The Rogers-Ramanujan journey from mathematics to physics

Geoffrey B. Campbell

Baxter2025 Conference,
The Australian National University,
Canberra, Australia.
12 noon, Thursday 11th September 2025.

ROGERS-RAMANUJAN IDENTITIES IN STATISTICAL MECHANICS

GEOFFREY B CAMPBELL

Abstract. We describe the story of the Rogers-Ramanujan identities; being known for 85 years and having about 130 pure mathematics proofs, suddenly entering physics when Rodney Baxter solved the Hard Hexagon Model in Statistical Mechanics in 1980. We next cover the accompanying proofs by George E Andrews of other related Baxter identities arisen of Rogers-Ramanujan type, leading into a new flourishing partnership of Physics and Mathematics. Our narrative goes into the subsequent 44 years, explaining the progress in physics and mathematical analysis. Finally we show some related crossovers with regard to the *Elliptic q-gamma function* and some *Vector Partition* generating functional equations; the latter of which may be new. The present paper is essentially chapter 11 of the author's 32 chapter book [23] to appear in June 2024.

In 1894 Leonard J. Rogers discovered

The Rogers–Ramanujan identities are

$$G(q) = \sum_{n=0}^{\infty} \frac{q^{n^2}}{(q;q)_n} = \frac{1}{(q;q^5)_\infty (q^4;q^5)_\infty} = 1 + q + q^2 + q^3 + 2q^4 + 2q^5 + 3q^6 + \cdots$$ (sequence A003114 in the OEIS)

and

$$H(q) = \sum_{n=0}^{\infty} \frac{q^{n^2+n}}{(q;q)_n} = \frac{1}{(q^2;q^5)_\infty (q^3;q^5)_\infty} = 1 + q^2 + q^3 + q^4 + q^5 + 2q^6 + \cdots$$ (sequence A003106 in the OEIS).

Here, $(a;q)_n$ denotes the q-Pochhammer symbol.

In mathematics, the **Rogers–Ramanujan identities** are two identities related to basic hypergeometric series and integer partitions. The identities were first discovered and proved by Leonard James Rogers (1894), and were subsequently rediscovered (without a proof) by Srinivasa Ramanujan some time before 1913. Ramanujan had no proof, but rediscovered Rogers's paper in 1917, and they then published a joint new proof (Rogers & Ramanujan 1919). Issai Schur (1917) independently rediscovered and proved the identities.

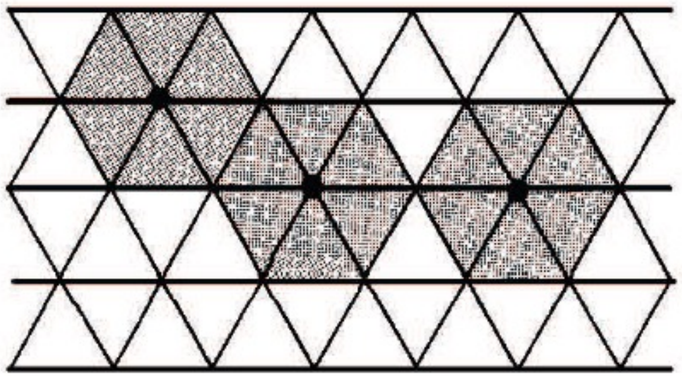
Think About It...

> "In 1980, Baxter found his beautiful solution to the hard-hexagon model of statistical mechanics. His treatment of this model is naturally divided into six regimes that depend on values taken by various parameters associated with the model. Then in truly astounding fashion it turns out that eight Rogers-Ramanujan type identities, all essentially known to Rogers ([33], [34]), are the fundamental keys for finding infinite product representations of the related statistical mechanics partition functions in **regimes I, III, and IV**."
>
> - *George E. Andrews, Proceedings of the National Academy of Sciences, 1981*

the model. For a lattice of N sites, the grand-partition function is

$$Z(z) = \sum_{n=0}^{n/3} z^n g(n, N) \tag{1.1}$$

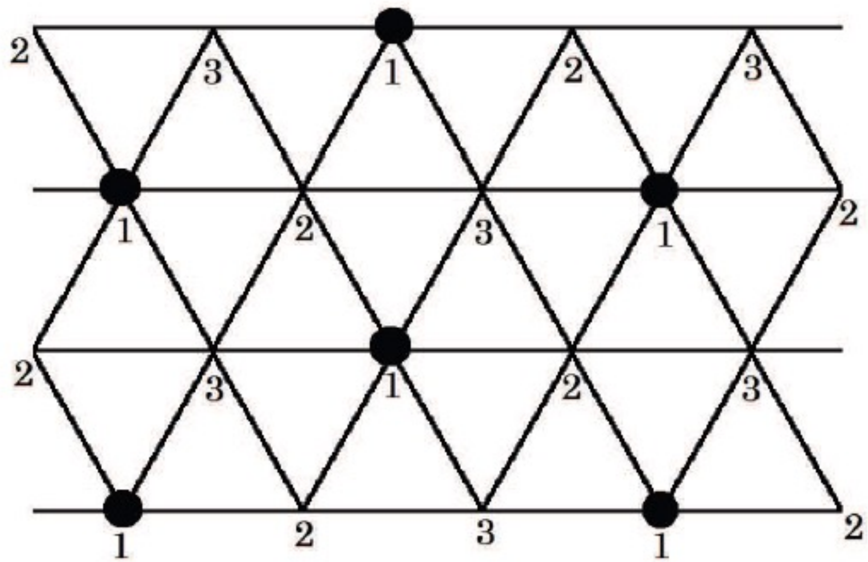
where $g(n, N)$ is the allowed number of ways of placing n particles on the lattice, and the sum is over all possible values of n. (Since no more than $\frac{1}{3}$ of the sites can be occupied, n takes values from 0 to $\frac{N}{3}$.)



A typical arrangement of particles (black circles) on the triangular lattice, such that no two particles are together or adjacent. The six faces round each particle are shaded: they form non-overlapping (ie. "hard") hexagons.

FIGURE 1. The particle arrangement for the Hard Hexagon Model.

The three sub-lattices of the triangular lattice: sub-lattice 1 consists of all sites of type 1, and similarly for sub-lattices 2 and 3. Adjacent sites lie on different sub-lattices. a close-packed arrangement of particles (black circles) is shown: all sites of one sub-lattice (in this case sub-lattice 1) are occupied, the rest are empty. (See figure 2)



The three sub-lattices of the triangular lattice: sub-lattice 1 consists of all sites of type 1, and similarly for sub-lattices 2 and 3.

FIGURE 2. Baxter's three sub-lattices of the triangular lattice.

1.2. What Baxter found in 1980. So, continuing our discussion of equation (1.1), we want to calculate Z, or rather the partition-function per site of the infinite lattice

$$\kappa = \lim_{N\to\infty} Z^{\frac{1}{N}} \tag{1.2}$$

as a function of the positive real variable z. This z is known as the 'activity'. This problem can be put into what is called 'spin'-type language by associating with each site i a variable σ_i. However, instead of the usual approach in statistical mechanics models of letting σ_i take values 1 and -1, we take the values 0 and 1: if the site (lattice point) is empty, then $\sigma_i = 0$; if it is full then $\sigma_i = 1$. Thus σ_i is the number of particles at site i: the 'occupation number'. Then (1.1) can be written as

$$Z = \sum_{\sigma} z^{(\sigma_1+\sigma_2+\ldots+\sigma_N)} \prod_{i,j}(1-\sigma_i\sigma_j), \tag{1.3}$$

where the product is over all edges (i,j) of the triangular lattice, and the sum is over all values (0 and 1) of all the occupation numbers $\sigma_1, \sigma_2, \ldots, \sigma_N$.

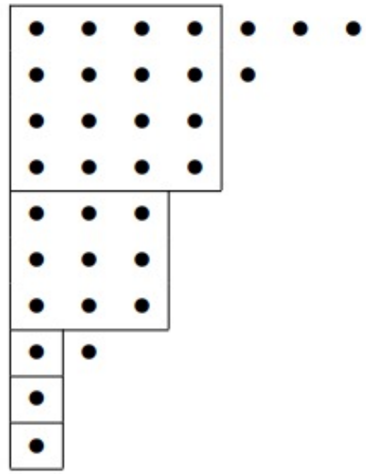
The shock to Number Theory of Baxter's work

Clearly, from the mathematical 'non-physicist' perspective, that this schematic diagram could lead to enumeration of well defined types of integer partitions was (back in 1980 at least) seen as a major departure from the more typical Ferrers Graphs, Durfee Squares and Plane Partition cube stack representations adopted in the literature by Andrews and others over the contemporary mathematical landscapes. As a mathematician already versed in the ideas of our earlier chapters, it becomes natural to ask how say, Bijection equivalences, q-series partition generating functions, and such, may become the same world as the statistical mechanics models involving assigned values of *spin charges* along a vertex or edge of a lattice of a certain structure, shape and form. We put such questions aside for now, as we want to present Baxter's extraordinary findings here. However, these questions of equivalences of theories and analyses are worthy of deeper understanding it seems.

7.3 SUCCESSIVE DURFEE SQUARES

We extend the idea of a Durfee Square of a partition, to having successive Durfee Squares in a partition.

Consider for example the partition $33 = 7+5+4+4+3+3+3+2+1+1$ of Ferrers graph



having successive Durfee squares of sides $4, 3, 1, 1, 1$.

So, adopting this generalization from one Durfee Square to many successive Durfee Squares, we can then find the generating function for partitions that have at most k successive Durfee Squares. For an account of the below process see Andrews and Eriksson [29]. Recall that we saw earlier that if $k = 1$ then

$$s_n(q) = \sum_{j=0}^{n} q^{j^2} \begin{bmatrix} m \\ j \end{bmatrix}_q .$$

k successive Durfee Squares. For an account of the below process see Andrews and Eriksson [29]. Recall that we saw earlier that if $k = 1$ then

$$s_n(q) = \sum_{j=0}^{n} q^{j^2} \begin{bmatrix} m \\ j \end{bmatrix}_q .$$

So for general k, we take account of k possible squares $j_1^2 \geq j_2^2 \geq j_3^2 \geq \cdots \geq j_k^2$ and we see how the portions of the partition to the right side of each Durfee Square contribute to the partition in this form. Hence the generating function for partitions with at most k successive Durfee Squares, $s_{k,n}(q)$ is given by

$$\begin{aligned} s_{k,n}(q) &= \sum_{n \geq j_1 \geq j_2 \geq \cdots \geq j_k \geq 0} q^{j_1^2 + j_2^2 + \cdots + j_k^2} \begin{bmatrix} n \\ j_1 \end{bmatrix}_q \begin{bmatrix} j_1 \\ j_2 \end{bmatrix}_q \begin{bmatrix} j_2 \\ j_3 \end{bmatrix}_q \cdots \begin{bmatrix} j_{k-1} \\ j_k \end{bmatrix}_q \\ &= \sum_{n \geq j_1 \geq j_2 \geq \cdots \geq j_k \geq 0} \frac{q^{j_1^2 + j_2^2 + \cdots + j_k^2} (q;q)_n}{(q;q)_{n-j_1} (q;q)_{n-j_2} (q;q)_{n-j_3} \cdots (q;q)_{n-j_{k-1}-j_k} (q;q)_{j_k}} . \end{aligned}$$

We next take $n \to \infty$, so the generating function for all partitions with at most k successive Durfee Squares is

$$s_{k,\infty}(q) = \sum_{n \geq j_1 \geq j_2 \geq \cdots \geq j_k \geq 0} \frac{q^{j_1^2+j_2^2+\cdots+j_k^2}(q;q)_n}{(q;q)_{n-j_1}(q;q)_{n-j_2}(q;q)_{n-j_3}\cdots(q;q)_{n-j_{k-1}-j_k}(q;q)_{j_k}}.$$

This enables us to state the

Theorem 7.2 *(see Andrews [16]) The number of partitions of* n*, each with at most* k *successive Durfee Squares, equals the number of partitions of* n *into parts not congruent to* $0, \pm(k+1)\,(mod 2k+3)$*. The generating function encoding this is*

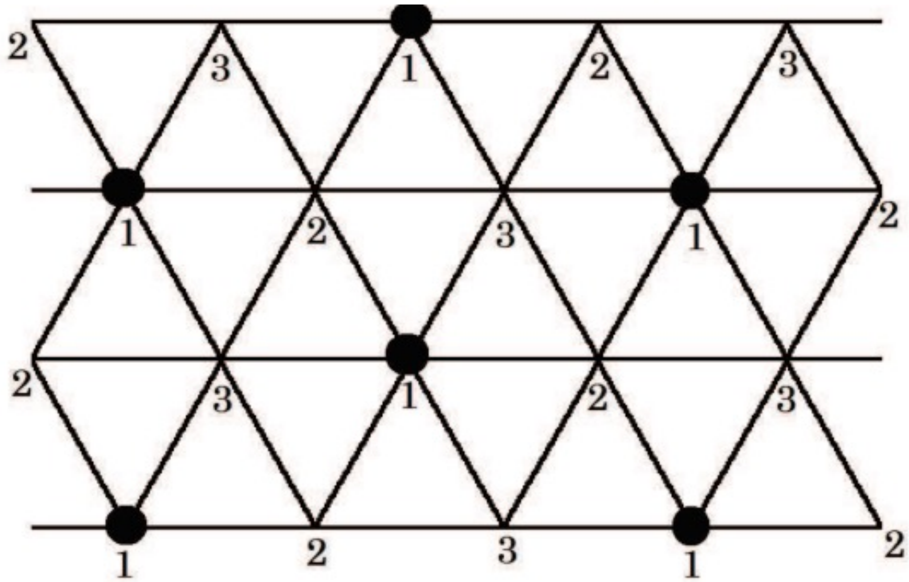
$$\sum_{n \geq j_1 \geq j_2 \geq \cdots \geq j_k \geq 0} \frac{q^{j_1^2+j_2^2+\cdots+j_k^2}(q;q)_n}{(q;q)_{n-j_1}(q;q)_{n-j_2}(q;q)_{n-j_3}\cdots(q;q)_{n-j_{k-1}-j_k}(q;q)_{j_k}} \qquad (7.3.1)$$

$$= \prod_{\substack{n=1 \\ n \not\equiv 0, \pm(k+1)\,(mod 2k+3);}}^{\infty} \frac{1}{(1-q^n)}.$$

The case of theorem 7.2 with $k=1$ is the first Rogers-Ramanujan identity. This remarkable proof was given by Bressoud and Zeilberger in [75] and a generalized approach using bijections by these authors is given in the next section.

In seeking to calculate the free energy in the hard hexagon model as depicted in figure 1, Baxter was led to consider a range of logical options or *Regimes*, which are here depicted as Baxter described in 1980.

The three sub-lattices of the triangular lattice: sub-lattice 1 consists of all sites of type 1, and similarly for sub-lattices 2 and 3. Adjacent sites lie on different sub-lattices. a close-packed arrangement of particles (black circles) is shown: all sites of one sub-lattice (in this case sub-lattice 1) are occupied, the rest are empty. (See figure 2)



The three sub-lattices of the triangular lattice: sub-lattice 1 consists of all sites of type 1, and similarly for sub-lattices 2 and 3.

FIGURE 2. Baxter's three sub-lattices of the triangular lattice.

1.2. **What Baxter found in 1980.** So, continuing our discussion of equation (1.1), we want to calculate Z, or rather the partition-function per site of the infinite lattice

$$\kappa = \lim_{N\to\infty} Z^{\frac{1}{N}} \tag{1.2}$$

as a function of the positive real variable z. This z is known as the 'activity'. This problem can be put into what is called 'spin'-type language by associating with each site i a variable σ_i. However, instead of the usual approach in statistical mechanics models of letting σ_i take values 1 and -1, we take the values 0 and 1: if the site (lattice point) is empty, then $\sigma_i = 0$; if it is full then $\sigma_i = 1$. Thus σ_i is the number of particles at site i: the 'occupation number'. Then (1.1) can be written as

$$Z = \sum_{\sigma} z^{(\sigma_1+\sigma_2+\ldots+\sigma_N)} \prod_{i,j} (1 - \sigma_i\sigma_j), \tag{1.3}$$

where the product is over all edges (i, j) of the triangular lattice, and the sum is over all values (0 and 1) of all the occupation numbers $\sigma_1, \sigma_2, ..., \sigma_N$.

We expect this model to undergo a phase transition from an homogeneous fluid state at low activity z to an inhomogeneous solid state at high activity z. To see this, divide the lattice into three sub-lattices $1, 2, 3$, so that no two sites of the same type are adjacent. Then there are three possible close-packed configurations of particles on the lattice: either all sites of type 1 are occupied, or all sites of type 2, or all sites of type 3. Suppose we fix the boundary sites as in the first possibility, i.e. all boundary sites of type 1 are full, and all other boundary sites are empty. Then for an infinite lattice the second and third possibilities give a negligible contribution to the sum-over-states in (1.3). Clearly, sites on different sub-lattices are no longer equivalent. Let ρ_r be the local density at a site of type r, given by,

$$\rho_r = \langle \sigma_l \rangle = Z^{-1} \sum_{\sigma} \sigma_l z^{(\sigma_1+\sigma_2+\ldots+\sigma_N)} \prod_{i,j} (1-\sigma_i \sigma_j), \tag{1.4}$$

where l is a site of type r.

When z is infinite, the system is close-packed with all sites of type 1 occupied, so $\rho_1, = 1$, $\rho_2 = \rho_3 = 0$. We can expand each ρ_r in inverse powers of z by considering successive perturbations of the close-packed state. For a site l deep inside a large lattice, this gives

$$\begin{aligned} \rho_1 &= 1 - z^{-1} - 5z^{-2} - 34z^{-3} - 267z^{-4} - 2037z^{-5} - \ldots \\ \rho_2 &= \rho_3 = z^{-2} + 9z^{-3} + 80z^{-4} + 965z^{-5} + \ldots \end{aligned} \tag{1.5}$$

The system is therefore not homogeneous, since ρ_1, ρ_2, ρ_3, are not all equal. This contrasts with the low-activity situation: starting from the state with all sites empty and successively introducing particles, we obtain

$$\rho_1 = \rho_2 = \rho_3 = z^1 - 7z^2 - 58z^3 - 519z^4 + 4856z^5 - \dots \tag{1.6}$$

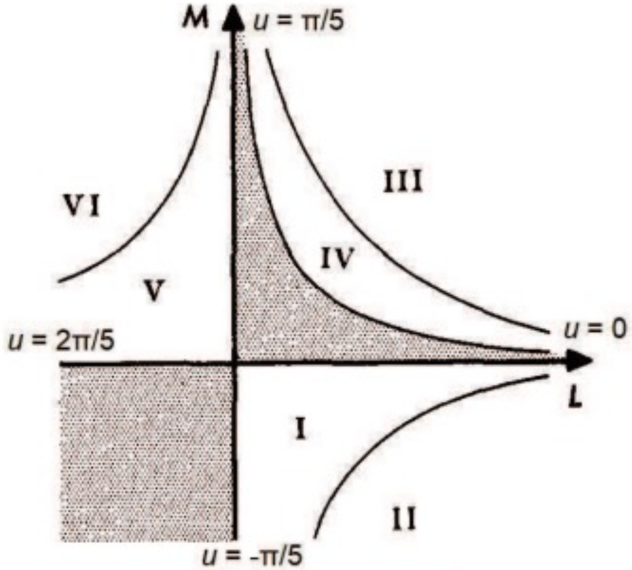
To all orders in this expansion it is true that $\rho_1 = \rho_2 = \rho_3$.

The system is therefore inhomogeneous for sufficiently large z, and homogeneous for sufficiently small z if the series converge. There must be a critical value z_c, of z above which the system ceases to be homogeneous. Since the homogeneous phase is typical of a fluid, and the ordered inhomogeneous phase is typical of a solid, the model can be said to undergo a fluid to solid transition at $z = z_c$.

1.3. **Baxter's 1980 numerical preliminaries.** Baxter then applied the Corner Transfer Matrix approach to the matrix eigenvalues $a_1, \ldots, a_{10}$, for the hard hexagon model with $z = 1$. The values were approximate, being calculated from finite truncations of the triangular lattice analogue of the matrix equation related to this. The eigenvalues occur in groups of comparable magnitude, and it is sensible to include all members of a group. For this reason the truncations used were 2×2, 3×3, 5×5, 7×7 and 10×10. Each a_i was given for successively larger truncations, and clearly each tended rapidly to a limit. This limit is its exact value for the infinite-dimensional corner transfer matrix. Refer to Baxter [12, Chapter 13] for how the Corner Transfer Matrix context applies.

Baxter was able to determine the required coefficients up to 30 terms of the power series *exact solution*. Equipped with knowledge from his famous eight-vertex model exact solution from a few years earlier, Baxter was able to put the hard hexagon proposed solution expansion into the form

$$z = -x \prod_{n=1}^{\infty} (1 - x^n)^{c_n}. \tag{1.7}$$



THE SIX REGIMES

I: $\Delta > \Delta_c, \quad q^2 < 0, -\pi/5 < u < 0,$

II: $0 < \Delta < \Delta_c, \quad q^2 > 0, -\pi/5 < u < 0,$

III: $-\Delta_c < \Delta < 0, q^2 > 0, \quad 0 < u < \pi/5,$

IV: $\Delta < -\Delta_c, \quad q^2 < 0, \quad 0 < u < \pi/5,$

V: $\Delta > \Delta_c, \quad q^2 < 0, \quad \pi/5 < u < 2\pi/5,$

VI: $0 < \Delta < \Delta_c, \quad q^2 > 0, \quad \pi/5 < u < 2\pi/5.$

The six regimes in the (L, M) plane, as listed at right. Shaded areas are unphysical, since $z = (1 - e^{-L})(1 - e^{-M})/(e^{L+M} - e^L - e^M)$ and $\Delta = z^{-1/2}(1 - z\,e^{L+M})$ for $\Delta = \Delta_1$ gives z therein to be negative. Regimes I, III, V are disordered, II and VI have triangular ordering, IV has square ordering. The system is critical on the (I, II), (III, IV) and (V, VI) boundaries, where $|\Delta| = \Delta_c$ the values of u on the (L, M) axes are indicated.

FIGURE 3. Baxter's six Regimes for the Hard Hexagon Model. The six regimes in the (L, M) plane. Shaded areas are not physical. Regimes I, III and V are disordered, and regimes II and VI have triangular ordering, while Regime IV has square ordering.

With this in mind he discovered that

$$\begin{aligned} c_1, c_2, \ldots, c_{29} &= 5, -5, -5, 5, 0, 5, -5, -5, 5, 0, \\ & \quad 5, -5, -5, 5, 0, 5, -5, -5, 5, 0, \\ & \quad 5, -5, -5, 5, 0, 5, -5, -5, 5. \end{aligned} \tag{1.8}$$

Baxter then was able to infer that

$$z = -x \left(\frac{H(x)}{G(x)} \right)^5, \tag{1.9}$$

where

$$G(x) = \prod_{n=1}^{\infty} \frac{1}{(1 - x^{5n-4})(1 - x^{5n-1})}, \tag{1.10}$$

$$H(x) = \prod_{n=1}^{\infty} \frac{1}{(1 - x^{5n-3})(1 - x^{5n-2})}. \tag{1.11}$$

The mathematical partition theorist will will recognise the products $G(x)$ and $H(x)$ as from the Rogers-Ramanujan identities.

So, Baxter followed through this Corner Transfer Matrix approach to calculate the coefficients and determine a range of identities for each of the Hard Hexagon Regime types of allowable spin configurations. What he derived, is what Andrews subsequently was able to prove on his visit to Baxter in Australia the following year. So, we have set the scene for Andrews proofs in 1981.

1.4. **What Andrews found in 1981.** So, after Baxter's solution to the Hard Hexagon Model in 1980, the opening to a new world of research in both theoretical physics and partition-theoretic mathematics was about to happen. George E. Andrews visited Baxter in 1981, and proved the identities arisen from the Regimes of Baxter's Hard Hexagon exact solutions.

In a landmark paper, Andrews [3] goes on then to say that Baxter found the following identities occurring for the six regimes, conspicuously missing (but later returning to) the more complicated Regime II, as follows:

Regime I

(1.12) $$\sum_{n=0}^{\infty} \frac{q^{n^2}}{(q)_n} = \frac{1}{\Pi(q, q^4; q^5)};$$

(1.13) $$\sum_{n=0}^{\infty} \frac{q^{n(n+1)}}{(q)_n} = \frac{1}{\Pi(q^2, q^3; q^5)};$$

Regime III

(1.14) $$\sum_{n=0}^{\infty} \frac{q^{\frac{n(3n-1)}{2}}}{(q)_n (q;q^2)_n} = \frac{\Pi(q^4, q^6, q^{10}; q^{10})}{(q)_\infty};$$

(1.15) $$\sum_{n=0}^{\infty} \frac{q^{\frac{3n(n+1)}{2}}}{(q)_n (q;q^2)_{n+1}} = \frac{\Pi(q^2, q^8, q^{10}; q^{10})}{(q)_\infty};$$

Regime IV

$$\sum_{n=0}^{\infty} \frac{q^{n(n+1)}}{(q)_{2n+1}} = \frac{\Pi(q^3, q^7, q^{10}; q^{10})\Pi(q^4, q^{16}; q^{20})}{(q)_\infty}; \tag{1.16}$$

$$\sum_{n=0}^{\infty} \frac{q^{n(n+1)}}{(q)_{2n}} = \frac{\Pi(q, q^9, q^{10}; q^{10})\Pi(q^8, q^{12}; q^{20})}{(q)_\infty}; \tag{1.17}$$

$$\sum_{n=0}^{\infty} \frac{q^{n^2}}{(q)_{2n}} = \frac{1}{\Pi(q^4, q^{16}; q^{20})(q; q^2)_\infty}; \tag{1.18}$$

$$\sum_{n=1}^{\infty} \frac{q^{n^2}}{(q)_{2n-1}} = \frac{q}{\Pi(q^8, q^{12}; q^{20})(q; q^2)_\infty}. \tag{1.19}$$

These use the standard notation of Slater [38],

$$(a)_n = (a; q)_n = \prod_{j=0}^{n-1} (1 - aq^j); \tag{1.20}$$

$$(a)_\infty = (a; q)_\infty = \prod_{j=0}^{\infty} (1 - aq^j); \tag{1.21}$$

"For about the first 85 years after their discovery interest in these identities and their generalizations was confined to mathematicians. Many ingenious proofs and relations applying combinatorics, basic hypergeometric functions and Lie algebras were discovered by MacMahon, Rogers, Schur, Ramanujan, Watson, Bailey, Slater, Gordon, Göllnitz, Andrews, Bressoud, Lepowsky and Wilson; so by 1980 there were over 130 isolated identities and several infinite families of identities known.

The entry of these identities into physics occurred in the early 1980s when Baxter [11], Andrews, Baxter and Forrester (see [3] and [27]), and the Kyoto group [24] encountered (6.1) and various generalizations in the computation of order parameters of certain lattice models of statistical mechanics.

A further glimpse of the relation to physics is seen in the development of conformal field theory by Belavin, Polyakov and Zamolodchikov [14] and the form of computation of characters of representations of Virasoro algebra by Kac [28], Feigin and Fuchs [25] and Rocha-Caridi [32]. The occurrence of (6.1) in this context led Kac [29] to suggest that *every modular invariant representation of Virasoro should produce a Rogers-Ramanujan type identity.*"

- *Alexander Berkovich and Barry M. McCoy, State University of New York, 1998*