Towards a theory of multivariate generating functions

Mark C. Wilson

University of Auckland

(joint with Robin Pemantle, Penn State)

http://www.cs.auckland.ac.nz/~mcw/mvGF/

INRIA Rocquencourt, 28 June 2004

Multivariate GFs - overview

- Often used as a technical device for lower-dimensional problems ("marking", cumulative GFs, auxiliary recurrence).
- Determining the GF in closed form is nontrivial even for linear constant coefficient recurrences (Bousquet-Mélou and Petkovšek; kernel method).
- Inverting the GF transform (coefficient extraction) is harder (what do asymptotics mean? phase transitions; geometry of singularities).
- Current theory is scanty, scattered in the literature (queueing theory, tilings, analysis of algorithms, . . .) and not always easy to use.

Inversion - some quotations

- (E. Bender, SIAM Review 1974) Practically nothing is known about asymptotics for recursions in two variables even when a GF is available. Techniques for obtaining asymptotics from bivariate GFs would be quite useful.
- (A. Odlyzko, Handbook of Combinatorics, 1995) A major difficulty in estimating the coefficients of mvGFs is that the geometry of the problem is far more difficult. . . . Even rational multivariate functions are not easy to deal with.
- (P. Flajolet/R. Sedgewick, Analytic Combinatorics Ch 9 draft) Roughly, we regard here a bivariate GF as a collection of univariate GFs . . . .

Our project

- Thoroughly investigate coefficient extraction for meromorphic $F(\mathbf{z}) := F(z_1, \ldots, z_{d+1})$ ("small singularities"). Amazingly little is known even about rational F in 2 variables.
- Goal 1: improve over all previous work in generality, ease of use, symmetry, computational effectiveness, uniformity of asymptotics. Create a theory!
- Goal 2: establish mvGFs as an area worth studying in its own right, a meeting place for many different areas, a common language. I am recruiting!

Notation and basic taxonomy

- $F(\mathbf{z}) = \sum a_{\mathbf{r}} \mathbf{z}^{\mathbf{r}} = G(\mathbf{z})/H(\mathbf{z})$ meromorphic in nontrivial polydisc in $\mathbb{C}^{d+1}$.
- $\mathcal{V} = \{\mathbf{z} | \, H(\mathbf{z}) = 0\}$ the singular variety of F.
- $\mathcal{T}(\mathbf{z}), \mathcal{D}(\mathbf{z})$ the torus, polydisc centred at $\mathbf{0}$ and containing $\mathbf{z}$.
- Note $\dim \mathcal{V} = 2d$, $\dim \mathcal{T} = d + 1$, $\dim \mathcal{D} = 2(d + 1)$. Geometry for $d > 0$ very different from $d = 0$.
- A point of $\mathcal{V}$ is strictly minimal (with respect to the usual partial order on moduli of coordinates) if $\mathcal{V} \cap \mathcal{D}(\mathbf{z}) = \{\mathbf{z}\}$. When $F \geq 0$, such points lie in the positive real orthant.
- A minimal point can be a smooth, multiple or cone point, depending on local geometry of $\mathcal{V}$.

Examples of each geometry

- (smooth points) The generic case. All problems of "Gaussian" type in analytic combinatorics (sequences, sums of independent random variables, many more). Airy-type problems.
- (multiple points) Simplest: H a product of distinct affine factors. For example, $F(\mathbf{z}) = \prod_i (1 - \sum_j a_{ij} z_j)^{-1}$ gives normalization constants of queueing networks.
- (cone points) GF for tilings of Aztec diamond (not given here). Aim to prove the Arctic Circle Theorem by direct GF analysis.

Outline of our method

- We use Cauchy integral formula; residue approximation in 1 variable; convert to Fourier-Laplace integral in remaining d variables; stationary phase method.
- Must specify a direction $\bar{\mathbf{r}} = \mathbf{r}/|\mathbf{r}|$ for asymptotics.
- To each minimal point $\mathbf{z}^*$ we associate a cone $\boldsymbol{\kappa}(\mathbf{z}^*)$ of directions. For smooth points of $\mathcal{V}$, $\boldsymbol{\kappa}$ collapses to a single ray represented by **dir**; for multiple points, $\boldsymbol{\kappa}$ is nontrivial.
- If $\bar{\mathbf{r}}$ is bounded away from $\boldsymbol{\kappa}(\mathbf{z}^*)$, then $|{\mathbf{z}^*}^{\mathbf{r}}\mathbf{a}_{\mathbf{r}}|$ decreases exponentially. We show that if $\bar{\mathbf{r}}$ is in $\boldsymbol{\kappa}(\mathbf{z}^*)$, then $(\mathbf{z}^*)^{-\mathbf{r}}$ is the right asymptotic order, and develop full asymptotic expansions, on a case-by-case basis.

Generic case theorem – smooth point

Theorem 1. Let $\mathbf{z}^*$ be a strictly minimal, simple pole of F. Then for $\bar{\mathbf{r}} = \mathbf{dir}(\mathbf{z}^*)$, there is a full asymptotic expansion

$$a_{\mathbf{r}} \sim (\mathbf{z}^*)^{-\mathbf{r}} \sum_{l \geq 0} C_l |\mathbf{r}|^{-(d+l)/k}.$$

The constants C_l and k depend analytically on derivatives of G and H at $\mathbf{z}^*$ of order at most l.

The expansion is uniform over compact sets of minimal poles with k and the vanishing order of G and H remaining constant. $\square$

Generically, $k = 2$ and we have Ornstein-Zernike ("central limit") behaviour. Airy phenomena occur when $k = 3$ for a given direction but $k = 2$ at neighbouring directions.

Specialization to dimension 2

Theorem. Suppose that $H(z, w)$ has a simple pole at $P = (1, 1)$ and is otherwise analytic in $|z| \leq 1, |w| \leq 1$. Define

$$Q(1, 1) = -a^2 b - ab^2 - a^2 z^2 H_{zz} - b^2 w^2 H_{ww} + ab H_{zw}$$

where $a = wH_w, b = zH_z$, all computed at P. Then when $r/s = b/a$

$$a_{rs} \sim \frac{G(1, 1)}{\sqrt{2\pi}} \sqrt{\frac{-a}{sQ(1, 1)}}.$$

The apparent lack of symmetry is illusory, since $r/s = b/a$. It is true *mutatis mutandis* for each smooth minimal point P.

Exemplifying Theorem 1

- Walks in integer lattice going $\uparrow, \rightarrow, \nearrow$. Here $F(x,y) = (1-x-y-xy)^{-1}$. Necessary condition for minimal point: $x(1+y) = \kappa y(1+x), \kappa \geq 0$. So minimal points are all smooth and in first quadrant.
- For r/s fixed, asymptotics are governed by the minimal point satisfying $1-x-y-xy=0, x(1+y)s = y(1+x)r$.
- Using these relations and the theorem we obtain to first order

$$a_{rs} \sim \left[\frac{\Delta - s}{r}\right]^{-r} \left[\frac{\Delta - r}{s}\right]^{-s} \sqrt{\frac{rs}{2\pi\Delta(r+s-\Delta)^2}}.$$

  where $\Delta = \sqrt{r^2+s^2}$.
- Extracting the diagonal ("Delannoy numbers") is easy:

$$a_{rr} \sim (3+2\sqrt{2})^r \frac{1}{4\sqrt{2}(3-2\sqrt{2})} r^{-1/2}.$$

New phenomena - multiple points

Theorem 2 Suppose that H has a transverse double pole at $(1,1)$ but is otherwise analytic in $|z| \leq 1, |w| \leq 1$. Let $\mathcal{H}$ denote the Hessian of H. Then for each compact subset K of the interior of $\boldsymbol{\kappa}(1,1)$, there is $c > 0$ such that

$$a_{rs} = \left(\frac{G(1,1)}{\sqrt{-\det \mathcal{H}(1,1)}} + O(e^{-c}) \right) \qquad \text{uniformly for } (r,s) \in K.$$

The uniformity breaks down near the walls of the cone, but we know the expansion on the boundary (in powers of Δ^{-1}).

There are other results for general d and multiplicity.

Exemplifying Theorem 2

An IID sequence of uniform $[0,1]$ random variables X is used to generate biased coin-flips as follows. If $\Pr(H) = p$, then $X \le p$ means heads and $X > p$ means tails.

The coins will be biased so that $p = 2/3$ for the first n flips, and $p = 1/3$ thereafter. A player desires to get r heads and s tails and is allowed to choose n. On average, how many choices of $n \le r + s$ will be winning choices?

The generating function is readily computed to be

$$F(z,w) = \frac{1}{(1 - \frac{1}{3}z - \frac{2}{3}w)(1 - \frac{2}{3}z - \frac{1}{3}w)} .$$

Here $(1,1)$ is a strictly minimal transverse double point. By Theorem 2 $a_{rs} = 3$ plus a correction which is exponentially small as $r, s \to \infty$ with $r/(r+s)$ staying in a compact subinterval of $(1/3, 2/3)$. For other values of $r/(r+s)$, Theorem 1 applies.

More complicated multiple point results

Suppose $\mathcal{V}$ is locally the intersection of $n+1$ sheets in dimension $d+1$ (like queueing example).

- If $n \geq d$, generically we have: $a_{\mathbf{r}}$ is piecewise polynomial with exponential error. There are finitely many subcones on each of which we get a different polynomial.
- If $n < d$, generically we have: $a_{\mathbf{r}}$ has expansion in descending powers of $|\mathbf{r}|$, starting with $(n-d)/2$.
- Actual results depend on rank of a certain matrix. All derived by analysis of Fourier-Laplace integrals. Explicit formulae are available.

Fourier-Laplace integrals

We are quickly led via $\mathbf{z} = \mathbf{z}^* e^{i\boldsymbol{\theta}}$ to large-λ analysis of integrals of the form

$$I(\lambda) = \int_D e^{-\lambda f(\mathbf{x})} \psi(\mathbf{x})\, dV(\mathbf{x})$$

where:

- $f(0) = 0$, $f'(0) = 0$ iff $\mathbf{r} \in \boldsymbol{\kappa}(\mathbf{z}^*)$.
- $\operatorname{Re} f \geq 0$; the phase f is analytic, the amplitude $\psi \in C^\infty$.
- D is an $(n+d)$-dimensional product of tori, intervals and simplices; dV the volume element.

Difficulties in analysis: interplay betwen exponential and oscillatory decay, nonsmooth boundary of simplex.

Sample reduction to F-L in simple case

Suppose $(1, 1)$ is a smooth or multiple strictly minimal point. Here C_a is the circle of radius a centred at 0, $R(z; s; \varepsilon)$ = residue sum in annulus, N a nbhd of 1.

$$\begin{aligned}
a_{rs} &= (2\pi i)^{-2} \int_{C_1} z^{-r-1} \int_{C_{1-\varepsilon}} w^{-s-1} F(z, w)\, dw\, dz \\
&= (2\pi i)^{-2} \int_N z^{-r-1} \left[\int_{C_{1+\varepsilon}} w^{-s-1} F(z, w) - 2\pi i R(z; s; \varepsilon) \right] dz \\
&\cong -(2\pi i)^{-1} \int_N z^{-r-1} R(z; s; \varepsilon)\, dz \\
&= (2\pi)^{-1} \int_N \exp(-ir\theta + \log(-R(z; s; \varepsilon))\, d\theta.
\end{aligned}$$

To proceed we need a formula for the residue sum.

Dealing with the residues

- In smooth case $R(z;\varepsilon) = v(z)^s \operatorname{Res}(F/w)_{|w=1/v(z)} := v(z)^s \phi(z)$. So above has the form
$$(2\pi)^{-1} \int_N \exp(-s(ir\theta/s - \log v(z) - \log(-\phi(z))\, d\theta.$$
- In multiple case there are $n+1$ poles in the ε-annulus and we use the following nice lemma:

  Let $h : \mathbb{C} \to \mathbb{C}$ and let μ be the normalized volume measure on $\mathcal{S}_n$. Then
$$\sum_{j=0}^{n} \frac{h(v_j)}{\prod_{r \neq j}(v_j - v_r)} = \int_{\mathcal{S}_n} h^{(n)}(\boldsymbol{\alpha}\mathbf{v})\, d\mu(\boldsymbol{\alpha}).$$

For each fixed direction r/s, previous slide's integral has the F-L form in $n+d$ dimensions. Introduction of the n-simplex $\mathcal{S}$ makes the F-L analysis harder.

Asymptotics for F-L integrals

Standard methods for such integrals:

- Stationary phase - localize to critical points of f. Use integration by parts. See Hörmander.
- Change of variable - away from critical points f can be locally taken as a coordinate. Differential forms approach. See AGV.
- Move the contour, using Cauchy apparatus. Apply Laplace.

Most (all?) authors require at least one of:

- f purely real or purely imaginary;
- smooth boundary, ψ vanishing near boundary;
- isolated stationary phase points.

Each of these is violated by some of our examples of interest.

Types of critical points arising generically

- Smooth: isolated stationary point, no simplex corners to worry about. Real part of phase has strict minimum at 0. Simple extension of Laplace method.
- Multiple, $n \geq d$: isolated stationary point. Real part of phase is zero on lower-dimensional subspace. Need good definition of critical point. Laplace doesn't work. Mostly in Hörmander.
- Multiple, $n \leq d$: stationary phase points form an affine subspace of the unit simplex. Not covered anywhere, and tricky.

For most directions the critical points are interior, but some are on boundary. They are generically quadratically nondegenerate. Many special cases can occur.

Low-dimensional examples of F-L integrals

- Typical smooth point example looks like

$$\int_{-1}^{1} e^{-\lambda(1+i)x^2}\, dx.$$

  Isolated nondegenerate critical point, exponential decay

- Simplest double point example looks roughly like

$$\int_{-1}^{1}\int_{0}^{1} e^{-\lambda(x^2+2ix^y)}\, dy\, dx.$$

  Note $\operatorname{Re} f = 0$ on $x = 0$ so rely on oscillation for smallness.

- Multiple point with $n = 2, d = 1$ gives integral like

$$\int_{-1}^{1}\int_{0}^{1}\int_{-x}^{x} e^{-\lambda(z^2+2izy)}\, dy\, dx\, dz.$$

Simplex corners now intrude, continuum of critical points.

Good points of the method

- Natural, fairly unified approach, reduces to F-L integral which can almost certainly be done.
- Extremely complicated Leray residue theory is avoided.
- There is an easily checked necessary condition for minimality.
- If $F \geq 0$ then minimal points exist for all directions of interest.
- In all other examples, there appears to be a ("topologically minimal") point that controls the asymptotics. The simple-minded contour moving must be replaced by a different homotopy.

Homological approach

- Residue theory in several complex variables is homological, difficult to make effective. Consider homology of $\mathbb{C}^{d+1} \setminus \mathcal{V}$.
- Recently, Y. Baryshnikov and R. Pemantle have generated asymptotic expansions when H is a product of affine factors (as in queueing examples). Uses stratified Morse theory.
- Can't get directions on boundary of cone by this method. Not (yet) generalized to nonaffine factors. Still assumes $F \geq 0$.

Work still required

- Complete analysis of F-L integrals in general case (large stationary phase set).
- How to find and classify minimal singularities algorithmically? Note: a minimal point is a Pareto optimum of the functions $|z_1|, \dots, |z_{d+1}|$.
- Computer algebra of multivariate asymptotic expansions.
- Patching together asymptotics at cone boundaries; uniformity, phase transitions.
- Expansions controlled by cone points? A more high-powered approach (e.g. resolution of singularities) may be needed.

References

- http://www.cs.auckland.ac.nz/˜mcw/Research/mvGF/
- Lars Hörmander, *The Analysis of Linear Partial Differential Operators, I.* Springer, 1983.
- V. Arnold, A. Varchenko, & S. Guseĭn-Zade, *Singularities of Differentiable Maps*, Birkhäuser 1988.
- M. Bousquet-Mélou & M. Petkovšek, 'Linear recurrences with constant coefficients: the multivariate case', *Discrete Math.* **225**, (2000), 51–75.