

Discrete Mathematics 132 (1994) 267-290

DISCRETE MATHEMATICS

Riordan arrays and combinatorial sums[☆]

Renzo Sprugnoli

Dipartimento di Sistemi e Informatica, Via Lombroso 6/17, I-50134 Firenze, Italy

Received 25 March 1991; revised 29 September 1992

Abstract

The concept of a Riordan array is used in a constructive way to find the generating function of many combinatorial sums. The generating function can then help us to obtain the closed form of the sum or its asymptotic value. Some examples of sums involving binomial coefficients and Stirling numbers are examined, together with an application of Riordan arrays to some walk problems.

Introduction

In 1978 Rogers [18] introduced the concept of *renewal array* as a generalization of the Pascal, Catalan and Motzkin triangles, and Kettle [14] used it to study other types of combinatorial triangles, especially those found in walk problems. More recently, Shapiro et al. [22] have examined and further generalized the same concept under the name of *Riordan array*, and have pointed out its connection with the 1-umbral calculus, as described in Roman [19] and others.

It is apparent, however, that the importance of the connection between Riordan arrays and combinatorial sums has been underestimated. As a result, our aim is to show how Riordan arrays allow us to find the generating function of many combinatorial sums. In turn, a generating function can be used either for finding a closed form for the sum or for determining its asymptotic value. The method is a constructive one in the sense that we do not have to know the value V of a sum $\sum_k f_k$ in advance to prove that $\sum_k f_k = V$ and given the sum, we can find out V, if it exists.

The traditional methods used for solving combinatorial sums (see, e.g., Riordan [17] or Comtet [6]) are nicely demonstrated by Knuth [15] and Graham et al. [10], where the authors show how to use the rules of binomial coefficients, Stirling numbers, and so on, in a skillful way.

^{*} This work was supported by the Italian Ministry for University and Scientific Research.

⁰⁰¹²⁻³⁶⁵X/94/\$07.00 © 1994—Elsevier Science B.V. All rights reserved SSDI 0012-365X (92)00570-6

In 1978, Gosper [9] discovered a general method for solving many combinatorial sums. His approach is excellent, but it is very difficult to perform it manually; it has been embodied in some systems of Computer Algebra, such as MACSYMA and MAPLE. The reader is also referred to Karr [12] for related concepts.

An approach using hypergeometric series was widely studied for sums only involving powers and factorials (and hence binomial coefficients). Andrews [1] and Ranjan Roy [16] made some important contributions on the subject, but final success was obtained by Zeilberger [26] and Wilf and Zeilberger [25] thanks to their concept of WZ-pair. This is a non-constructive method which certifies that an identity $\sum_k f_k = V$ is valid. An impressive quantity of combinatorial identities was proved in this way.

In an earlier work, Wilf [23] had proposed what he calls the "snake oil" method, a constructive technique for proving combinatorial identities involving sums. Basically, the method is a generating function approach to the problem, and consists in expressing the generating function of a sum as a sum of sums. By inverting the order of summation, it is often possible to obtain a simpler expression from which a closed form of the original sum can be deduced.

The "snake oil" method implies the transformation of generating functions. A somewhat similar approach was developed by Egorychev [7] with his "integral representation" of sums. In the present paper we try to combine the work of Rogers, Shapiro, Wilf and Egorychev (with a constant eye to Knuth) to answer the following question: what are the conditions under which a combinatorial sum can be solved by transforming the generating functions? The concept of a Riordan array seems to be an adequate answer to the problem.

We follow the structure of Rogers' paper [18] and begin with some general properties of Riordan arrays. In Sections 2 and 3 we study Riordan arrays related to binomial coefficients. In Section 4, we give a rather general theory of coloured walks on the line; finally, in Section 5, we propose some applications of the Riordan array concept to Stirling numbers.

1. Riordan arrays

Let $\mathscr{F} = \mathbb{R}[t]$ be the ring of the formal power series with real coefficients in some indeterminate t. If $f(t) \in \mathscr{F}$, $f(t) = \sum_{k=0}^{\infty} f_k t^k$, the order $\omega(f(t))$ of f(t) is the smallest integer k for which $f_k \neq 0$. \mathscr{F}_r is the set of exactly r order formal power series. As is well-known, an $f(t) \in \mathscr{F}$ is invertible if and only if $f(t) \in \mathscr{F}_0$. If $\{f_k\}_{k \in \mathbb{N}}$ is a sequence of real numbers, the formal power series $f(t) = \sum_{k=0}^{\infty} f_k t^k$ is called the generating function of the sequence, and we write $f(t) = \mathscr{G}_t \{f_k\}_{k \in \mathbb{N}} = \mathscr{G}\{f_k\}$. A Riordan array is a couple D = (d(t), h(t)), in which $d(t), h(t) \in \mathscr{F}$; if $h(t) \in \mathscr{F}_0$, the Riordan array is called proper. We are mainly interested in the sequence of functions $\{d_k(t)\}_{k \in \mathbb{N}}$ iteratively defined by

$$\begin{cases} d_0(t) = d(t), \\ d_k(t) = d(t)(th(t))^k. \end{cases}$$
(1.1)

These functions define an infinite triangle $\{d_{n,k} | k, n \in \mathbb{N}, k \leq n\}$, in which $d_{n,k} = [t^n]d_k(t)$ and therefore the functions $d_k(t)$ are called the *column generating functions* of the Riordan array, which is usually identified by the triangle. Another way of characterizing the Riordan array is to consider the bivariate generating function of the triangle:

$$d(t,w) = \sum_{k=0}^{\infty} d(t)(th(t))^{k} w^{k} = \frac{d(t)}{1 - twh(t)}.$$
(1.2)

A common example of a Riordan array is the Pascal triangle for which we have d(t) = h(t) = 1/(1-t). By (1.2) we find the well-known bivariate generating function $d(t, w) = (1 - t - wt)^{-1}$.

From our point of view, the most relevant property is the fact that the sums involving the rows of a Riordan array can be performed by operating a suitable transformation on a generating function and by then extracting a coefficient from the resulting function. In fact, if D = (d(t), h(t)) is a Riordan array and f(t) is the generating function of the sequence $\{f_i\}_{i \in \mathbb{N}}$ then we have

$$\sum_{k=0}^{\infty} d_{n,k} f_{k} = \sum_{k=0}^{\infty} [t^{n}] d_{k}(t) [y^{k}] f(y) = [t^{n}] \sum_{k=0}^{\infty} d(t) (th(t))^{k} [y^{k}] f(y)$$
$$= [t^{n}] d(t) \sum_{k=0}^{\infty} [y^{k}] f(y) (th(t))^{k} = [t^{n}] d(t) f(th(t)).$$

Because of its importance, this property is stated as the following theorem.

Theorem 1.1. Let D = (d(t), h(t)) be a Riordan array and let f(t) be the generating function of the sequence $\{f_i\}_{i \in \mathbb{N}}$. Then we have

$$\sum_{k=0}^{\infty} d_{n,k} f_k = [t^n] d(t) f(th(t)).$$
(1.3)

In the case of the Pascal triangle, the relation (1.3) is known as "Euler transformation" and reads $\sum {n \choose k} f_k = (1-t)^{-1} f(t(1-t)^{-1})$. Considering the generating functions $\mathscr{G}\{1\} = (1-t)^{-1}$, $\mathscr{G}\{(-1)^{-k}\} = (1+t)^{-1}$ and $\mathscr{G}\{k\} = t(1-t)^{-2}$, we immediately obtain:

$$\sum d_{n,k} = [t^n] \frac{d(t)}{1 - th(t)}$$
 (row sums),

$$\sum (-1)^k d_{n,k} = [t^n] \frac{d(t)}{1 + th(t)}$$
 (alternating row sums),

$$\sum k d_{n,k} = [t^n] \frac{td(t)h(t)}{(1 - th(t))^2}$$
 (weighted row sums).

Moreover, by observing that $\hat{D} = (d(t), th(t))$ is a Riordan array, whose rows are the diagonals of D, we have

$$\sum d_{n-k,k} = [t^n] \frac{d(t)}{1-t^2 h(t)} \quad \text{(diagonal sums)}.$$

Obviously, this last observation can be generalized to find the generating function of any sum $\sum d_{n-sk,k}$ for every $s \ge 1$. We obtain well-known results for the Pascal triangle. In particular, if F_n is the *n*th Fibonacci number, for the diagonal sums we have

$$\sum_{k=0}^{\infty} \binom{n-k}{k} = [t^n] \frac{1}{1-t} \frac{1}{1-t^2(1-t)^{-1}} = [t^n] \frac{1}{1-t-t^2} = [t^{n+1}] \frac{t}{1-t-t^2} = F_{n+1}.$$

Another general result can be obtained by means of two sequences $\{f_k\}_{k\in\mathbb{N}}, \{g_k\}_{k\in\mathbb{N}}$ and their generating functions f(t), g(t). For p=1, 2, ... the general element of the Riordan array $(f(t), t^{p-1})$ is

$$d_{n,k} = [t^n] f(t)(t^p)^k = [t^{n-pk}] f(t) = f_{n-pk}.$$

Hence, by Theorem 1.1, we have

$$\sum_{k=0}^{n} f_{n-pk} g_{k} = [t^{n}] f(t) [g(y) | y = t^{p}] = [t^{n}] f(t) g(t^{p}).$$

This can be called the *rule of generalized convolution* since it reduces to the usual convolution rule for p=1. Suppose, for example, that we wish to sum one out of every three powers of 2, starting with 2^n and going down to the lowest integer exponent ≥ 0 , we have

$$S_{n} = \sum_{k=0}^{\lfloor n/3 \rfloor} 2^{n-3k} = [t^{n}] \frac{1}{1-2t} \frac{1}{1-t^{3}} \sim 2^{n} \frac{1}{1-1/8} = \frac{2^{n+3}}{7};$$

we actually have $S_n = \lfloor 2^{n+3}/7 \rfloor$.

In a sense, Theorem 1.1 is a characterization of Riordan arrays and we can also prove a sort of inverse property.

Theorem 1.2. Let $\{d_{n,k} | n, k \in \mathbb{N}, k \leq n\}$ be an infinite triangle such that for every sequence $\{f_k\}_{k \in \mathbb{N}}$ we have $\mathscr{G}\{\sum d_{n,k}f_k\} = d(t)f(th(t))$, where f(t) is the generating function of the sequence and d(t), h(t) are two formal power series not depending on f(t). The triangle defined by the Riordan array (d(t), h(t)) coincides with $\{d_{n,k}\}$.

Proof. For any $k \in \mathbb{N}$, take the sequence which is 0 everywhere except in the kth element $f_k = 1$. The corresponding generating function is $f(t) = t^k$ and we have $\sum_{i=0}^{\infty} d_{n,i} f_i = d_{n,k}$. Hence, according to the theorem's hypotheses, we find $\mathscr{G}_t \{d_{n,k}\}_{n \in \mathbb{N}} = d_k(t) = d(t)(th(t))^k$, and this corresponds to (1.1) for every k = 1, 2, ...

271

Proper Riordan arrays play a very important role in this approach. Let us consider a Riordan array D = (d(t), h(t)), which is not proper. Since h(0) = 0, an s > 0 exists such that $h(t) = h_s t^s + h_{s+1} t^{s+1} + \cdots$ and $h_s \neq 0$. If we define $\hat{h}(t) = h_s + h_{s+1} t + \cdots$, then $\hat{h}(t) \in \mathscr{F}$ and $\hat{h}(0) = h_s \neq 0$. Consequently, the Riordan array $\hat{D} = (d(t), \hat{h}(t))$ is proper and the rows of D can be seen as the *s*-diagonals $\{\hat{d}_{n-sk} | k \ge 0\}$ of \hat{D} . Fortunately, for proper Riordan arrays, Rogers [18] has found an important characterization: every element $d_{n+1,k+1}, n, k \in \mathbb{N}$, can be expressed as a linear combination of the elements in the preceding row, i.e.,

$$d_{n+1,k+1} = a_0 d_{n,k} + a_1 d_{n,k+1} + a_2 d_{n,k+2} + \dots = \sum_{j=0}^{\infty} a_j d_{n,k+j}.$$
 (1.4)

The sum is actually finite and the sequence $A = \{a_j\}_{j \in \mathbb{N}}$ is fixed. It is called the *A*-sequence of the Riordan array and, as Rogers has shown, it only depends on h(t):

$$h(t) = A(th(t)), \tag{1.5}$$

if A(t) is the generating function of the sequence A. By using the Lagrange inversion formula, Rogers [18] has also shown that the A-sequence determines an infinite triangle as a proper Riordan array in the following way.

Theorem 1.3. Let $\{d_{n,k} | k, n \in \mathbb{N}, k \leq n\}$ be an infinite triangle such that $d_{n,n} \neq 0, \forall n \in \mathbb{N}$, and for which the relation (1.4) holds true for some sequence $A = \{a_j\}_{j \in \mathbb{N}}, a_0 \neq 0$. Then D is a Riordan array (d(t), h(t)), where d(t) is the generating function of the sequence $\{d_{n,0}\}_{n \in \mathbb{N}}$, and h(t) is the unique solution of h(t) = A(th(t)) with $h(0) \in \mathbb{R} \setminus \{0\}$.

The A-sequence of the Pascal triangle is the solution A(y) of the functional equation 1/(1-t) = A(t/(1-t)). The simple substitution $y = t(1-t)^{-1}$ gives A(y) = 1 + y, corresponding to the well-known recurrence $\binom{n+1}{k+1} = \binom{n}{k} + \binom{n}{k+1}$. At this point, we realize that we could have started with this recurrence relation and directly found A(y) = 1 + y. h(t) is defined by (1.5) as the solution of h(t) = 1 + th(t), and this immediately gives $h(t) = (1-t)^{-1}$. Furthermore, since the first column is $\{1, 1, 1, ...\}$, the Pascal triangle corresponds to the Riordan array $((1-t)^{-1}, (1-t)^{-1})$.

Finally, it is possible to consider *improper Riordan arrays* (d(t), h(t)), for which (1.1) is changed to:

$$\begin{cases} d_0(t) = d(t), \\ d_k(t) = d(t)(h(t))^k. \end{cases}$$

If $h(0) \neq 0$, the improper Riordan arrays correspond to quadrangular arrays and Theorem 1.1 remains valid under more restrictive conditions, such as when the sequence $\{f_k\}_{k \in \mathbb{N}}$ is actually finite and hence f(t) reduces to a polynomial.

2. Binomial coefficients of the form $\binom{n+ak}{m+bk}$

272

Let us consider the binomial coefficient $\binom{n+ak}{m+bk}$, where a, b are two parameters and k is a non-negative integer variable. Depending if we consider n a variable and m a parameter, or vice versa, we have two different infinite arrays: $\{d_{n,k}\}$ of $\{\hat{d}_{m,k}\}$, whose elements depend on the parameters a, b, m or a, b, n, respectively. In either case we have two Riordan arrays and therefore, we can apply Theorem 1.2 to find the value of many sums.

Theorem 2.1. Let $d_{n,k}$ and $\hat{d}_{m,k}$ be as above. If b > a and b-a is an integer, then $D = \{d_{n,k}\}$ is a Riordan array. If b < 0 is an integer then $\hat{D} = \{\hat{d}_{m,k}\}$ is a Riordan array. We have

$$D = \left(\frac{t^m}{(1-t)^{m+1}}, \frac{t^{b-a-1}}{(1-t)^b}\right), \qquad \hat{D} = \left((1+t)^n, \frac{t^{-b-1}}{(1+t)^{-a}}\right).$$

Proof. By using well-known properties of binomial coefficients, we find

$$\begin{aligned} d_{n,k} &= \binom{n+ak}{m+bk} \\ &= \binom{n+ak}{n-m+ak-bk} = \binom{-n-ak+n-m+ak-bk-1}{n-m+ak-bk} (-1)^{n-m+ak-bk} \\ &= \binom{-m-bk-1}{(n-m)+(a-b)k} (-1)^{n-m+ak-bk} = [t^{n-m+ak-bk}] \frac{1}{(1-t)^{m+1+bk}} \\ &= [t^n] \frac{t^m}{(1-t)^{m+1}} \left(\frac{t^{b-a}}{(1-t)^b}\right)^k, \\ \hat{d}_{m,k} &= \binom{n+ak}{m+bk} = [t^{m+bk}] (1+t)^{n+ak} = [t^m] (1+t)^n (t^{-b}(1+t)^a)^k. \end{aligned}$$

The theorem now directly follows from (1.1). \Box

For m=a=0 and b=1 we again find the Riordan array of the Pascal triangle. For b=a, $d_{n,k}$ is an improper Riordan array, and so $\hat{d}_{m,k}$ is for b=0. In such cases, Theorem 1.1 can only be applied under the restrictions described at the end of Section 1.

The sum (1.3) takes on two specific forms, which are worth being stated explicitly:

$$\sum {\binom{n+ak}{m+bk}} f_k = [t^n] \frac{t^m}{(1-t)^{m+1}} f\left(\frac{t^{b-a}}{(1-t)^b}\right) \quad b > a,$$
(2.1)

R. Sprugnoli / Discrete Mathematics 132 (1994) 267-290

$$\sum {\binom{n+ak}{m+bk}} f_k = [t^m](1+t)^n f(t^{-b}(1+t)^a) \quad b \le -1.$$
(2.2)

273

If *m* and *n* are independent of each other, these relations can also be stated as generating function identities. The binomial coefficient $\binom{n+ak}{m+bk}$ is so general that a large number of combinatorial sums can be solved by means of formulas (2.1) and (2.2). Some examples follow in which we use the notation [f(y)|y=g(t)] as a linearization of the more common one $f(y)|_{y=g(t)}$ to denote substitution f(g(t)).

Let us begin with a very simple case. By Theorem 2.1, the binomial coefficient $\binom{n-k}{m}$ corresponds to the Riordan array $(t^m/(1-t)^{m+1}, 1)$, therefore, by the formula concerning the row sums, we have

$$\sum \binom{n-k}{m} = [t^n] \frac{t^m}{(1-t)^{m+1}} \frac{1}{1-t} = [t^{n-m}] \frac{1}{(1-t)^{m+2}} = \binom{n+1}{m+1}.$$
(2.3)

The sum $\sum {\binom{n+k}{m+2k}} {\binom{2k}{k}} (-1)^k / (k+1)$ is a more interesting example. From the generating function of the Catalan numbers, we immediately have

$$\mathscr{G}\left\{\binom{2k}{k}\frac{(-1)^k}{k+1}\right\} = (\sqrt{1+4t}-1)/2t.$$

Hence,

$$\begin{split} \sum \binom{n+k}{m+2k} \binom{2k}{k} \frac{(-1)^k}{k+1} &= \left[t^n\right] \frac{t^m}{(1-t)^{m+1}} \left[\frac{\sqrt{1+4y}-1}{2y}\right] y = \frac{t}{(1-t)^2} \\ &= \left[t^{n-m}\right] \frac{1}{(1-t)^{m+1}} \left(\sqrt{1-\frac{4t}{(1-t)^2}} - 1\right) \frac{(1-t)^2}{2t} \\ &= \left[t^{n-m}\right] \frac{1}{(1-t)^m} = \binom{n-1}{m-1}. \end{split}$$

In the sum $\sum {\binom{z+1}{2k+1}} {\binom{z-2k}{n-k}} 2^{2k+1}$ we use the bisection formulas for series (see, for example Riordan [17]). Since the generating function for ${\binom{z+1}{k}} 2^k$ is $(1+2t)^{z+1}$, we have

$$\mathscr{G}\left\{\binom{z+1}{2k+1}2^{2k+1}\right\} = \frac{1}{2\sqrt{t}}((1+2\sqrt{t})^{z+1}-(1-2\sqrt{t})^{z+1}).$$

By now applying formula (2.2) we find

$$\Sigma {\binom{z+1}{2k+1}} {\binom{z-2k}{n-k}} 2^{2k+1}$$

= $[t^n](1+t)^z \left[\frac{(1+2\sqrt{y})^{z+1} - (1-2\sqrt{y})^{z+1}}{2\sqrt{y}} \right] y = \frac{t}{(1+t)^2} \right]$

R. Sprugnoli / Discrete Mathematics 132 (1994) 267-290

$$= [t^{n}](1+t)^{z+1} \frac{(1+t+2\sqrt{t})^{z+1} - (1+t-2\sqrt{t})^{z+1}}{2\sqrt{t}(1+t)^{z+1}}$$

$$\stackrel{(*)}{=} [t^{2n+1}](1+t)^{2z+2} = \binom{2z+2}{2n+1}$$

in which we used the bisection rule backwards in (*), since $(1+t\pm 2\sqrt{t})^{z+1} = (1\pm\sqrt{t})^{2z+2}$.

If $\binom{k}{m}$ denotes a Stirling number of the second kind, the corresponding generating function is $t^{m}/(1-t)(1-2t)\cdots(1-mt)$. Hence, by applying formula (2.1), we find

$$\sum {\binom{n}{k}} {\binom{k}{m}} = [t^n] \frac{1}{1-t} \frac{t^m}{(1-t)^m (1-t/(1-t)) \cdots (1-mt/(1-t))}$$
$$= [t^{n+1}] \frac{t^{m+1}}{(1-t)(1-2t) \cdots (1-(m+1)t)} = {\binom{n+1}{m+1}}.$$

We solve the following sum by using (2.2):

$$\sum {\binom{2n-2k}{m-k}} {\binom{n}{k}} (-2)^k = [t^m] (1+t)^{2n} \left[(1-2y)^n \left| \xi y = \frac{t}{(1+t)^2} \right] \right]$$
$$= [t^m] (1+t)^{2n} \frac{(1+2t+t^2-2t)^n}{(1+t)^{2n}} = [t^m] (1+t^2)^n = {\binom{n}{m/2}}$$

in which the binomial coefficient is to be taken as zero for m odd.

In the sum $\sum {\binom{n+k}{k}} {\binom{n}{k}} k(-1)^{k-1}$ we have $f(t) = \mathscr{G}\left\{ {\binom{n}{k}} k(-1)^{k-1} \right\} = nt(1-t)^{n-1}$, and ${\binom{n+k}{k}} = {\binom{n+k}{n}}$. We apply formula (2.2) having b = 0. This means that the Riordan array is improper, but since the generating function f(t) is a polynomial, our rule is still applicable:

$$\sum {\binom{n+k}{n}} {\binom{n}{k}} k(-1)^{k-1} = [t^n](1+t)^n [ny(1-y)^{n-1} | y=1+t]$$

= $n[t^n](1+t)^{n+1}(-t)^{n-1} = (-1)^{n-1} n[t^1](1+t)^{n+1}$
= $(-1)^{n-1} n(n+1).$

The examples are almost infinite and can become very complicated. It is worth examining at least one case in which we do not obtain a closed form but whose generating function allows us to find the asymptotic value of the sum. The following is a typical example. Let us determine the asymptotic value of the sum:

$$S_{n} = \sum {\binom{n}{k}}^{2} p^{k} q^{n-k} = [t^{n}](1+pt)^{n}(1+qt)^{n},$$

or, more in general, the coefficient $[t^n](1+\alpha t+\beta t^2)^n$, where p, q, α, β may also be complex. Given the obvious fact that $[t^n]f(at) = a^n[t^n]f(t)$, we use the rules of

convolution and (2.1) and obtain

$$S_{n} = \sum {\binom{n}{n-k}} q^{n-k} {\binom{n}{k}} p^{k} = [t^{n}] \frac{1}{1-qt} \left[(1+py)^{n} \middle| y = \frac{t}{1-qt} \right]$$
$$= [t^{n}] \frac{(1+(p-q)t)^{n}}{(1-qt)^{n+1}} = \sum {\binom{2n-k}{n-k}} q^{n-k} {\binom{n}{k}} (p-q)^{k}.$$

Now we observe that $\binom{2n-k}{n}\binom{n}{k} = \binom{2n-2k}{n-k}\binom{2n-2k}{n-k}$; by writing h for n-k and by using (2.1) again:

$$S_{n} = \sum {\binom{n+h}{2h}} (p-q)^{n-h} {\binom{2h}{h}} q^{h}$$

= $[t^{n}] \frac{1}{(1-(p-q)t)} \left[\frac{1}{\sqrt{1-4qy}} \middle| y = \frac{t}{(1-(p-q)t)^{2}} \right]$
= $[t^{n}] \frac{1}{\sqrt{1-2(p+q)t+(p-q)^{2}t^{2}}}$
= $[t^{n}] (1-(\sqrt{p}+\sqrt{q})^{2}t)^{-1/2} (1-(\sqrt{p}-\sqrt{q})^{2}t)^{-1/2}.$

The last two expressions do not depend on *n* and therefore they represent the generating function of S_n . The very last expression can be used to determine the asymptotic value of S_n by means of Darboux's method. Let us suppose that $(\sqrt{p}+\sqrt{q})^{-2}$ has a smaller modulus than $(\sqrt{p}-\sqrt{q})^{-2}$; then the second factor is analytic in $t = (\sqrt{p}+\sqrt{q})^{-2}$ and can be developed in a Taylor series around this point. By extracting the coefficient of t^n in the resulting expression, we eventually find,

$$S_{n} \sim \frac{(\sqrt{p} + \sqrt{q})^{2n+1}}{2\sqrt{\pi n}\sqrt{pq}} \times \left(1 + \frac{p+q-4\sqrt{pq}}{16n\sqrt{pq}} + \frac{pq+3\sqrt{pq}(\sqrt{p} - \sqrt{q})^{2} + 9(\sqrt{p} - \sqrt{q})^{4}}{128n^{2}pq}\right).$$

Obviously, if $(\sqrt{p} - \sqrt{q})^{-2}$ has a smaller modulus, then we would proceed in an analogous way by developing the first factor.

Other Riordan arrays can be found by using Theorems 2.1 and 2.2 and the rule:

$$\frac{\alpha \pm \beta}{\beta} \binom{\alpha}{\beta} = \binom{\alpha}{\beta} \pm \binom{\alpha-1}{\beta-1} \qquad (\alpha \neq 0, \text{ if } -\text{ is considered}).$$

For example, by writing $\binom{n+k}{2k}$ as $\binom{n+k}{n-k}$ we easily find

$$\frac{2n}{n+k}\binom{n+k}{n-k} = \binom{n+k}{n-k} + \binom{n+k-1}{n-k-1} = \binom{n+k}{2k} + \binom{n-1+k}{2k}.$$

Hence, by (2.1) we have

$$\sum \frac{2n}{n+k} \binom{n+k}{n-k} f_k = \sum \binom{n+k}{2k} f_k + \sum \binom{n-1+k}{2k} f_k$$
$$= [t^n] \frac{1}{1-t} f\left(\frac{t}{(1-t)^2}\right) + [t^{n-1}] \frac{1}{1-t} f\left(\frac{t}{(1-t)^2}\right)$$
$$= [t^n] \frac{1+t}{1-t} f\left(\frac{t}{(1-t)^2}\right). \tag{2.4}$$

This proves that the infinite triangle of the elements $\frac{2n}{n+k} \binom{n+k}{2k}$ is a proper Riordan array and many typical identities can be proved by means of (2.4). For example,

$$\sum \frac{2n}{n+k} \binom{n+k}{n-k} \binom{2k}{k} (-1)^k = [t^n] \frac{1+t}{1-t} \left[\frac{1}{\sqrt{1+4y}} \middle| y = \frac{t}{(1-t)^2} \right] = [t^n] 1 = \delta_{n0},$$

$$\sum \frac{2n}{n+k} \binom{n+k}{n-k} \binom{2k}{k} \frac{(-1)^k}{k+1} = [t^n] \frac{1+t}{1-t} \left[\frac{\sqrt{1+4y}-1}{2y} \middle| y = \frac{t}{(1-t)^2} \right]$$

$$= [t^n] (1+t) = \delta_{n0} + \delta_{n1}.$$

The following is quite a different case. Let $f(t) = \mathscr{G} \{ f_k \}$ and

$$F(t) = \mathscr{G}\left\{\frac{f_k}{k}\right\} = \int \frac{f(t) - f_0}{t} dt$$

with F(0) = 0. Obviously, we have

$$\binom{n-k}{k} = \frac{n-k}{k} \binom{n-k-1}{k-1},$$

except for k=0, when the left-hand side is 1 and the right-hand side is not defined. Hence, by (2.1)

$$\sum \frac{n}{n-k} \binom{n-k}{k} f_k = f_0 + n \sum_{k=1}^{\infty} \binom{n-k-1}{k-1} \frac{f_k}{k} = f_0 + n [t^n] F\left(\frac{t^2}{1-t}\right).$$
(2.5)

This formula gives an immediate proof of Hardy's identity:

$$\sum \frac{n}{n-k} \binom{n-k}{k} (-1)^k = [t^n] \log \frac{1+t}{1+t^3} = [t^n] \left(\log \frac{1}{1+t^3} - \log \frac{1}{1+t} \right)$$
$$= \begin{cases} (-1)^n 2/n, & \text{if } 3 \setminus n, \\ (-1)^{n-1}/n, & \text{else.} \end{cases}$$

We also immediately obtain

$$\sum \frac{1}{n-k} \binom{n-k}{k} = (\phi^n + \hat{\phi}^n)/n,$$

where ϕ is the golden ratio and $\hat{\phi} = -\phi^{-1}$. The reader can generalize formula (2.5) by using the change of variable $t \rightarrow pt$ and then he can prove some well-known formulas such as Riordan's "very old identity" [17, p. 58] or a generalization of Hardy's identity:

$$\sum \frac{n}{n-k} \binom{n-k}{k} (a+b)^{n-2k} (-ab)^k = a^n + b^n,$$

$$\sum \frac{n}{n-k} \binom{n-k}{k} x^{n-2k} (-1)^k = \frac{(x+\sqrt{x^2-4})^n + (x-\sqrt{x^2-4})^n}{2^n}.$$

3. Binomial coefficients of the form $\binom{2n+ak}{n+bk}$

When *m* depends on *n*, the formulas of the previous section are no longer appropriate and can only be applied in a few special cases. In this section we examine the binomial coefficients of the form $\binom{2n+ak}{n+bk}$, in which *a* and *b* are two integer parameters, and find out the values of *a* and *b* corresponding to Riordan arrays. In the latter case we also determine the explicit form of the couple (d(t), h(t)), and are therefore able to find the generating function relative to every sum $\sum \binom{2n+ak}{n+bk} f_k$.

In Section 2, we used a direct method for finding the Riordan array relative to the binomial coefficients considered. We now use the A-sequence to determine the form of the Riordan arrays regarding two particular cases which then help us generalize the result.

We start out with a simple identity:

$$\binom{2n}{n-k-1} + 2\binom{2n}{n-k} + \binom{2n}{n-k+1} = \binom{2n+1}{n-k} + \binom{2n+1}{n-k+1} = \binom{2n+2}{n+1-k}.$$

If we set $d_{n,k} = \binom{2n}{n-k}$, it means that $d_{n+1,k} = d_{n,k-1} + 2d_{n,k} + d_{n,k+1}$ and, by Theorem 1.3, this proves that $\{d_{n,k}\}$ is a Riordan array, having an A-sequence (1, 2, 1) or $A(t) = 1 + 2t + t^2$. By (1.5), the function h(t) is the solution of $h(t) = 1 + 2th(t) + t^2h^2(t)$, with $h(0) \in \mathbb{R} \setminus \{0\}$. Since for k = 0, $d_{n,0} = \binom{2n}{n}$, the triangular array of the elements $\binom{2n}{n-k}$ is the proper Riordan array:

$$\left\{ \binom{2n}{n-k} \right\} = \left(\frac{1}{\sqrt{1-4t}}, \frac{1-2t-\sqrt{1-4t}}{2t^2} \right).$$
(3.1)

Analogously, the following sum can be closed by formula (2.3):

$$\sum_{j} \binom{2n - (k-1) - j}{n - (k-1) - j} = \sum_{j} \binom{2n - k + 1 - j}{n} = \binom{2n - k + 2}{n+1} = \binom{2(n+1) - k}{n+1}.$$

If we set $d_{n,k} = \binom{2n-k}{n-k}$, it means that $d_{n+1,k} = d_{n,k-1} + d_{n,k} + d_{n,k+1} + \cdots$. Again, by Theorem 1.3, $\{d_{n,k}\}$ is a Riordan array having an A-sequence $(1, 1, 1, \ldots)$, or $A(t) = (1-t)^{-1}$. By (1.5), the function h(t) is the solution of $h(t) = (1-th(t))^{-1}$ with $h(0) \neq 0$, and we obtain the proper Riordan array

$$\left\{ \binom{2n-k}{n-k} \right\} = \left(\frac{1}{\sqrt{1-4t}}, \frac{1-\sqrt{1-4t}}{2t} \right). \tag{3.2}$$

Let us now give two non-trivial examples. First, the identity of Ruskey [20] is proved by using (3.2)

$$\sum {\binom{2n-k}{n-k}} (k+1)2^k = [t^n] \frac{1}{\sqrt{1-4t}} \left[\frac{1}{(1-2y)^2} \middle| y = \frac{1-\sqrt{1-4t}}{2} \right]$$
$$= [t^n] \frac{1}{\sqrt{1-4t}} \frac{1}{1-4t} = [t^n](1-4t)^{-3/2}$$
$$= {\binom{-3/2}{n}} (-4)^n = (2n+1) {\binom{2n}{n}}.$$

We then prove the identity of Van Ebbenhorst-Tengbergen (see Egorychev [7])

$$\sum \binom{k+m-1}{k} \binom{2n}{n-k} \binom{k}{m} = \binom{2n-1}{n} \binom{n}{m} = \frac{1}{2} \binom{2n}{n} \binom{n}{m},$$
(3.3)

by means of (3.1). First, we take a binomial coefficient out of the sum by using the identity

$$\binom{k+m-1}{k}\binom{k}{m} = (-1)^k \binom{-m}{k}\binom{k}{m} = \binom{2m-1}{m}\binom{m+k-1}{k-m}.$$

We can now observe that $\mathscr{G}\left\{\binom{m+k-1}{k-m}\right\} = t^m(1-t)^{-2m}$. Hence,

$$\sum {\binom{2n}{n-k}} {\binom{m+k-1}{k-m}} = [t^n] \frac{1}{\sqrt{1-4t}} \left[\left(\frac{y}{(1-y)^2} \right)^m \middle| y = \frac{1-2t-\sqrt{1-4t}}{2t} \right].$$

After performing the substitution and simplifying, we find

$$[t^{n}]\frac{1}{\sqrt{1-4t}}\frac{t^{m}}{(1-4t)^{m}}=[t^{n-m}]\frac{1}{(1-4t)^{m+1/2}}$$

$$= \binom{-m-1/2}{n-m} (-4)^{n-m} = \binom{n-1/2}{n-m} 4^{n-m}.$$

The closed form of the sum is therefore $\binom{2m-1}{m}\binom{n-1/2}{n-m}4^{n-m}$, and this is proved equivalent to (3.3) by using the Γ function to express the binomial coefficients.

Many other identities can be proved, from a simple one like $\sum {\binom{2n-k}{n-k}} 2^k = 4^n$ to a more complicated one like $\sum {\binom{2n}{n-k}}/k = \frac{1}{2} {\binom{2n}{n}} H_n$. Now, to come to the general problem, we begin with the binomial coefficients $d_{n,k}^{(a)} = {\binom{2n+ak}{n-k}}$ and prove that every $\{d_{n,k}^{(a)}\}$ is a proper Riordan array. Formulas (3.1) and (3.2) are two particular cases of this. We first obtain the following general result on Riordan arrays.

Theorem 3.1. If D = (d(t), h(t)) is a Riordan array, the s-section array $\{d_{n,sk}\}$ is the Riordan array

$$D^{(s)} = \{d_{n,sk}\} = (d(t), t^{s-1}h^{s}(t))$$

Proof. The array $\{d_{n,sk}\}$ is composed of one out of every s column in the Riordan array $\{d_{n,k}\}$, i.e., $\{d_{n,sk}\}$ contains the columns

$$d(t) \quad d(t)t^{s}h^{s}(t) \quad d(t)t^{2s}h^{2s}(t) \quad d(t)t^{3s}h^{3s}(t) \quad \dots$$

which can be written as

$$d(t) \quad d(t)t(t^{s-1}h^{s}(t)) \quad d(t)t^{2}(t^{s-1}h^{s}(t))^{2} \quad d(t)t^{3}(t^{s-1}h^{s}(t))^{3} \quad \dots$$

and the theorem directly follows from definition (1.1). \Box

We can now observe that if $\{\binom{2n+ak}{n-k}\}$ is a Riordan array, then it is proper because the binomial coefficient is different from zero only for $0 \le k \le n$. If we consider the (a+1)-diagonals of the array, this new array consists of the elements $d_{n-(a+1)k,k} = \binom{2n-(a+2)k}{n-(a+2)k}$ and hence it may be considered to be the (a+2)-section of the Riordan array defined by $\binom{2n-k}{n-k}$. It is therefore a Riordan array and by Theorem 3.1 and formula (3.2) we have

$$\{d_{n-(a+1)k,k}\} = \left(\frac{1}{\sqrt{1-4t}}, t^{a+1}\left(\frac{1-\sqrt{1-4t}}{2t}\right)^{a+2}\right).$$

Finally, we must "push up" each column k of this array (a + 1)k positions in order to obtain the elements of the array $\{d_{n,k}^{(a)}\}$, and this is done by dividing h(t) by t^{a+1} . This proves that the array is indeed a Riordan array and

$$\left\{ \begin{pmatrix} 2n+ak\\n-k \end{pmatrix} \right\} = \left(\frac{1}{\sqrt{1-4t}}, \left(\frac{1-\sqrt{1-4t}}{2t} \right)^{a+2} \right).$$
(3.4)

In Table 1 we list a few cases corresponding to a = -3, -2, -1, 0, 1, 2.

Table 1 A few cases of Riordan arrays

$$\left\{ \binom{2n+2k}{n-k} \right\} = \left(\frac{1}{\sqrt{1-4t}}, \frac{1-4t+2t^2-(1-2t)\sqrt{1-4t}}{2t^4} \right)$$

$$\left\{ \binom{2n+k}{n-k} \right\} = \left(\frac{1}{\sqrt{1-4t}}, \frac{1-3t-(1-t)\sqrt{1-4t}}{2t^3} \right)$$

$$\left\{ \binom{2n}{n-k} \right\} = \left(\frac{1}{\sqrt{1-4t}}, \frac{1-2t-\sqrt{1-4t}}{2t^2} \right)$$

$$\left\{ \binom{2n-k}{n-k} \right\} = \left(\frac{1}{\sqrt{1-4t}}, \frac{1-\sqrt{1-4t}}{2t} \right)$$

$$\left\{ \binom{2n-2k}{n-k} \right\} = \left(\frac{1}{\sqrt{1-4t}}, 1 \right)$$

$$\left\{ \binom{2n-3k}{n-k} \right\} = \left(\frac{1}{\sqrt{1-4t}}, \frac{1+\sqrt{1-4t}}{2} \right)$$

As an example, let us now try to compute the sum $\sum {\binom{2n+2k}{n-k}} 2^k$. By Table 1, we have

$$\sum {\binom{2n+2k}{n-k}} 2^k = [t^n] \frac{1}{\sqrt{1-4t}} \left[\frac{1}{1-2y} \right| y = \frac{1-4t+2t^2-(1-2t)\sqrt{1-4t}}{2t^3} \right]$$
$$= [t^n] \frac{(1-2t)\sqrt{1-4t}+1-4t+2t^2-t^3}{\sqrt{1-4t}(2-12t+4t^2-t^3)}.$$

Because of this generating function, it is not very likely that the sum has a (simple) closed form. However, by solving the third-degree equation $2-12t+4t^2-t^3=0$, we find that the generating function has a simple pole at

$$t = \frac{4}{3} - \sqrt[3]{\frac{125}{27}} + \sqrt{\frac{875}{27}} - \sqrt[3]{\frac{125}{27}} - \sqrt{\frac{875}{27}} \approx 0.1766039807,$$

which is the singularity of the smallest modulus. Hence, we have

$$\sum \binom{2n+2k}{n-k} 2^k \approx 0.685792972794 \times (5.6623865217)^n.$$

Finally, let us come to the general case $\binom{2n+ak}{n+bk}$. We distinguish several subcases. (1) Let $a \ge 0$. Then:

(a) if b < 0, then let us write $\binom{2n+ak}{n+bk} = \binom{2n+ak}{n-ck}$, and imagine the latter to be the (c-1)-diagonals of the Riordan array of the binomial coefficient $\binom{2n+(a+2c-2)k}{n-k}$, which has the previously described standard form of formula

(3.4). Hence, we have

$$\left\{ \begin{pmatrix} 2n+ak\\ n-ck \end{pmatrix} \right\} = \left(\frac{1}{\sqrt{1-4t}}, t^{c-1} \left(\frac{1-\sqrt{1-4t}}{2t} \right)^{a+2c} \right);$$
(3.5)

- (b) if $0 \le b \le a$, the array is not triangular and, at most, can be considered to be an improper Riordan array;
- (c) if b > a, by symmetry we have: $\binom{2n+ak}{n+bk} = \binom{2n+ak}{n+(a-b)k}$ and, since a-b < 0, we are reduced to case (1a).
- (2) If a < 0, the rule of symmetry is no longer applicable since (^{2n-qk}) is always defined except for a negative p. Hence, (^{2n+ak}_{n+bk}), for a < 0 and b≥0, is defined for every value of k and does not correspond to any triangular array, whereas (^{2n+ak}_{n-(b-a)k}) is only defined for k≤n/(b-a). Consequently, for b<0 we have the same result as in case (1a). Note that (^{2n-5k}_{n-2k}) and (^{2n-5k}_{n-3k}) define two different Riordan arrays, even though they are identical as binomial coefficients.

We conclude by applying (3.5) to the following alternating sum:

$$\sum {\binom{2n}{n-3k}} (-1)^k = [t^n] \frac{1}{\sqrt{1-4t}} \left[\frac{1}{1+y} \middle| y = t^3 \left(\frac{1-\sqrt{1-4t}}{2t} \right)^6 \right]$$
$$= [t^n] \left(\frac{1}{2\sqrt{1-4t}} + \frac{1-t}{2(1-3t)} \right) = \frac{1}{2} {\binom{2n}{n}} + 3^{n-1} + \frac{1}{6} \delta_{n,0}$$

(see [17, p. 86] for the non-alternating sum, which can be proved analogously).

4. Coloured walks

We now examine the combinatorial problem concerning the walks on the line. Let us define the *discrete line* as the set of the integer points on the real line \mathbb{R} . A *walk on the line* is a sequence of steps starting from the origin. There are three kinds of steps:

- (a) right steps, which go from the point x to the point x+1;
- (b) "sur-place" steps, which remain on the point x;
- (c) left steps, which go from the point x to the point x-1.

A coloured walk is a walk in which every kind of step can assume different colours; we denote by a, b, c ($a > 0, b, c \ge 0$) the number of colours the right, sur-place and left steps can be. We discuss *complete* coloured walks, i.e., walks without any restrictions, and *positive* coloured walks, i.e., walks that never go to the left of the origin. The *length* of a walk is the number of its steps, and we denote by $d_{n,k}$ the number of coloured walks which have length n and reach a distance k from the origin, i.e., the last step ends at position k on the discrete line. A coloured walk problem is any (counting) problem corresponding to coloured walks; a problem is called symmetric iff a = c.

We want to point out that our considerations are by no means limited to the walks on the line. Many combinatorial problems can be proven to be equivalent to some walk problems; bracketing problems are a typical example of this. Moreover, several kinds of walks on the integral bidimensional lattice \mathbb{Z}^2 can be reduced to walks on the line; right, left and sur-place steps are identified with east, north and diagonal steps, respectively, or other types of steps. In particular, positive walks correspond to the very important class of *underdiagonal* walks. A vast amount of literature exists on walk problems, and the reader is referred to Chapter 5 in Goulden and Jackson's book [11], while Barcucci et al. papers [4, 3] are specifically related to the following considerations.

Let us consider $d_{n+1,k+1}$, i.e., the number of coloured walks of length n+1 reaching the distance k+1 from the origin. We can observe that each walk can be obtained in a unique way as: (i) a walk of length n reaching the distance k from the origin, followed by any of the a right steps; (ii) a walk of length n reaching the distance k+1 from the origin, followed by any of the b sur-place steps; (iii) a walk of length n reaching the distance k+2 from the origin, followed by any of the c left steps. Hence we have: $d_{n+1,k+1} = ad_{n,k} + bd_{n,k+1} + cd_{n,k+2}$. This proves that A = (a, b, c) is the A-sequence of $\{d_{n,k}\}$ and therefore $\{d_{n,k}\}$ is a Riordan array and is proper since, $d_{n,k} = 0$ for k > n and $d_{n,n} = a^n$, for every n. This significant fact can be stated as:

Theorem 4.1. Let $d_{n,k}$ be the number of coloured walks of length n reaching a distance k from the origin; then the infinite triangle $\{d_{n,k}\}$ is a proper Riordan array.

The Pascal, Catalan and Motzkin triangles define walking problems that have different values of a, b, c. When c = 0, it is easily proved that $d_{n,k} = \binom{n}{k} a^k b^{n-k}$ and so we end up with the Pascal triangle. Consequently, we assume c > 0. For any given triple (a, b, c) we obtain one type of array from complete walks and another from positive ones. However, the function h(t), that only depends on the A-sequence, is the same in both cases, and we can find it by means of formula (1.5). In fact, $A(t) = a + bt + ct^2$ and h(t) is the solution of the equation $h(t) = a + bth(t) + ct^2 h^2(t)$, with $h(0) \neq 0$:

$$h(t) = \frac{1 - bt - \sqrt{1 - 2bt + b^2 t^2 - 4act^2}}{2ct^2}.$$
(4.1)

The radicand $1-2bt+(b^2-4ac)t^2 = (1-(b+2\sqrt{ac})t)(1-(b-2\sqrt{ac})t)$ is simply denoted by Δ . Obviously, h(t) has two algebraic singularities at $t=(b+2\sqrt{ac})^{-1}$ and $t=(b-2\sqrt{ac})^{-1}$, and since a, b, c are non-negative, the former has a smaller modulus. It is worth noting that this singularity takes on the value $(b+2a)^{-1}$ in the symmetric case.

Let us now focus our attention on positive walks. If we consider $d_{n+1,0}$, we observe that every walk returning to the origin can only be obtained from another walk returning to the origin followed by any sur-place step or a walk ending at distance 1 from the origin followed by any left step. Hence, we have $d_{n+1,0} = bd_{n,0} + cd_{n,1}$ and in the column generating functions this corresponds to d(t) - 1 = btd(t) + ctd(t)h(t).

283

From this relation we easily find d(t) = (1/a)h(t), and therefore by (4.1) the Riordan array of positive coloured walks is:

$$\{d_{n,k}\} = \left(\frac{1-bt-\sqrt{\Delta}}{2act^2}, \frac{1-bt-\sqrt{\Delta}}{2ct^2}\right).$$

$$(4.2)$$

In current literature, major importance is usually given to the following three quantities:

(i) the number of walks returning to the origin; this is $d_n = [t^n]d(t)$, for every n;

(ii) the total number of walks of length *n*; this is $\alpha_n = \sum d_{n,k}$, i.e., the value of the row sums of the Riordan array;

(iii) the average distance from the origin of all the walks of length *n*; this is $\delta_n = \sum k d_{n,k}$, which is the weighted row sums of the Riordan array, divided by α_n .

For (i) we could find a complete asymptotic development for d_n , but we only give the main term here. By using Theorem 4 in Bender [5], we find

$$d_{n} = [t^{n}]d(t) \sim -\left[-\frac{\sqrt{1-(b-2\sqrt{ac})t}}{2ac}\Big|t = \frac{1}{b+2\sqrt{ac}}\right][t^{n+2}]\sqrt{1-(b+2\sqrt{ac})t}$$
$$= \frac{1}{2ac}\sqrt{\frac{4\sqrt{ac}}{b+2\sqrt{ac}}}\binom{1/2}{n+2}(b+2\sqrt{ac})^{n+2}$$
$$\sim \frac{(b+2\sqrt{ac})^{n+1.5}}{\sqrt[4]{(ac)^{3}(2n+3)}\sqrt{\pi(n+2)}}$$

With regard to points (ii) and (iii) above, the formulas for the row sums and the weighted row sums given in Section 1 allow us to find the generating functions $\alpha(t)$ of the total number α_n of positive walks of length *n*, and $\delta(t)$ of the total distance δ_n of these walks from the origin

$$\alpha(t) = \frac{1}{2at} \frac{1 - (b + 2a)t - \sqrt{\Delta}}{(a + b + c)t - 1},$$
(4.3)

$$\delta(t) = \frac{1}{4at} \left(\frac{1 - (b + 2a)t - \sqrt{\Delta}}{(a + b + c)t - 1} \right)^2.$$
(4.4)

The asymptotic value of α_n and δ_n can be evaluated easily. However, we would like to point out an aspect of the dominating singularity of the two generating functions. Besides the algebraic singularities $t = (b + 2\sqrt{ac})^{-1}$ and $t = (b - 2\sqrt{ac})^{-1}$, we now possibly have a simple pole at $t = (a+b+c)^{-1}$. From $(\sqrt{a}-\sqrt{c})^2 \ge 0$, we have $a+c \ge 2\sqrt{ac}$ or $(a+b+c)^{-1} \le (b+2\sqrt{ac})^{-1}$. Surprisingly enough, the numerator of $\delta(t)$ annihilates for $t = (a+b+c)^{-1}$, but this pole dominates in $\alpha(t)$, unless a=c (for more details, see Barcucci et al. [3]). In the symmetric case, formulas (4.3) and (4.4) simplify as follows:

$$\alpha(t) = \frac{1}{2at} \left(\sqrt{\frac{1 - (b - 2a)t}{1 - (b + 2a)t}} - 1 \right),$$

$$\delta(t) = \frac{1}{2at} \left(\frac{1 - bt}{1 - (b + 2a)t} - \sqrt{\frac{1 - (b - 2a)t}{1 - (b + 2a)t}} \right).$$

The alternating row sums and the diagonal sums sometimes have some combinatorial significance as well, and so they can be treated in the same way (see, for example, Barcucci et al. [2]).

The study of complete walks follows the same lines and we only have to derive the form of the corresponding Riordan array, which is

$$\{d_{n,k}\} = \left(\frac{1}{\sqrt{\Delta}}, \frac{1 - bt - \sqrt{\Delta}}{2ct^2}\right). \tag{4.6}$$

The proof is as follows. Since a complete walk can go to the left of the origin, the array $\{d_{n,k} | n, k \in \mathbb{N}, k \leq n\}$ is only the right part of an infinite triangle, in which k can also assume the negative values. By following the logic of the proof of Theorem 4.1 we see that the generating function of the *n*th row is $((c/w) + b + aw)^n$, and therefore, the bivariate generating function of the extended triangle is

$$d(t,w) = \sum \left(\frac{c}{w} + b + aw\right)^n t^n = \frac{1}{1 - (aw + b + c/w)t}.$$

If we expand this expression by partial fractions, we obtain

$$d(t,w) = \frac{1}{\sqrt{\Delta}} \left[\frac{1}{1 - \frac{1 - bt - \sqrt{\Delta}}{2ct}w} - \frac{1}{1 - \frac{1 - bt + \sqrt{\Delta}}{2ct}w} \right]$$

$$=\frac{1}{\sqrt{\Delta}}\left[\frac{1}{1-\frac{1-bt-\sqrt{\Delta}}{2ct}w}+\frac{1-bt-\sqrt{\Delta}}{2at}\frac{1}{w}\frac{1}{1-\frac{1-bt-\sqrt{\Delta}}{2at}\frac{1}{w}}\right].$$

The first term represents the right part of the extended triangle and this corresponds to $k \ge 0$, whereas the second term corresponds to the left part (k < 0). We are interested

in the right part, and the expression can be written as,

$$\frac{1}{\sqrt{\Delta}} \frac{1}{1 - \frac{1 - bt - \sqrt{\Delta}}{2ct} w} = \frac{1}{\sqrt{\Delta}} \sum \left(\frac{1 - bt - \sqrt{\Delta}}{2ct} \right)^k w^k,$$

which by (1.1) is equivalent to (4.6).

5. Stirling numbers

Given any two analytic functions or formal power series d(t) and h(t), we can obtain a Riordan array (d(t), h(t)), and therefore the number of possibilities is infinite. In general, the arbitrary choice of d(t) and h(t) (or A(t)) is limited by combinatorial considerations. In this final section, we describe some simple applications of the Riordan array concept to Stirling numbers. This is particularly important because identities involving Stirling numbers cannot be treated by methods related to hypergeometric functions, such as Gosper's algorithm or WZ-pairs.

Properly, neither kind of the Stirling number triangles (see also Kemeny [13]) constitutes a Riordan array but two well-known related quantities produce proper Riordan arrays. In accordance with Graham et al. [10] and Wilf [24], we have the following generating functions:

$$\mathscr{G}\left\{\frac{m!}{n!}\begin{bmatrix}n\\m\end{bmatrix}\right\} = \left(\log\frac{1}{1-t}\right)^m, \qquad \mathscr{G}\left\{\frac{m!}{n!}\begin{bmatrix}n\\m\end{bmatrix}\right\} = (e^t-1)^m.$$

By definition (1.1), these functions correspond to proper Riordan arrays:

Stirling numbers of the first kind $\left(1, \frac{1}{t}\log\frac{1}{1-t}\right),$ (5.1)

Stirling numbers of the second kind
$$\left(1, \frac{e^{t}-1}{t}\right)$$
. (5.2)

The two fundamental properties of Stirling numbers follow:

$$\sum {n \brack k} {k \atop m} (-1)^{n-k} = (-1)^n \frac{n!}{m!} \sum \frac{k!}{n!} {n \brack k} \frac{m!}{k!} {k \atop m} (-1)^k$$
$$= (-1)^n \frac{n!}{m!} [t^n] \left[(e^{-y} - 1)^m | y = \log \frac{1}{1-t} \right]$$
$$= (-1)^n \frac{n!}{m!} [t^n] (-t)^m = \delta_{nm}$$

and $\sum {\binom{n}{k}} {\binom{k}{m}} (-1)^{n-k} = \delta_{nm}$. An apparently less-known identity on Stirling numbers is

$$\sum \begin{bmatrix} n \\ k \end{bmatrix} \begin{Bmatrix} k \\ m \end{Bmatrix} = \frac{n!}{m!} \sum \frac{k!}{n!} \begin{bmatrix} n \\ k \end{bmatrix} \frac{m!}{k!} \begin{Bmatrix} k \\ m \end{Bmatrix} = \frac{n!}{m!} [t^n] \begin{bmatrix} (e^y - 1)^m | y = \log \frac{1}{1 - t} \end{bmatrix}$$
$$= \frac{n!}{m!} [t^n] \frac{t^m}{(1 - t)^m} = \frac{n!}{m!} [t^{n-m}] \frac{1}{(1 - t)^m} = \frac{n!}{m!} \binom{n-1}{m-1}.$$

As one can see in Table 337 in Graham et al. [10], formulas (5.1) and (5.2) are by no means the only Riordan arrays related to Stirling numbers. Strangely enough, the functions $t/(1-e^{-t})$ and $t/\log(1+t)$ turn out to be the A-sequences of the Riordan arrays (5.1) and (5.2). In fact, if we apply relation (1.5) to the h-function of (5.1), we have

$$\frac{1}{t}\log\frac{1}{1-t} = A\left(\log\frac{1}{1-t}\right).$$

By setting $y = \log(1-t)^{-1}$ or $t = (e^y - 1)/e^y$, we have $A(y) = ye^y/(e^y - 1) = y/(1-e^{-y})$. In a similar way, we start out from $t^{-1}(e^t - 1) = A(e^t - 1)$ and by setting $y = e^t - 1$ or $t = \log(1+y)$, we find $A(y) = y/\log(1+y)$.

Two typical results concern the row sums of the Stirling triangles; by using $\mathscr{G}{1/k!} = e^t$, we have

$$\sum {n \choose k} = n! \sum \frac{k!}{n!} {n \choose k} \frac{1}{k!} = n! [t^n] \left[e^y | y = \log \frac{1}{1-t} \right] = n! [t^n] \frac{1}{1-t} = n!,$$

$$\sum {n \choose k} = n! \sum \frac{k!}{n!} {n \choose k} \frac{1}{k!} = n! [t^n] [e^y | y = e^t - 1] = n! [t^n] \exp(e^t - 1) = \mathscr{B}_n,$$

in which \mathscr{B}_n is the *n*th Bell number, whose exponential generating function is just $\exp(e^t - 1)$. In fact, the Bell numbers \mathscr{B}_n , the ordered Bell numbers \mathscr{O}_n and the Bernoulli numbers B_n are strictly related to the Stirling numbers because of their generating functions, all of which involve the exponential or the logarithmic functions. The ordered Bell numbers \mathscr{O}_n are defined as the total number of ordered partitions of the integers 1 through *n*, and hence $\mathscr{O}_n = \sum {n \atop k} k!$. The exponential generating function of the \mathscr{O}_n is easily found by using the formula for row sums presented in Section 1:

$$\frac{\mathcal{O}_n}{n!} = \sum \frac{k!}{n!} \begin{cases} n \\ k \end{cases} = [t^n] \frac{1}{1 - (e^t - 1)} = [t^n] \frac{1}{2 - e^t}$$

From the generating function we obtain the asymptotic value $\mathcal{O}_n \sim \frac{1}{2} n! L^{n+2}$, in which $L = \log_2 e$ (see Wilf [24]). The following identity is worth observing:

$$\sum {n \brack k} \mathcal{O}_k = n! \sum \frac{k!}{n!} {n \brack k} \frac{\mathcal{O}_k}{k!} = n! [t^n] \left[\frac{1}{2 - e^y} \middle| y = \log \frac{1}{1 - t} \right]$$

R. Sprugnoli / Discrete Mathematics 132 (1994) 267-290

$$= n! [t^{n}] \frac{1}{2 - 1/(1 - t)} = n! [t^{n}] \frac{1 - t}{1 - 2t} = n! (2^{n} - 2^{n-1}) = n! 2^{n-1}.$$

For $\sum {n \brack k} \mathscr{B}_k$ we cannot find a closed form, but we have

$$\sum \begin{bmatrix} n \\ k \end{bmatrix} \mathscr{B}_k = n! \sum \frac{k!}{n!} \begin{bmatrix} n \\ k \end{bmatrix} \frac{\mathscr{B}_k}{k!} = n! [t^n] \left[\exp(e^y - 1) | y = \log \frac{1}{1 - t} \right]$$
$$= n! [t^n] \exp \frac{t}{1 - t} \stackrel{(*)}{=} n! \sum \binom{n - 1}{k - 1} \frac{1}{k!} = \sum \binom{n - 1}{k - 1} \frac{n!}{k!}.$$

Note that the first expression in the second line is thought to be the result of applying (1.3) both to e^t and the Riordan array $\{d_{n,k}\} = (1, 1/(1-t))$. The value of $d_{n,k}$ is found by means of (1.1):

$$d_{n,k} = [t^n] \frac{t^k}{(1-t)^k} = [t^{n-k}] \frac{1}{(1-t)^k} = \binom{-k}{n-k} (-1)^{n-k} = \binom{n-1}{k-1};$$

from this equality (*) follows immediately. It is possible to derive the asymptotic value (see Bender [5] or Flajolet et al. [8]) from the generating function $\exp(t/(1-t))$:

$$\sum \begin{bmatrix} n \\ k \end{bmatrix} \mathscr{B}_k = \sum \binom{n-1}{k-1} \frac{n!}{k!} \sim n! \frac{e^{2\sqrt{n}}}{2\sqrt{\pi e n^{3/4}}}.$$

A simple property is

$$\sum {\binom{n}{k}} \frac{(-1)^{k}k!}{k+1} = n! \sum \frac{k!}{n!} {\binom{n}{k}} \frac{(-1)^{k}}{k+1} = n! [t^{n}] \left[-\frac{1}{y} \log \frac{1}{1+y} \middle| y = e^{t} - 1 \right]$$
$$= n! [t^{n}] \frac{t}{e^{t} - 1} = B_{n}.$$

A more complex result is

$$\sum {m \atop k} {k+1 \atop p} \frac{(-1)^{k+1-p}}{k+1} = (-1)^p \frac{m!}{p!} \sum \frac{k!}{m!} {m \atop k} \frac{p!}{(k+1)!} {k+1 \atop p} (-1)^{k+1}$$
$$= (-1)^p \frac{m!}{p!} [t^m] \left[\frac{1}{y} \left(\log \frac{1}{1+y} \right)^p \middle| y = e^t - 1 \right]$$
$$= (-1)^p \frac{m!}{p!} [t^m] \frac{(-t)^p}{e^t - 1} = \frac{m!}{p!} [t^{m-p+1}] \frac{t}{e^t - 1}$$
$$= \frac{m!}{p!} \frac{B_{m-p+1}}{(m-p+1)!} = \frac{1}{m+1} {m+1 \choose p} B_{m-p+1}$$

Knuth [15, III] calls the following identity "curious":

$$\sum {\binom{n}{k}} k! (-2)^{-k} = \frac{2B_{n+1}(1-2^{n+1})}{(n+1)},$$

i.e.,

$$\sum {\binom{n}{k}} k! \left(-\frac{1}{2}\right)^{k} = n! \sum \frac{k!}{n!} {\binom{n}{k}} \left(-\frac{1}{2}\right)^{k} = n! [t^{n}] \left[\frac{1}{1+y/2}\right] y = e^{t} - 1 \\ = 2n! [t^{n}] \frac{1}{e^{t}+1} = 2n! [t^{n}] \left(\frac{1}{e^{t}-1} - \frac{1}{e^{2t}-1}\right) \\ = 2n! \frac{B_{n+1} - 2^{n+1} B_{n+1}}{(n+1)!}.$$

On the contrary, we believe that it is not so curious from the point of view of Riordan arrays.

Another identity involving the Bernoulli numbers is

$$\sum {n \brack k} B_k = n! \sum \frac{k!}{n!} {n \brack k} \frac{B_k}{k!} = n! [t^n] \left[\frac{y}{e^y - 1} \middle| y = \log \frac{1}{1 - t} \right]$$
$$= n! [t^n] \frac{1 - t}{t} \log \frac{1}{1 - t} = n! [t^{n+1}] (1 - t) \log \frac{1}{1 - t}$$
$$= n! \left(\frac{1}{n+1} - \frac{1}{n} \right) = -\frac{(n-1)!}{n+1}.$$

The Stirling numbers of the second kind have entirely different implications:

$$S_{n} = \sum {\binom{n}{k}} B_{k} = n! \sum \frac{k!}{n!} {\binom{n}{k}} \frac{B_{k}}{k!} = n! [t^{n}] \left[\frac{y}{e^{y} - 1} \middle| y = e^{t} - 1 \right]$$
$$= n! [t^{n}] \frac{e^{t} - 1}{\exp(e^{t} - 1) - 1}.$$

This function's dominating singularities are simple poles at $t=\frac{1}{2}\log(1+4\pi^2)\pm i \arctan(2\pi)$. Salvy [21] computed the asymptotic value of the sum S_n using $\Lambda_{\Upsilon}\Omega$ (see Flajolet et al. [8]):

$$S_{n} \sim \frac{-4\pi}{\sqrt{1+4\pi^{2}}} (\log^{2} \sqrt{1+4\pi^{2}} + \operatorname{arctg}^{2}(2\pi))^{-(n+1)/2}$$
$$\times \sin\left((n+1)\operatorname{arctg}\left(\frac{2\operatorname{arctg}(2\pi)}{\log(1+4\pi^{2})}\right) + \operatorname{arctg}(2\pi)\right)$$
$$+ O((\log^{2} \sqrt{1+16\pi^{2}} + \operatorname{arctg}^{2}(4\pi))^{n/2}).$$

The relation defining the ordered Bell numbers can be amplified in various ways. For example,

$$\sum {\binom{n}{k}} (k+1)! = n! \sum \frac{k!}{n!} {\binom{n}{k}} (k+1) = n! [t^n] \left[\frac{1}{(1-y)^2} \middle| y = e^t - 1 \right]$$
$$= n! [t^n] \left(\frac{1}{2-e^t} \right)^2 = n! \sum \frac{\mathcal{O}_k}{k!} \frac{\mathcal{O}_{n-k}}{(n-k)!} = \sum {\binom{n}{k}} \mathcal{O}_k \mathcal{O}_{n-k} = \hat{b}_n$$

The numbers \hat{b}_n are related to the sum of binomial coefficient inverses; Comtet [6, p. 294] gives the values of \hat{b}_n up to n = 10, but the last three values are wrong. Since the generating function $(2 - e^t)^{-2}$ has a double pole at $t = \log 2$, we can easily find the asymptotic value:

$$\hat{b}_n \sim \frac{1}{4} n! (n+1+\log 2) L^{n+2} \quad (L=\log_2 e)$$

In a similar way we find

$$\hat{c}_n = \sum \left\{ {n \atop k} (k-1)! = n! [t^n] \left[\log \frac{1}{1-y} \middle| y = e^t - 1 \right] = n! [t^n] \log \frac{1}{2-e^t}.$$

By using the rule $[t^n]f(t) = (1/n)[t^{n-1}]f'(t)$ and observing that $e^t/(2-e^t)$ has a simple pole at $t = \log 2$, we obtain

$$\hat{c}_n = (n-1)! [t^{n-1}] \frac{e^t}{2-e^t} \sim (n-1)! L^n \quad (L = \log_2 e).$$

References

- [1] G. Andrews, Applications of basic hypergeometric functions, SIAM Rev. 16 (1974) 441-484.
- [2] E. Barcucci, R. Pinzani and R. Sprugnoli, The Motzkin family, Pure Math. Appl. A 2 (1991) 249-279.
- [3] E. Barcucci, R. Pinzani and R. Sprugnoli, The random generation of underdiagonal walks, in: 4th Congress on Formal Power Series and Algebraic Combinatorics, Montreal, 1992.
- [4] E. Barcucci, R. Pinzani and R. Sprugnoli, The random generation of directed animals, Theoret. Comput. Sci., to appear.
- [5] E.A. Bender, Asymptotic methods in enumeration, SIAM Rev. 16 (1974) 485-515.
- [6] L. Comtet, Advanced Combinatorics (Reidel, Dordrecht, 1974).
- [7] G.P. Egorychev, Integral Representation and the Computation of Combinatorial Sums, Amer. Math. Soc. Translations, Vol. 59 (Amer. Math. Soc., Providence, RI, 1984).
- [8] Ph. Flajolet, B. Salvy and P. Zimmermann, Automatic average-case analysis of algorithms. Theoret. Comput. Sci. 79 (1991), 37-109.
- [9] R.W. Gosper, Jr., Decision procedure for indefinite hypergeometric summation, Proc. Nat. Acad. Sci. USA, 75 (1978) 40-42.
- [10] R.L. Graham, D.E. Knuth and O. Patashnik, Concrete Mathematics (Addison-Wesley, Reading, MA, 1988).
- [11] I.P. Goulden and D.M. Jackson, Combinatorial Enumeration (Wiley, New York, 1983).
- [12] M. Karr, Summation in finite terms, J. ACM 28 (1981) 305-350.
- [13] J.G. Kemeny, Matrix representation for combinatorics, J. Combin. Theory 36 (1984) 279-306.

- [14] S.G. Kettle, Families enumerated by the Schröder-Etherington sequence and a renewal array it generates, in: Combinatorial Mathematics X, Lecture Notes in Math. 1036 (Springer, Berlin) 244-274.
- [15] D.E. Knuth, The Art of Computer Programming, Vol. I-III (Addison-Wesley, Reading, MA, 1968-1973).
- [16] Roy Ranjan, Binomial identities and hypergeometric series, Amer. Math. Monthly 97 (1987) 36-46.
- [17] J. Riordan, Combinatorial Identities (Wiley, New York, 1968).
- [18] D.G. Rogers, Pascal triangles, Catalan numbers and renewal arrays. Discrete Math. 22 (1978) 301-310.
- [19] S. Roman, The Umbral Calculus (Academic Press, New York, 1984).
- [20] F. Ruskey, On the average shape of binary trees, SIAM J. Algebraic Discrete Methods I 1 (1980) 43-49.
- [21] B. Salvy, Private communication, 1992.
- [22] L.V. Shapiro, S. Getu, W.-J. Woan and L. Woodson, The Riordan Group. Discrete Appl. Math. 34 (1991) 229–239.
- [23] H.S. Wilf, The "Snake Oil" method for proving combinatorial identities, Surveys in Combinatorics, 1989 (Cambridge University Press, Cambridge, 1990) 208-217.
- [24] H.S. Wilf, Generating functionology (Academic Press, New York, 1990).
- [25] H.S. Wilf and D. Zeilberger, Rational functions certify combinatorial identities. J. AMS 3 (1990) 147-158.
- [26] D. Zeilberger, A fast algorithm for proving terminating hypergeometric identities, Discrete Math. 80 (1990) 207-211.