

# LAGRANGE INVERSION

IRA M. GESSEL\*

ABSTRACT. We give a survey of the Lagrange inversion formula, including different versions and proofs, with applications to combinatorial and formal power series identities.

## 1. INTRODUCTION

The Lagrange inversion formula is one of the fundamental formulas of combinatorics. In its simplest form it gives a formula for the power series coefficients of the solution  $f(x)$  of the function equation  $f(x) = xG(f(x))$  in terms of coefficients of powers of  $G$ . Functional equations of this form often arise in combinatorics, and our interest is in these applications rather than in other areas of mathematics.

There are many generalizations of Lagrange inversion: multivariable forms [28],  $q$ -analogues [22, 23, 25, 71] noncommutative versions [6, 7, 23, 56] and others [29, 43, 45]. In this paper we discuss only ordinary one-variable Lagrange inversion, but in greater detail than elsewhere in the literature.

In section 2 we give a thorough discussion of some of the many different forms of Lagrange inversion, prove that they are equivalent to each other, and work through some simple examples involving Catalan and ballot numbers. We address a number of subtle issues that are overlooked in most accounts of Lagrange inversion (and which some readers may want to skip). In sections 3 we describe applications of Lagrange inversion to identities involving binomial coefficients, Catalan numbers, and their generalizations. In section 4, we give several proofs of Lagrange inversion, some of which are combinatorial.

A number of exercises giving additional results are included.

An excellent introduction to Lagrange inversion can be found in Chapter 5 of Stanley's *Enumerative Combinatorics*, Volume 2. Other expository accounts of Lagrange inversion can be found in Hofbauer [35], Bergeron, Labelle, and Leroux [5, Chapter 3], Sokal [68], and Merlini, Sprugnoli, and Verri [51].

**1.1. Formal power series.** Although Lagrange inversion is often presented as a theorem of analysis (see, e.g., Whittaker and Watson [76, pp. 132–133]), we will work only with formal power series and formal Laurent series. A good account of formal power series can be found in Niven [55]; we sketch here some of the basic facts. Given a coefficient ring  $C$ , which for us will always be an integral domain containing the rational numbers, the ring  $C[[x]]$  of *formal power series* in the variable  $x$  with coefficients in  $C$  is the set of all “formal sums”  $\sum_{n=0}^{\infty} c_n x^n$ , where  $c_n \in C$ , with termwise addition and multiplication defined as one would expect using distributivity:  $\sum_{n=0}^{\infty} a_n x^n \cdot \sum_{n=0}^{\infty} b_n x^n = \sum_{n=0}^{\infty} c_n x^n$ , where  $c_n = \sum_{i=0}^n a_i b_{n-i}$ . Differentiation of formal power series is also defined termwise. A series  $\sum_{n=0}^{\infty} c_n x^n$  has a multiplicative inverse if and only if  $c_0$  is invertible in  $C$ . We may also consider the ring of

---

*Date:* June 6, 2016.

*formal Laurent series*  $C((x))$  whose elements are formal sums  $\sum_{n=n_0}^{\infty} c_n x^n$  for some integer  $n_0$ , i.e., formal sums  $\sum_{n=-\infty}^{\infty} c_n x^n$  in which only finitely many negative powers of  $x$  have nonzero coefficients. Henceforth will omit the word “formal” and speak of power series and Laurent series.

We can iterate the power series and Laurent series ring constructions, obtaining, for example the ring  $C((x))[[y]]$  of power series in  $y$  whose coefficients are Laurent series in  $x$ . In any (possibly iterated) power series or Laurent series ring we will say that a set  $\{f_\alpha\}$  of series is *summable* if for any monomial  $m$  in the variables, the coefficient of  $m$  is nonzero in only finitely many  $f_\alpha$ . In this case the sum  $\sum_\alpha f_\alpha$  is well-defined and we will say that  $\sum_\alpha f_\alpha$  is summable. If we write  $\sum_\alpha f_\alpha$  as an iterated sum, then the order of summation is irrelevant. If  $f(x) = \sum_n c_n x^n$  is a Laurent series in  $C((x))$  and  $u \in C$ , where  $C$  may be a power series or Laurent series ring, then we say that the substitution of  $u$  for  $x$  is *admissible* if  $f(u) = \sum_n c_n u^n$  is summable, and similarly for multivariable substitutions. Admissible substitutions are homomorphisms. If  $u$  is a power series or Laurent series  $g(x)$  then  $f(g(x))$ , if summable, is called the *composition* of  $f$  and  $g$ . If  $f(x) = c_1 x + c_2 x^2 + \dots$ , where  $c_1$  is invertible in  $C$ , then there is a unique power series  $g(x) = c_1^{-1} x + \dots$  such that  $f(g(x)) = x$ ; this implies that  $g(f(x)) = x$ . We call  $g(x)$  the *compositional inverse* of  $f(x)$  and write  $g(x) = f(x)^{(-1)}$ . For simplicity, we will always assume that if  $f(x) = c_1 x + c_2 x^2 + \dots$ , where  $c_1 \neq 0$  then  $c_1$  is invertible. (Since  $C$  is an integral domain, we can always adjoin  $c_1^{-1}$  to  $C$  if necessary.)

The iterated power series rings  $C[[x]][[y]]$  and  $C[[y]][[x]]$  are essentially the same, in that both consist of all sums  $\sum_{m,n \geq 0} c_{m,n} x^m y^n$ . We may therefore write  $C[[x, y]]$  for either of these rings. However, the iterated Laurent series  $C((x))[[y]]$ ,  $C((y))[[x]]$ , and  $C[[y]]((x))$  are all different: in the first we have  $(x - y)^{-1} = \sum_{n=0}^{\infty} y^n / x^{n+1}$ , in the second we have  $(x - y)^{-1} = -\sum_{n=0}^{\infty} x^n / y^{n+1}$ , and in the third  $x - y$  is not invertible.

It is sometimes convenient to work with power series in infinitely many variables; for example, we may consider the power series  $\sum_{n=0}^{\infty} r_n t^n$  where the  $r_n$  are independent indeterminates. Although we don't give a formal definition of these series, they behave, in our applications, exactly as expected.

We use the notation  $[x^n] f(x)$  to denote the coefficient of  $x^n$  in the Laurent series  $f(x)$ . An important fact about the coefficient operator that we will use often, without comment, is that  $[x^n] x^k f(x) = [x^{n-k}] f(x)$ .

The binomial coefficient  $\binom{a}{k}$  is defined to be  $a(a-1)\dots(a-k+1)/k!$  if  $k$  is a nonnegative integer and 0 otherwise. Thus the binomial theorem  $(1+x)^a = \sum_{k=0}^{\infty} \binom{a}{k} x^k$  holds for all  $a$ .

## 2. THE LAGRANGE INVERSION FORMULA

**2.1. Forms of Lagrange inversion.** We will give several proofs of the Lagrange inversion formula in section 4. Here we state several different forms of Lagrange inversion and show that they are equivalent.

**Theorem 2.1.1.** *Let  $R(t)$  be a power series not involving  $x$ . Then there is a unique power series  $f = f(x)$  such that  $f(x) = xR(f(x))$ , and for any Laurent series  $\phi(t)$  and  $\psi(t)$  not involving  $x$  and any integer  $n$  we have*

$$[x^n] \phi(f) = \frac{1}{n} [t^{n-1}] \phi'(t) R(t)^n, \text{ where } n \neq 0, \quad (2.1.1)$$

$$[x^n] \phi(f) = [t^n] (1 - tR'(t)/R(t)) \phi(t) R(t)^n, \quad (2.1.2)$$

$$\phi(f) = \sum_n x^n [t^n] (1 - xR'(t)) \phi(t) R(t)^n, \quad (2.1.3)$$

$$[x^n] \frac{\psi(f)}{1 - xR'(f)} = [t^n] \psi(t) R(t)^n, \quad (2.1.4)$$

$$[x^n] \frac{\psi(f)}{1 - fR'(f)/R(f)} = [t^n] \psi(t) R(t)^n. \quad (2.1.5)$$

We show here these formulas are equivalent in the sense that any one of them is easily derivable from any other; proofs of these formulas are given in section 4. It is clear that (2.1.4) and (2.1.5) are equivalent since  $x = f/R(f)$ . Taking  $\psi(t) = (1 - tR'(t)/R(t))\phi(t)$  shows that (2.1.2) and (2.1.5) are equivalent.

To derive (2.1.3) from (2.1.4), we rewrite (2.1.4) as

$$\frac{\psi(f)}{1 - xR'(f)} = \sum_n x^n [t^n] \psi(t) R(t)^n. \quad (2.1.6)$$

Until now, we have assumed that  $\phi(t)$  and  $\psi(t)$  do not involve  $x$ . We leave it to the reader to see that in (2.1.6) this assumption can be removed. Then (2.1.3) follows from (2.1.6) by setting  $\psi(t) = (1 - xR'(t))\phi(t)$ , and similarly (2.1.6) follows from (2.1.3).

Although we allow  $R$  to be an arbitrary power series in Theorem 2.1.1, if  $R$  has constant term 0 then  $f(x) = 0$ , so we may assume now that  $R$  has a nonzero constant term, and thus  $f$  and  $R(f)$  are nonzero. Then the equation  $f(x) = xR(f(x))$  may be rewritten as  $f/R(f) = x$ . So if we set  $g(t) = t/R(t)$  then we have  $g(f) = x$ , and thus  $g = f^{(-1)}$ . It is sometimes convenient to rewrite the formulas of Theorem 2.1.1 using  $g$ , rather than  $R$ . Since  $1 - tR'(t)/R(t) = tg'(t)/g(t)$ , formula (2.1.2) takes on a slightly simpler form (which will be useful later on) when expressed in terms of  $g$  rather than  $R$ :

$$[x^n] \phi(f) = [t^{n-1}] \phi(t) \frac{g'(t)}{g(t)} \left( \frac{t}{g(t)} \right)^n = [t^{-1}] \frac{\phi(t)g'(t)}{g(t)^{n+1}}, \quad (2.1.7)$$

For future use, we note also the corresponding form for (2.1.5):

$$[x^{n-1}] \frac{\psi(f)}{fg'(f)} = [t^n] \psi(t) \left( \frac{t}{g(t)} \right)^n = [t^0] \frac{\psi(t)}{g(t)^n}. \quad (2.1.8)$$

To show that (2.1.1) and (2.1.2) are equivalent, using (2.1.7) in place of (2.1.2), we show that

$$[t^{-1}] \frac{\phi'(t)}{g(t)^n} = n [t^{-1}] \frac{\phi(t)g'(t)}{g(t)^{n+1}} \quad (2.1.9)$$

But

$$\frac{\phi'(t)}{g(t)^n} - n \frac{\phi(t)g'(t)}{g(t)^{n+1}} = \frac{d}{dt} \frac{\phi(t)}{g(t)^n},$$

and the coefficient of  $t^{-1}$  in the derivative of any Laurent series is 0, so (2.1.9) follows. This shows that (2.1.1) and (2.1.2) are equivalent if  $n \neq 0$ . If  $\phi$  is a power series, then the coefficient of  $x^0$  in  $\phi(f)$  is simply the constant term in  $\phi$ , but if  $\phi$  is a more general Laurent series, then the constant term in  $\phi(f)$  is not so obvious, and cannot be determined by (2.1.1). In equation (2.2.8) we will give a formula for the constant term in  $\phi(f)$  for all  $\phi$ .

The case  $\phi(t) = t^k$  of (2.1.1), with  $R(t) = t/g(t)$  may be written

$$[x^n] f^k = \frac{k}{n} [t^{n-k}] \left( \frac{t}{g(t)} \right)^n = \frac{k}{n} [t^{-k}] g(t)^{-n}. \quad (2.1.10)$$

In other words, if  $g = f^{(-1)}$ , and for all integers  $k$ ,  $f^k = \sum_n a_{n,k} x^n$  and  $g^k = \sum_n b_{n,k} x^n$  then

$$a_{n,k} = \frac{k}{n} b_{-k,-n} \quad (2.1.11)$$

for  $n \neq 0$ . Equation (2.1.11) is known as the *Schur-Jabotinsky theorem*. (See Schur [66, equation (10)] and Jabotinsky [37, Theorem II].)

**2.2. Polynomials.** We now give a slightly more general form of Lagrange inversion based on the fact that if two polynomials agree at infinitely many values than they are identically equal. This will imply that to prove our Lagrange formulas for all  $n$ , it is sufficient to prove them in the case in which  $n$  is a positive integer. (Some proofs require this restriction.)

By linearity, the formulas of Section 2.1 are implied by the special cases in which  $\phi(t)$  and  $\psi(t)$  are of the form  $t^k$  for some integer  $k$ , and these special cases (especially  $k = 0$  and  $k = 1$ ) are particularly important. These special cases of (2.1.1), (2.1.4), and (2.1.5) are especially useful and may be written

$$[x^n] f^k = \frac{k}{n} [t^{n-k}] R(t)^n, \text{ where } n \neq 0, \quad (2.2.1)$$

$$[x^n] \frac{f^k}{1 - xR'(f)} = [t^{n-k}] R(t)^n. \quad (2.2.2)$$

$$[x^n] \frac{f^k}{1 - fR'(f)/R(f)} = [t^{n-k}] R(t)^n. \quad (2.2.3)$$

In these formulas, let us assume that  $R(t)$  has constant term 1. (It is not hard to modify our approach to take care of the more general situation in which the constant term of  $R(t)$  is invertible.) Then the coefficient of  $x$  in  $f(x)$  is 1, so  $f(x)/x$  has constant term 1. If we set  $n = m + k$  in (2.2.1), (2.2.2), and (2.2.3) then the results may be written

$$[x^m](f/x)^k = \frac{k}{m+k} [t^m] R(t)^{m+k}, \text{ where } m+k \neq 0, \quad (2.2.4)$$

$$[x^m] \frac{(f/x)^k}{1 - xR'(f)} = [t^m] R(t)^{m+k} \quad (2.2.5)$$

$$[x^m] \frac{(f/x)^k}{1 - fR'(f)/R(f)} = [t^m] R(t)^{m+k}. \quad (2.2.6)$$

It is easy to see that in each of these equations, for fixed  $m$  both sides are polynomials in  $k$ . Thus if these equalities hold whenever  $k$  is a positive integer, then they hold as identities of polynomials in  $k$ . Moreover, although (2.2.4) is invalid for  $k = -m$ , if  $m > 0$  we may take the limit as  $k \rightarrow -m$  with l'Hôpital's rule to obtain

$$[x^m](f/x)^{-m} = [x^0] f^{-m} = -m [t^m] \log R. \quad (2.2.7)$$

(Note that (2.2.7) does not hold for  $m = 0$ .) By linearity, (2.2.7) yields a supplement to (2.1.1) that takes care of the case  $n = 0$ :

$$[x^0] \phi(f) = [t^0] \phi(t) + [t^{-1}] \phi'(t) \log R. \quad (2.2.8)$$

We can also differentiate (2.2.4) with respect to  $k$  and then set  $k = 0$  to obtain

$$[x^m] \log(f/x) = \frac{1}{m} [t^m] R(t)^m, \text{ for } m \neq 0. \quad (2.2.9)$$

Returning to (2.2.1)–(2.2.3), we see that if they hold when  $n$  and  $k$  are positive integers, then they also hold when  $n$  and  $k$  are arbitrary integers. (Note that if  $n < k$  then everything is zero.)

**2.3. A simple example: Catalan numbers.** The Catalan numbers  $C_n$  may be defined by the equation

$$c(x) = 1 + xc(x)^2 \quad (2.3.1)$$

for their generating function  $c(x) = \sum_{n=0}^{\infty} C_n x^n$ . The quadratic equation (2.3.1) has two solutions,  $(1 \pm \sqrt{1-4x})/(2x)$ , but only the minus sign gives a power series, so

$$c(x) = \frac{1 - \sqrt{1-4x}}{2x}.$$

Unfortunately (2.3.1) is not of the form  $f(x) = xR(f(x))$ , so we cannot apply directly any of the versions of Lagrange inversion that we have seen so far.

One way to apply Lagrange inversion is to set  $f(x) = c(x) - 1$ , so that  $f = x(1+f)^2$ . We may then apply Theorem 2.1.1 to the case  $R(t) = (1+t)^2$ . The equation  $f = x(1+f)^2$  has the solution

$$f(x) = c(x) - 1 = xc(x)^2 = \frac{1 - \sqrt{1-4x}}{2x} - 1.$$

Then (2.1.1) with  $\phi(t) = (1+t)^k$  gives for  $n > 0$

$$\begin{aligned} [x^n] c(x)^k &= [x^n] (1+f)^k = \frac{1}{n} [t^{n-1}] k(1+t)^{k-1}(1+t)^{2n} \\ &= \frac{k}{n} [t^{n-1}] \binom{2n+k-1}{n-1}. \end{aligned}$$

Thus since the constant term in  $c(x)^k$  is 1, we have

$$c(x)^k = 1 + \sum_{n=1}^{\infty} \frac{k}{n} \binom{2n+k-1}{n-1} x^n.$$

The sum may also be written

$$c(x)^k = \sum_{n=0}^{\infty} \frac{k}{2n+k} \binom{2n+k}{n} x^n = \sum_{n=0}^{\infty} \frac{k}{n+k} \binom{2n+k-1}{n} x^n. \quad (2.3.2)$$

These formulas are valid for all  $k$  except where  $n = -k/2$  in the first sum in (2.3.2) or  $n = -k$  in the second sum in (2.3.2). These coefficients are called *ballot numbers*, and for  $k = 1$  (2.3.2) gives the usual formula for the Catalan numbers,  $C_n = \frac{1}{n+1} \binom{2n}{n}$ .

Equation (2.1.2) with  $\phi(t) = (1+t)^k$  gives a formula for the ballot number as a difference of two binomial coefficients,

$$\begin{aligned} [x^n] c(x)^k &= [t^n] \frac{1-t}{1+t} (1+t)^k (1+t)^{2n} \\ &= [t^n] (1-t)(1+t)^{2n+k-1} \\ &= \binom{2n+k-1}{n} - \binom{2n+k-1}{n-1}, \end{aligned}$$

and equation (2.1.3) with  $\phi(t) = (1+t)^k$  gives another such formula,

$$\begin{aligned} [x^n] c(x)^k &= \sum_n x^n [t^n] ((1+t)^{2n+k} - 2x(1+t)^{2n+k+1}) \\ &= \sum_n x^n \left[ \binom{2n+k}{n} - 2 \binom{2n+k-1}{n-1} \right]. \end{aligned}$$

Finally, (2.2.9) gives

$$[x^m] \log(f/x) = [x^m] 2 \log c(x) = \frac{1}{m} [t^m] (1+t)^{2m} = \frac{1}{m} \binom{2m}{m},$$

so

$$\log c(x) = \sum_{m=1}^{\infty} \frac{1}{2m} \binom{2m}{m} x^m.$$

Since  $R(t) = (1+t)^2$ , we have  $R'(t) = 2(1+t)$ , so  $1 - xR'(f) = 1 - 2xc(x) = \sqrt{1-4x}$ . Thus (2.1.4), with  $\psi(t) = (1+t)^k$  gives

$$[x^n] \frac{c(x)^k}{\sqrt{1-4x}} = [t^n] (1+t)^k (1+t)^{2n} = \binom{2n+k}{k},$$

so

$$\frac{c(x)^k}{\sqrt{1-4x}} = \sum_{n=0}^{\infty} \binom{2n+k}{k} x^n. \quad (2.3.3)$$

Equating coefficients of  $x^n$  in  $c(x)^k c(x)^l = c(x)^{k+l}$ , and using (2.3.2), gives the convolution identity

$$\sum_{i+j=n} \frac{k}{2i+k} \binom{2i+k}{i} \cdot \frac{l}{2j+l} \binom{2j+l}{j} = \frac{k+l}{2n+k+l} \binom{2n+k+l}{n}.$$

Similarly using (2.3.2) and (2.3.3) we get

$$\sum_{i+j=n} \frac{k}{2i+k} \binom{2i+k}{i} \binom{2j+l}{j} = \binom{2n+k+l}{n}.$$

These convolution identities are special cases of identities discussed in section 3.3.

**Exercise 2.3.1.** Derive these formulas for  $c(x)$  in other ways by applying Lagrange inversion to the equations  $f = x/(1-f)$  and  $f = x(1+f^2)$ .

**2.4. A generalization.** There is another way to apply Lagrange inversion to the equation  $c(x) = 1 + xc(x)^2$  that, while very simple, has far-reaching consequences. Consider the equation

$$F = z(1 + xF^2) \quad (2.4.1)$$

where we think of  $F$  as a power series in  $z$  with coefficients that are polynomials in  $x$ . We may apply (2.2.1) to (2.4.1) to get

$$[z^n] F^k = \frac{k}{n} [t^{n-k}] (1 + xt^2)^n.$$

The right side is 0 unless  $n - k$  is even, and for  $n = 2m + k$  we have

$$[z^{2m+k}] F^k = \frac{k}{2m+k} [t^{2m}] (1 + xt^2)^{2m+k} = \frac{k}{2m+k} \binom{2m+k}{m} x^m.$$

Thus we have (if  $k$  is an integer but not a negative even integer)

$$F^k = \sum_{m=0}^{\infty} \frac{k}{2m+k} \binom{2m+k}{m} x^m z^{2m+k}. \quad (2.4.2)$$

Now let  $c$  be the result of setting  $z = 1$  in  $F$  (an admissible substitution), so by (2.4.2), we have

$$c^k = \sum_{m=0}^{\infty} \frac{k}{2m+k} \binom{2m+k}{m} x^m$$

and by (2.4.1) we have  $c = 1 + xc^2$ . Moreover, as we have seen before,  $c = 1 + xc^2$  has a unique power series solution.

The same idea works much more generally, but we must take care that the substitution is admissible. For example, we can solve  $f = x(1 + f)$  by Lagrange inversion, but we cannot set  $x = 1$  in the solution.

The case in which the coefficients of  $R(t)$  are indeterminates is easy to deal with.

**Theorem 2.4.1.** *Suppose that  $R(t) = \sum_{n=0}^{\infty} r_n t^n$ , where the  $r_n$  are indeterminates. Then there is a unique power series  $f$  satisfying  $f = R(f)$ . If  $\phi(t)$  is a power series then*

$$\phi(f) = \phi(0) + \sum_{n=1}^{\infty} \frac{1}{n} [t^{n-1}] \phi'(t) R(t)^n, \quad (2.4.3)$$

and for any Laurent series  $\phi(t)$  and  $\psi(t)$  we have

$$\phi(f) = [t^0] \phi(t) + [t^{-1}] \phi'(t) \log(R/r_0) + \sum_{n \neq 0} \frac{1}{n} [t^{n-1}] \phi'(t) R(t)^n, \quad (2.4.4)$$

$$\phi(f) = \sum_n [t^n] (1 - tR'(t)/R(t)) \phi(t) R(t)^n = \sum_n [t^n] (1 - R'(t)) \phi(t) R(t)^n, \quad (2.4.5)$$

$$\frac{\psi(f)}{1 - R'(f)} = \frac{\psi(f)}{1 - fR'(f)/R(f)} = \sum_n [t^n] \psi(t) R(t)^n. \quad (2.4.6)$$

In (2.4.5) and (2.4.6) the sum is over all integers  $n$ .

*Proof.* These formulas follow from equations (2.1.1) to (2.1.5) on making the admissible substitution  $x = 1$ , where for (2.4.4) we have included the correction term given by (2.2.8), modified to take into account that the constant term of  $R(t)$  is  $r_0$  rather than 1. Uniqueness follows by equating coefficients of the monomials  $r_0^{i_0} r_1^{i_1} \cdots$  on both sides of  $f = R(f)$ , which gives a recurrence that determines them uniquely.  $\square$

We would like to relax the requirement in Theorem 2.4.1 that the  $r_n$  be indeterminates. To do this, we can take any of the formulas of Theorem 2.4.1 and apply any admissible substitution for the  $r_n$ . For example, the following result, while not the most general possible, is sometimes useful.

**Theorem 2.4.2.** *Suppose that  $R(t) = \sum_{n=0}^{\infty} r_n t^n$ , where the coefficients lie in some power series ring  $C[[u_1, u_2, \dots]]$ , and that each  $r_n$  with  $n > 0$  is divisible by some  $u_i$ . Then there is a unique power series  $f$  satisfying  $f = R(f)$ , and formulas (2.4.3) to (2.4.6) hold.  $\square$*

We note that more generally, any admissible substitution will yield a solution of  $f = R(f)$  to which these formulas hold, but uniqueness is not guaranteed. For example, the equation  $f = x + yf^2$  has the unique power series solution,

$$f = \frac{1 - \sqrt{1 - 4xy}}{2y} = \sum_{n=0}^{\infty} \frac{1}{n+1} \binom{2n}{n} x^{n+1} y^n.$$

The admissible substitution  $y = 1$  gives  $f = \frac{1}{2}(1 - \sqrt{1 - 4x})$  as a power series solution of the equation  $f = x + f^2$ . However, the equation  $f = x + f^2$  has another power series solution,  $f = \frac{1}{2}(1 + \sqrt{1 - 4x})$ .

**Exercise 2.4.3.** The equation  $f = x + f^2$  has two power series solutions,  $f = \frac{1}{2}(1 \pm \sqrt{1 - 4x})$ . However, according to Theorem 2.1.1, the equivalent equation  $f = x/(1 - f)$  has only one power series solution. Explain the discrepancy.

**2.5. Explicit formulas for the coefficients.** It is sometimes useful to have an explicit formula for the coefficients of  $f^k$  where  $f = xR(f)$ . With  $R(t) = \sum_{n=0}^{\infty} r_n t^n$ , if we expand  $R(t)^n$  by the multinomial theorem then (2.2.1) gives

$$f^k = \sum_{\substack{n_0+n_1+\dots=n \\ n_1+2n_2+3n_3+\dots=n-k}} k \frac{(n-1)!}{n_0! n_1! \dots} r_0^{n_0} r_1^{n_1} \cdots x^n. \quad (2.5.1)$$

We might also want to express the coefficients of  $f^k$  in terms of the coefficients of  $g = f^{(-1)}$ . Suppose that  $g(x) = x - g_2 x^2 - g_3 x^3 - \cdots$ , where the minus signs and the assumption that the coefficient of  $x$  in  $g$  is 1 make our formula simpler with no real loss of generality. Then (2.2.1) gives

$$[x^m] f^k = \frac{k}{m} [t^{m-k}] \left( \frac{1}{1 - g_2 t - g_3 t^2 - \dots} \right)^m.$$

Expanding by the binomial theorem and simplifying gives

$$f^k = \sum_{\substack{n_2+n_3+\dots=n-m \\ n_2+2n_3+\dots=m-k}} k \frac{(n-1)!}{m! n_2! n_3! \dots} g_2^{n_2} g_3^{n_3} \cdots x^m$$



where the sum is over all  $m, n$ , and  $n_2, n_3, \dots$  satisfying the two subscripted equalities. If we replace  $m$  with  $n_0$  then we may write the formula as

$$f^k = \sum_{\substack{n_0+n_2+n_3+\dots=n \\ 2n_2+3n_3+\dots=n-k}} k \frac{(n-1)!}{n_0! n_2! n_3! \dots} x^{n_0} g_2^{n_2} g_3^{n_3} \dots \quad (2.5.2)$$

and we see that the coefficients here are exactly the same as the coefficients in (2.5.1) (with  $n_1 = 0$ ).

**Exercise 2.5.1.** Explain the connection between (2.5.1) and (2.5.2) without using Lagrange inversion.

**2.6. Derivative formulas.** Lagrange inversion, especially in its analytic formulations, is often stated in terms of derivatives. We give here several derivative forms of Lagrange inversion.

**Theorem 2.6.1.** Let  $G(t) = \sum_{n=0}^{\infty} g_n t^n$ , where the  $g_i$  are indeterminates. Then there is a unique power series  $f$  satisfying

$$f = x + G(f)$$

and for any power series  $\phi(t)$  and  $\psi(t)$  we have

$$\phi(f) = \sum_{m=0}^{\infty} \frac{d^m}{dx^m} \left( \phi(x) (1 - G'(x)) \frac{G(x)^m}{m!} \right), \quad (2.6.1)$$

$$\phi(f) = \phi(x) + \sum_{m=1}^{\infty} \frac{d^{m-1}}{dx^{m-1}} \left( \phi'(x) \frac{G(x)^m}{m!} \right), \quad (2.6.2)$$

and

$$\frac{\psi(f)}{1 - G'(f)} = \sum_{m=0}^{\infty} \frac{d^m}{dx^m} \left( \psi(x) \frac{G(x)^m}{m!} \right). \quad (2.6.3)$$

*Proof.* We first prove (2.6.3). By (2.4.6) we have

$$\frac{\psi(f)}{1 - G'(f)} = \sum_{n=0}^{\infty} [t^n] \psi(t) (x + G(t))^n = \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} [t^n] \binom{n}{m} x^{n-m} \psi(t) G(t)^m.$$

So to prove (2.6.3), it suffices to prove that

$$\frac{d^m}{dx^m} \left( \psi(x) \frac{G(x)^m}{m!} \right) = \sum_{n=0}^{\infty} [t^n] \binom{n}{m} x^{n-m} \psi(t) G(t)^m.$$

We show more generally, that for any power series  $\alpha(t)$  we have

$$\frac{d^m}{dx^m} \frac{\alpha(x)}{m!} = \sum_{n=0}^{\infty} [t^n] \binom{n}{m} x^{n-m} \alpha(t). \quad (2.6.4)$$

But by linearity, it is sufficient to prove (2.6.4) for the case  $\alpha(t) = t^j$ , where both sides are equal to  $\binom{j}{m} x^{j-m}$ .

Next, (2.6.1) follows from (2.6.3) by taking  $\psi(t) = (1 - G'(t))\phi(t)$ .

Finally, we derive (2.6.2) from (2.6.1). Writing  $D$  for  $d/dx$ ,  $\phi$  for  $\phi(x)$  and  $G$  for  $G(x)$ , we have

$$D^m \left( \phi \frac{G^m}{m!} \right) = D^{m-1} \left( \frac{D(\phi G^m)}{m!} \right) = D^{m-1} \left( \phi' \frac{G^m}{m!} + \phi G' \frac{G^{m-1}}{(m-1)!} \right).$$

Thus

$$D^{m-1} \left( \phi' \frac{G^m}{m!} \right) = D^m \left( \phi \frac{G^m}{m!} \right) - D^{m-1} \left( \phi G' \frac{G^{m-1}}{(m-1)!} \right),$$

so

$$\phi + \sum_{m=1}^{\infty} D^{m-1} \left( \phi' \frac{G^m}{m!} \right) = \sum_{m=0}^{\infty} D^m \left( \phi \frac{G^m}{m!} \right) - \sum_{m=0}^{\infty} D^m \left( \phi G' \frac{G^m}{m!} \right). \quad \square$$

As before, applying an admissible substitution allows more general coefficients to be used. As an application of these formulas, let us take  $G(x) = zH(x)$  and consider the formula

$$\phi(f) \cdot \frac{\psi(f)}{1 - zH'(f)} = \frac{\phi(f)\psi(f)}{1 - zH'(f)}.$$

Applying (2.6.2) and (2.6.3) to the left side and (2.6.3) to the right, and then equating coefficients of  $z^n$  gives the convolution identity

$$\sum_{m=0}^n \binom{n}{m} \frac{d^{m-1}}{dx^{m-1}} (\phi'(x)H(x)^m) \frac{d^{n-m}}{dx^{n-m}} (\psi(x)H(x)^{n-m}) = \frac{d^n}{dx^n} (\phi(x)\psi(x)H(x)^n) \quad (2.6.5)$$

and similarly, expanding  $\phi(f)\psi(f)$  in two ways using (2.6.2) gives

$$\sum_{m=0}^n \binom{n}{m} \frac{d^{m-1}}{dx^{m-1}} (\phi'(x)H(x)^m) \frac{d^{n-m-1}}{dx^{n-m-1}} (\psi'(x)H(x)^{n-m}) = \frac{d^{n-1}}{dx^{n-1}} ((\phi(x)\psi(x))' H(x)^n). \quad (2.6.6)$$

Here  $d^{m-1}\phi'(x)/dx$  for  $m = 0$  is to be interpreted as  $\phi(x)$ . Formulas (2.6.5) and (2.6.6) were found by Cauchy [10]; a formula equivalent to (2.6.5) had been found earlier by Pfaff [58]. A detailed historical discussion of these identities and generalizations has been given by Johnson [40]; see also Chu [12, 14] and Abel [2]. Our approach to these identities has also been given by Huang and Ma [36].

**Exercise 2.6.2.** With the notation of Theorem 2.6.1, show that for any positive integer  $k$ ,

$$G(f)^k = \sum_{m=0}^{\infty} \frac{k}{(m+k)} \frac{d^m}{dx^m} G(x)^{m+k}.$$

### 3. APPLICATIONS

In this section we describe some applications of Lagrange inversion.

**3.1. A rational function expansion.** It is surprising that Lagrange inversion can give interesting results when the solution to the equation to be solved is rational.

We consider the equation

$$f = 1 + a + abf,$$

with solution

$$f = \frac{1+a}{1-ab}.$$

We apply (2.4.6) with  $R(t) = 1 + a + abt$  and  $\psi(t) = t^r(1 + bt)^s$ . Here we have  $1 + bf = (1 + b)/(1 - ab)$  and  $1 - R'(f) = 1 - ab$ . Then

$$\begin{aligned} \frac{(1 + a)^r(1 + b)^s}{(1 - ab)^{r+s+1}} &= \sum_n [t^n] t^r (1 + bt)^s (1 + a + abt)^n \\ &= \sum_{n,i} [t^n] \binom{n}{i} a^i t^r (1 + bt)^{s+i} \\ &= \sum_{n,i,j} [t^n] \binom{n}{i} a^i \binom{s+i}{j} b^j t^{r+j} \\ &= \sum_{i,j} \binom{r+j}{i} \binom{s+i}{j} a^i b^j. \end{aligned}$$

For another approach to this identity, see Gessel and Stanton [26].

**3.2. The tree function.** In applying Lagrange inversion, the nicest examples are those in which the series  $R(t)$  has the property that there is a simple formula for the coefficients of  $R(t)^n$ , and these simple formulas usually come from the exponential function or the binomial theorem. In this section we discuss the simplest case, in which  $R(t) = e^t$ . Later, in section 3.5, we discuss a more complicated example involving the exponential function.

Let  $T(x)$  be the power series satisfying

$$T(x) = xe^{T(x)}.$$

Equivalently,  $T(x) = (xe^{-x})^{(-1)}$  and thus  $T(xe^{-x}) = x$ . Then by the properties of exponential generating functions (see, e.g., Stanley [70, Chapter 5]),  $T(x)$  is the exponential generating function for rooted trees and  $e^{T(x)} = T(x)/x$  is the exponential generating function for forests of rooted trees. We shall call  $T(x)$  the *tree function* and we shall call  $F(x) = e^{T(x)}$  the *forest function*. The tree function is closely related to the Lambert  $W$  function [16, 17] which may be defined by  $W(x) = -T(-x)$ . Although the Lambert  $W$  function is better known, we will state our results in terms of the tree and forest functions.

Applying (2.2.1) with  $R(t) = e^t$  gives

$$T(x) = \sum_{n=1}^{\infty} n^{n-1} \frac{x^n}{n!} \tag{3.2.1}$$

and more generally,

$$\frac{T(x)^k}{k!} = \sum_{n=k}^{\infty} kn^{n-k-1} \binom{n}{k} \frac{x^n}{n!} \tag{3.2.2}$$

for all positive integers  $k$ , and

$$F(x)^k = e^{kT(x)} = \sum_{n=0}^{\infty} k(n+k)^{n-1} \frac{x^n}{n!} \tag{3.2.3}$$

for all  $k$ . Equation (3.2.2) implies that there are  $kn^{n-k-1} \binom{n}{k}$  forests of  $n$  rooted trees on  $n$  vertices and equation (3.2.3) implies that there are  $k(n+k)^{n-1}$  forests with vertex set  $\{1, 2, \dots, n+k\}$  in which the roots are  $1, 2, \dots, k$ .

An interesting special case of (3.2.3) is  $k = -1$ , which may be rearranged to

$$F(x) = \left(1 - \sum_{n=1}^{\infty} (n-1)^{n-1} \frac{x^n}{n!}\right)^{-1}. \quad (3.2.4)$$

Equation (3.2.4) may be interpreted in terms of prime parking functions [70, Exercise 5.49f, p. 95; Solution, p. 141].

Applying (2.2.6) gives

$$\frac{F(x)^k}{1 - T(x)} = \sum_{n=0}^{\infty} (n+k)^n \frac{x^n}{n!}. \quad (3.2.5)$$

A. Lacasse [46, p. 90] conjectured an identity that may be written as

$$U(x)^3 - U(x)^2 = \sum_{n=0}^{\infty} n^{n+1} \frac{x^n}{n!}, \quad (3.2.6)$$

where

$$U(x) = \sum_{n=0}^{\infty} n^n \frac{x^n}{n!}.$$

Proofs of Lacasse's conjecture were given by Chen et al. [11], Prodinger [59], Sun [73], and Younsi [80]. We will prove (3.2.6) by showing that both sides are equal to  $T(x)/(1 - T(x))^3$ .

To do this, we first note that the right side of (3.2.6) is

$$\frac{d}{dx} \sum_{n=0}^{\infty} (n-1)^n \frac{x^n}{n!} = \frac{d}{dx} \frac{e^{-T(x)}}{1 - T(x)} = \frac{e^{-T(x)} T(x) T'(x)}{(1 - T(x))^2}. \quad (3.2.7)$$

Differentiating (3.2.1) with respect to  $x$  gives

$$T'(x) = \sum_{n=0}^{\infty} (n+1)^n \frac{x^n}{n!},$$

which by (3.2.5) is equal to  $e^{T(x)}/(1 - T(x))$ . Thus (3.2.7) is equal to  $T(x)/(1 - T(x))^3$ . But by (3.2.5),  $U(x) = 1/(1 - T(x))$  from which it follows easily that  $U(x)^3 - U(x)^2 = T(x)/(1 - T(x))^3$ .

The series  $T(x)$  and  $U(x)$  were studied by Zvonkine [79], who showed that  $D^k U(x)$  and  $D^k U(x)^2$ , where  $D = xd/dx$ , are polynomials in  $U(x)$ .

The series  $\sum_{n=0}^{\infty} (n+k)^{n+m} x^n/n!$ , where  $m$  is an arbitrary integer, can be expressed in terms of  $T(x)$ . (The case  $k = 0$  has been studied by Smiley [67].) We first deal with the case in which  $m$  is negative.

**Theorem 3.2.1.** *Let  $l$  be a positive integer. Then for some polynomial  $p_l(u)$  of degree  $l - 1$ , with coefficients that are rational functions of  $k$ , we have*

$$\sum_{n=0}^{\infty} (n+k)^{n-l} \frac{x^n}{n!} = e^{kT(x)} p_l(T(x)). \quad (3.2.8)$$

The first three polynomials  $p_l(u)$  are

$$\begin{aligned} p_1(u) &= \frac{1}{k}, \\ p_2(u) &= \frac{1}{k^2} - \frac{u}{k(k+1)}, \\ p_3(u) &= \frac{1}{k^3} - \frac{(2k+1)}{k^2(k+1)^2}u + \frac{u^2}{k(k+1)(k+2)}. \end{aligned}$$

Before proving Theorem 3.2.1 let us check (3.2.8) for  $l = 1$  and  $l = 2$ . The case  $l = 1$  is equivalent to (3.2.3). For  $l = 2$ , we have

$$e^{kT(x)} = F(x)^k = k \sum_{n=0}^{\infty} (n+k)^{n-1} \frac{x^n}{n!}$$

and

$$\begin{aligned} e^{kT(x)}T(x) &= F(x)^k \cdot xF(x) = xF(x)^{k+1} \\ &= x(k+1) \sum_{n=0}^{\infty} (n+k+1)^{n-1} \frac{x^n}{n!} \\ &= (k+1) \sum_{n=0}^{\infty} n(n+k)^{n-2} \frac{x^n}{n!}. \end{aligned}$$

Thus

$$\begin{aligned} e^{kT(x)} \left( \frac{1}{k^2} - \frac{T(x)}{k(k+1)} \right) &= \frac{1}{k} \sum_{n=0}^{\infty} \left( (n+k)^{n-1} - n(n+k)^{n-2} \right) \frac{x^n}{n!} \\ &= \sum_{n=0}^{\infty} (n+k)^{n-2} \frac{x^n}{n!}. \end{aligned}$$

The general case can be proved in a similar way. (See Exercise 3.2.4.) However, it is instructive to take a different approach, using finite differences (cf. Gould [31]), that we will use again in section 3.3.

Let  $s$  be a function defined on the nonnegative integers. The *shift operator*  $E$  takes  $s$  to the function  $Es$  defined by  $(Es)(n) = s(n+1)$ . We denote by  $I$  the *identity operator* that takes  $s$  to itself and by  $\Delta$  the *difference operator*  $E - I$ , so  $(\Delta s)(n) = s(n+1) - s(n)$ . It is easily verified that if  $s$  is a polynomial of degree  $d > 0$  with leading coefficient  $L$  then  $\Delta s$  is a polynomial of degree  $d - 1$  with leading coefficient  $dL$ , and if  $s$  is a constant then  $\Delta s = 0$ . The  $k$ th difference of  $s$  is the function  $\Delta^k s$ . Thus if  $s$  is a polynomial of degree  $d$  with leading coefficient  $L$  then  $\Delta^k s = 0$  for  $k > d$  and  $\Delta^d s$  is the constant  $d!L$ .

Since the operators  $E$  and  $I$  commute, we can expand  $\Delta^k = (E - I)^k$  by the binomial theorem to obtain

$$(\Delta^k s)(n) = \sum_{i=0}^k (-1)^{k-i} \binom{k}{i} s(n+i).$$

We may summarize the result of this discussion in the following lemma.

**Lemma 3.2.2.** *Let  $s$  be a polynomial of degree  $d$  with leading coefficient  $L$ . Then*

$$\sum_{i=0}^k (-1)^{k-i} \binom{k}{i} s(n+i)$$

*is 0 if  $k > d$  and is the constant  $d!L$  for  $k = d$ .* □

*Proof of Theorem 3.2.1.* If we set  $x = ue^{-u}$  in Theorem 3.2.1 and use the fact that  $T(ue^{-u}) = u$  we see that Theorem (3.2.1) is equivalent to the formula

$$e^{-ku} \sum_{n=0}^{\infty} (n+k)^{n-l} \frac{(ue^{-u})^n}{n!} = p_l(u).$$

We have

$$\begin{aligned} e^{-ku} \sum_{n=0}^{\infty} (n+k)^{n+m} \frac{(ue^{-u})^n}{n!} &= \sum_{n=0}^{\infty} (n+k)^{n+m} \frac{u^n}{n!} e^{-(n+k)u} \\ &= \sum_{j=0}^{\infty} \frac{u^j}{j!} \sum_{n=0}^j (-1)^{j-n} \binom{j}{n} (n+k)^{j+m} \end{aligned} \tag{3.2.9}$$

If  $m = -l$  is a negative integer and  $j \geq l$  then  $(n+k)^{j+m} = (n+k)^{j-l}$  is a polynomial in  $n$  of degree less than  $j$ , so the inner sum in (3.2.9) is 0. Thus (3.2.8) follows, with

$$p_l(u) = \sum_{j=0}^{l-1} \frac{u^j}{j!} \sum_{n=0}^j (-1)^{j-n} \binom{j}{n} (n+k)^{-(l-j)}. \quad \square$$

We cannot set  $k = 0$  in (3.2.8), since the  $n = 0$  term on the left is  $k^{-l}$ . However it is not hard to evaluate  $\sum_{n=1}^{\infty} n^{n-l} x^n / n!$ .

**Theorem 3.2.3.** *Let  $q_l(u)$  be the result of setting  $k = 1$  in  $p_l(u)$ . Then*

$$\sum_{n=1}^{\infty} n^{n-l} \frac{x^n}{n!} = T(x) q_l(T(x)).$$

*Proof.* We have

$$\sum_{n=1}^{\infty} n^{n-l} \frac{x^n}{n!} = x \sum_{n=0}^{\infty} (n+1)^{n-l} \frac{x^n}{n!}$$

By (3.2.8) with  $k = 1$  this is

$$xe^{T(x)} q_l(T(x)) = T(x) q_l(T(x)). \quad \square$$

**Exercise 3.2.4.** Prove Theorem 3.2.1 by finding a formula for the coefficient  $t_j(n)$  of  $x^n/n!$  in  $e^{kT(x)} T(x)^j$  and showing that  $(n+k)^{n-l}$  can be expressed as a linear combination, with coefficients that are rational functions of  $k$ , of  $t_0(n), \dots, t_{l-1}(n)$ .

Next we consider  $\sum_{n=0}^{\infty} (n+k)^{n+m} x^n / n!$  where  $m$  is a nonnegative integer. (We evaluated the case  $k = 0, m = 1$  in our discussion of Lacasse's conjecture.)

**Theorem 3.2.5.** Let  $m$  be a nonnegative integer. Then there exists a polynomial  $r_m(u, k)$ , with integer coefficients, of degree  $m$  in  $u$  and degree  $m$  in  $k$ , such that

$$\sum_{n=0}^{\infty} (n+k)^{n+m} \frac{x^n}{n!} = e^{kT(x)} \frac{r_m(T(x), k)}{(1-T(x))^{2m+1}}.$$

The first three polynomials  $r_m(u, k)$  are

$$\begin{aligned} r_0(u, k) &= 1, \\ r_1(u, k) &= k + (1-k)u, \\ r_2(u, k) &= k^2 + (1+3k-2k^2)u + (2-3k+k^2)u^2. \end{aligned}$$

*Proof sketch.* We give here a sketch of a proof that tells us something interesting about the polynomials  $r_m(u, k)$ ; for a more direct approach see Exercise 3.2.6.

As in the proof of Theorem 3.2.1 we set  $x = ue^{-u}$  and consider the sum on the left side of (3.2.9).

Following Carlitz [9], we define the *weighted Stirling numbers of the second kind*  $R(n, j, k)$  by

$$R(n, j, k) = \frac{1}{j!} \sum_{i=0}^j (-1)^{j-i} \binom{j}{i} (k+i)^n,$$

so that

$$\sum_{n=0}^{\infty} R(n, j, k) \frac{x^n}{n!} = e^{kx} \frac{(e^x - 1)^j}{j!}. \quad (3.2.10)$$

(For  $k = 0$ ,  $R(n, j, k)$  reduces to the ordinary Stirling number of the second kind  $S(n, j)$ .) Equation (3.2.10) implies that the  $R(n, j, k)$  is a polynomial in  $k$  with integer coefficients. Then (3.2.9) is equal to  $\sum_{j=0}^{\infty} R(j+m, j, k) u^j$ . It is not hard to show that for fixed  $m$ ,  $R(j+m, j, k)$  is a polynomial in  $j$  of degree  $2m$ .

Thus

$$\sum_{j=0}^{\infty} R(j+m, j, k) u^j = \frac{r_m(u, k)}{(1-u)^{2m+1}}$$

for some polynomial  $r_m(u, k)$  of degree at most  $2m$ .

We omit the proof that  $r_m(u, k)$  actually has degree  $m$  in  $u$ . □

For  $k = 1$ , the coefficients of  $r_m(u, 1)$  are positive integers, sometimes called *second-order Eulerian numbers*; see, for example, [32, p. 270] and [24].

**Exercise 3.2.6.** Give an inductive proof of Theorem 3.2.5 using the fact that if  $W(m, k) = \sum_{n=0}^{\infty} (n+k)^{n+m} x^n/n!$  then  $dW(m, k)/dx = W(m+1, k+1)$ .

We can get convolution identities by applying (3.2.3) and (3.2.5) to

$$\frac{F(x)^{k+l}}{1-T(x)} = F(x)^k \frac{F(x)^l}{1-T(x)}$$

and

$$F(x)^{k+l} = F(x)^k F(x)^l.$$

The first identity yields

$$(n+k+l)^n = \sum_{i=0}^n \binom{n}{i} k(i+k)^{i-1} (n-i+l)^{n-i}.$$

and the second yields

$$(k+l)(n+k+l)^{n-1} = \sum_{i=0}^n \binom{n}{i} k(i+k)^{i-1} l(n-i+l)^{n-i-1}.$$

Note that these are identities of polynomials in  $k$  and  $l$ . If we set  $k = x$  and  $l = y - n$  in the first formula we get the nicer looking

$$(x+y)^n = \sum_{i=0}^n \binom{n}{i} x(x+i)^{i-1} (y-i)^{n-i}. \quad (3.2.11)$$

Replacing  $x$  with  $x/z$  and  $y$  with  $y/z$  in (3.2.11), and multiplying through by  $z^n$  gives the homogeneous form

$$(x+y)^n = \sum_{i=0}^n \binom{n}{i} x(x+iz)^{i-1} (y-iz)^{n-i} \quad (3.2.12)$$

which was proved by N. H. Abel in 1826 [1]. Note that for  $z = 0$ , (3.2.12) reduces to the binomial theorem. Riordan [64, pp. 18–27] gives a comprehensive account of Abel's identity and its generalizations, though he does not use Lagrange inversion.

**Exercise 3.2.7.** (Chu [13].) Prove Abel's identity (3.2.11) using finite differences. (Start by expanding  $(y-i)^{n-i} = [(x+y) - (x+i)]^{n-i}$  by the binomial theorem.)

**3.3. Fuss-Catalan numbers.** The Fuss-Catalan (or Fuß-Catalan) numbers of order  $p$  (also called generalized Catalan numbers) are the numbers  $\frac{1}{pn+1} \binom{pn+1}{n} = \frac{1}{(p-1)n+1} \binom{pn}{n}$ , which reduce to Catalan numbers for  $p = 2$ . They were first studied by N. Fuss in 1791 [21]. As we shall see, they are the coefficients of the power series  $c_p(x)$ , satisfying the functional equation

$$c_p(x) = 1 + xc_p(x)^p, \quad (3.3.1)$$

or equivalently,

$$c_p(x) = \frac{1}{1 - xc_p(x)^{p-1}},$$

as was shown using Lagrange inversion by Liouville [49].

An account of these generating functions can be found Graham, Knuth, and Patashnik [32, pp. 200–204].

It follows easily from (3.3.1) that

$$\begin{aligned} c_p(x) - 1 &= xc_p(x)^p = \left( \frac{x}{(1+x)^p} \right)^{\langle -1 \rangle}, \\ xc_p(x)^{p-1} &= (x(1-x)^{p-1})^{\langle -1 \rangle}, \\ xc_p(x^{p-1}) &= (x-x^p)^{\langle -1 \rangle}, \end{aligned} \quad (3.3.2)$$

and

$$c_p'(x) = \frac{c_p(x)^p}{1 - px c_p(x)^{p-1}}.$$



Lagrange inversion gives

$$c_p(x)^k = \sum_{n=0}^{\infty} \frac{k}{pn+k} \binom{pn+k}{n} x^n \quad (3.3.3)$$

for all  $k$ . With  $R(t) = 1 + xt^p$  we have  $R'(t) = pxt^{p-1}$  so

$$1 - R'(c_p(x)) = 1 - pxc_p(x)^{p-1} = 1 - p(c_p(x) - 1)/c_p(x) = 1 - p + pc_p(x)^{-1},$$

and thus by (2.4.6),

$$\sum_{n=0}^{\infty} \binom{pn+k}{n} x^n = \frac{c_p(x)^k}{1 - pxc_p(x)^{p-1}} = \frac{c_p(x)^{k+1}}{1 - (p-1)(c_p(x) - 1)}. \quad (3.3.4)$$

Equivalently,

$$\sum_{n=0}^{\infty} \binom{pn+k}{n} \left( \frac{x}{(1+x)^p} \right)^n = \frac{(1+x)^{k+1}}{1 - (p-1)x}$$

and

$$\sum_{n=0}^{\infty} \binom{pn+k}{n} (x(1-x)^{p-1})^n = \frac{1}{(1-px)(1-x)^k}.$$

The convolution identities obtained from (3.3.3) and (3.3.4), known as Rothe-Hagen identities [65, 33, 30] are

$$\sum_{i+j=n} \frac{k}{pi+k} \binom{pi+k}{i} \cdot \frac{l}{pj+l} \binom{pj+l}{j} = \frac{k+l}{pn+k+l} \binom{pn+k+l}{n}.$$

and

$$\sum_{i+j=n} \frac{k}{pi+k} \binom{pi+k}{i} \binom{pj+l}{j} = \binom{pn+k+l}{n}.$$

**Exercise 3.3.1.** Show that  $c_{-p}(x) = 1/c_{p+1}(-x)$ .

**Exercise 3.3.2.** Prove that  $c_{p+q}(x) = c_p(xc_{p+q}(x)^q)$  (a) combinatorially (b) algebraically (c) using Lagrange inversion.

**Exercise 3.3.3.** Prove that

$$(xc_p(x^a)^b)^{\langle -1 \rangle} = xc_{ab-p+1}(-x^a)^b$$

(a) algebraically (b) using Lagrange inversion. In particular, as noted by Dennis Stanton, if  $f(x) = xc(x)^3$  then  $f(x)^{\langle -1 \rangle} = -f(-x)$ .

**Exercise 3.3.4.** (Mansour and Sun [50, Example 5.6], Sun [72].) Show that

$$\frac{1}{1-x} c_3 \left( \frac{x^2}{(1-x)^3} \right) = c_2(x).$$

**Exercise 3.3.5.** Prove (3.3.3) and (3.3.4) by finite differences.

**Exercise 3.3.6.** (Chu [13].) Prove the Hagen-Rothe identities by finite differences.

Next, we prove Jensen's formula [39]

$$\sum_{l=0}^n \binom{j+pl}{l} \binom{r-pl}{n-l} = \sum_{i=0}^n \binom{j+r-i}{n-i} p^i. \quad (3.3.5)$$

By (3.3.4) we have

$$\begin{aligned} \sum_{l=0}^{\infty} \binom{pl+j}{l} x^l \sum_{m=0}^{\infty} \binom{pm+k}{m} x^m &= \frac{c_p(x)^{j+k}}{(1-pxc_p(x)^{p-1})^2} \\ &= \frac{c_p(x)^{j+k}}{1-pxc_p(x)^{p-1}} \sum_{i=0}^{\infty} p^i x^i c_p(x)^{(p-1)i} \\ &= \sum_{i=0}^{\infty} \frac{p^i x^i c_p(x)^{j+k+(p-1)i}}{1-pxc_p(x)^{p-1}} \\ &= \sum_{i=0}^{\infty} p^i x^i \sum_{m=0}^{\infty} \binom{pm+j+k+(p-1)i}{m} x^m \\ &= \sum_{n=0}^{\infty} x^n \sum_{i=0}^n \binom{pn+j+k-i}{n-i} p^i. \end{aligned}$$

Equating coefficients of  $x^n$  on both sides gives

$$\sum_{l=0}^n \binom{pl+j}{l} \binom{p(n-l)+k}{n-l} = \sum_{i=0}^n \binom{pn+j+k-i}{n-i} p^i.$$

Setting  $k = r - pn$  gives (3.3.5).

We also have analogues of Theorems 3.2.1 and 3.2.5 for Fuss-Catalan numbers.

**Theorem 3.3.7.** *Let  $i$  and  $j$  be nonnegative integers with  $i < j$ . Then*

$$\sum_{n=0}^{\infty} \frac{(pn+i)!}{n!((p-1)n+j)!} \frac{x^n}{(1+x)^{pn+i+1}}$$

is a polynomial  $u_{i,j}(x)$  in  $x$  of degree  $j-i-1$ .

*Proof.* We have

$$\begin{aligned} \sum_{n=0}^{\infty} \frac{(pn+i)!}{n!((p-1)n+j)!} \frac{x^n}{(1+x)^{pn+i+1}} &= \sum_{n=0}^{\infty} \frac{(pn+i)!}{n!((p-1)n+j)!} x^n \sum_{l=0}^{\infty} (-1)^l \binom{pn+i+l}{l} x^l \\ &= \sum_{m=0}^{\infty} x^m \sum_{n=0}^m \frac{(pn+i)!}{n!((p-1)n+j)!} (-1)^{m-n} \binom{(p-1)n+i+m}{m-n}. \end{aligned}$$

For  $m \geq j-i$ , the coefficient of  $x^m$  may be rearranged to

$$\frac{(m-(j-i))!}{m!} \sum_{n=0}^m (-1)^{m-n} \binom{m}{n} \binom{(p-1)n+i+m}{m-(j-i)}.$$

The sum is the  $m$ th difference of a polynomial of degree less than  $m$  and is therefore 0. For  $m = j - i - 1$ , the coefficient of  $x^{j-i-1}$  reduces to

$$\frac{1}{m!} \sum_{n=0}^m (-1)^{m-n} \binom{m}{n} \frac{1}{(p-1)n+j},$$

which is nonzero by the well-known identity

$$\sum_{n=0}^m (-1)^n \binom{m}{n} \frac{a}{n+a} = \binom{m+a}{a}^{-1},$$

so the degree of the polynomial is not less than  $j - i - 1$ . □

The first few values of these polynomials are

$$\begin{aligned} u_{i,i+1}(x) &= \frac{1}{i+1} \\ u_{i,i+2}(x) &= \frac{1}{(i+1)(i+2)} - \frac{p-1}{(i+2)(p+i+1)}x \\ u_{i,i+3}(x) &= \frac{1}{(i+1)(i+2)(i+3)} - \frac{(p-1)(p+2i+4)}{(i+2)(i+3)(p+i+1)(p+i+2)}x \\ &\quad + \frac{(p-1)^2}{(i+3)(p+i+2)(2p+i+1)}x^2 \end{aligned}$$

As a simple example of Theorem 3.3.7, the number of 2-stack-sortable permutations of  $\{1, 2, \dots, n\}$  is

$$a_n = 2 \frac{(3n)!}{(n+1)!(2n+1)!} = 4 \frac{(3n)!}{n!(2n+2)!}$$

(see [57, Sequence A000139]), so by Theorem 3.3.7, with  $p = 3$ ,  $i = 0$ , and  $j = 2$ ,  $\sum_{n=0}^{\infty} a_n x^n / (1+x)^{3n+1}$  is a polynomial of degree 1, which is easily computed to be  $2 - x$ . Then by (3.3.2), we find that

$$\sum_{n=0}^{\infty} a_n x^n = 3c_3(x) - c_3(x)^2,$$

which can be checked directly from (3.3.3).

There is a result similar to Theorem 3.3.7 for  $i \geq j$ , which we state without proof.

**Theorem 3.3.8.** *Let  $i$  and  $j$  be nonnegative integers with  $i \geq j$ . Then*

$$(1 - (p-1)x)^{2(i-j)+1} \sum_{n=0}^{\infty} \frac{(pn+i)!}{n!((p-1)n+j)!} \frac{x^n}{(1+x)^{pn+i+1}}$$

*is a polynomial in  $x$  of degree at most  $i - j$ .* □

**Exercise 3.3.9.** Prove Theorem 3.3.8.

**3.4. Narayana and Fuss-Narayana numbers.** The Narayana numbers may be defined by  $N(n, i) = \frac{1}{n} \binom{n}{i} \binom{n}{i-1}$  for  $n \geq 1$ . They have many combinatorial interpretations, in terms of Dyck paths, ordered trees, binary trees, and noncrossing partitions.

It is not hard to see from the formula for Narayana numbers that  $N(n, i) = N(n, n+1-i)$ . A generating function for the Narayana that exhibits this symmetry is given by the solution to the equation

$$f = (1 + xf)(1 + yf). \quad (3.4.1)$$

Lagrange inversion gives

$$\begin{aligned} f^k &= \sum_{n=0}^{\infty} \frac{k}{n} [t^{n-k}] (1 + xt)^n (1 + yt)^n \\ &= \sum_{n=0}^{\infty} \frac{k}{n} [t^{n-k}] \sum_{i,j} \binom{n}{i} \binom{n}{j} x^i y^j t^{i+j} \\ &= \sum_{i,j=0}^{\infty} \frac{k}{i+j+k} \binom{i+j+k}{i} \binom{i+j+k}{j} x^i y^j. \end{aligned}$$

In particular

$$\begin{aligned} f &= \sum_{i,j=0}^{\infty} \frac{1}{i+j+1} \binom{i+j+1}{i} \binom{i+j+1}{i+1} x^i y^j \\ &= \sum_{n=1}^{\infty} \sum_{i=0}^{n-1} N(n, i+1) x^i y^{n-i-1}. \end{aligned}$$

Equation (3.4.1) can be solved explicitly to give

$$f = \frac{1 - x - y - \sqrt{(1 - x - y)^2 - 4xy}}{2xy}.$$

**Exercise 3.4.1.** Prove that

$$\frac{(1 + xf)^r (1 + yf)^s}{\sqrt{(1 - x - y)^2 - 4xy}} = \sum_{i,j} \binom{r+i+j}{i} \binom{s+i+j}{j} x^i y^j$$

and

$$\begin{aligned} (1 + xf)^r (1 + yf)^s &= \sum_{i,j} \left[ \binom{r+i+j-1}{i} \binom{s+i+j}{j} - \binom{r+i+j}{i} \binom{s+i+j-1}{j-1} \right] x^i y^j \\ &= \sum_{i,j} \frac{rs + ri + sj}{(r+i+j)(s+i+j)} \binom{r+i+j}{i} \binom{s+i+j}{j} x^i y^j. \end{aligned}$$

The first formula is equivalent to a well-known generating function for Jacobi polynomials; see Carlitz [8].

We may generalize (3.4.1) to

$$f = (1 + x_1 f)^{r_1} (1 + x_2 f)^{r_2} \cdots (1 + x_m f)^{r_m}, \quad (3.4.2)$$

for which Lagrange inversion gives

$$f^k = \sum_{n=k}^{\infty} \sum_{i_1+\dots+i_m=n-k} \frac{k}{n} \binom{r_1 n}{i_1} \cdots \binom{r_m n}{i_m} x_1^{i_1} \cdots x_m^{i_m}. \quad (3.4.3)$$

These numbers reduce to Catalan numbers for  $k = m = 1, r_1 = 2$  and to Narayana numbers for  $k = 1, m = 2, r_1 = r_2 = 1$ . For  $k = 1, m = 2, r_1 = 1$  they are sometimes called Fuss-Narayana numbers; see Armstrong [3], Cigler [15], Edelman [19], Eu and Fu [20], and Wang [75]. The numbers for  $k = 1, r_i = 1$  for all  $i$  have been called generalized Fuss-Narayana numbers by Lenczewski and Sałapata [48]; they have also been studied by Edelman [19], Stanley [69], and Xu [78]. The case  $k = 1, m = 3, r_2 = -r_1, r_3 = 1$  of these numbers was considered by Krattenthaler, [44, equation (31)]. We note that if  $r_1 = \cdots = r_m$ , then  $f$  is a symmetric function of  $x_1, \dots, x_m$ , and this symmetric function arises in the study of algebraic aspects of parking functions [69].

If we set  $r_i = -s_i$  in (3.4.2) and replace  $x_i$  with  $-x_i$ , and  $f$  with  $g$ , then (3.4.2) becomes

$$g = \frac{1}{(1 - x_1 g)^{s_1} (1 - x_2 g)^{s_2} \cdots (1 - x_m g)^{s_m}}, \quad (3.4.4)$$

and, with the formula  $\binom{-a}{i} = (-1)^i \binom{a+i-1}{i}$ , (3.4.3) becomes

$$g^k = \sum_{n=k}^{\infty} \sum_{i_1+\dots+i_m=n-k} \frac{k}{n} \binom{s_1 n + i_1 - 1}{i_1} \cdots \binom{s_m n + i_m - 1}{i_m} x_1^{i_1} \cdots x_m^{i_m}. \quad (3.4.5)$$

These numbers reduce to Catalan numbers for  $k = m = s_1 = 1$ . For  $s_1 = \cdots = s_m = 1$  they have been considered by Aval [4] and (for  $k = 1$ ) by Stanley [69].

Of special interest are the cases of (3.4.2) that reduce to a quadratic equation, since in these cases there are simple explicit formulas for  $f$ . If we take  $m = 1, r_1 = r_2 = 1$ , and  $r_3 = -1$  then with a change of variable names and one sign we have

$$f = \frac{(1 + xf)(1 + yf)}{1 - zf},$$

with the solution

$$f = \frac{1 - x - y - \sqrt{(1 - x - y)^2 - 4xy - 4z}}{2(xy + z)},$$

and (3.4.3) gives

$$f^k = \sum_{n=k}^{\infty} \sum_{i_1+i_2+i_3=n-k} \frac{k}{n} \binom{n}{i_1} \binom{n}{i_2} \binom{n+i_3-1}{i_3} x^{i_1} y^{i_2} z^{i_3}.$$

Another case of (3.4.2) that reduces to a quadratic is  $f = \sqrt{(1 + xf)(1 + yf)}$ ; we leave the details to the reader.

**3.5. Raney's equation.** G. Raney [62] considered the equation

$$f = \sum_{m=1}^{\infty} A_m e^{B_m f}, \quad (3.5.1)$$

in which  $f$  is a power series in the indeterminates  $A_m$  and  $B_m$ . He used Prüfer's correspondence to give a combinatorial derivation of a formula for the coefficients in this power series.

We can use Lagrange inversion to give a formula for the coefficients of  $f$ .

**Theorem 3.5.1.** *Let  $f$  be the power series in  $A_m$  and  $B_m$  satisfying (3.5.1), and let  $k$  be a positive integer. Let  $i_1, i_2, \dots$  and  $j_1, j_2, \dots$  be nonnegative integers, only finitely many of which are nonzero. If  $i_1 + i_2 + \dots = k + j_1 + j_2 + \dots$  then the coefficient of  $A_1^{i_1} A_2^{i_2} \dots B_1^{j_1} B_2^{j_2} \dots$  in  $f^k$  is*

$$k \frac{(i_1 + i_2 + \dots - 1)!}{i_1! i_2! \dots} \frac{i_1^{j_1} i_2^{j_2} \dots}{j_1! j_2! \dots}$$

and if  $i_1 + i_2 + \dots \neq k + j_1 + j_2 + \dots$  then the coefficient is zero.

*Proof.* Applying equation (2.4.3) to (3.5.1) gives

$$\begin{aligned} f^k &= \sum_{n=k}^{\infty} \frac{k}{n} [t^{n-k}] \left( \sum_m A_m e^{B_m t} \right)^n \\ &= \sum_{n=k}^{\infty} \frac{k}{n} [t^{n-k}] \sum_{i_1+i_2+\dots=n} \frac{n!}{i_1! i_2! \dots} A_1^{i_1} A_2^{i_2} \dots e^{i_1 B_1 t} e^{i_2 B_2 t} \dots \\ &= \sum_{n=k}^{\infty} \frac{k}{n} [t^{n-k}] \sum_{i_1+i_2+\dots=n} \frac{n!}{i_1! i_2! \dots} A_1^{i_1} A_2^{i_2} \dots \sum_{j_1} \frac{(i_1 B_1 t)^{j_1}}{j_1!} \sum_{j_2} \frac{(i_2 B_2 t)^{j_2}}{j_2!} \dots \\ &= \sum_{n=k}^{\infty} \sum_{\substack{i_1+i_2+\dots=n \\ j_1+j_2+\dots=n-k}} k \frac{(n-1)!}{i_1! i_2! \dots} A_1^{i_1} A_2^{i_2} \dots B_1^{j_1} B_2^{j_2} \dots \frac{i_1^{j_1} i_2^{j_2} \dots}{j_1! j_2! \dots}, \end{aligned}$$

and the formula follows. □

A combinatorial derivation of Raney's formula has also been given by D. Knuth [42, Section 2.3.4.4].

## 4. PROOFS

In this section we give three several proofs of the Lagrange inversion formula.

**4.1. Residues.** The simplest proof of Lagrange inversion is due to Jacobi [38]. We define the *residue*  $\text{res } f(x)$  of a Laurent series  $f(x) = \sum_n f_n x^n$  to be  $f_{-1}$ .

Jacobi proved the following change of variables formula for residues:

**Theorem 4.1.1.** *Let  $f$  be a Laurent series and let  $g(x) = \sum_{n=1}^{\infty} g_n x^n$  be a power series with  $g_1 \neq 0$ . Then*

$$\text{res } f(x) = \text{res } f(g(x))g'(x).$$

*Proof.* By linearity, it is sufficient to prove the formula when  $f(x) = x^k$  for some integer  $k$ . If  $k \neq -1$  then  $\text{res } x^k = 0$  and

$$\text{res } g(x)^k g'(x) = \text{res } \frac{d}{dx} g(x)^{k+1} / (k+1) = 0,$$

since the residue of a derivative is 0.

If  $k = -1$  then  $\text{res } x^k = 1$  and

$$\begin{aligned} \text{res } g(x)^k g'(x) &= \text{res } g'(x)/g(x) = \text{res } \frac{g_1 + 2g_2x + \cdots}{g_1x + g_2x^2 + \cdots} \\ &= \text{res } \frac{1}{x} \cdot \frac{g_1 + 2g_2x + \cdots}{g_1 + g_2x + \cdots} = 1. \end{aligned} \quad \square$$

Jacobi's paper [38] contains a multivariable generalization of Theorem 4.1.1; see also Gessel [28] and Xin [77].

Now let  $f(x)$  and  $g(x)$  be compositional inverses. Then for any Laurent series  $\phi$ ,

$$[x^n] \phi(f) = \text{res } \frac{\phi(f)}{x^{n+1}} = \text{res } \frac{\phi(f(g))g'}{g^{n+1}} = \text{res } \frac{\phi(x)g'}{g^{n+1}}.$$

This is equation (2.1.7), which we have already seen is equivalent to the other forms of Lagrange inversion.

**Exercise 4.1.2.** Let  $f$  be a Laurent series and let  $g(x) = \sum_{n=m}^{\infty} g_n x^n$  be a Laurent series with  $g_m \neq 0$ . Show that if  $f(g(x))$  is well-defined as a Laurent series then

$$m \text{res } f = \text{res } f(g(x))g'(x).$$

**Exercise 4.1.3.** (Hirzebruch [34]; see also Kneezel [41].) Use the change of variables formula (Theorem 4.1.1) to show that the unique power series  $f(x)$  satisfying

$$\text{res} \left( \frac{f(x)}{x} \right)^n = 1$$

for all  $n \geq 1$  is  $f(x) = x/(1 - e^{-x})$ .

**4.2. Induction.** In this proof and the next we consider the equation  $f = xR(f)$ , where  $R(t)$  is a power series. If  $R(t)$  has no constant term then  $f = 0$  and the formulas are trivial. So we may assume that  $R(t)$  has a nonzero constant term, so  $f$  exists and is unique, since it is the compositional inverse of  $x/R(x)$ .

We now give an inductive proof of (2.1.2): for any power series  $\phi(t)$ ,

$$[x^n] \phi(f) = [t^n] \left( 1 - \frac{tR'(t)}{R(t)} \right) \phi(t)R(t)^n. \quad (4.2.1)$$

(As noted in section 2, this implies that (4.2.1) holds more generally when  $\phi(t)$  is a Laurent series.)

We first take care of the case in which  $\phi(t) = 1$ . The case  $\phi(t) = 1$ ,  $n = 0$  is trivial. If  $\phi(t) = 1$  and  $n > 0$ , we have

$$\begin{aligned} \left( 1 - \frac{tR'(t)}{R(t)} \right) R(t)^n &= R(t)^n - tR'(t)R(t)^{n-1} \\ &= R(t)^n - \frac{t}{n} \frac{d}{dt} R(t)^n. \end{aligned} \quad (4.2.2)$$

Now for any power series  $u(t)$ ,

$$[t^n] \left( u(t) - \frac{t}{n} u'(t) \right) = 0.$$

With(4.2.2), this proves (4.2.1) for  $\phi(t) = 1$ ,  $n > 0$ .

Now we prove the formula (4.2.1) by induction on  $n$ . It is clear that (4.2.1) holds for  $n = 0$ . Now let us suppose that for some nonnegative integer  $m$ , (4.2.1) holds for all  $\phi$  when  $n = m$ . We now want to show that (4.2.1) holds for all  $\phi$  when  $n = m + 1$ . By linearity and the case  $\phi(t) = 1$ , it is enough to prove (4.2.1) for  $n = m + 1$  and  $\phi(t) = t^k$ , where  $k \geq 1$ . In this case we have

$$\begin{aligned} [x^{m+1}] f^k &= [x^{m+1}] f^{k-1} \cdot xR(f) \\ &= [x^m] f^{k-1} R(f) \\ &= [t^m] \left(1 - \frac{tR'(t)}{R(t)}\right) t^{k-1} R(t) R(t)^m \\ &= [t^{m+1}] \left(1 - \frac{tR'(t)}{R(t)}\right) t^k R(t)^{m+1}. \end{aligned}$$

**4.3. Factorization.** Another proof is based on a version of the “factor theorem”: if  $f = xR(f)$  then  $t - f$  divides  $t - xR(t)$ . This proof is taken from Gessel [27] but it is similar to Lagrange’s original proof [47].

We will prove (2.2.1), which gives a formula for  $f^k$  where  $f = f(x)$  satisfies  $f = xR(f)$ .

First we recall Taylor’s theorem for power series: if  $P(t)$  is a power series in  $t$ , and  $\alpha$  is an element of the coefficient ring, then

$$P(t) = \sum_{n=0}^{\infty} \frac{(t - \alpha)^n}{n!} P^{(n)}(\alpha), \quad (4.3.1)$$

as long as this sum is summable. (The case  $P(t) = t^m$  of (4.3.1) is just the binomial theorem, and the general case follows by linearity.)

Now let us apply (4.3.1) with  $P(t) = t - xR(t)$  and  $\alpha = f$ , where  $f = xR(f)$  so that  $P(f) = 0$ . Then we have

$$\begin{aligned} t - xR(t) &= 0 + (t - f)(1 - xR'(f)) + (t - f)^2 S(x, t) \\ &= (t - f)Q(x, t), \end{aligned} \quad (4.3.2)$$

where  $S(x, t)$  is a power series and  $Q(x, t)$  is a power series with constant term 1.

Equation (4.3.2) is an identity in the ring  $C[[x, t]]$ , which is naturally embedded in the ring  $C((t))[[x]]$  of power series in  $x$  with coefficients that are Laurent series in  $t$ . In this ring, series like  $\sum_{n=0}^{\infty} (x/t)^n$  are allowed, even though they have infinitely many negative powers of  $t$ , since the coefficient of any power of  $x$  is a Laurent series in  $t$ . We now do some computations in  $C((t))[[x]]$ .

By (4.3.2), we have  $1 - xR(t)/t = (1 - f/t)Q(x, t)$ . Since  $xR(t)/t$  and  $f/t$  are divisible by  $x$  and  $Q(x, t)$  is a power series in  $x$  and  $t$  with constant term 1, we may take logarithms to obtain

$$-\log(1 - xR(t)/t) = -\log(1 - f/t) - \log Q(x, t). \quad (4.3.3)$$

Note that  $\log Q(x, t)$  is a power series in  $x$  and  $t$ , and so has no negative powers of  $t$ . Now we equate coefficients of  $x^n t^{-k}$  on both sides of (4.3.3) where  $n$  and  $k$  are both positive integers. On the left we have  $[t^{-k}](R(t)/t)^n/n = [t^{n-k}] R(t)^n/n$  and on the right we have  $[x^n] f^k/k$ . Thus,

$$[x^n] f^k = \frac{k}{n} [t^{n-k}] R(t)^n,$$



which is (2.2.1).

**Exercise 4.3.1.** (Gessel [27].) Derive (2.2.2) similarly.

**4.4. Combinatorial proofs.** There are several different combinatorial proofs of Lagrange inversion. They all interpret the solution  $f$  of  $f = xR(f)$  or  $f = R(f)$  as counting certain trees. Here  $f$  may be interpreted as either an ordinary or exponential generating function and thus different types of trees may be involved.

In ordinary generating function proofs,  $f$  will count unlabeled *ordered trees* (also called *plane trees*), which are rooted trees in which the children of each vertex are linearly ordered. (See Figure 1.) More generally,  $f^k$  will count  $k$ -tuples of ordered trees, which we also call

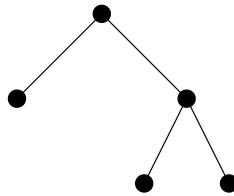


FIGURE 1. An ordered tree

*forests of ordered trees.*

If  $f = xR(f)$ , where  $R(t) = \sum_{i=0}^{\infty} r_i t^i$ , then the coefficient of  $x^n$  in  $f$  is the sum of the weights of the ordered trees with a total of  $n$  vertices, where the weight of a tree is the product of the weights of its vertices and the weight of a vertex with  $i$  children is  $r_i$ . (For example, the tree of Figure 1 has weight  $r_0^3 r_2$ .) So to give a combinatorial proof of formula (2.2.1), we show that the sum of the weights of all  $k$ -tuples of ordered trees with a total of  $n$  vertices is  $(k/n) [t^{n-k}] R(t)^n$ , or equivalently by (2.5.1), the number of  $k$ -tuples of ordered trees with  $n_i$  vertices with  $i$  children, for each  $i$ , is equal to

$$\frac{k}{n} \binom{n}{n_0, n_1, n_2, \dots} \quad (4.4.1)$$

if

$$n = \sum_k n_i \quad \text{and} \quad n - k = \sum_k i n_i, \quad (4.4.2)$$

and is zero otherwise. W. T. Tutte [74] gave the case  $k = 1$  of (4.4.1), which he derived (in a roundabout way) from Lagrange inversion.

We can also work with exponential generating functions. One way to do this is consider the equation  $f = x \sum_{i=0}^{\infty} s_i f^i / i!$  where  $x$  is the exponential variable and the  $s_n$  are weights. Then by the properties of exponential generating functions (see, e.g., [70, Chapter 5])  $f$  counts labeled rooted trees where a vertex with  $i$  children is weighted  $s_i$ . More precisely, the coefficient of  $x^n / n!$  in  $f^k / k!$  is the sum of the weights of all forests of  $k$  rooted trees with vertex set  $[n] = \{1, 2, \dots, n\}$ . Then Lagrange inversion in the form (2.5.1) is equivalent to assertion that the number of forests of  $k$  rooted trees with vertex set  $[n]$  in which  $n_i$  vertices have  $i$  children is equal to

$$\frac{(n-1)!}{(k-1)! 0!^{n_0} 1!^{n_1} 2!^{n_2} \dots} \binom{n}{n_0, n_1, \dots} \quad (4.4.3)$$

with the same conditions on  $n_0, n_1, \dots$  as before. (See Stanley [70, p. 30, Corollary 5.3.5].)

Another exponential generating function approach is to consider the equation

$$f = \sum_{i=0}^{\infty} s_i \frac{f^i}{i!}$$

where we work with exponential generating functions in the variables  $s_0, s_1, \dots$ . Then by the properties of multivariable exponential generating functions, the coefficient of

$$\frac{s_0^{n_0} s_1^{n_1}}{n_0! n_1!} \dots$$

in  $f^k/k!$  is the number of forests of  $k$  rooted trees in which for each  $i$ ,  $n_i$  vertices have  $i$  children and these vertices are labeled  $1, 2, \dots, n_i$ . (Since vertices with the same label have different numbers of children, there are no nontrivial label-preserving automorphisms of such trees.) For example, such a forest with  $k = 2$ ,  $n_0 = 5$ ,  $n_2 = 3$ , and  $n_i = 0$  for  $i \notin \{0, 2\}$  is shown in Figure 2. Then Lagrange inversion in the form (2.5.1) (with  $x = 1$ ) is equivalent to

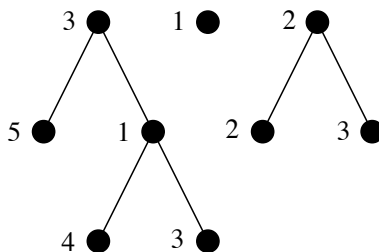


FIGURE 2. A forest with  $n_0 = 5$ ,  $n_2 = 3$

the assertion that the number of such forests is

$$\frac{(n-1)!}{(k-1)! 0!^{n_0} 1!^{n_1} 2!^{n_2} \dots}, \quad (4.4.4)$$

where  $n = n_0 + n_1 + \dots$  and  $n_1 + 2n_2 + 3n_3 + \dots = n - k$ .

**4.5. Raney's proof.** The earliest combinatorial proof of Lagrange inversion is that of Raney [61]. We sketch here a proof that is based on Raney's though the details are different. (See also Stanley [70, pp. 31–35 and 39–40].) We define the *suffix code*  $c(T)$  for an ordered tree  $T$  to be a sequence of nonnegative integers defined recursively: If the root  $r$  of  $T$  has  $j$  children, and the trees rooted at the children of  $r$  are  $T_1, \dots, T_j$ , then  $c(T)$  is the concatenation  $c(T_1) \cdots c(T_j)j$ . More generally, the suffix code for a  $k$ -tuple  $(T_1, \dots, T_k)$  of ordered trees is the concatenation  $c(T_1) \cdots c(T_k)$ . We defined the *reduced code* of a tree or  $k$ -tuple of trees to be the sequence obtained from the suffix code by subtracting 1 from each entry.

For example the suffix code of the  $k$ -tuple of trees in Figure 3 is 00201 and the reduced code is  $\bar{1}\bar{1}1\bar{1}0$ , where  $\bar{1}$  denotes  $-1$ .

The following lemma can be proved by induction:

**Lemma 4.5.1.**

- (i) *A forest of ordered trees is uniquely determined by its reduced code.*

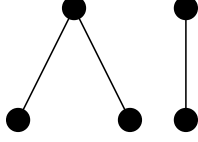


FIGURE 3. A forest of ordered trees

- (ii) A sequence  $a_1 a_2 \cdots a_n$  of integers greater than or equal to  $-1$  is the reduced code of an ordered  $k$ -forest if and only if  $a_1 + \cdots + a_n = -k$  and  $a_1 + \cdots + a_i$  is negative for  $i = 1, \dots, n$ .  $\square$

We also need a lemma, due to Raney, that generalizes the “cycle lemma” of Dvoretzky and Motzkin [18]. It can be proved by induction, or in other ways. (See, e.g., Stanley [70, pp. 32–33].)

**Lemma 4.5.2.** *Let  $a_1 \cdots a_n$  be a sequence of integers greater than or equal to  $-1$  with sum  $-k < 0$ . Then there exactly  $k$  integers  $i$ , with  $1 \leq i \leq n$ , such that the sequence  $a_i \cdots a_n a_1 \cdots a_{i-1}$  has all partial sums negative.*

We can now prove Lagrange inversion. We want to prove that the sum of the weights of all  $k$ -forests with  $n$  vertices is  $k/n$  times

$$[t^{n-k}] R(t)^n = [t^{-k}] \left( \frac{R(t)}{t} \right)^n,$$

where  $R(t) = \sum_{n=0}^{\infty} s_n t^n$ . Let us define the weight of a sequence  $a_1 \cdots a_n$  of integers to be the product  $s_{a_1+1} \cdots s_{a_n+1}$ . Then by Lemma 4.5.1, the sum of the weights of all  $k$ -forests with  $n$  vertices is the sum of the weights of all sequences of integers of length  $n$ , with entries greater than or equal to  $-1$ , with sum  $-k$ , and with all partial sums negative. It is clear that  $[t^{-k}] (R(t)/t)^n$  is the sum of the weights of all sequences of integers of length  $n$ , with entries greater than or equal to  $-1$ , and with sum  $-k$ . But by Lemma 4.5.2, a proportion  $k/n$  of these sequences have all partial sums negative.

**4.6. Proofs by labeled trees.** We can derive Lagrange inversion by counting labeled trees with the following result, which seems to have first been proved by Moon [52].

**Theorem 4.6.1.** *Let  $n$  be a positive integer and let  $d_1, d_2, \dots, d_n$  be positive integers. Then the number of (unrooted) trees with vertex set  $[m]$  in which vertex  $i$  has degree  $d_i$  is the multinomial coefficient*

$$\binom{m-2}{d_1-1, \dots, d_m-1} \tag{4.6.1}$$

if  $\sum_{i=1}^m d_i = 2(m-1)$  and is 0 otherwise.

We will prove Theorem 4.6.1 a little later; we first look at some of its consequences. A corollary of Theorem 4.6.1 allows us to count forests of rooted trees:

**Corollary 4.6.2.** *Let  $e_1, e_2, \dots, e_n$  be nonnegative integers with  $e_1 + e_2 + \cdots = n - k$ , and let  $k$  be a positive integer. Then the number of forests of  $k$  rooted trees, with vertex set  $[n]$ , in which vertex  $i$  has  $e_i$  children is*

$$\binom{n-1}{k-1, e_1, \dots, e_n}.$$

*Proof.* Let  $T$  be a tree on  $[n+1]$  in which vertex  $n+1$  has degree  $k$ , and vertex  $i$  has degree  $e_i+1$  for  $1 \leq i \leq n$  (so that  $\sum_{i=1}^n e_i = n-k$ ). Removing vertex  $n+1$  from  $T$  and rooting the resulting component trees at the neighbors of  $n+1$  in  $T$  gives a forest  $F$  of  $k$  rooted trees in which vertex  $i$  has  $e_i$  children, and this operation gives a bijection from trees on  $[n+1]$  in which vertex  $n+1$  has degree  $k$  and vertex  $i$  has degree  $e_i+1$  for  $1 \leq i \leq n$  to the set of rooted forests of  $k$  trees on  $[n]$  in which vertex  $i$  has  $e_i$  children. The result then follows from Lemma (4.6.1).  $\square$

It follows from Corollary 4.6.2 that the number of forests of  $k$  rooted trees in which for each  $i$ ,  $n_i$  vertices have  $i$  children and these vertices are labeled  $1, 2, \dots, n_i$  is given by (4.4.4), which as we saw earlier, is equivalent to Lagrange inversion.

Now to count  $k$ -forests on  $[n]$  in which  $n_i$  vertices have  $i$  children, we first assign the number of children to each element of  $[n]$ , which can be done in  $\binom{n}{n_0, n_1, \dots}$  ways. For each assignment, the number of trees is

$$\frac{(n-1)!}{(k-1)! 0!^{n_0} 1!^{n_1} \dots}$$

Multiplying these factors gives (4.4.3).

We now present sketches of three proofs of Theorem 4.6.1. They all depend on the fact that a tree with at least two vertices must have at least one leaf (vertex with no children).

The first proof uses Prüfer's correspondence [60]. Given a tree  $T_1$  with vertex set  $[n]$ , let  $a_1$  be the least leaf of  $T_1$  and let  $b_1$  be the unique neighbor of  $a_1$  in  $T_1$ . Let  $T_2$  be the result of removing  $a_1$  and its incident edge from  $T_1$ . Let  $a_2$  be the least leaf of  $T_2$  and let  $b_2$  be its neighbor in  $T_2$ . Continue in this way to define  $b_3, \dots, b_{n-2}$ . Then the *Prüfer code* of  $T_1$  is the sequence  $(b_1, b_2, \dots, b_{n-2})$ . It can be shown that the map that takes a tree to its Prüfer code is a bijection from the set of trees with vertex set  $[n]$ , for  $n \geq 2$ , to the set of sequences  $b_1 b_2 \dots b_{n-2}$  of elements of  $[n]$ , with the property that a vertex of degree  $d$  appears  $d-1$  times in the Prüfer code. Then, as noted by Moon [52], Theorem 4.6.1 is an immediate consequence, since the multinomial coefficient counts Prüfer codes of trees in which vertex  $i$  has degree  $d_i$ .

The second proof is by induction, and is due to Moon [53] (see also [54, p. 13]). For  $m \geq 2$ , let  $T(m; d_1, \dots, d_m)$  be the number of trees on  $[m]$  in which vertex  $i$  has degree  $d_i$  and let  $U(m; d_1, \dots, d_m)$  be the multinomial coefficient (4.6.1), which is 0 if any  $d_i$  is less than 1 or if  $\sum_i d_i \neq 2(n-1)$ .

Clearly  $T(m; d_1, \dots, d_m)$  is equal to  $U(m; d_1, \dots, d_m)$  for  $m = 2$ . Now suppose that  $n > 2$  and that  $T(m; d_1, \dots, d_m)$  is equal to  $U(m; d_1, \dots, d_m)$  for  $m = n-1$  and all choices of  $d_1, \dots, d_m$ .

We next observe that if  $d_n = 1$  then

$$T(n; d_1, \dots, d_n) = \sum_{i=1}^{n-1} T(n-1; d_1, \dots, d_i-1, \dots, d_{n-1}) \quad (4.6.2)$$

since every tree in which  $n$  is a leaf is obtained by joining vertex  $n$  with an edge to some vertex of a tree on  $[n-1]$ , increasing by 1 the degree of this edge and leaving the other degrees unchanged. Then by the inductive hypothesis and a well-known recurrence for multinomial coefficients,  $T(n; d_1, \dots, d_n)$  is equal to  $U(n; d_1, \dots, d_n)$ . We have assumed that  $d_n = 1$ , but

since  $T(n; d_1, \dots, d_n)$  is symmetric in  $d_1, \dots, d_n$  and every tree has at least one leaf, the result holds without this assumption.

Rényi [63] (see also Moon [54, p. 13]) gave an elegant variation of this proof. He showed by induction that for  $m \geq 2$  the number of trees on  $[m]$  with degree sequence  $(d_1, \dots, d_m)$ , i.e., trees in which vertex  $i$  has degree  $d_i$ , is the coefficient of  $x_1^{d_1-1} \cdots x_m^{d_m-1}$  in  $(x_1 + \cdots + x_m)^{m-2}$ . The result clearly holds for  $m = 2$ , so suppose that  $n > 2$  and that the result holds when  $m = n - 1$ . To show that it holds for  $m = n$ , we note that every tree on  $[n]$  has at least one leaf, so by symmetry, we may assume, without loss of generality, that  $d_n = 1$ . Thus we need only show that the number of trees on  $[n]$  with degree sequence  $(d_1, \dots, d_{n-1}, 1)$  is the coefficient of  $x_1^{d_1-1} \cdots x_{n-1}^{d_{n-1}-1}$  in

$$(x_1 + \cdots + x_{n-1})^{n-2} = (x_1 + \cdots + x_{n-1}) \cdot (x_1 + \cdots + x_{n-1})^{n-3}.$$

But every tree on  $[n]$  in which vertex  $n$  is a leaf is obtained by joining vertex  $n$  by an edge to some vertex of a tree on  $[n - 1]$ , increasing by 1 the degree of this edge and leaving the other degrees unchanged. Thus by the induction hypothesis, the contribution from trees in which vertex  $n$  is joined to vertex  $i$  is  $x_i(x_1 + \cdots + x_{n-1})^{n-3}$  and the result follows.

**Acknowledgment.** I would like to thank Sateesh Mane and an anonymous referee for helpful comments.

#### REFERENCES

- [1] N. H. Abel, *Beweis eines Ausdrucks, von welchem die Binomial-Formel ein einzelner Fall ist*, J. Reine Angew. Math. **1** (1826), 159–160.
- [2] Ulrich Abel, *A generalization of the Leibniz rule*, Amer. Math. Monthly **120** (2013), no. 10, 924–928.
- [3] Drew Armstrong, *Generalized noncrossing partitions and combinatorics of Coxeter groups*, Mem. Amer. Math. Soc. **202** (2009), no. 949, x+159.
- [4] Jean-Christophe Aval, *Multivariate Fuss-Catalan numbers*, Discrete Math. **308** (2008), no. 20, 4660–4669.
- [5] F. Bergeron, G. Labelle, and P. Leroux, *Combinatorial Species and Tree-like Structures*, Encyclopedia of Mathematics and its Applications, vol. 67, Cambridge University Press, Cambridge, 1998, Translated from the 1994 French original by Margaret Readdy. With a foreword by Gian-Carlo Rota.
- [6] Christian Brouder, Alessandra Frabetti, and Christian Krattenthaler, *Non-commutative Hopf algebra of formal diffeomorphisms*, Adv. Math. **200** (2006), no. 2, 479–524.
- [7] Jean-Paul Bultel, *Combinatorial properties of the noncommutative Faà di Bruno algebra*, J. Algebraic Combin. **38** (2013), no. 2, 243–273.
- [8] L. Carlitz, *The generating function for the Jacobi polynomial*, Rend. Sem. Mat. Univ. Padova **38** (1967), 86–88.
- [9] ———, *Weighted Stirling numbers of the first and second kind. I*, Fibonacci Quart. **18** (1980), no. 2, 147–162.
- [10] Augustin-Louis Cauchy, *Exercices de mathématiques, vol. 1*, De Bure frères, 1826, Application du calcul des résidus à la sommation de plusieurs suites. Oeuvres complètes d’Augustin Cauchy. Série 2, tome 6, pp. 62–73.
- [11] William Y. C. Chen, Janet F. F. Peng, and Harold R. L. Yang, *Decomposition of triply rooted trees*, Electron. J. Combin. **20** (2013), no. 2, Paper 10, 10 pp.
- [12] Wenchang Chu, *Leibniz inverse series relations and Pfaff-Cauchy derivative identities*, Rend. Mat. Appl. (7) **29** (2009), no. 2, 209–221.
- [13] ———, *Elementary proofs for convolution identities of Abel and Hagen-Rothe*, Electron. J. Combin. **17** (2010), no. 1, Note 24, 5 pp.
- [14] ———, *Derivative inverse series relations and Lagrange expansion formula*, Int. J. Number Theory **9** (2013), no. 4, 1001–1013.

- [15] J. Cigler, *Some remarks on Catalan families*, European J. Combin. **8** (1987), no. 3, 261–267.
- [16] R. M. Corless, G. H. Gonnet, D. E. G. Hare, D. J. Jeffrey, and D. E. Knuth, *On the Lambert W function*, Adv. Comput. Math. **5** (1996), no. 4, 329–359.
- [17] Robert M. Corless, David J. Jeffrey, and Donald E. Knuth, *A sequence of series for the Lambert W function*, Proceedings of the 1997 International Symposium on Symbolic and Algebraic Computation (Kihei, HI), ACM, New York, 1997, pp. 197–204.
- [18] A. Dvoretzky and Th. Motzkin, *A problem of arrangements*, Duke Math. J. **14** (1947), 305–313.
- [19] Paul H. Edelman, *Chain enumeration and noncrossing partitions*, Discrete Math. **31** (1980), no. 2, 171–180.
- [20] Sen-Peng Eu and Tung-Shan Fu, *Lattice paths and generalized cluster complexes*, J. Combin. Theory Ser. A **115** (2008), no. 7, 1183–1210.
- [21] Nicolao Fuss, *Solutio quaestionis, quot modis polygonum  $n$  laterum in polygona  $m$  laterum, per diagonales resolvi queat*, Nova Acta Academiae Sci. Petropolitanae **9** (1791), 243–251.
- [22] Adriano M. Garsia, *A  $q$ -analogue of the Lagrange inversion formula*, Houston J. Math. **7** (1981), no. 2, 205–237.
- [23] Ira Gessel, *A noncommutative generalization and  $q$ -analog of the Lagrange inversion formula*, Trans. Amer. Math. Soc. **257** (1980), no. 2, 455–482.
- [24] Ira Gessel and Richard P. Stanley, *Stirling polynomials*, J. Combinatorial Theory Ser. A **24** (1978), no. 1, 24–33.
- [25] Ira Gessel and Dennis Stanton, *Applications of  $q$ -Lagrange inversion to basic hypergeometric series*, Trans. Amer. Math. Soc. **277** (1983), no. 1, 173–201.
- [26] ———, *Short proofs of Saalschütz’s and Dixon’s theorems*, J. Combin. Theory Ser. A **38** (1985), no. 1, 87–90.
- [27] Ira M. Gessel, *A factorization for formal Laurent series and lattice path enumeration*, J. Combin. Theory Ser. A **28** (1980), no. 3, 321–337.
- [28] ———, *A combinatorial proof of the multivariable Lagrange inversion formula*, J. Combin. Theory Ser. A **45** (1987), no. 2, 178–195.
- [29] Ira M. Gessel and Gilbert Labelle, *Lagrange inversion for species*, J. Combin. Theory Ser. A **72** (1995), no. 1, 95–117.
- [30] H. W. Gould, *Final analysis of Vandermonde’s convolution*, Amer. Math. Monthly **64** (1957), 409–415.
- [31] ———, *Euler’s formula for  $n$ th differences of powers*, Amer. Math. Monthly **85** (1978), no. 6, 450–467.
- [32] Ronald L. Graham, Donald E. Knuth, and Oren Patashnik, *Concrete Mathematics: A Foundation for Computer Science*, second ed., Addison-Wesley Publishing Company, Reading, MA, 1994.
- [33] S. J. Hagen, Johann G., *Synopsis der Hoheren Mathematik: Arithmetische und Algebraische Analyse*, vol. 1, Felix L. Dames, Berlin, 1891.
- [34] F. Hirzebruch, *The signature theorem: reminiscences and recreation*, Prospects in mathematics (Proc. Sympos., Princeton Univ., Princeton, N.J., 1970), Princeton Univ. Press, Princeton, N.J., 1971, pp. 3–31. Ann. of Math. Studies, No. 70.
- [35] Josef Hofbauer, *Lagrange-inversion*, Sémin. Lothar. Combin. **6** (1982), Art. B06a, <http://www.emis.de/journals/SLC/opapers/s06hofbauer.html>.
- [36] Jianfeng Huang and Xinrong Ma, *Two elementary applications of the Lagrange expansion formula*, J. Math. Res. Appl. **35** (2015), no. 3, 263–270.
- [37] Eri Jabotinsky, *Representation of functions by matrices. Application to Faber polynomials*, Proc. Amer. Math. Soc. **4** (1953), 546–553.
- [38] C. G. J. Jacobi, *De resolutione aequationum per series infinitas*, Journal für die reine und angewandte Mathematik **6** (1830), 257–286, Gesammelte Werke, vol. 6, pp. 26–61, G. Reimer, Berlin (1891), reprinted by Chelsea Publishing Company, New York (1969).
- [39] J. L. W. V. Jensen, *Sur une identité d’Abel et sur d’autres formules analogues*, Acta Math. **26** (1902), 307–318.
- [40] Warren P. Johnson, *The Pfaff/Cauchy derivative identities and Hurwitz type extensions*, Ramanujan J. **13** (2007), no. 1–3, 167–201.
- [41] Dan Kneezel, *Hirzebruch’s motivation of the Todd class*, MathOverflow, <http://mathoverflow.net/q/60478> (version: 2011-04-03).

- [42] Donald E. Knuth, *The Art of Computer Programming. Vol. 1: Fundamental Algorithms*, third ed., Addison-Wesley, Reading, MA, 1997.
- [43] Ch. Krattenthaler, *Operator methods and Lagrange inversion: a unified approach to Lagrange formulas*, Trans. Amer. Math. Soc. **305** (1988), no. 2, 431–465.
- [44] Christian Krattenthaler, *The  $F$ -triangle of the generalised cluster complex*, Topics in Discrete Mathematics, Algorithms Combin., vol. 26, Springer, Berlin, 2006, pp. 93–126.
- [45] Gilbert Labelle, *Some new computational methods in the theory of species*, Combinatoire énumérative (Montreal, Que., 1985/Quebec, Que., 1985), Lecture Notes in Math., vol. 1234, Springer, Berlin, 1986, pp. 192–209.
- [46] Alexandre Lacasse, *Bornes PAC-Bayes et algorithmes d'apprentissage*, Ph.D. thesis, Université Laval, Québec, 2010.
- [47] Joseph Louis Lagrange, *Nouvelle méthode pour résoudre les équations littérales par le moyen des séries*, Mémoires de l'Académie Royale des Sciences et Belles-Lettres de Berlin **24** (1770), 251–326, Oeuvres complète, tome 3, Paris, 1867, 5–73.
- [48] Romuald Lenczewski and Rafał Salapata, *Multivariate Fuss-Narayana polynomials and their application to random matrices*, Electron. J. Combin. **20** (2013), no. 2, Paper 41, 14 pp.
- [49] J. Liouville, *Remarques sur un mémoire de N. Fuss*, J. Math. Pures Appl. **8** (1843), 391–394.
- [50] Toufik Mansour and Yidong Sun, *Bell polynomials and  $k$ -generalized Dyck paths*, Discrete Appl. Math. **156** (2008), no. 12, 2279–2292.
- [51] D. Merlini, R. Sprugnoli, and M. C. Verri, *Lagrange inversion: when and how*, Acta Appl. Math. **94** (2006), no. 3, 233–249 (2007).
- [52] J. W. Moon, *The second moment of the complexity of a graph*, Mathematika **11** (1964), 95–98.
- [53] ———, *Enumerating labelled trees*, Graph Theory and Theoretical Physics, Academic Press, London, 1967, pp. 261–272.
- [54] ———, *Counting Labelled Trees*, Canadian Mathematical Monographs, No. 1, Canadian Mathematical Congress, Montreal, Que., 1970.
- [55] Ivan Niven, *Formal power series*, Amer. Math. Monthly **76** (1969), 871–889.
- [56] Jean-Christophe Novelli and Jean-Yves Thibon, *Noncommutative symmetric functions and Lagrange inversion*, Adv. in Appl. Math. **40** (2008), no. 1, 8–35.
- [57] *The On-Line Encyclopedia of Integer Sequences*, published electronically at <http://oeis.org>, 2016.
- [58] J. F. Pfaff, *Allgemeine Summation einer Reihe, worinn höhere Differenziale vorkommen*, Archiv der reinen und angewandten Mathematik **1** (1795), 337–47.
- [59] Helmut Prodinger, *An identity conjectured by Lacasse via the tree function*, Electron. J. Combin. **20** (2013), no. 3, Paper 7, 3 pp.
- [60] Heinz Prüfer, *Neuer Beweis eines Satzes über Permutationen*, Arch. Math. Phys **27** (1918), 742–744.
- [61] George N. Raney, *Functional composition patterns and power series reversion*, Trans. Amer. Math. Soc. **94** (1960), 441–451.
- [62] ———, *A formal solution of  $\sum_{i=1}^{\infty} A_i e^{B_i X} = X$* , Canad. J. Math. **16** (1964), 755–762.
- [63] Alfréd Rényi, *On the enumeration of trees*, Combinatorial Structures and their Applications (Proc. Calgary Internat. Conf., Calgary, Alta., 1969), Gordon and Breach, New York, 1970, pp. 355–360.
- [64] John Riordan, *Combinatorial Identities*, John Wiley & Sons, Inc., New York-London-Sydney, 1968.
- [65] Henrico Augusto Rothe, *Formulae De Serierum Reversione Demonstratio Universalis Signis Localibus Combinatorio-Analyticorum Vicariis Exhibita*, Sommer, Leipzig, 1793.
- [66] Issai Schur, *On Faber polynomials*, Amer. J. Math. **67** (1945), 33–41.
- [67] Leonard M. Smiley, *Completion of a rational function sequence of Carlitz*, 2000, [arXiv:math/0006106](https://arxiv.org/abs/math/0006106) [math.CO].
- [68] Alan D. Sokal, *A ridiculously simple and explicit implicit function theorem*, Sémin. Lothar. Combin. **61A** (2009/11), Art. B61Ad, 21 pp.
- [69] Richard P. Stanley, *Parking functions and noncrossing partitions*, Electron. J. Combin. **4** (1997), no. 2, Research Paper 20, 4 pp., The Wilf Festschrift (Philadelphia, PA, 1996).
- [70] ———, *Enumerative Combinatorics. Vol. 2*, Cambridge Studies in Advanced Mathematics, vol. 62, Cambridge University Press, Cambridge, 1999, With a foreword by Gian-Carlo Rota and appendix 1 by Sergej Fomin.

- [71] Dennis Stanton, *Recent results for the  $q$ -Lagrange inversion formula*, Ramanujan revisited (Urbana-Champaign, Ill., 1987), Academic Press, Boston, MA, 1988, pp. 525–536.
- [72] Yidong Sun, *A simple bijection between binary trees and colored ternary trees*, Electron. J. Combin. **17** (2010), no. 1, Note 20, 5.
- [73] ———, *A simple proof of an identity of Lacasse*, Electron. J. Combin. **20** (2013), no. 2, Paper 11, 3 pp.
- [74] William T. Tutte, *The number of planted plane trees with a given partition*, Amer. Math. Monthly **71** (1964), no. 3, 272–277.
- [75] Chao-Jen Wang, *Applications of the Goulden-Jackson Cluster Method to Counting Dyck paths by Occurrences of Subwords*, Ph.D. thesis, Brandeis University, 2011.
- [76] E. T. Whittaker and G. N. Watson, *A Course of Modern Analysis*, Third Edition, Cambridge University Press, 1920.
- [77] Guoce Xin, *A residue theorem for Malcev-Neumann series*, Adv. in Appl. Math. **35** (2005), no. 3, 271–293.
- [78] Dapeng Xu, *Generalizations of two-stack-sortable permutations*, 2002, [arXiv:math/0209313\[math.CO\]](#).
- [79] Malik Younsi, *An algebra of power series arising in the intersection theory of moduli spaces of curves and in the enumeration of ramified coverings of the sphere*, 2004, [arXiv:math/0403092\[math.AG\]](#).
- [80] ———, *Proof of a combinatorial conjecture coming from the PAC-Bayesian machine learning theory*, 2012, [arXiv:1209.0824\[math.CO\]](#).

DEPARTMENT OF MATHEMATICS, BRANDEIS UNIVERSITY, WALTHAM, MA 02453-2700  
*E-mail address:* [gessel@brandeis.edu](mailto:gessel@brandeis.edu)