Multivariate Asymptotics for Products of Large Powers with Applications to Lagrange Inversion

Edward A. Bender
Department of Mathematics
University of California, San Diego
La Jolla, CA 92093-0112, USA
ebender@ucsd.edu

L. Bruce Richmond
Department of Combinatorics and Optimization
University of Waterloo
Waterloo, Ontario N2L 3G1, Canada
lbrichmond@watdragon.uwaterloo.ca

Submitted: March 27, 1998
Revised: January 5, 1999
Accepted: January 12, 1999

Abstract

An asymptotic estimate is given for the coefficients of products of large powers of generating functions. This theorem and another local limit theorem which is useful for conditioning are applied to various combinatorial enumeration problems that involve multivariate Lagrange inversion.

1991 AMS Class. No. Primary: 41A63 Secondary: 05A16, 05C05, 41A60

1. Introduction

If $f(0) \neq 0$ has a (possibly formal) power series expansion at 0, the equation $w = xf(w)$ determines the power series $w(x)$. Two forms of the Lagrange inversion formula are

$$g_n = [x^n]\, g(w) = [x^n] \Big(g(x) f(x)^n \{(1 - xf'(x)/f(x)\} \Big) \tag{1}$$
$$= (1/n)[x^n] \big(xg'(x) f(x)^n \big), \tag{2}$$

where $[x^n]\, h(x)$ denotes the coefficient of the monomial x^n in the power series $h(x)$. We obtained asymptotics for g_n from (2) for some types of formal power series [6]. When f has a nonzero radius of convergence, various authors have studied the asymptotics of $[x^n]\, g(w)$ using three basic approaches:

- *Exact Formula.* Using (2), obtain an exact formula g_n. This is often either a simple expression or a summation with alternating signs. Obtain asymptotics from the exact formula. This has been used only sporadically.
- *Singularity Analysis.* Determine the nature of the singularities of w by looking at $xf(w) - w = 0$. They are usually square root branch points due to the vanishing of $\partial(xf(w) - w)/\partial w$. Obtain asymptotics by what is essentially Darboux's Theorem. For a systematic discussion of this approach, see Sprugnoli and Verri [24].
- *Contour Integration.* Using the Cauchy Residue Theorem, one can estimate g_n from (2). Since $f(x)^n$ leads to an integral with a simple dominant term it suffices to use a circle. For a systematic discussion of this approach, see Gardy [13].

One can easily include other variables in (2) by simply thinking of the coefficients of f, g, and w as involving the new variables. Furthermore, there are extensions of Lagrange inversion to several functions and other variables can be included in these as well.

Recently Drmota [12] treated a system of functional equations using singularity analysis. His results can be applied to multifunction Lagrange inversion when $g(w_1, w_2, \ldots, w_d) = w_i$ for some i. Not all cases of interest have this form, a prime example being map enumeration.

The asymptotics of rooted convex polyhedra by faces and vertices (two equations with no extra variables) were studied by us [7] using singularity analysis and later by Bender and Wormald [11] using an exact formula. Rooted maps on general surfaces were dealt with in a similar manner by us and Canfield [4].

In this paper, we are concerned with the asymptotic behavior of the coefficients of large powers of multivariate generating functions and their application to multivariate Lagrange inversion. In Theorem 2 of [5] we studied coefficients of large powers of a single multivariate function using a contour integration approach. In Theorem 2.1 below, we extend this to products of powers of several functions when the exponents tend to infinity in such a way that their ratios are bounded. When there is only one power, Theorem 2.1 is essentially contained in Theorem 2 of [5], but we believe the conditions here are more easily verified than those in [5]. From a probabilistic viewpoint, our concern is with local limit theorems (estimates of coefficients) rather than central limit theorems (estimates for averages of coefficients). One could certainly obtain a central limit theorem extending Theorem 1 of [5]; however, more general central limit theorems have been obtained by Hwang [17] in the case of two variables. Hwang also studies the rate of convergence (which we do not) and points out that the central limit theorem we would derive would have a convergence rate of $O(n^{-1/2})$.

In the next section we state and prove Theorem 2.1, our theorem for products of powers. In Section 3 we explain how the theorem applies to Lagrange inversion of a single function and discuss the problem of conditioning on some of the indices. This is useful when one wishes to study combinatorial objects conditioned on things such as "size" or number of "components." Section 4 illustrates applies these ideas to specific enumeration problems. Since neither conditioning nor Lagrange inversion applications were discussed in [5], the material in Sections 3 and 4 is new even though Theorem 2.1 follows from Theorem 2 of [5] in this case. In Section 5, we recall Lagrange inversion formulas for several functions and show how the product of powers theorem can be applied to these formulas. We also prove a local limit theorem that is needed to continue the discussion of conditioning. Section 6 contains examples of specific applications. Although Theorem 2.1 leads to Langrange inversion asymptotics for many functions g; maps present a difficulty which we can resolve only in the single variable situation. This is explained in Section 6. In the final section, we indicate some research problems suggested by the limitations of our approach.

We thank Donatella Merlini and Renzo Sprugnoli for helpful conversations.

2. A Limit Theorem for Products of Powers

Let $\mathbb{Z}$ denote the integers. For a set V of vectors, let $\mathcal{A}(V)$ be the additive abelian group generated by V. Bold face letters denote vectors, $\mathbf{x}^{\mathbf{n}} = x_1^{n_1} x_2^{n_2} \cdots$, $|\mathbf{x}|$ denotes the vector whose components are $|x_i|$, and $\|\mathbf{x}\|$ denotes the length of $\mathbf{x}$. As already noted $[\mathbf{x}^{\mathbf{n}}]\, h(\mathbf{x})$ denotes the coefficient of $\mathbf{x}^{\mathbf{n}}$ in the power series $h(\mathbf{x})$.

Let $\mathbf{m}(f(\mathbf{z}))$ and $B(f(\mathbf{z}))$ be the vector and matrix given by

$$\mathbf{m}(f(\mathbf{z}))_i = \frac{\partial(\log f)}{\partial(\log z_i)} \quad \text{and} \quad B(f(\mathbf{z}))_{i,j} = \frac{\partial^2(\log f)}{\partial(\log z_i)\,\partial(\log z_j)}.$$

In all cases, the logarithms are real for real positive $\mathbf{z}$. (This is possible since our functions are positive reals for such $\mathbf{z}$.) Note that

$$\mathbf{m}(\mathbf{f}^{\mathbf{n}}) = \textstyle\sum n_i \mathbf{m}(f_i) \quad \text{and} \quad B(\mathbf{f}^{\mathbf{n}}) = \textstyle\sum n_i B(f_i). \tag{3}$$

Theorem 2.1. *Let* $\mathbf{u}$ *denote an* l*-dimensional vector over the complex numbers, let* R *be a compact subset of* $(0,\infty)^l$ *with nonempty interior, and let* C *be the set of complex vectors* $\mathbf{u}$ *with* $|\mathbf{u}| \in R$*. Suppose* $f_j(\mathbf{u})$ $(1 \le j \le d)$ *and* $h(\mathbf{u})$ *are such that*

(a) h *and the* f_j *are analytic in* C *and strictly positive in* R*;*

(b) in R*, the* $B(f_j)$ *are positive semidefinite and* $\sum_{j=1}^d B(f_j)$ *is nonsingular;*

(c) in C*,* $|f_j(\mathbf{u})| \le f_j(|\mathbf{u}|)$ *with equality for all* j *only in* R*.*

Fix $\delta > 0$ *and let* $n = \sum n_i$*. Then we have*

$$[\mathbf{u}^{\mathbf{k}}]\big(h(\mathbf{u})\mathbf{f}(\mathbf{u})^{\mathbf{n}}\big) = \frac{h(\mathbf{r})\mathbf{f}(\mathbf{r})^{\mathbf{n}}\big\{\exp(-\mathbf{t}B^{-1}\mathbf{t}'/2) + o(1)\big\}}{\sqrt{\det(2\pi B)}\,\mathbf{r}^{\mathbf{k}}} \quad \text{as} \quad n \to \infty \tag{4}$$

uniformly for $\mathbf{n} \in [n\delta, n/\delta]^d$ *and* $\mathbf{r} \in R$*, where* $\mathbf{i} = \mathbf{m}(\mathbf{f}(\mathbf{r})^{\mathbf{n}})$*,* $B = B(\mathbf{f}(\mathbf{r})^{\mathbf{n}})$*,* $\mathbf{t} = \mathbf{k} - \mathbf{i}$*, and* $\mathbf{t}'$ *denotes transpose.*

If, for all i*,* $f_i(\mathbf{u}) = \sum a_i(\mathbf{k})\mathbf{u}^{\mathbf{k}}$ *where* $a_i(\mathbf{k}) \ge 0$ *for all* $\mathbf{k}$ *and*

$$\Lambda(\mathbf{f}) = \mathcal{A}\big\{\mathbf{k} - \mathbf{j} \mid a_i(\mathbf{k})a_i(\mathbf{j}) \ne 0 \text{ for some } i\big\} = \mathbb{Z}^l, \tag{5}$$

then conditions (b) and (c) are satisfied. Frequently $a_i(\mathbf{0}) > 0$ *for all* i*, in which case (5) becomes* $\mathcal{A}\big\{\mathbf{k} \mid a_i(\mathbf{k}) > 0 \text{ for some } i\big\} = \mathbb{Z}^l$*.*

Proof: Note that $B(\mathbf{f}^{\mathbf{1}}) = \sum B(f_j)$. Since

$$B(\mathbf{f}^{\mathbf{n}}) = n\delta B(\mathbf{f}^{\mathbf{1}}) + \sum (n_i - \delta n) B(f_i),$$

$\mathbf{n} \in [n\delta, n/\delta]^d$, and the $B(f_i)$ are positive semidefinite, it follows that

$$\mathbf{t}B(\mathbf{f}^{\mathbf{n}})\mathbf{t}' \ge n\delta\mathbf{t}B(\mathbf{f}^{\mathbf{1}})\mathbf{t}'.$$

Since the domain of $\mathbf{r}$ is compact and $B(\mathbf{f}^{\mathbf{1}})$ is positive definite, it follows that $B(\mathbf{f}^{\mathbf{n}})/n$ is positive definite in a uniform sense; that is, there is a constant C such that $\mathbf{t}B(\mathbf{f}(\mathbf{r})^{\mathbf{n}})\mathbf{t}' \ge nC\mathbf{t}\mathbf{t}'$ for all $\mathbf{r} \in R$, all $\mathbf{n} \in [n\delta, n/\delta]^d$, and all $\mathbf{t}$.

The proof of (4) follows that of Theorem 2 of [5] almost exactly:

Expand the logarithm of $h(\mathbf{z})\mathbf{f}(\mathbf{z})^{\mathbf{n}}$ in a Taylor series about $\mathbf{r}$, keeping quadratic terms and a third-order error estimate. Use the Cauchy Residue Theorem with the contours $|z_i| = r_i$ to estimate the desired coefficient, with (b), (c), and the uniform positive definiteness of $B(\mathbf{f}^{\mathbf{n}})/n$ ensuring that

$$\left|\frac{h(\mathbf{z})\mathbf{f}(\mathbf{z})^{\mathbf{n}}}{h(\mathbf{r})\mathbf{f}(\mathbf{r})^{\mathbf{n}}}\right| = O\big(\exp(-C\theta^2 n)\big)$$

uniformly for some $C > 0$ and $\theta = \max(|\arg z_j|)$. See [5] for details.

We now prove the claims concerning (5). Since f_j has a power series with nonnegative coefficients:

(i) The first part of (c) holds.

(ii) By Rényi's number 2 on p.341 of [23], the first part of (b) holds.

(iii) $\mathbf{f^1}$ has a power series with nonnegative coefficients $a(\mathbf{k})$ and

$$\Lambda(\mathbf{f^1}) = \mathcal{A}\big\{\mathbf{k}-\mathbf{j} \mid a(\mathbf{k})a(\mathbf{j}) \neq 0\big\} = \Lambda(\mathbf{f})$$

Since $\Lambda(\mathbf{f}) = \mathbb{Z}^l$, it follows from (iii) that $|\mathbf{f^1}(\mathbf{u})| = \mathbf{f^1}(|\mathbf{u}|)$ if and only if $\mathbf{u} = |\mathbf{u}|$ and so the proof of (c) is complete. The second half of (b) follows from Lemma 6 in [10], with the matrix T in in that paper being the 1×1 matrix $\mathbf{f^1}$ and $\mathcal{A}^{(s)}_{i,j} = \mathcal{A}^{(1)}_{1,1} = \Lambda(\mathbf{f^1})$. ∎

For Theorem 2.1 to give more than an asymptotic upper bound, the exponential in (4) must not be $o(1)$. In other words, we must have $|\mathbf{t}| = O(n^{1/2})$. Thus the domain of useful $\mathbf{k}$ is asymptotically the same as the domain of $\mathbf{i}$. The latter depends on the problem and becomes evident only by calculation; however, we can describe the typical situation. Let $Z(\mathbf{n})$ be the set of $\mathbf{j}$ such that $[\mathbf{u^j}]\big(h(\mathbf{u})\mathbf{f}(\mathbf{u})^{\mathbf{n}}\big) = 0$. It usually suffices to require that $\mathbf{i}$ be at least ϵn from $Z(\mathbf{n})$, where $\epsilon > 0$ is an arbitrary constant. In particular, all components of $\mathbf{i}$ will be at least ϵn.

Theorem 2.1 can be strengthened in at least two ways:

(a) The function h can depend on $\mathbf{n}$ so long as its partials through second order are uniformly $o(n)$.

(b) It may happen that the lattice $\Lambda(\mathbf{f})$ in (5) is a proper sublattice of $\mathbb{Z}^l$ rather than all of $\mathbb{Z}^l$. A theorem still exists, but it requires multisection as discussed in [10].

We have omitted these from the theorem because they are relatively rare and add complications.

3. Lagrange Inversion of One Function

How does Theorem 2.1 apply to Lagrange inversion of a single function? Since (1) and (2) deal with formal power series over a commutative ring of characteristic zero, we are free to include extra variables $\mathbf{y}$ in the coefficients of g, f and w. Thus, if

$$w(x,\mathbf{y}) = xf(w(x,\mathbf{y}),\mathbf{y}) \quad \text{with} \quad f(0,\mathbf{y}) \neq 0,$$

we have $[x^n\mathbf{y^j}]\, g(w(x,\mathbf{y}),\mathbf{y}) = [\mathbf{y^j}]\, g_n$ where g_n as in (1). Apply Theorem 2.1 with

$$d = 1, \quad \mathbf{n} = (n), \quad \mathbf{f} = (f), \quad \mathbf{u} = (x,\mathbf{y}), \quad \mathbf{k} = (n,\mathbf{j}),$$

and h the remaining factors in (1) or (2) after f^n is removed. We start the indexing of $\mathbf{k}$, $\mathbf{i}$, and $\mathbf{t}$ at zero so that $k_s = j_s$ for $s > 0$. For greatest accuracy in estimating the coefficient of $\mathbf{u^k}$, one would normally set $\mathbf{t} = \mathbf{0}$, that is, $\mathbf{i} = \mathbf{k}$. The equation for $\mathbf{i}$ is then

$$n = n\frac{r_0}{f(r_0,\mathbf{r})}\,\frac{\partial f(r_0,\mathbf{r})}{\partial r_0} \quad \text{and} \quad j_s = n\frac{r_s}{f(r_0,\mathbf{r})}\,\frac{\partial f(r_0,\mathbf{r})}{\partial r_s} \quad \text{for} \quad s > 0.$$

Regarding the first of these as an equation in n, it has a nontrivial solution if and only if

$$1 - \frac{r_0}{f(r_0,\mathbf{r})}\frac{\partial f(r_0,\mathbf{r})}{\partial r_0} = 0; \tag{6}$$

that is, the last factor in (1) vanishes. Hence h fails to satisfy the $h > 0$ condition in Theorem 2.1 and so we must use (2):

$$[x^n\mathbf{y}^{\mathbf{j}}]\, g(w,\mathbf{y}) = [x^n\mathbf{y}^{\mathbf{j}}]\left(x\frac{\partial g(x,\mathbf{y})}{\partial x} f(x,\mathbf{y})^n\right).$$

Conditioning. In addition to providing asymptotics, (4) provides a local limit theorem for $\mathbf{j}$ as $n \to \infty$. One obtains a normal distribution by setting $\mathbf{r} = (r_0, \mathbf{1})$, choosing r_0 so that $d\log(f(r_0,\mathbf{1})/d\log r_0 = 1$, and conditioning on the zeroth component of $\mathbf{i}$ being $k_0 = n$. To condition, one drops the zeroth component of $\mathbf{t}$ and the corresponding row and column of B^{-1}. The latter corresponds to replacing the $l \times l$ "covariance" matrix B with the inverse of the lower $(l-1)\times(l-1)$ block of B^{-1}, say C. One can compute C directly from the block matrix formula found on pp.25–26 of [22]:

$$\begin{bmatrix} B_{1,1} & B_{1,2} \\ B'_{1,2} & B_{2,2} \end{bmatrix}^{-1} = \begin{bmatrix} * & * \\ * & C^{-1} \end{bmatrix}, \quad \text{where} \quad C = B_{2,2} - B'_{1,2}(B_{1,1})^{-1}B_{1,2}. \tag{7}$$

In this case $B_{1,1}$ is 1×1.

One can condition on a set of variables that includes n. In this case, we set $r_i = 1$ if we are *not* conditioning on j_i. The remaining components of the equation $(n,\mathbf{j}) = \mathbf{m}(\mathbf{r})$, including the zeroth, are used to solve for r_0 and the remaining r_i. Again the indices of the variables being conditioned on are dropped from $\mathbf{t}$ and B^{-1}. Equation (7) still applies, but $B_{1,1}$ is no longer 1×1 since it is indexed by all the variables on which we are conditioning. Since the asymptotics obtained from (4) is uniform, so are the asymptotics for limiting distributions, provided $(r_0,\mathbf{r}) \in R$ and all components of $\mathbf{j}$ lie in $[n\delta, n/\delta]$. Of course $(r_0,\mathbf{r})$ varies as $n \to \infty$ unless the conditioned components grow at a rate proportional to n.

It is possible to condition on a set of variables that does *not* contain n.

This is more complex. Rather than discuss it here, we treat the general case in the context of multiple Lagrange inversion in Section 5.

Summing over variables, which is roughly the complement of conditioning, is also discussed in Section 5.

4. Examples of Inversion of One Function

We now turn to examples of single function inversion.

Example 4.1. (Noncrossing Partitions) A *noncrossing partition* of a set of integers is a set partition such that there are no integers $a_1 < b_1 < a_2 < b_2$ with a_i in one block and b_i in another block. Kreweras [18] showed that the number of noncrossing partitions with s_m blocks of size m is

$$a(n,\mathbf{s}) = \frac{(n)_{k-1}}{s_1!\, s_2! \cdots}, \quad \text{where } k = \sum s_m \text{ and } n = \sum m s_m.$$

Asymptotic results can be obtained by summing this formula over appropriate indices. Alternatively, one can study the ordinary generating function $A(x,\mathbf{z})$ for noncrossing partitions with z_m keeping track of s_m and x keeping track of n (the size of the set). By the argument leading to (6.2) of Beissinger [3],

$$A(x,\mathbf{z}) = 1 + \sum_m (xA(x,\mathbf{z}))^m z_m.$$

With $w(x,\mathbf{z}) = xA(x,\mathbf{z})$, we have

$$w(x,\mathbf{z}) = xf(w,\mathbf{z}) \ \text{ where } f(w,\mathbf{z}) = 1 + \sum_m z_m w^m.$$

By specializing z_m to 0, 1, and a finite set of indeterminates we can count various noncrossing partitions. For example,

$$z_1 = y_1, \quad z_m = y_2 \text{ for } m \in M \subseteq \{2,3,\ldots\}, \quad \text{and} \quad z_m = 0 \text{ otherwise},$$

counts noncrossing partitions whose blocks are singletons or have sizes in M, keeping track of the number of each type. To verify (5), fix $m_0 \in M$ and note that

$$\begin{aligned}\mathcal{A}\big(\{(1,1,0),\,(m,0,1) : m \in M\}\big) \\ &= \mathcal{A}\big(\{(1,1,0),\ (m-m',0,0),\ (m_0,0,1) : m,m' \in M\}\big) \\ &= \mathcal{A}\big(\{(1,1,0),\ (g,0,0),\ (m_0,0,1)\}\big),\end{aligned}$$

where $g = \gcd\{m-m' : m,m' \in M\}$. Hence (5) holds if and only if $g=1$. Thus

$$f(x,\mathbf{y}) = 1 + y_1x + y_2S_0, \quad \text{where } \ S_i = S_i(x) = \sum_{m\in M} m^i x^m,$$

and so

$$(n,k_1,k_2) = \mathbf{m} = (n/f)\big(y_1x + y_2S_1,\ y_1x,\ y_2S_0\big).$$

Since we want the zeroth component of $\mathbf{m}$ to be n, $f = y_1x + y_2S_1$ and so

$$y_2S_1 = 1 + y_2S_0 \ \text{ at } (x,y_1,y_2) = \mathbf{r}.$$

After some calculations,

$$B = \frac{n}{f^2}\begin{bmatrix} fy_2(S_2-S_1) & 0 & f \\ 0 & y_1x(f-y_1x) & -y_1xy_2S_0 \\ f & -y_1xy_2S_0 & y_2S_0(1+y_1x)\end{bmatrix} \ \text{ at } (x,y_1,y_2) = \mathbf{r}.$$

These equations can be used in the theorem to obtain asymptotics.

With k_1 the number of singleton blocks and k_2 the number of other blocks, we can get a local limit theorem for the distribution of (k_1, k_2) as $n \to \infty$ when noncrossing partitions of $\underline{n}$ are selected at random. To do this, we set $r_1 = r_2 = 1$ and use (7) to obtain the covariance matrix. It follows that the joint distribution of $(k_1 - n\mu_1)\big/\sqrt{n}$ and $(k_2 - n\mu_2)\big/\sqrt{n}$ is asymptotically normal with covariance matrix C where

$$S_i = S_i(r_0), \quad S_1 = 1 + S_0 \text{ determines } r_0, \quad f = r_0 + S_1 = 1 + r_0 + S_0,$$

$$\mu_1 = r_0/f, \quad \mu_2 = S_0/f,$$

and

$$C = \begin{bmatrix} r_0 S_1/f^2 & -r_0 S_0/f^2 \\ -r_0 S_0/f^2 & (1+r_0)S_0/f^2 \end{bmatrix} - \begin{bmatrix} 0 \\ 1/f \end{bmatrix} \frac{f}{S_2 - S_1} \begin{bmatrix} 0 & 1/f \end{bmatrix}.$$

For example, when M is the set of primes, $r_0 = 0.5580260$,

$$\mu_1 = 0.263674, \quad \mu_2 = 0.263815, \quad \text{and} \quad C = \begin{bmatrix} 0.194150 & -0.069561 \\ -0.069561 & 0.067667 \end{bmatrix}. \quad \blacksquare$$

Example 4.2. (Powers of an Inversion) Suppose $w(x, \mathbf{y}) = xf(w, \mathbf{y})$. How do the coefficients of $[x^n]\, w^k$ behave as $k \to \infty$ with n? Meir and Moon [19] studied the case when $\mathbf{y}$ was absent because $w(x) = xf(w(x))$ is associated with a variety of labeled and unlabeled tree enumerations and w^k counts forests with k components. The introduction of $\mathbf{y}$ allows us to keep track of additional information (such as vertex degrees), but we can still follow Meir and Moon's approach. Furthermore, when $\mathbf{y}$ is absent, we obtain their result. Since $g(w) = w^k$, Meir and Moon observed that Lagrange inversion gives

$$[x^n]\, w(x, \mathbf{y})^k = (1/n)[x^n]\left(xkx^{k-1} f(x, \mathbf{y})^n\right) = (k/n)[x^{n-k}]\, f(x, \mathbf{y})^n.$$

One can now apply Theorem 2.1 to obtain asymptotics. The zeroth component of $\mathbf{m}$ gives the equation

$$\frac{n-k}{n} f(x, \mathbf{y}) = x \frac{\partial f(x, \mathbf{y})}{\partial x}. \tag{8}$$

It follows that $\frac{n-k}{n}$ must be bounded away from 0 and so the value of k must be restricted to $1 \le k \le \alpha n$ where $\alpha < 1$. If (8) has a solution $(r_0, \mathbf{r})$ when $k = 0$ and if the power series for f has nonnegative coefficients, letting r_0 decrease toward 0 produces a solution for the same $\mathbf{r}$ and all larger values of k. In particular, when $\mathbf{r} = \mathbf{1}$, one obtains a local limit theorem for the distribution of the variables counted by $\mathbf{y}$, with means and covariance matrix proportional to n and their values depending on the value of k/n. $\blacksquare$

Example 4.3. (Plane Trees by Vertex Degree) A planted plane tree is a rooted plane tree in which the root has degree 1. If x counts nonroot vertices and y_k counts nonroot vertices of degree k, then the generating function satisfies

$$w(x, \mathbf{y}) = x \sum_{k \ge 0} z_{k+1} w^k. \tag{9}$$

Goulden and Jackson [15, Sec.2.7.7] obtain the formula

$$[x^n \mathbf{y}^{\mathbf{k}}]\, w(x,\mathbf{y}) = \frac{(n-1)!}{\prod k_i!},$$

provided $\sum k_i = n$ and $\sum ik_i = 2n-1$, and is zero otherwise. where the last factor is a multinomial coefficient If one wishes to keep track of only a few degrees, say those in a finite set D, summing this formula could be impractical. In Exercise 2.7.2 of [15], Goulden and Jackson obtain formulas when D is a singleton or a pair of degree. The former is an alternating sum and the latter an alternating double sum.

Specializing (9) by setting $y_k = 1$ for $k \notin D$, we can apply the theorem with $h = x$ and

$$f = \sum_{k \notin D} x^{k-1} + \sum_{k \in D} y_k x^{k-1} = \frac{1}{1-x} + \sum_{k\in D} (y_k - 1)x^{k-1} \tag{10}$$

Since f has positive coefficients, we now verify (5). If k is the jth element in D, let $\mathbf{e}_k$ be the unit vector whose jth component is 1 and let $i \notin D$ be fixed. Since $\gcd\{i-j \mid i,j \notin D\} = 1$ when D is finite,

$$\begin{aligned}\mathcal{A}\Big\{\mathbf{k}-\mathbf{j} \;\Big|\; a(\mathbf{k})a(\mathbf{j}) \neq 0\Big\} &= \mathcal{A}\Big\{(i-j,\mathbf{0}),\ \ (k-j,\mathbf{e}_k) \;\Big|\; i,j \notin D,\ k \in D\Big\} \\ &= \mathcal{A}\Big\{(1,\mathbf{0}),\ \ (k-j,\mathbf{e}_k) \;\Big|\; j \notin D,\ k \in D\Big\} \\ &= \mathbb{Z}^{1+|D|}.\end{aligned}$$

One easily computes that the component of $\mathbf{m}$ associated with x is

$$m_0 = \frac{n}{f}\left(\frac{x}{(1-x)^2} + \sum_{k\in D}(k-1)(y_k-1)x^{k-1}\right), \tag{11}$$

and that associated with y_k is

$$m_k = n y_k x^{k-1}/f. \tag{12}$$

After some calculation, and using the fact that $m_0 = n$ for Lagrange inversion, we find that

$$\begin{aligned} b_{0,0} &= \frac{n}{f}\left(\frac{x(1+x)}{(1-x)^3} + \sum_{k\in D}(k-1)^2(y_k-1)x^{k-1}\right) - 1, \\ b_{0,k} &= (k-2)m_k/f, \\ b_{k,k} &= m_k(1-m_k/n), \\ b_{k,j} &= -m_k m_j/n, \quad k \neq j. \end{aligned}$$

We can use the theorem to obtain either asymptotics or a local limit theorem.

To obtain asymptotics, we want $\mathbf{m}$ to give the number of vertices of each type so that then $\mathbf{t} = \mathbf{0}$ in (4), which will give the greatest accuracy. The values of r_0

and r_k are given by setting $x = r_0$ and $y_k = r_k$ and then combining (10), (11), and (12): With $\mu_k = m_k/n$, the fraction of vertices of degree k, we have

$$\frac{1}{1-r_0} + \sum_{k\in D}(\mu_k - r_0^{k-1}) = \frac{r_0}{(1-r_0)^2} + \sum_{k\in D}(k-1)(\mu_k - r_0^{k-1}),$$

which can be solved numerically for r_0 once D and the fractions μ_k are given. Then $r_k = \mu_k/r_0^{k-1}$. Using these values in the formulas for $b_{i,j}$ and then in (4) with $\mathbf{t} = \mathbf{0}$ gives the asymptotics.

The local limit theorem is easily obtained since we simply set $y_k = r_k = 1$ for $k \in D$ and $x = r_0$. This leads to

$$r_0 = 1/2, \quad f = 2, \quad \mu_k = 2^{-k}, \quad b_{0,0} = 2n.$$

Using (7), we obtain

$$\frac{c_{k,k}}{n} = \mu_k - \mu_k^2\left(1 + \frac{(k-2)^2}{2}\right) \quad \text{and} \quad \frac{c_{k,j}}{n} = -\mu_k\,\mu_j\left(1 + \frac{(k-2)(j-2)}{2}\right).$$

We could equally well have looked at out-degrees in simply generated families of trees. In that case, (10) becomes

$$f = \sum_{k\ge 1} f_k x^k + \sum_{k\in D} f_k(y_k - 1)x^k,$$

and the analysis proceeds as above. In particular, when D is a singleton set, we recover Theorem 1(i) of Meir and Moon [20]. ∎

Example 4.4. (3-Connected Rooted Maps) The asymptotics for 3-connected rooted maps by number of edges were determined by Tutte [25]. We use Mullin and Schellenberg's parameterization [21]. They found that the generating function with $x^m y^n$ counting 3-connected rooted planar maps with $m+1$ vertices and $n+1$ faces is

$$p(x,y) = \left(\frac{1}{1+x} + \frac{1}{1+y} - 1\right)xy - \frac{rs}{(r+s+1)^3},$$

where $r = x(s+1)^2$ and $s = y(r+1)^2$. Setting $x = y$ and $r = s$, we obtain the generating function by number of edges since $E = V + F - 2$ by Euler's relation. For asymptotic purposes, we can ignore the first part of $p(x,y)$ and look at

$$[x^n]\left(\frac{-r^2}{(1+2r)^3}\right) \quad \text{where} \quad r = x(1+r)^2.$$

We have

$$\begin{aligned}[x^n]\left(\frac{-r^2}{(1+2r)^3}\right) &= n^{-1}[x^n]\left(x\left(\frac{-x^2}{(1+2x)^3}\right)'(x+1)^{2n}\right)\\ &= n^{-1}[x^n]\left(\frac{2x^2(x-1)}{(1+2x)^4}(x+1)^{2n}\right).\end{aligned} \tag{13}$$

The fraction in the last line is $h(x)$. Since condition (6) becomes $1 - \frac{2r_0}{1+r_0} = 0$, $h(r_0) = 0$ and Theorem 2.1 fails to apply. Fortunately, we can rewrite the last line of (13) in the form of (1) and then use (2):

$$\begin{aligned} n^{-1}[x^n]\left(\frac{2x^2(x-1)}{(1+2x)^4}(x+1)^{2n}\right) &= n^{-1}[x^n]\left(\frac{-2x^2(x+1)}{(1+2x)^4}(x+1)^{2n}\left(1-\frac{2x}{1+x}\right)\right) \\ &= n^{-2}[x^n]\left(x\left(\frac{-2x^2(1+x)}{(1+2x)^4}\right)'(x+1)^{2n}\right) \\ &= n^{-2}[x^n]\left(\frac{2x^2(2x^2+x-2)}{(1+2x)^5}(x+1)^{2n}\right). \end{aligned}$$

Noting that $r_0 = 1$ and putting all this in the theorem we find that the number of rooted 3-connected maps with n edges is asymptotic to $2^{2n+1} \,/\, 3^5\sqrt{n\pi}\, n^2$. ∎

5. Lagrange Inversion of Several Functions

Suppose we have

$$w_i(\mathbf{x},\mathbf{y}) = x_i f_i(\mathbf{w}(\mathbf{x},\mathbf{y}),\mathbf{y}) \quad \text{for} \quad 1 \le i \le d \tag{14}$$

and want $[\mathbf{x^n y^j}]\, g(\mathbf{w},\mathbf{y})$. The two forms of Lagrange inversion for several equations that parallel (1) and (2), respectively, are [9]

$$[\mathbf{x^n y^j}]\, g(\mathbf{w}(\mathbf{x},\mathbf{y}),\mathbf{y})$$
$$= [\mathbf{x^n y^j}]\left\{ g(\mathbf{x},\mathbf{y})\, \mathbf{f}(\mathbf{x},\mathbf{y})^{\mathbf{n}} \left\| \delta_{i,j} - \frac{x_i}{f_j(\mathbf{x},\mathbf{y})} \frac{\partial f_j(\mathbf{x},\mathbf{y})}{\partial x_i} \right\| \right\} \tag{15}$$
$$= \frac{1}{\prod n_i}\, [\mathbf{x^n y^j}] \sum_{\mathcal{T}} \mathbf{x^1} \frac{\partial(g, f_1^{n_1}, \ldots, f_d^{n_d})}{\partial \mathcal{T}}, \tag{16}$$

where $\mathbf{1}$ is the vector of all ones, $\delta_{i,j}$ is the Kronecker delta,

- $\|D\|$ is the determinant of the $d \times d$ matrix D,
- the vector $(g, f_1^{n_1}, \ldots, f_d^{n_d})$ has indices $0, \ldots, d$,
- $\mathcal{T}$ runs over all trees with vertex set indexed on $0, \ldots, d$ and edges directed toward 0, and
- for a directed graph $\mathcal{D}$ whose vertex set V is the indices of $\mathbf{h}$ and $\mathbf{x}$

$$\frac{\partial \mathbf{h}}{\partial \mathcal{D}} = \prod_{j \in V} \left\{ \left(\prod_{(i,j)\in E} \frac{\partial}{\partial x_i} \right) h_j(\mathbf{x}) \right\},$$

where E is the edge set of D.

To apply the theorem, we let $\mathbf{u} = (\mathbf{x}, \mathbf{y})$, $\mathbf{k} = \mathbf{i} = (\mathbf{n}, \mathbf{j})$, and we let h be what is left in (15) or (16) after removing $\mathbf{f}^{\mathbf{n}}$. By Theorem 2.1,

$$[\mathbf{x}^{\mathbf{n}}\mathbf{y}^{\mathbf{j}}]\, g(\mathbf{w}(\mathbf{x},\mathbf{y}),\mathbf{y}) \sim \frac{h(\mathbf{r})\mathbf{f}(\mathbf{r})^{\mathbf{n}}}{\sqrt{\det(2\pi B)}\;\mathbf{r}^{\mathbf{k}}}, \tag{17}$$

where $B = B(\mathbf{f}(\mathbf{r})^{\mathbf{n}})$ and $\mathbf{k} = \mathbf{m}\big(\mathbf{f}^{\mathbf{n}}\big)$. The last equation can be written

$$\mathbf{n}L(\mathbf{r}) = \mathbf{0} \quad \text{where} \quad L_{i,j}(\mathbf{x}) = \delta_{i,j} - \frac{x_j}{f_i}\frac{\partial f_i}{\partial x_j}, \tag{18}$$

where $\delta_{i,j}$ is the Kronecker delta. Thinking of $\mathbf{n}$ as unknown, we see the condition for this set of equations to have a nontrivial solution is precisely that the determinant in (15) vanish, and so h will violate condition (a) in Theorem 2.1. Hence we use (16) rather than (15). This is the multiple inversion case of what happened with (1).

We now turn our attention to conditioning. Since the discussion is somewhat involved, you may wish to read Example 6.1 beforehand.

To discuss conditioning, we first need an appropriate local limit theorem. It will be simpler not to distinguish between the variables $\mathbf{x}$ and $\mathbf{y}$. This can be done by supplementing (14) with the additional equations $w_i = y_i$, which means $f_i = 1$. In this way, we eliminate references to $\mathbf{y}$ and incorporate $\mathbf{j}$ in $\mathbf{n}$. Since the new $f_i = 1$, (17) and (18) are still valid and $L_{\text{new}} = L_{\text{old}} \oplus I$, where I is an identity matrix indexed for the added Lagrange equations $w_i = y_i$.

Theorem 5.1. *Let $\mathbf{r}$ be the solution to $\mathbf{i}L(\mathbf{r}) = \mathbf{0}$. Under the assumptions of Theorem 2.1, with the extra variables $\mathbf{y}$ eliminated as described above, we have*

$$[\mathbf{x}^{\mathbf{k}}]\,\big(h(\mathbf{x})\mathbf{f}(\mathbf{x})^{\mathbf{k}}\big) = \frac{h(\mathbf{r})\mathbf{f}(\mathbf{r})^{\mathbf{k}}\big\{\exp\big(-\mathbf{t}LB^{-1}L'\mathbf{t}'/2\big) + o(1)\big\}}{\sqrt{\det(2\pi B)}\;\mathbf{r}^{\mathbf{k}}} \tag{19}$$

uniformly as $\|\mathbf{k}\| \to \infty$, where $L = L(\mathbf{r})$, $B = B(\mathbf{f}(\mathbf{r})^{\mathbf{k}})$, and $\mathbf{t} = \mathbf{k} - \mathbf{i}$. Since $\mathbf{i}L = \mathbf{0}$, we can replace $\mathbf{t}$ by $\mathbf{k}$.

Equation (19) does not give a distribution becausee $\det(L) = 0$ implies that $LB^{-1}L'$ is singular. It turns out that conditioning leads to a nonsingular matrix.

Proof: Let $n = \|\mathbf{k}\|$. Let $\mathbf{r}^*$ be the solution to $\mathbf{k}L = \mathbf{0}$. Let $\mathbf{s}$ and $\mathbf{s}^*$ be the the componentwise logarithms of $\mathbf{r}$ and $\mathbf{r}^*$, respectively. Let B denote $B(\mathbf{f}(\mathbf{r})^{\mathbf{k}})$. Note that $B = O(n)$ and $\det(B)$ is of order $n^{\dim(B)}$, where the latter follows from (a) B is positive definite, (b) $B(f_i)$ is positive semidefinite, and (c) k_i/n is bounded away from 0. It follows that $B^{-1} = O(1/n)$. These results also hold for $B(\mathbf{f}(\mathbf{r}^*)^{\mathbf{k}})$.

For now, we assume that $\|\mathbf{t}L\| \le n^{3/5}$. By Taylor series with remainder,

$$\mathbf{k}L(\mathbf{r}^*) = \mathbf{k}L(\mathbf{r}) + (\mathbf{s}^* - \mathbf{s})B + O\left(\|\mathbf{s}^* - \mathbf{s}\|^2 n\right). \tag{20}$$

Since $\mathbf{k}L(\mathbf{r}^*) = \mathbf{0}$, it follows from the above that

$$\|\mathbf{s}^* - \mathbf{s}\| = O(n^{3/5})B^{-1} = O(n^{-2/5})$$

and so, from (20),

$$\begin{aligned}\mathbf{s}^* - \mathbf{s} &= -\mathbf{k}L(\mathbf{r}^*)B^{-1} + O(n^{-4/5}n)B^{-1}\\ &= -\mathbf{t}L(\mathbf{r}^*)B^{-1} + O(n^{-4/5}) = O(n^{-2/5}).\end{aligned} \tag{21}$$

Since the components of $\mathbf{r}$ are bounded away from 0 and ∞, we also have $\mathbf{r}^* - \mathbf{r} = O(n^{-2/5})$. We now consider the expansion of the logarithm of the ratio

$$\frac{h(\mathbf{r})\mathbf{f}(\mathbf{r})^{\mathbf{k}}/\mathbf{r}^{\mathbf{k}}}{\sqrt{\det\bigl(2\pi B(\mathbf{f}(\mathbf{r})^{\mathbf{k}})\bigr)}} \Bigg/ \frac{h(\mathbf{r}^*)\mathbf{f}(\mathbf{r}^*)^{\mathbf{k}}/(\mathbf{r}^*)^{\mathbf{k}}}{\sqrt{\det\bigl(2\pi B(\mathbf{f}(\mathbf{r}^*)^{\mathbf{k}})\bigr)}} \tag{22}$$

in a Taylor series about $\mathbf{s}^*$. Note that

$$h(\mathbf{r}^*) = h(\mathbf{r}) + O(\|\mathbf{r}^* - \mathbf{r}\|) = h(\mathbf{r})(1 + O(n^{-2/5})$$

since h is bounded away from 0. Since

$$B = B(\mathbf{f}(\mathbf{r}^*)^{\mathbf{k}}) + O(n\|\mathbf{r} - \mathbf{r}^*\|) = B(\mathbf{f}(\mathbf{r}^*)^{\mathbf{k}}) + O(n^{3/5}).$$

and $B^{-1} = O(1/n)$, it follows that

$$B(\mathbf{f}(\mathbf{r})^{\mathbf{k}})^{-1}B(\mathbf{f}(\mathbf{r}^*)^{\mathbf{k}}) = I + O(B^{-1}n\|\mathbf{r} - \mathbf{r}^*\|) = I + O(n^{-2/5}). \tag{23}$$

Finally, with $\Delta\mathbf{s} = \mathbf{s} - \mathbf{s}^*$,

$$\begin{aligned}\log\left(\frac{\mathbf{f}(\mathbf{r})^{\mathbf{k}}(\mathbf{r}^*)^{\mathbf{k}}}{\mathbf{f}(\mathbf{r}^*)^{\mathbf{k}}\mathbf{r}^{\mathbf{k}}}\right) &= -\mathbf{k}L(\mathbf{f}(\mathbf{r}^*))(\Delta\mathbf{s})' + \tfrac{1}{2}\Delta\mathbf{s}B(\mathbf{f}(\mathbf{r}^*)^{\mathbf{k}})(\Delta\mathbf{s})' + O\left(n\|\Delta\mathbf{s}\|^3\right)\\ &= \tfrac{1}{2}\Delta\mathbf{s}B(\mathbf{f}(\mathbf{r}^*)^{\mathbf{k}})(\Delta\mathbf{s})' + O(n^{-1/5}),\end{aligned}$$

since $\mathbf{k}L(\mathbf{r}^*) = \mathbf{0}$. Combining this with (21),

$$\log\left(\frac{\mathbf{f}(\mathbf{r})^{\mathbf{k}}(\mathbf{r}^*)^{\mathbf{k}}}{\mathbf{f}(\mathbf{r}^*)^{\mathbf{k}}\mathbf{r}^{\mathbf{k}}}\right) = \tfrac{1}{2}\mathbf{t}L(\mathbf{r})B^{-1}B(\mathbf{f}(\mathbf{r}^*)^{\mathbf{k}})B^{-1}L(\mathbf{r})'\mathbf{t}' + O(n^{-1/5}). \tag{24}$$

From (23),

$$B^{-1}B(\mathbf{f}(\mathbf{r}^*)^{\mathbf{k}})B^{-1} = B^{-1}(I + O(n^{-2/5})) = B^{-1} + O(n^{-7/5})$$

and $B(\mathbf{f}(\mathbf{r}^*)^{\mathbf{k}})^{-1} = B^{-1} + O(n^{-7/5})$. Combining the various equations and (4) with $\mathbf{i} = \mathbf{k}$ and the fact that $\mathbf{t}L(\mathbf{r}) = O(n^{3/5})$, we obtain (19).

We now assume that $\|\mathbf{t}L\| \geq n^{3/5}$. In this case, the exponential in (19) is $o(1)$. Thus, it suffices to prove that (22) tends to ∞ with n. The compactness of R and the nonsingularity of B assure us that $h(\mathbf{r})/h(\mathbf{r}^*)$ and the ratios of the square roots in (22) are bounded away from 0 and ∞. Thus it suffices to consider the ratio $F(\mathbf{s})/F(\mathbf{s}^*)$ where

$$F(\mathbf{u}) = \mathbf{f}(\mathbf{x})^{\mathbf{k}}/\mathbf{x}^{\mathbf{k}}, \quad u_i = \log x_i, \quad \text{and} \quad \mathbf{x} \in R.$$

We may assume that the region R, viewed in $\mathbf{s}$ coordinates is convex: Other than the compactness requirement, the main restriction on R was that the various power series converge. If $\sum a_{\mathbf{n}} \exp(\mathbf{n}\cdot\mathbf{s})$ and $\sum a_{\mathbf{n}} \exp(\mathbf{n}\cdot\mathbf{s}^*)$ converge absolutely, then so does $\sum a_{\mathbf{n}} \exp\big(\mathbf{n}\cdot(\lambda\mathbf{s} + (1-\lambda)\mathbf{s}^*)\big)$ because, for series with nonnegative terms,

$$\begin{aligned}\sum A_{\mathbf{n}}^{\lambda} B_{\mathbf{n}}^{1-\lambda} &\le \sum \max(A_{\mathbf{n}}, B_n)^{\lambda} \max(A_{\mathbf{n}}, B_n)^{1-\lambda} \\ &= \sum \max(A_n, B_n) \le \sum A_n + \sum B_n.\end{aligned}$$

By standard calculus, the gradient and matrix of second derivatives of F are $\nabla F = \mathbf{k}L(\mathbf{x})$ and $B(\mathbf{f}(\mathbf{x})^{\mathbf{k}})$. Since B is positive definite, F is convex and so has just one minimum, which is given by the solution $\mathbf{s}^*$ to $\nabla F(\mathbf{u}) = \mathbf{0}$. Think of $\mathbf{k}$ as fixed and $\mathbf{i}$ as variable. (Thus $\mathbf{s}$ is also variable.) Using (24) for $\|\mathbf{t}L\| = n^{3/5}$, we see that $F(\mathbf{s}^*) = o(F(\mathbf{s}))$ uniformly for such values. From the convexity of F, it follows that $F(\mathbf{s}^*) = o(F(\mathbf{s}))$ for $\|\mathbf{t}L\| \ge n^{3/5}$. ∎

Conditioning and Summing. Let $\mathcal{C}$ be the indices of variables on which we are conditioning and $\mathcal{N}$ the remaining indices. One uses the equation $\mathbf{i}L = \mathbf{0}$ to determine r_j for $j \in \mathcal{C}$ and i_j for $j \in \mathcal{N}$. In addition, one has the equations $f_j(\mathbf{r})/r_j = 1$ for $j \in \mathcal{N}$. The latter equations guarantee that i_j will be chosen to be at the peak of the distribution when $j \in \mathcal{N}$. One extracts the diagonal submatrix of $LB^{-1}L'$ that is indexed by $\mathcal{N}$. If we condition on *all* the original $\mathbf{x}$, then $r_j = 1$ for $j \in \mathcal{N}$ (since $f_j = 1$) and the relevant portion of L is the identity matrix so we are just extracting a diagonal submatrix of B^{-1}; all of which is as described in Section 3 for the single Lagrange equation. In particular, (7) applies with $B_{1,1}$ indexed by $\mathcal{C}$ and $B_{2,2}$ and C indexed by $\mathcal{N}$.

The theorem can also be used for summing over variables. Suppose we have a formula as in (19) and want to sum over certain components of $\mathbf{t}$. Let the set of indices be $\mathcal{S}$ The corresponding $\mathbf{i}$ components must be chosen so that we are the peak; that is, $f_j(\mathbf{r})/r_j = 1$ for $j \in \mathcal{S}$. Partition $M = LB^{-1}L'$ into a 2×2 block matrix where the subscript 1 refers to elements of $\mathcal{S}$. In summing (19), the tail will be negligible because of the convexity discussed in the last paragraph of the theorem's proof. Thus we are faced with the problem of summing $\exp(-z/2)$ over $\mathbf{t}_1$ where

$$\begin{aligned} z &= (\mathbf{t}_1, \mathbf{t}_2)\begin{bmatrix} M_{1,1} & M_{1,2} \\ M_{2,1} & M_{2,2}\end{bmatrix}(\mathbf{t}_1, \mathbf{t}_2)' \\ &= (\mathbf{t}_1 + \mathbf{t}_2 T)M_{1,1}(\mathbf{t}_1 + \mathbf{t}_2 T)' + \mathbf{t}_2(M_{2,2} - TM_{1,1}T')\mathbf{t}_2', \end{aligned}$$

and T satisfies $TM_{1,1} = M_{2,1}$. (We have used the fact that M is symmetric.) In order to be able to carry out the summation, $M_{1,1}$ must be nonsingular. In that case, we obtain a factor of $\left(\det(2\pi M_{1,1}^{-1})\right)^{1/2}$ from the summation and the exponential becomes

$$\exp\left(-\mathbf{t}_2(M_{2,2} - TM_{1,1}T')\mathbf{t}_2'/2\right),$$

which is still singular if M is singular. If M is not singular, the above matrix is the inverse of the lower righthand block of M^{-1}.

(In this case, an alternative proof can be found in Section 3.4 of Press [22].) Conditioning, before or after summing, will normally make the matrix nonsingular.

Before conditioning and/or summing, one may wish to make a linear change of variables. For example, if m_k is the number of vertices of type k, one might introduce the coordinate $n = \sum m_k$, the total number of vertices, and condition on it. After changing coordinates, the rules for selecting $\mathbf{r}$ and $\mathbf{i}$ differ somewhat, but the underlying principles are the same: If $\mathbf{n} = \mathbf{m}A$ gives the old coordinates $\mathbf{n}$ in terms of the new coordinates $\mathbf{m}$, then

$$\frac{\mathbf{f}(\mathbf{r})^{\mathbf{n}}}{\mathbf{r}^{\mathbf{n}}} = \frac{\mathbf{f}^*(\mathbf{r})^{\mathbf{m}}}{(\mathbf{r}^*)^{\mathbf{m}}},$$

where $\log \mathbf{f}^* = (\log \mathbf{f})A$ and $\log \mathbf{r}^* = (\log \mathbf{r})A$. The earlier rules now apply with $\mathbf{f}^*(\mathbf{r})$, $\mathbf{r}^*$, and $\mathbf{i}A^{-1}$ in place of $\mathbf{f}(\mathbf{r})$, $\mathbf{r}$, and $\mathbf{i}$. (This is illustrated in the next example.)

6. Examples of Inversion of Several Functions

We now turn to examples of inversion of more than one function. Because of the intimate connection between Lagrange inversion and tree enumeration, it is not surprising that Lagrange inversion is often ideal for studying tree enumeration questions. For a single function, this can be seen in the examples of the Section 4. Good [14] may have been the first to point this out for inversion of several functions. In this connection, see also Goulden and Jackson [15]. As the examples in this section illustrate, Lagrange inversion of several functions leads to more complicated calculations than occur for a single function. Thus one should reduce the inversion problem to a single function when possible.

Example 6.1. (Plane Rooted Colored Trees) As usual, the set of colors is finite. We consider situations in which local conditions determine the possible colors of a vertex. Similar methods apply to rooted labeled colored trees (no longer planar). Examples of situations that can be dealt with in this manner are:

- A vertex must have a different color from its children (proper coloring).
- The children of a vertex must have distinct colors.
- A vertex with grandchildren must have the same color as a grandchild.

The more complex the conditions and the greater the number of colors, the greater the number of equations that must be inverted, and so the more complicated the calculations.

We begin by considering trees with green and red vertices. Our local condition will be that a nonleaf green vertex must have exactly one red child and two green children, while a nonleaf red vertex must have exactly one child of each color with the left one being red. Associate the subscript 1 with green and 2 with red. Let t_i keep track of number of vertices of color i and let w_i be the generating function for trees by root color. The conditions translate to

$$\begin{aligned} w_1(\mathbf{x}) &= x_1 f_1(\mathbf{w}), \quad \text{where} \quad f_1(\mathbf{w}) = 1 + 3w_1(\mathbf{x})^2 w_2(\mathbf{x}), \\ w_2(\mathbf{x}) &= x_2 f_2(\mathbf{w}), \quad \text{where} \quad f_2(\mathbf{w}) = 1 + w_1(\mathbf{x}) w_2(\mathbf{x}). \end{aligned}$$

Lagrange inversion gives us $c(\mathbf{k})$, the number of trees with k_1 green and k_2 red vertices:

$$c(\mathbf{k}) = [\mathbf{x}^{\mathbf{k}}]\,(w_1 + w_2).$$

There are 3 directed trees in the sum over $\mathcal{T}$. Their edge sets are

$$E_1 = \{(1,0),\ (2,1)\}, \quad E_2 = \{(2,0),\ (1,2)\}, \quad \text{and} \quad E_3 = \{(1,0),\ (2,0)\}. \tag{25}$$

Thus

$$\begin{aligned} c(\mathbf{k}) &= (k_1k_2)^{-1}[\mathbf{x}^{\mathbf{k}}]\left\{ x_1x_2 \left(\frac{\partial(x_1+x_2)}{\partial x_1} \frac{\partial f_1(\mathbf{x})^{k_1}}{\partial x_2} f_2(\mathbf{x})^{k_2} \right.\right. \\ &\qquad \left.\left. + \frac{\partial(x_1+x_2)}{\partial x_2} \frac{\partial f_2(\mathbf{x})^{k_2}}{\partial x_1} f_1(\mathbf{x})^{k_1} + \frac{\partial^2(x_1+x_2)}{\partial x_1 \partial x_2} f_1(\mathbf{x})^{k_1} f_2(\mathbf{x})^{k_2} \right) \right\} \\ &= (k_1k_2)^{-1}[\mathbf{x}^{\mathbf{k}}]\left\{ x_1x_2 \left(\frac{3k_1x_1^2}{f_1(\mathbf{x})} f_1(\mathbf{x})^{k_1} f_2(\mathbf{x})^{k_2} + \frac{k_2x_2}{f_2(\mathbf{x})} f_1(\mathbf{x})^{k_1} f_2(\mathbf{x})^{k_2} \right) \right\}. \end{aligned}$$

We apply Theorem 2.1 with $\mathbf{n} = (k_1, k_2)$ and $h = 3x_1^3x_2/f_1(\mathbf{x}) + x_1x_2^2/f_2(\mathbf{x})$. We set $\mathbf{k} = \mathbf{i}$ and $\nu = \mathbf{0}$ in (4). The equation $\mathbf{i} = \mathbf{m}(\mathbf{f}(\mathbf{r})^{\mathbf{i}})$ is

$$\begin{aligned} k_1 &= \frac{6r_1^2r_2}{1+3r_1^2r_2} k_1 + \frac{r_1r_2}{1+r_1r_2} k_2 \\ k_2 &= \frac{3r_1^2r_2}{1+3r_1^2r_2} k_1 + \frac{r_1r_2}{1+r_1r_2} k_2 \end{aligned} \tag{26}$$

and (4) becomes

$$c(\mathbf{k}) \sim \left(\frac{3r_1^3r_2}{k_2(1+3r_1^2r_2)} + \frac{r_1r_2^2}{k_1(1+r_1r_2)} \right) \frac{(1+3r_1^2r_2)^{k_1}(1+r_1r_2)^{k_2}}{\sqrt{\det(2\pi B)}\, r_1^{k_1} r_2^{k_2}}, \tag{27}$$

where $B = k_1B(f_1) + k_2B(f_2)$. To verify (5), note that

$$\Lambda(\mathbf{f}) = \mathcal{A}\{(2,1),\ (1,1)\} = \mathbb{Z}^2.$$

It remains to compute $\mathbf{r}$, verify that $h > 0$, compute B, and, perhaps, simplify (27), details of which we omit. One can check that, with $1 < \rho = k_1/k_2 < 2$,

$$r_1 = \frac{(\rho-1)^2}{3(2-\rho)} \quad \text{and} \quad r_2 = \frac{3(2-\rho)^2}{(\rho-1)^3} \tag{28}$$

r_1 and r_2 are positive and (26) is satisfied. Hence the theorem applies if, for some $\epsilon > 0$, we have $1 + \epsilon \le k_1/k_2 \le 2 - \epsilon$ as $\|\mathbf{k}\| \to \infty$.

Let $n = k_1 + k_2$. After some calculations

$$\begin{aligned}
B(f_1) &= \frac{3r_1^2 r_2}{(1+3r_1^2 r_2)^2}\begin{bmatrix}4 & 2\\ 2 & 1\end{bmatrix} = \frac{\rho-1}{\rho^2}\begin{bmatrix}4 & 2\\ 2 & 1\end{bmatrix}\\
B(f_2) &= \frac{r_1 r_2}{(1+r_1 r_2)^2}\begin{bmatrix}1 & 1\\ 1 & 1\end{bmatrix} = (2-\rho)(\rho-1)\begin{bmatrix}1 & 1\\ 1 & 1\end{bmatrix}\\
B(\mathbf{f^n}) &= k_1 B(f_1) + k_2 B(f_2) = \frac{n\rho}{\rho+1}B(f_1) + \frac{n}{\rho+1}B(f_2)\\
&= \frac{n(\rho-1)}{\rho(\rho+1)}\begin{bmatrix}4+2\rho-\rho^2 & 2+2\rho-\rho^2\\ 2+2\rho-\rho^2 & 1+2\rho-\rho^2\end{bmatrix}.
\end{aligned}$$

One can eliminate $\mathbf{r}$ from (27) by using (28), and $\mathbf{k}$ by using $\mathbf{k} = n\left(\frac{\rho}{\rho+1}, \frac{1}{\rho+1}\right)$.

We now obtain a local limit theorem for the number of red vertices in trees having a fixed number of vertices.

To do so, we use (19), change coordinates from $\mathbf{k}$ to (n, k_2), choose i_1/i_2 so that we are at a peak, and condition on n. Since the change of coordinates is given by

$$(k_1, k_2) = (n, k_2)A \quad \text{where} \quad A = \begin{bmatrix}1 & 0\\ -1 & 1\end{bmatrix}, \tag{29}$$

$\mathbf{t}LB^{-1}L'\mathbf{t}' = \mathbf{u}(ALB^{-1}L'A')\mathbf{u}'$, where $\mathbf{u} = (n - i_1 - i_2,\ k_2 - i_2)$. To condition on $u_1 = 0$, we delete the first row and column of $ALB^{-1}L'A'$, which leaves a single number, say $1/n\sigma^2$. Thus the number of trees satisfies

$$c(n, k_2) = \frac{h(\mathbf{r})(f_1/r_1)^n\{\exp(-(k_2-i_2)^2/2n\sigma^2) + o(1)\}}{\sqrt{\det(2\pi B)}}\left(\frac{f_2 r_1}{f_1 r_2}\right)^{k_2}.$$

To be a peak as a function of k_2, the last factor must be 1; that is $f_1/r_1 = f_2/r_2$. With Maple's help we found that $\rho = 1.73473$. Using this in the preceding formula, we obtain

$$c(n, k_2) = \frac{C_1 C_2^n\{\exp(-(k_2-n\mu)^2/2n\sigma^2) + o(1)\}}{2\pi n^2 C_3},$$

where $C_1 = 1.00829$, $C_2 = 2.55726$, $C_3 = 0.105061$, $\mu = 0.36567$, and $\sigma = 0.110205$. (The extra factor of n in the denominator is due to $h(\mathbf{r})$.) Of course, since the local limit theorem states that k_2 is normally distributed with mean $n\mu$ and variance $n\sigma^2$, it does not require the various C_i values.

We now turn our attention to the last problem raised at the start of this example, namely, a vertex with grandchildren must have the same color as a grandchild. We want to count trees according to the number of vertices of each color. This problem has a new feature: with each vertex we must keep track of its color and the set of colors of its children, with the empty set arising when a vertex has no children. As a result, we consider generating functions $T_{c,S}(\mathbf{x})$ for trees where c is the root color and S is the set of children's colors. This leads to functional equations of the form

$$T_{c,S} = x_c\, f_{c,S}(\mathbf{T}),$$

which are inappropriate for Lagrange inversion. Consider instead

$$T_{c,S} = x_{c,S}\, f_{c,S}(\mathbf{T}).$$

We may think of $x_{c,S}$ as keeping track of $k_{c,S}$, the number of vertices of color c whose children's colors are S. After applying (19) to obtain asymptotics, we change coordinates to $k_c = \sum_T k_{c,T}$ and $k_{c,S}$ with $S \neq \emptyset$, then we sum over all values of $k_{c,S}$ (with $S \neq \emptyset$) to obtain a result in terms of just the k_c. The change of coordinates is done as in (29). The method for summation is described at the end of Section 5. After changing coordinates, the condition for setting the components of $\mathbf{r}$ becomes "$f_{c,S}(\mathbf{r})/r_{c,S}$ must be independent of S." We omit the details. ■

Example 6.2. (Plane Trees by Vertex Degree, Continued) Counting planted plane trees by vertex degree was treated in Example 4.3. We now consider a multiple Lagrange equation approach so that one can compare the two approaches.

Let the sets S_i be a finite partition of a subset of $\{1, 2, \ldots\}$ with $1 \in S_1$. (We need 1 to have leaves.)

Let y_i keep track of the number of vertices having degree in S_i.

To follow the idea of Example 4.3, we introduce one further variable x keeping track of all vertices except the root and write the single equation

$$T(x, \mathbf{y}) = xf(T(x, \mathbf{y}), \mathbf{y}) \quad \text{where} \quad f(w, \mathbf{y}) = \sum_i y_i \left(\sum_{k \in S_i} w^{k-1} \right),$$

which is dealt with as follows: The equations

$$n = \frac{nx\partial f(x, \mathbf{y})/\partial x}{f(x, \mathbf{y})} \quad \text{and} \quad k_i = \frac{ny_i \partial f(x, \mathbf{y})/\partial y_i}{f(x, \mathbf{y})} \tag{30}$$

must be solved for $x = r_0$ and $\mathbf{y} = \mathbf{r}$ if we have values of n and $\mathbf{k}$ in mind. If we want to study the distribution of $\mathbf{k}$ conditioned on n, we must set $\mathbf{r} = \mathbf{1}$ and solve the first equation from (30),

$$1 = \frac{r_0 \partial f(r_0, \mathbf{1})/\partial r_0}{f(r_0, \mathbf{1})} \tag{31}$$

for r_0. The remaining equations in (30) determine $\mathbf{k}/n$, the value of $\mathbf{k}$ that gives the peak of the normal. We can then proceed with conditioning as in the previous example.

If we do not introduce x, it is natural to introduce $T_i(\mathbf{y})$, the enumerator for trees where the degree of the root's son is in S_i. We have

$$T_i(\mathbf{y}) = y_i f_i(T(\mathbf{y})), \quad \text{where} \quad f_i(T) = \sum_{k\in S_i} T^{k-1} \quad \text{and} \quad T(\mathbf{y}) = \sum_i T_i(\mathbf{y}).$$

This leads to the equations

$$k_i = \sum_j k_j r_i f_j'(r)/f_j(r), \quad \text{where} \quad r = \sum_j r_j.$$

Hence r_i is proportional to k_i and so $r_i = rk_i/n$, which is chosen so that

$$1 = \sum_j r_j f_j'(r)/f_j(r). \tag{32}$$

If we want to get a distribution by conditioning on $\sum k_i = n$, then, as in the previous example, we want $r_j/f_j(r)$ to be independent of j. This together with (32) gives us $\dim \mathbf{r}$ equations in the same number of unknowns; however, that does not seem to be as easily solved as (31). ∎

Example 6.3. (3-Connected Rooted Maps, Continued) Continuing Example 4.4, we now consider enumeration of 3-connected rooted maps by vertices and faces. As noted in that example, we want

$$[\mathbf{x}^\mathbf{n}]\left(\frac{-w_1w_2}{(1+w_1+w_2)^3}\right) \quad \text{where} \quad w_1 = x_1(1+w_2)^2 \quad \text{and} \quad w_2 = x_2(1+w_1)^2.$$

We must sum over 3 trees, namely those in (25). The tree with E_3 produces a term that is smaller by a factor on the order of n than those for E_1 and E_2. The sum over the trees with E_1 and E_2 lead to a term which vanishes when we set $\mathbf{x}$ to the $\mathbf{r}$ according to Theorem 2.1. Paralleling Example 4.4, we could write this sum S as $(S/D)\times D$, where D is the determinant in (15), and then use the fact that (15) equals (16) to get a more tractable formula. This approach requires that S/D be well behaved as we approach $\mathbf{r}$. In Example 4.4, this was the case since we had cancellation between S and D. In this case there is no such cancellation and, based on Maple calculations, the limit of S/D as we approach $\mathbf{r}$ depends on how $\mathbf{r}$ is approached.

Consequently, we are unable to proceed. ∎

7. Unsolved Problems

Examples 4.4 and 6.3 raise the issue that the function $g(\mathbf{w})$ can cause problems. We were able to avoid them in the first example but not in the second. If one attempts to enumerate all rooted maps a similar problem arises. There is also a new problem in that case: One can enumerate all maps on general surfaces [1, 4] and the vanishing determinant problem causes difficulty on the projective plane even in the one variable case because of a branch point. Can the approach in this paper be extended to such situations or must one use singularity analysis? If the latter, is there a useful general formulation that will deal with problems of this sort?

Suppose additional variables appear in g but not in h. In particular, suppose $h = h(w)$ and $g = g(w, y)$ and g does not have a singularity at $w = r$. If $g(r, y)$ has finite radius of convergence and has nice singularities there, then the arguments in [10] may be usable. If $g(r, y)$ is entire, this approach breaks down. For example, if $w = xf(w)$ is the functional equation for the exponential generating function of a regular family of labelled trees, then $(1 - w(x))^{-y}$ is the enumerator for functional digraphs such that removal of cyclic edges results in a forest of trees enumerated by w and y keeps track of the number of cycles. When $w(r) < 1$, $(1 - w(r))^{-y}$ is a Hayman admissible function [16]. We are not aware of any multivariate singularity analysis that combines algebraic singularities and Hayman-admissibility. We discussed multivariate Hayman admissibility in [8].

We believe it should be possible to extend Drmota's functional equation results [12] to remove the limitation that $g(\mathbf{w})$ be w_i for some i.. Also, one should be able to eliminate the conditioning on n in his Theorem 1, thereby obtaining a result like our Theorem 5.1, probably with his $I - F_{\mathbf{y}}$ and $F_{\mathbf{y},\mathbf{y}}$ playing a role akin to our L and B. We have not attempted to develop these ideas and currently have no plans to do so.

References

[1] D. Arquès, Une relation fonctionelle nouvelle sur les cartes planaire pointées, *J. Combin. Theory Ser. B* **39** (1985) 27–424.

[2] D. Arquès, Relations fonctionelles et dénombrement des cartes pointées sur le tore, *J. Combin. Theory Ser. B* **43** (1987) 253–274.

[3] J. S. Beissinger, The enumeration of irreducible combinatorial objects, *J. Combin. Theory Ser. A* **38** (1985) 143–169.

[4] E. A. Bender, E. R. Canfield, and L. B. Richmond, The asymptotic number of rooted maps on a surface II: Enumeration by vertices and faces, *J. Combin. Theory Ser. A* **63** (1993) 318–329.

[5] E. A. Bender and L. B. Richmond, Central and local limit theorems applied to asymptotic enumeration II: Multivariate generating functions, *J. Combin. Theory Ser. A* **34** (1983) 255–265.

[6] E. A. Bender and L. B. Richmond, An asymptotic expansion for the coefficients of some formal power series II: Lagrange inversion, *Discrete Math.* **50** (1984) 135–142.

[7] E. A. Bender and L. B. Richmond, The asymptotic enumeration of rooted convex polyhedra, *J. Combin. Theory Ser. B* **36** (1984) 276–283.

[8] E. A. Bender and L. B. Richmond, Admissible functions and asymptotics for labeled structures by number of components, *Electron. J. Combin.* **3** (1996) R34.

[9] E. A. Bender and L. B. Richmond, A multivariate Lagrange inversion formula for asymptotic calculations, *Electron. J. Combin.* submitted.

[10] E. A. Bender, L. B. Richmond, and S. G. Williamson, Central and local limit theorems applied to asymptotic enumeration III: Matrix recursions, *J. Combin. Theory Ser. A* **35** (1983) 263–278.

[11] E. A. Bender and N. C. Wormald, The number of rooted convex polyhedra, *Canad. Math. Bull.* **31** (1988) 99–102.

[12] M. Drmota, Systems of functional equations, *Random Structures and Algorithms* **10** (1997) 103–124.

[13] D. Gardy, Some results on the asymptotic behaviour of coefficients of large powers of functions, *Discrete Math.* **139** (1995) 189–217.

[14] I. J. Good, The generalization of Lagrange's expansion and the enumeration of trees, *Proc. Cambridge Phil. Soc.* **61** (1965) 499–517.

[15] I. P. Goulden and D. M. Jackson, *Combinatorial Enumeration*, J. Wiley and Sons, New York (1983).

[16] W. K. Hayman, A generalisation of Stirling's formula, *J. Reine Angew. Math.* **196** (1956) 67–95.

[17] H. K. Hwang, ??? quasipowers ??? *Europ. J. Combin.* **19??** (1998) ???–???.

[18] G. Kreweras, Sur les partitions non croisées d'un cycle, *Discrete Math.* **1** (1972) 333–350.

[19] A. Meir and J. W. Moon, The asymptotic behaviour of coefficients of powers of certain generating functions, *Europ. J. Combin.* **11** (1990) 581–587.

[20] A. Meir and J. W. Moon, On nodes of given out-degree in random trees, 213–222 in J. Nešetřil and M. Fiedler (eds.), *Fourth Czechoslovakian Symposium on Combinatorics, Graphs and Complexity*, Elsevier (1992).

[21] R. C. Mullin and P. J. Schellenberg, The enumeration of c-nets via quadrangulations, *J. Combin. Theory* **3** (1968) 259–276.

[22] S. J. Press, *Applied Multivariate Analysis*, Holt, Rinehart and Winston, New York (1972).

[23] A. Rényi, *Probability Theory*, North-Holland, Amsterdam (1970).

[24] R. Sprugnoli and C. M. Verri, Asymptotics for Lagrange inversion, *Pure Math. Appl.* **5** (1994) 79–104.

[25] W. T. Tutte, A census of planar maps, *Canad. J. Math.* **15** (1963) 249–271.