k,n} = \frac{\begin{vmatrix} A_{n-k} & \cdots & A_{n-1} & A_n \\ \Delta A_{n-k} & \cdots & \Delta A_{n-1} & \Delta A_n \\ \Delta A_{n-k+1} & \cdots & \Delta A_n & \Delta A_{n+1} \\ \vdots & & \vdots & \vdots \\ \Delta A_{n-1} & & \Delta A_{n+k-2} & \Delta A_{n+k-1} \end{vmatrix}}{\begin{vmatrix} 1 & & 1 & 1 \\ \Delta A_{n-k} & \cdots & \Delta A_{n-1} & \Delta A_n \\ \Delta A_{n-k+1} & \cdots & \Delta A_n & \Delta A_{n+1} \\ \vdots & & & \\ \Delta A_{n-1} & \cdots & \Delta A_{n+k-2} & \Delta A_{n+k-1} \end{vmatrix}}, \tag{1.1}$$

where Δ is the (forward) difference operator $\Delta A_i = A_{i+1} - A_i$.

Under appropriate conditions $\lim_{n \to \infty} B_{k,n} = L$, and the convergence of $B_{k,n}$ to L is more rapid than that of A_n.

We shall be concerned with one aspect of Shanks' transform: namely, his applications to number theory [10, p. 34; 9, 11]. In particular, if $A_n^{-1} = \prod_{j=1}^n (1-q^j)$, then

$$B_{n,n}^{-1} = 1 + \sum_{j=1}^{n} (-1)^j q^{j(3j-1)/2}(1+q^j), \qquad [10; \text{p. } 34, \text{Eq. } (1.2.3)] \tag{1.2}$$

and if $A_n^{-1} = \prod_{j=1}^n (1-q^{2j})/(1-q^{2j-1})$, then

$$B_{n,n}^{-1} = \sum_{j=0}^{2n} q^{j(j+1)/2}, \qquad [9; \text{p. } 749]. \tag{1.3}$$

Thus Shanks' transform not only increases convergence in these two instances, but it also provides resulting sequences which yield immediately the famous corollaries:

$$\prod_{n=1}^{\infty} (1-q^n) = 1 + \sum_{j=1}^{\infty} (-1)^j q^{j(3j-1)/2}(1+q^j), \tag{1.4}$$

Euler's pentagonal number theorem [2; p. 11, Corollary 1.7], and

$$\prod_{n=1}^{\infty} \frac{(1-q^{2n})}{(1-q^{2n-1})} = \sum_{j=0}^{\infty} q^{j(j+1)/2}, \tag{1.5}$$

Gauss's theorem [2; p. 23, Eq. (2.2.13)].

A number of natural questions arise. First, what's going on here anyway? Also, what is the extent to which Shanks' transform will apply to other

q-products? We study this question analytically in Sections 2 and 3. We rely on the close relationship of Shanks' transform to Padé approximants [10; p. 21], and reveal that the little q-Jacobi polynomials and the Padé approximants related to their moment generating function lie behind (1.2) and (1.3). In Section 4 we consider and make rigorous Shanks' application of his transform to a sequence related to the Sieve of Eratosthenes. In Section 5 we present the combinatorial aspects of the identities arising from our analytic studies.

2. Shanks' Transform and Padé Approximants

Shanks himself proved the main result on which our work is based. Namely he proved [10; Theorem VI, p. 22] that if $A_n = \sum_{i=0}^{n} c_i z^i$, then $B_{k,n}$ is the Padé approximant $[k/n]$ to

$$f(z) = \sum_{i=0}^{\infty} c_i z^i. \tag{2.1}$$

We are using the notation of [4] for Padé approximants. In addition, Wynn [14; p. 88] (see also [6, 13]) has effectively made possible the explicit construction of the Padé approximants for the case of (2.1) when

$$c_i = \prod_{j=0}^{i-1} \frac{A - q^{j+l}}{B - q^{j+m}}. \tag{2.2}$$

In fact the relevant orthogonal polynomials discovered by Wynn [14; p. 88, Eq. (46)] are the little q-Jacobi polynomials first studied by Hahn [5]. The important details about these polynomials were presented in [3]. Given this information about the little q-Jacobi polynomials, we can make explicit the Padé approximants for Wynn's series (2.1) whose coefficients are given by (2.2). We shall restrict our construction to Shanks' sequence $B_{n,n}$ so that we can see the general series acceleration phenomenon explicitly.

The following lemma is a reformulation of standard results related to Padé approximants for moment generating functions.

Lemma [4; Sect. 5.3]. *Let $p_n(x)$ be a family of orthogonal polynomials on $[a, b]$ relative to the distribution $d\omega$, where*

$$p_n(x) = \sum_{j=0}^{n} c_j(n)\, x^j. \tag{2.3}$$

Define a sequence A_n *by*

$$A_n = A_0 + \varepsilon \int_a^b \frac{1-(\lambda x)^n}{1-\lambda x}\, d\omega. \tag{2.4}$$

Then in the notation of (1.1)

$$B_{n,n} = A_0 + \varepsilon \int_a^b \frac{d\omega}{1-\lambda x} - \frac{\varepsilon \lambda^n}{p_n(\lambda^{-1})} \int_a^b x^n p_n(x)\, d\omega \tag{2.5}$$

$$= \frac{1}{p_n(\lambda^{-1})} \sum_{i=0}^{n} A_{n+i} c_i(n)\, \lambda^{-i}, \tag{2.6}$$

provided $\det(A_{i+j-2})_{n\times n}$ *exists and is nonzero for each n.*

Remark. We note that $\int_a^b d\omega/(1-\lambda x)$ is the moment generating function for the distribution $d\omega$, and A_n is merely a linear shift of the nth partial sum of this function.

Proof. For any sequence y_i, we recall

$$\Delta y_i = y_{i+1} - y_i.$$

Returning to (1.1) we note that $B_{nn} = 1/x_0$, where $x_0, x_1, \ldots, x_n$ are defined by the system of $n+1$ equations:

$$A_j x_0 + \Delta A_j x_1 + \Delta A_{j+1} x_2 + \cdots + \Delta A_{j+n-1} x_n = 1, \qquad 0 \leqslant j \leqslant n. \tag{2.7}$$

This system is of course equivalent to

$$A_j\, \Delta x_0 + A_{j+1}\, \Delta x_i + \cdots + A_{j+n}\, \Delta x_n = -1, \qquad 0 \leqslant j \leqslant n \tag{2.8}$$

(where $x_{n+1} \equiv 0$, so $\Delta x_n = -x_n$), which is in turn equivalent to

$$\begin{cases} \Delta A_j\, \Delta x_0 + \Delta A_{j+1}\, \Delta x_1 + \cdots + \Delta A_{j+n}\, \Delta x_n = 0, & 0 \leqslant j < n, \\ A_n\, \Delta x_0 + A_{n+1}\, \Delta x_1 + \cdots + A_{2n}\, \Delta x_n = -1. \end{cases} \tag{2.9}$$

Furthermore each of these systems has a unique solution because $\det(A_{i+j-2}) \neq 0$ for each n by hypothesis.

The solution of (2.9) is given by

$$\Delta x_i = c_i(n)\, \lambda^{-i}/\alpha_n, \tag{2.10}$$

where

$$\begin{aligned} \alpha_n &= -p_n(\lambda^{-1})\, A_0 + \varepsilon \int_a^b \frac{x^n \lambda^n p_n(x) - p_n(\lambda^{-1})}{1-\lambda x}\, d\omega \\ &= -\sum_{i=0}^{n} A_{n+i} c_i(n)\, \lambda^{-i}. \end{aligned} \tag{2.11}$$

The orthogonality property of the $p_n(x)$ and the fact that

$$\Delta A_n = \varepsilon \lambda^n \int_a^b x^n\, d\omega \tag{2.12}$$

is the appropriate multiple of the nth moment of the distribution yields the first n equations of (2.9), and α_n is chosen to force the nth equation to work.

Finally we note that

$$B_{n,n} = \frac{1}{x_0} = \frac{-1}{\sum_{i=0}^{n} \Delta x_i} = -\frac{\alpha_n}{p_n(\lambda^{-1})}; \tag{2.13}$$

thus (2.5) and (2.6) now follow from (2.11). ∎

3. Application to the Little q-Jacobi Polynomials

In the lemma of Section 2, let us take

$$A_0 = 1, \tag{3.1}$$

$$\lambda = q, \tag{3.2}$$

$$\varepsilon = \frac{-aq(b/a)_\infty (aq)_\infty}{(1-q)(q)_\infty (bq)_\infty}, \tag{3.3}$$

$$p_n(x) = \sum_{j=0}^{n} \frac{(q^{-n})_j (bq^{n+1})_j (xq)^j}{(q)_j (aq)_j}. \tag{3.4}$$

where, using standard notation

$$(A)_n = (A; q)_n = \prod_{j=0}^{n-1} (1 - Aq^j), \tag{3.5}$$

and

$$(A)_\infty = (A; q)_\infty = \lim_{n\to\infty} (A)_n. \tag{3.6}$$

The $p_n(x)$ are the little q-Jacobi polynomials of Hahn [5]. The relevant distribution is discrete and consists of weights w_i at the points $q^i (i=0, 1,...)$, where

$$w_i = \frac{a^i (q^{i+1})_\infty \, q^i (1-q)}{(bq^{i+1}/a)_\infty}, \qquad [3; \text{p. 13, Eq. (3.8)}]. \tag{3.7}$$

Hence in this instance

$$\begin{aligned} A_n &= A_0 + \varepsilon \sum_{i=0}^{\infty} \frac{(1-(q^{i+1})^n)}{(1-q^{i+1})} \cdot \frac{a^i (q^{i+1})_\infty \, q^i (1-q)}{(bq^{i+1}/a)_\infty} \\ &= 1 - \frac{(aq)_\infty}{(bq)_\infty} \sum_{i=0}^{\infty} \frac{(1-q^{(i+1)n}) \, a^{i+1} q^{i+1} (b/a)_{i+1}}{(q)_{i+1}} \\ &= 1 - \frac{(aq)_\infty}{(bq)_\infty} \sum_{i=0}^{\infty} \frac{(1-q^{in}) \, a^i q^i (b/a)_i}{(q)_i} \\ &= 1 - 1 + \frac{(aq)_n}{(bq)_n} \\ &= \frac{(aq)_n}{(bq)_n}. \end{aligned} \tag{3.8}$$

Therefore by line 2 of (2.11),

$$\begin{aligned} \alpha_n &= -\sum_{i=0}^{n} \frac{(aq)_{n+i}}{(bq)_{n+i}} \cdot \frac{(q^{-n})_i (bq^{n+1})_i}{(q)_i (aq)_i} \\ &= -\frac{(aq)_n}{(bq)_n} \sum_{i=0}^{n} \frac{(q^{-n})_i (aq^{n+1})_i}{(q)_i (aq)_i} \\ &= -\frac{(aq)_n}{(bq)_n} \lim_{c \to \infty} \sum_{i=0}^{n} \frac{(q^{-n})_i (aq^{n+1})_i (c)_i \, q^i}{(q)_i (aq)_i (cq)_i} \\ &= -\frac{(aq)_n}{(bq)_n} \lim_{c \to \infty} \frac{(aq/c)_n (q^{-n})_n}{(aq)_n (q^{-n}/c)_n} \end{aligned}$$

(by the q-Pfaff–Saalschutz summation, [2; p. 38 (3.3.12)])

$$\begin{aligned} &= -\frac{(q^{-n})_n}{(bq)_n} \\ &= \frac{(-1)^{n+1} q^{-n(n+1)/2} (q)_n}{(bq)_n}. \end{aligned} \tag{3.9}$$

These observations now leave us in position to prove our main result.

THEOREM 1. *If $A_n=(aq)_n/(bq)_n$, then Shanks' transform B_{nn} given by* (1.1) *is*

$$\frac{1}{B_{n,n}}=\frac{(-1)^n q^{n(n+1)/2}(bq)_n}{(q)_n}\sum_{i=0}^{n}\frac{(q^{-n})_i(bq^{n+1})_i}{(q)_i(aq)_i} \tag{3.10}$$

$$=\frac{(bq)_n}{(aq)_n}\sum_{i=0}^{n}\frac{(b/a)_i\, a^i q^{i(n+1)}}{(q)_i} \tag{3.11}$$

$$=1+\sum_{j=1}^{n}\frac{(bq)_{j-1}(1-bq^{2j})(b/a)_j\, q^{j^2}a^j}{(q)_j(aq)_j}. \tag{3.12}$$

Proof. Eq. (3.10) follows immediately from (3.9) and (2.13). Identity (3.12) requires Watson's q-analog of Whipple's theorem [12; p. 100, Eq. (3.4.1.5)]:

$$1+\sum_{j=0}^{N}\frac{(\alpha q)_{j-1}(1-\alpha q^{2j})(\beta)_j(\gamma)_j(\delta)_j(\varepsilon)_j(q^{-N})_j}{(q)_j\left(\frac{\alpha q}{\beta}\right)_j\left(\frac{\alpha q}{\gamma}\right)_j\left(\frac{\alpha q}{\delta}\right)_j\left(\frac{\alpha q}{\varepsilon}\right)_j(\alpha q^{N+1})_j}\left(\frac{\alpha^2 q^{N+2}}{\beta\gamma\delta\varepsilon}\right)^j$$

$$=\frac{(\alpha q)_N\left(\frac{\alpha q}{\delta\varepsilon}\right)_N}{\left(\frac{\alpha q}{\varepsilon}\right)_N\left(\frac{\alpha q}{\delta}\right)_N}\sum_{j=0}^{N}\frac{\left(\frac{\alpha q}{\beta\gamma}\right)_j(\delta)_j(\varepsilon)_j(q^{-N})_j\, q^j}{(q)_j\left(\frac{\alpha q}{\beta}\right)_j\left(\frac{\alpha q}{\gamma}\right)_j\left(\frac{\delta\varepsilon q^{-N}}{\alpha}\right)_j}. \tag{3.13}$$

In (3.13) set $N=n$, $\delta=bq^{n+1}$, $\gamma=b/a$, $\alpha=b$, and let ε, $\beta\to\infty$. This yields the identity of (3.10) with (3.12). To obtain (3.11) we note by (3.10) that

$$\frac{1}{B_{n,n}}=\frac{(bq)_n}{(q^{-n})_n}\lim_{e\to 0}\sum_{i=0}^{n}\frac{(q^{-n})_i(bq^{n+1})_i\left(\frac{ae}{b}\right)_i}{(q)_i(aq)_i(e)_i}$$

$$=\frac{(bq)_n}{(q^{-n})_n}\cdot\frac{(q^{-n})_n}{(aq)_n}\sum_{i=0}^{n}\frac{(b/a)_i\, a^i q^{(n+1)i}}{(q)_i}$$

(by [8; p. 174, Eq. (10.1)])

$$=\frac{(bq)_n}{(aq)_n}\sum_{i=0}^{n}\frac{(b/a)_i\, a^i q^{(n+1)i}}{(q)_i}. \quad \blacksquare \tag{3.14}$$

We conclude this section by noting that the identities Shanks derived via his transform in [9, 11] are merely special cases of portions of Theorem 1.

If we set $b=1$ and let $a\to 0$, then (3.11) and (3.12) yield

$$(q)_n \sum_{i=0}^{n} \frac{(-1)^i q^{ni+i(i+1)/2}}{(q)_i} = 1 + \sum_{j=1}^{n} (-1)^j q^{j(3j-1)/2}(1+q^j), \quad (3.15)$$

which is Shanks' result in [9], see also [7].

If we replace q by q^2 and then set $b=1$, $a=q^{-1}$ in (3.11) and (3.12) we obtain

$$\frac{(q^2;q^2)_n}{(q;q^2)_n} \sum_{i=0}^{n} \frac{(q;q^2)_i\, q^{i(2n+1)}}{(q^2;q^2)_i} = \sum_{s=0}^{2n} q^{s(s+1)/2}, \quad (3.16)$$

which is the main result of Shanks' in [11].

4. Shanks' Sieve of Eratosthenes Application

Let us quote Shanks [10; p. 35] as he considers "...a sequence akin to the Eratosthenes Sieve. Let the decimal number $A_n (n=1, 2,...)$

$A_1 = 0.111111111111111....$

$A_2 = 0.1212121212121212....$

$A_3 = 0.1222131222131222....$

$A_4 = 0.1223131322141223....$

$A_5 = 0.1223231323141233....$

$A_6 = 0.1223241323151233....$

represent the number of pebbles in contiguous boxes after n stages of the following operation. First a pebble is dropped into every box, then a pebble is dropped into every second box, then into every third box, etc. It is seen that the ultimate population in box n is $d(n)$, the number of divisors of n. The number A_n[for $n<48$] yields $d(n)$ correctly up to but not beyond the first n integers...".

As Shanks then notes he is considering

$$A_n = \sum_{j=1}^{n} \frac{q^j}{1-q^j},$$

for $q=\frac{1}{10}$, and, as in the other cases we have considered, his transform greatly speeds convergence.

Surprisingly (perhaps not) this application is explained with little difficulty using Theorem 1. First we observe that if A_n is replaced by $\alpha A_n + \beta$

then $B_{k,n}$ becomes $\alpha B_{k,n}+\beta$; this is immediate from (1.1) or from the lemma of Section 2. Let us consider then

$$A_n=\frac{(q)_n/(bq)_n-1}{b-1}, \tag{4.1}$$

which is a linear shift of $A_n=(q)_n/(bq)_n$. By Theorem 1, Eq. (3.12), the related $B_{n,n}$ is given by

$$B_{n,n}=\frac{-\sum_{j=1}^{n}(bq)_{j-1}(1-bq^{2j})(b)_j\,q^{j^2}/(q)_j^2}{(b-1)(1+\sum_{j=1}^{n}(bq)_{j-1}(1-bq^{2j})(b)_j\,q^{j^2}/(q)_j^2)}. \tag{4.2}$$

Now when $b\to 1$, we see that

$$A_n\to\frac{d}{db}\left(\frac{(q)_n}{(bq)_n}\right)_{b=1}=\sum_{j=1}^{n}\frac{q^j}{1-q^j}, \tag{4.3}$$

while

$$B_{nn}\to\sum_{j=1}^{n}\frac{q^{j^2}(1+q^j)}{1-q^j} \tag{4.4}$$

Thus while A_n differs from A_∞ at q^{n+1}, B_{nn} differs from A_∞ at q^{n^2+2n+1}; this is in complete agreement with Shanks' numerical table [10; p. 35].

We should add that this result is not terribly surprising. It is merely the analytic counterpart of the fact that we can determine $d(n)$ by counting twice all the divisors of n less than $\sqrt{n}$ and adding 1 if n is a perfect square.

5. Combinatorial Aspects

Identity (3.15) has been proved in an elegant combinatorial manner by Knuth and Paterson [7]. Knuth and Paterson study the Franklin involution F very carefully. We begin by recounting their definition of F.

Let Π be a partition of n into m parts, so that $\Pi=\{\alpha_1,\ldots,\alpha_m\}$ for some integers $\alpha_1\geqslant\cdots\geqslant\alpha_m>0$, where $\alpha_1+\cdots+\alpha_m=n$. We shall write

$$\Sigma(\Pi)=n,\qquad \nu(\Pi)=m,\qquad \lambda(\Pi)=\alpha_1,$$
$$\omega(\Pi)=|\{\alpha_i|\ \alpha_i\equiv 1\ (\mathrm{mod}\ 2)\}|, \tag{5.1}$$

for the sum, number of parts, largest part and number of odd parts of Π, respectively; if Π is the empty set, we let $\Sigma(\Pi)=\nu(\Pi)=\lambda(\Pi)=\omega(\Pi)=0$.

We alo define the *base* $b(\Pi)$ and *slope* $\sigma(\Pi)$ when Π has distinct parts as follows:

$$\beta(\Pi)=\min\{j \mid j\in\Pi\}=\alpha_m, \qquad \sigma(\Pi)=\min\{j \mid \lambda(\Pi)-j\notin\Pi\}. \quad (5.2)$$

Note that if Π is nonempty we have

$$\lambda(\Pi)\geqslant\beta(\Pi)+\nu(\Pi)-1 \qquad \text{and} \qquad \nu(\Pi)\geqslant\sigma(\Pi). \quad (5.3)$$

The Franklin involution F is a mapping from the set of all partitions with distinct parts to the set of all partitions with distinct parts. The partition $F(\Pi)$ corresponding to Π under Franklin's transformation F is obtained as follows:

(i) If $\beta(\Pi)\leqslant\sigma(\Pi)$ and $\beta(\Pi)<\nu(\Pi)$, remove the smallest part, $\beta(\Pi)$, and increase each of the largest $\beta(\Pi)$ parts by one.

(ii) If $\beta(\Pi)>\sigma(\Pi)$ and $\sigma(\Pi)<\nu(\Pi)$ or $\sigma(\Pi)\neq\beta(\Pi)-1$, decrease each of the largest $\sigma(\Pi)$ parts by one and append a new smallest part, $\sigma(\Pi)$.

(iii) Otherwise $F(\Pi)=\Pi$. [This case holds if and only if Π is empty or $\sigma(\Pi)=\nu(\Pi)\leqslant\beta(\Pi)\leqslant\sigma(\Pi)+1$.]

It is not difficult to verify that F is an involution on the set of partitions with distinct parts, i.e., that

$$F(F(\Pi))=\Pi \quad (5.4)$$

for all such Π. Moreover, we shall see that on a certain set of partitions, the involution F is sign reversing.

For each $l\geqslant 0$ there is exactly one partition Π such that $\lambda(\Pi)=l$ and $F(\Pi)=\Pi$. We shall denote this fixed point of the mapping by f_l; it has $[(l+1)/2]$ consecutive parts,

$$f_l=\{l, l-1,..., [l/2]+1\}. \quad (5.5)$$

Let

$$\mathscr{A}=\{f_0, f_1, f_2,...\} \quad (5.6)$$

be the set of all such partitions.

Finally if S is any set of partitions, we define the *generating function* of S by formula

$$G(S; x, y, z, v)=\sum_{\Pi\in S} x^{\Sigma(\Pi)}y^{\lambda(\Pi)}z^{\nu(\Pi)}v^{\omega(\Pi)}. \quad (5.7)$$

All of the above notation is taken essentially from Knuth and Paterson [7]. In their paper, they use the Franklin mapping F to establish nicely our (3.15).

We point out that the other portion of Theorem 1 when $b=1$ and $a=0$ (namely (3.10) = (3.12)) follows from this approach as well:

$$\sum_{j=0}^{k} (-1)^{k-j} q^{\binom{k-j+1}{2}} \begin{bmatrix} k \\ j \end{bmatrix} (q^{k+1})_j = \sum_{i=-k}^{k} (-1)^i q^{i(3i-1)/2} \tag{5.8}$$

where

$$\begin{bmatrix} A \\ B \end{bmatrix} = \begin{cases} \dfrac{(q)_A}{(q)_B (q)_{A-B}}, & 0 \leqslant B \leqslant A \\ 0, & \text{otherwise.} \end{cases} \tag{5.9}$$

To see (5.8), we let $\mathscr{D}_j$ denote the set of partitions into distinct parts with exactly $k-j$ parts of size $\leqslant k$, while all other parts lie in $\{k+1, k+2,..., k+j\}$. Immediately from well-known properties of the Gaussian polynomials $[{A \atop B}]$ [2, Chap. 3], we see that

$$G(\mathscr{D}_j; q, -1, 1, 1) = (-1)^{k-j} q^{\binom{k-j+1}{2}} \begin{bmatrix} k \\ j \end{bmatrix} (q^{k+1})_j. \tag{5.10}$$

Clearly also for $\mathscr{A}_k = \{f_0, f_1,..., f_{2k}\}$

$$G(\mathscr{A}_k; q, -1, 1, 1) = \sum_{i=-k}^{k} (-1)^i q^{i(3i-1)/2}. \tag{5.11}$$

If we define $\mathscr{D} = \bigcup_{i=0}^{k} \mathscr{D}_i$, then it is clear that we need only show

$$\sum_{i=0}^{k} G(\mathscr{D}_i; q, -1, 1, 1) = G(\mathscr{D}; q, -1, 1, 1) = G(\mathscr{A}_k; q, -1, 1, 1). \tag{5.12}$$

Obviously $\mathscr{A} \subseteq \mathscr{D}$.

Noting that for any $\Pi \in \mathscr{D} - \mathscr{A}_k$, F has the sign reversing property

$$(-1)^{\lambda(\Pi)} = -(-1)^{\lambda(F(\Pi))},$$

we see that (5.12) will be proved if we can show that $F(\Pi) \in \mathscr{D} - \mathscr{A}_k$ when $\Pi \in \mathscr{D} - \mathscr{A}_k$. Suppose $\Pi \in \mathscr{D}_j$. If $\beta(\Pi) \leqslant \sigma(\Pi)$ then $F(\Pi) \in \mathscr{D}_{j+1}$ unless the parts larger than k are of the form $k+l, k+l-1,..., k+1$ and one of the parts $\leqslant k$ is k itself, in which case $F(\Pi) \in \mathscr{D}_{j+2}$. If $\beta(\Pi) > \sigma(\lambda)$, then $F(\Pi) \in \mathscr{D}_{j-1}$ unless the parts larger than k are of the form $k+l,..., k+1$. In this case it is clear that $F(\Pi) \in \mathscr{D}_{j-2}$ unless $l=j$. If $l=j$ then $\sigma(\Pi) \geqslant j$, and we must have $\sigma(\Pi)=j$, $\beta(\Pi)=j+1$ since $\beta(\Pi)\leqslant j+1$ by definition of $\mathscr{D}_j$. But $\beta(\Pi)=j+1$ implies that the parts $\leqslant k$ are $k, k-1,..., j+1$, so $F(\Pi)=k$, a contradiction. Therefore $l\neq j$, and so $F(\Pi)\in\mathscr{D}-\mathscr{A}_k$. This proves (5.12) and thus (5.8).

We may also derive the identity of (3.11) with (3.12) by adapting the Durfee rectangle proof of the Rogers–Fine identity [1; Sect. 4]. If in (3.11) and (3.12) we replace q by q^2, then set $a=t^2q^2$ and then set $b=-at^2q^3$, we obtain the following equivalent assertion:

$$\frac{(-at^2q^3;q^2)_{n+1}}{(t^2q^2;q^2)_{n+1}}\sum_{i=0}^{n}\frac{(-aq;q^2)_i\,t^{2i}q^{2i(n+2)}}{(q^2;q^2)_i}$$
$$=\sum_{j=0}^{n}\frac{(-at^2q^3;q^2)_j(-aq;q^2)_j(1+at^2q^{4j+3})\,q^{2j^2+2j}t^{2j}}{(q^2;q^2)_j(t^2q^2;q^2)_{j+1}}. \tag{5.13}$$

In order to make the argument of [1; Sect. 4] more comprehensible we introduce the modulus 2 representation of partitions due to MacMahon [2; p. 13]. This is a modified Ferrers graph in which each even part $2M$ is represented by a row of M 2's and each odd part $2M+1$ is represented by a row of M 2's followed by a 1. Thus the representation of $8+7+6+6+3+2$ is

$$\begin{matrix}2&2&2&2\\2&2&2&1\\2&2&2&\\2&2&2&\\2&1&&\\2&&&\end{matrix}\quad, \tag{5.14}$$

while the representation of $8+6+3+1$ is

$$\begin{matrix}2&2&2&2\\2&2&2&\\2&1&&\\1&&&\end{matrix} \tag{5.15}$$

The rectangle argument of [1; Sect. 4] can be recast as an examination of the modulus 2 representation of the set partitions $\mathscr{E}$ which have no repeated odd parts and even largest part. In particular, we look for the largest rectangle of 2's and 1's in the representation with N columns and

$N+1$ rows; we call this the *RF-rectangle of order N*. The only way a 1 can appear in the *RF*-rectangle of a partition from $\mathscr{E}$ is in the lower right-hand corner. Thus in the case (5.14) $N=3$ and no 1 appears in the *RF*-rectangle; in the case (5.15) $N=2$ and 1 does appear.

The subsection I of [1; Sect. 4] considers those partitions in $\mathscr{E}$ with *RF*-rectangle of order N and no 1's in *RF*-rectangle. The generating function derived there combinatorially is

$$\frac{(-aq;q^2)_N(-at^2q^3;q^2)_N\,t^{2N}q^{2N^2+2N}}{(q^2;q^2)_N(t^2q^2;q^2)_{N+1}}. \tag{5.16}$$

The subsection II of [1; Sect. 4] considers those partitions in $\mathscr{E}$ with *RF*-rectangle of order N and a 1 in the lower right-hand corner of the rectangle. The resulting generating function is

$$\frac{(-aq;q^2)_N(-at^2q^3;q^2)_N\,at^{2N+2}q^{2N^2+6N+3}}{(q^2;q^2)_N(t^2q^2;q^2)_{N+1}} \tag{5.17}$$

Adding all instances of (5.16) and (5.17) for $N\leqslant n$, we find that the right-hand side of (5.13) is just

$$G(\mathscr{E}_n;q,t,1,a)$$

where $\mathscr{E}_n$ is the subset of $\mathscr{E}$ containing those partions with *RF*-rectangle of order $\leqslant n$.

On the other hand there is another way of obtaining $G(\mathscr{E}_n;q,t,1,a)$. Namely we can classify each partition Π in $\mathscr{E}_n$ by the largest rectangle of 2's contained in the modulus 2 representation of Π with i columns and $n+2$ rows. We say such partitions are in $\mathscr{E}_n(i)$.

Now

$$\frac{(-at^2q^3;q^2)_{n+1}}{(t^2q^2;q^2)_{n+1}} \tag{5.18}$$

is the generating function $G(\Lambda_n;q,1,t^2,a)$ where Λ_n is the set of partitions in which each even part is $\leqslant 2n+2$, each odd is $\leqslant 2n+3$, no odds are repeated and 1 does not appear. By considering the conjugates of elements of Λ_n relative to the 2 modulus representation (i.e., read columns rather than rows of the representation), we see that

$$G(\Lambda_n;q,1,t^2,a)=G(\Lambda_n';q,t,1,a) \tag{5.19}$$

where Λ_n' is the set of partitions in which the largest part is even, no odds are repeated, there are at most $n+2$ parts and if there are $n+2$ parts the smallest is 1.

Hence

$$G(\mathscr{E}_n(i); q, t, 1, a) = \frac{(-aq; q^2)_i}{(q^2; q^2)_i} \cdot t^{2i} q^{2i(n+2)} \cdot G(A'_n; q, t, 1, a) \quad (5.20)$$

where

$$\frac{(-aq; q^2)_i}{(q^2; q^2)_i}$$

generates the portion of each partition in $\mathscr{E}_n(i)$ below the rectangle

$$n+2 \text{ terms} \left\{ \underbrace{\begin{matrix} 2 & 2 & 2 & \cdots & 2 \\ 2 & 2 & 2 & & 2 \\ 2 & 2 & 2 & \cdots & 2 \\ \vdots & \vdots & \vdots & & \vdots \\ 2 & 2 & 2 & \cdots & 2 \\ 2 & 2 & 2 & \cdots & 2 \end{matrix}}_{i \text{ terms}} \right.$$

and $G(A'_n; q, t, 1, a)$ generates the portion to the right. Combining (5.18)–(5.20) and summing on $i = 0$ to n, we find that the left-hand side of (5.13) also equals $G(\mathscr{E}_n; q, t, 1, a)$ thus (5.13) is established.

6. Conclusion

Our main object of explaining the mechanism behind Shanks' transform for $(aq)_n/(bq)_n$ has been accomplished. However, some questions remain. First, are there other families of orthogonal polynomials that lead to number-theoretic surprises like (1.2), (1.3), or (4.4)? Second, can a general study be made of the little q-Jacobi polynomials based on extending the interpretations we have given in Section 5?

While Shanks' method was primarily prepared for numerical acceleration of convergence and was, therefore, viewed as an analytic process; nonetheless, the problems we have considered can be completely dealt with in rings of formal power series. Let $R[[q]]$ be a ring of formal power series with $S_n(q)(n = 1, 2, 3, \ldots)$ and $S(q)$ all in this ring. We define $M(n)$ by the identity $S(q) - S_n(q) = cq^{M(n)} + dq^{M(n)+1} + \cdots$ where $c \neq 0$. We may say that $S_n(q)$ converges to $S(q)$ if $M(n) \to \infty$ with n. We may accelerate convergence by constructing a new sequence $\tilde{S}_n(q)$ with related $\tilde{M}(n)$ such that $\tilde{M}(n) > M(n)$ for n sufficiently large. In the actual cases we have considered $\tilde{M}(n)/M(n) \sim \lambda n$ as $n \to \infty$.

References

1. G. E. Andrews, Two theorems of Gauss and allied identities proved arithmetically, *Pacific J. Math.* **41** (1972), 563–578.
2. G. E. Andrews, "The Theory of Partitions, Encyclopedia of Mathematics and Its Applications," Vol. 2 (G.-C. Rota, Ed.), Addison–Wesley, Reading, Mass. 1976.
3. G. E. Andrews and R. Askey, "Enumeration of partitions: the role of Eulerian series and q-orthogonal polynomials," from Higher Combinatorics (M. Aigner, Ed.), Reidel, Dordrecht/Boston, pp. 3–26, 1977.
4. G. A. Baker and P. Graves–Morris, "Padé Approximants, Part I: Basic Theory, Encyclopedia of Mathematics and Its Applications," Vol. 13 (G.-C. Rota, Ed.), Addison–Wesley, Reading, Mass., 1981.
5. W. Hahn, Über Orthogonal polynome die q-Differenzengleichungen genügen, *Math. Nachr.* **2** (1949), 4–34.
6. W. Hahn, On a special Padé table, *Indian J. Math.* **2** (1960), 67–71.
7. D. E. Knuth and M. S. Paterson, Identities from partition involutions, *Fibonacci Quart.* **16** (1978), 198–212.
8. D. B. Sears, On the transformation theory of basic hypergeometric functions, *Proc. London Math. Soc.* (2) **53** (1951), 158–80.
9. D. Shanks, A short proof of an identity of Euler, *Proc. Amer. Math. Soc.* **2** (1951), 747–749.
10. D. Shanks, Nonlinear transformations of divergent and slowly convergent sequences, *J. Math. Phys.* **34** (1955), 1–42.
11. D. Shanks, Two theorems of Gauss, *Pacific J. Math.* **8** (1958), 609–612.
12. L. J. Slater, "Generalized Hypergeometric Functions," Cambridge Univ. Press, London/New York, 1966.
13. A. van de Sluis, Orthogonal polynomials and hypergeometric series, *Canad. J. Math.* **10** (1959), 592–612.
14. P. Wynn, A general system of orthogonal polynomials, *Quart. J. Math. Oxford Ser.* (2), **18** (1967), 81–96.