CHARACTERISTIC POINTS OF RECURSIVE SYSTEMS

JASON BELL, STANLEY BURRIS, AND KAREN YEATS

ABSTRACT. Characteristic points have been a primary tool in the study of a generating function defined by a single recursive equation. We investigate the proper way to adapt this tool when working with multi-equation recursive systems.

1. Introduction and Preliminaries

Recursively defined generating functions play a major role in combinatorial enumeration, for example (see [7]) when studying lattice animals, context free languages, random walks on free groups, directed walks in the plane, colored trees, and Boolean expression trees. The subject started in 1857 with Cayley’s [4] fundamental equation for the number $t_n$ of rooted unlabeled trees:

$$\sum_{n \geq 1} t_n x^n = x \cdot \prod_{n \geq 1} (1 - x^n)^{-t_n}. \hspace{1cm} (1)$$

He used this to recursively calculate (with some errors) the first dozen values of $t_n$, and later applied his method to recursively enumerate certain kinds of chemical compounds.

The generating function for trees is $T(x) = \sum_{n \geq 1} t_n x^n$. In 1937 Pólya (see [16]) converted (1) into

$$T(x) = x \cdot \exp \left( \sum_{m \geq 1} T(x^m)/m \right), \hspace{1cm} (2)$$

a form to which he was able to apply analytic techniques to find asymptotics for the $t_n$, namely he proved

$$t_n \sim C \rho^{-n} n^{-3/2} \hspace{1cm} (3)$$

Date: May 15, 2009.
where $\rho$ is the radius of convergence of $T(x)$, and $C$ a positive constant. A similar result held for the various classes of chemical compounds studied by Cayley. Although the function $T(x)$ was not expressible in terms of well-known functions, nonetheless Pólya showed how to determine $C$ and $\rho$ directly from (2). Pólya’s methods were applied to nearly regular classes of trees in 1948 by Otter [15].

In 1974 Bender [1], following Pólya’s ideas, formulated a general result for how to determine the radius of convergence $\rho$ of a power series $T(x)$ defined by a functional equation $F(x, y) = 0$. Bender’s hypotheses guaranteed that $\rho$ was positive and finite, and that $\tau := T(\rho)$ was also finite. His method was simply to find $(\rho, \tau)$ among the solutions $(a, b)$ (called characteristic points) of the characteristic system

$$F(x, y) = 0$$
$$\frac{\partial F}{\partial y}(x, y) = 0.$$  

A decade later Canfield [3] found a gap in the hypotheses of Bender’s formulation when there were several characteristic points. In the case of a polynomial functional equation, Canfield sketched a method (based on analyzing the singularities of an algebraic curve and how the branches relate to one another) to determine which of the characteristic points gives the radius of convergence of the solution $y = T(x)$.

In the late 1980s Meir and Moon [13] focused on a special case of Canfield’s work, namely when $F(x, y) = 0$ is of the form $y = G(x, y)$, where $G(x, y)$ is a power series with nonnegative coefficients. The interesting cases were such that setting $T(x) = G(x, T(x))$, with $T(x)$ an indeterminate power series, gave a recursive determination of the coefficients of $T(x)$. One advantage of their restricted form of recursive equation was that there could be at most one characteristic point. This formulation was adopted by Odlyzko in his 1995 survey paper [14] as well as in the recent book [7] of Flajolet and Sedgewick. These publications have focused on characteristic points in the interior of the domain of convergence of $G(x, y)$, in the context of proving that $\rho$ is a square root singularity of the solution $y = T(x)$. If $(\rho, \tau)$ is on the boundary of the domain of $G(x, y)$ then $\rho$ may not be a square-root singularity of $T(x)$.

In [2] we found this law so ubiquitous among naturally defined classes of trees defined by a single equation that we referred to it as the universal law for rooted trees.
Most of the areas of application listed in the first sentence of this introduction actually require a recursive system of equations

\[
\begin{align*}
y_1 &= G_1(x, y_1, \ldots, y_m) \\
\vdots \\
y_m &= G_m(x, y_1, \ldots, y_m),
\end{align*}
\]

written more briefly as \( y = G(x, y) \). (A precise definition of the systems considered in this paper is given in §2.) This rich area of enumeration has been rather slow in its development. In the 1970s Berstel and Soittola (see [7] V.3) carried out a thorough analysis of enumerating the words in a regular language using recursive systems of equations that were linear in \( y_1, \ldots, y_m \). However it was not until the 1990s that publications started appearing that used multi-equation non-linear systems. Following the trend with single recursion equations \( y = G(x, y) \), the focus has been on systems \( y = G(x, y) \) where the \( G_i(x, y) \) are power series with non-negative coefficients.

In 1993 Lalley [10] considered polynomial systems in his study of random walks on free groups. In 1997 Woods [17] used one particular system to analyze the asymptotic densities of monadic second-order definable classes of trees in the class of all trees. In the same year Drmota [5] extended Lalley’s results to power series systems. Lalley’s and Drmota’s results were for a wide range of irreducible systems, that is, systems in which each variable \( y_i \) (eventually) depends on any variable \( y_j \). An irreducible system of the kind they studied behaves in some ways like a single equation system, for example, the standard solution \( y_i = T_i(x) \) is such that all the \( T_i(x) \) have the same finite positive radius \( \rho \), the \( \tau_i := T_i(\rho) \) are all finite, and the asymptotics for the coefficients of \( T_i(x) \) is of the Pólya form \( C_i \rho^{-n} n^{-3/2} \).

Thus, as has been the case with single equation systems, it is desirable to find the radius of convergence \( \rho \) even though the solutions \( T_i(x) \) may be fairly intractable. The natural method was to extend the definition of the characteristic system from a single equation to a system of equations by adding the determinant of the Jacobian of the system, set equal to zero to, to the original system. The solutions of such a characteristic system will again be called characteristic points.

Under suitable conditions one can find \((\rho, \tau)\) among the characteristic points. To date, however, the necessary study of characteristic points \((a, b)\) for systems, so that one can locate \((\rho, \tau)\), has been essentially non-existent. Filling this void is the
The major goal of this paper. In December, 2007, one of the authors discovered that in the systems studied by Drmota it was possible that there could be more than one characteristic point — this was communicated to Flajolet and appears as an example in [7] (p. 484). The main objective of this paper is to give conditions to locate \((\rho, \tau)\) among the characteristic points, if indeed \((\rho, \tau)\) is a characteristic point. A review of and improvements to the theory of the single equation case (see Proposition 29) are also given. It turns out that even if there is a characteristic point of a system \(y = G(x, y)\) in the interior of the domain of \(G(x, y)\), one cannot claim that the asymptotics for the coefficients of the solutions \(T_i(x)\) will be of the above Pólya form (see Examples 16, 17).

Before proceeding with the study of such systems some background material is need.

1.1. The extended nonnegative real numbers.

Extend the usual operations on \([0, \infty)\) to \([0, \infty]\) in the obvious way as follows:

\[
c + \infty = \infty \quad \text{for } c \in [0, \infty]
\]
\[
c \cdot \infty = \infty \quad \text{for } c \in (0, \infty]
\]
\[
\sum_n c_n = \begin{cases} 
\text{the usual infinite sum} & \text{if all } c_n \in [0, \infty) \\
\infty & \text{if some } c_n = \infty.
\end{cases}
\]

Here the usual infinite sum is \(\infty\) if the series diverges.

1.2. Formal power series in several variables. This section gives the essential definitions that lay the foundations for working with formal power series in several variables. The standard number systems are:

\[\text{In 1997 Drmota [5] appears to claim that having a characteristic point in the interior of the domain would lead to Pólya asymptotics—however these examples show this not to be the case. In his 2009 book [6] this hypothesis is replaced with one regarding minimal characteristic points, which seems somewhat at odds with our Lemma [20] which says that the characteristic points form an antichain with the characteristic point \((a, b)\) of interest having the largest value of \(a\) among the characteristic points. Theorem [26] of §6.1 is a restatement of Drmota’s result which makes it clear which characteristic point is of interest, namely the one (if it exists) such that the Jacobian of } \text{G}(x, y) \text{has 1 as its largest real eigenvalue.} \]
the set $\mathbb{N} = \{0, 1, \ldots\}$ of nonnegative integers, the set $\mathbb{Q}$ of rational numbers, the set $\mathbb{R}$ of real numbers, and the set $\mathbb{C}$ of complex numbers.

For the linearly ordered set $\mathbb{R}$ of real numbers one has the posets of real-valued functions on $X$, where the partial ordering is given by $f \leq g$ if $f(x) \leq g(x)$ for all $x \in X$. Familiar examples are:

(a) $n$-vectors $\mathbf{v} = (v_1, \ldots, v_n)$, by setting $X = \{1, \ldots, n\}$

(b) $m \times n$-matrices $M$, by setting $X = \{1, \ldots, m\} \times \{1, \ldots, n\}$

(c) formal power series in $k$-variables $A(x_1, \ldots, x_k)$ by setting $X = \mathbb{N}^k$. In this case a function $a$ from $\mathbb{N}^k$ to $\mathbb{R}$ provides the coefficients, and one writes

$$A(x) := \sum_{i \in \mathbb{N}^k} a(i) x^i$$

A matrix (or vector) $M$ of real numbers is non-negative (written $M \geq 0$) if each entry is non-negative, and positive (written $M > 0$) if each entry is positive. A power series $A(x)$ is non-negative (written $A(x) \geq 0$) if each coefficient is non-negative.

1.2.1. Composition of formal power series.

For power series $A(w_1, \ldots, w_m)$ and $B_\ell(x)$, $1 \leq \ell \leq m$, where the constant term of each $B_\ell$ is zero, that is, $b_\ell(0) = 0$, define the formal composition

$$C(x) := A(B_1(x), \ldots, B_m(x))$$

by defining the coefficient function as follows:

$$c(i) := \sum_{j \geq 0} \left[x^i\right] a(j) \cdot B_1(x)^{j_1} \cdots B_m(x)^{j_m}$$

Requiring that the constant term of the $B_\ell(x)$ be 0 guarantees that for each $i$ only finitely many terms in this sum are nonzero. Consequently $C(x)$ is indeed a formal power series.

1.2.2. The function defined by a formal power series.

A power series $A(x)$ in $k$ variables defines a partial function, also denoted $A(x)$, on $\mathbb{R}^k$ (or $\mathbb{C}^k$) by setting

$$A(c) := \sum_{n \geq 0} \sum_{i_1 + \cdots + i_k = n} a(i) c^i \quad (c \in \mathbb{R}^k)$$
whenever the sum converges.

For \( A(x) \) a nonnegative power series in \( k \) variables, and for \( c \in [0, \infty]^k \), \( A(c) = \infty \) if the series (5) diverges, that is, if

\[
\lim_{n \to \infty} \sum_{j \leq n} \sum_{i_1 + \cdots + i_k = j} a(i)c^i = \infty.
\]

When one specializes to nonnegative power series, and focuses on nonnegative arguments \( c \), then the general study of power series functions \( A(x) \) becomes much easier, in part because one does not have to worry about rearrangements in (5) affecting convergence, or the value to which it converges, but also because one is dealing with a function which is continuous, and is monotone nondecreasing in each variable. A nonnegative power series \( A(x) \) in \( k \) variables defines a left-continuous function from \([0, \infty]^k\) to \([0, \infty]\), that is, if \( a_j \to b^- \) in \([0, \infty]^k\) then \( A(a_j) \to A(b) \).

1.2.3. The derivatives of a formal power series. Derivatives of [nonnegative] formal power series give [nonnegative] formal power series:

\[
\frac{\partial A(x)}{\partial x_j} := \sum_{i \geq 0} i_j a(i) x_1^{i_1} \cdots x_j^{i_j} \cdots x_k^{i_k}.
\]

The notation \( A_{x_j} \) is also used for the partial derivative \( \partial A/\partial x_j \).

Working with the formal derivative of formal power series instead of the analytical derivative will impact the discussion of characteristic points. To see why this might be the case just note, for example, that the power series \( A(x) = \sum x^n/n^3 \) has the formal derivative \( A'(x) = \sum x^{n-1}/n^2 \) which converges at \( x = 1 \), whereas the analytical derivative does not exist at \( x = 1 \).

1.2.4. Holomorphic functions and a law of permanence.

A complex-valued function \( f(x) \) of several complex variables is holomorphic at \( c \) if it is continuous and differentiable in a neighborhood of \( c \). The notation \([a, b]\) is short for \([a_1, b_1] \times \cdots \times [a_k, b_k]\).

**Proposition 1** (A Law of Permanence for Functional Equations).

Suppose \( A(x), B(x, y) \geq 0 \). If there is an \( \varepsilon > 0 \) such that

\[
A(x) = B(x, A(x)) < \infty \text{ for } x \in [0, \varepsilon]
\]
then 
\[ A(x) = B(x, A(x)) \quad \text{for } x \in [0, \infty]. \]

If furthermore \(a > 0\) and \(A(a) < \infty\) then

\[ A(x) = B(x, A(x)) \quad \text{for } |x_i| \leq a_i, \ 1 \leq i \leq k \]

and \(A(x)\) is holomorphic for \(|x_i| < a_i, \ 1 \leq i \leq k\).

Proof. This is a special case of Hille’s law of permanence for functional equations given in §10.7 of Vol. 2, [9].

1.3. The Perron-Frobenius theory of nonnegative matrices. The key to the main results of this paper are some simple observations based on the well-known Perron-Frobenius theory of nonnegative matrices that was developed ca. 1910.

**Proposition 2.** Let \(M\) be a nonnegative nonzero \(k \times k\) matrix with real entries.

(a) \(M\) has a real eigenvalue.

(b) The largest real eigenvalue \(\Lambda(M)\) is positive and is given by

\[ \Lambda(M) = \max_{x > 0} \min_{1 \leq i \leq k} \frac{(Mx)_i}{x_i}. \]

(c) \(\Lambda(M)\) is a simple root of the characteristic polynomial \(p_M(\lambda) = \det(\lambda I - M)\).

(d) The eigenspace belonging to \(\Lambda(M)\) is 1-dimensional, generated by a unique positive normalized eigenvector \(v_M\). (Normalized means the sum of the entries is 1).

Proof. (See §2 of Gantmacher [8].)

Note that Proposition 2(b) implies that for some \(x > 0\) one has \(\Lambda(M)\) equal to

\[ \min_{1 \leq i \leq k} \frac{(Mx)_i}{x_i}. \]

**Corollary 3.**

(a) A positive \(k \times k\) matrix \(M, k \geq 2\), has all diagonal entries \(<\Lambda(M)\).

(b) \(\Lambda(X)\) is a nondecreasing function on the set of nonnegative matrices, that is, \(M_1 \leq M_2\) implies \(\Lambda(M_1) \leq \Lambda(M_2)\). Furthermore if every row [column] sum of \(M_1\) is less than the corresponding row [column] sum of \(M_2\) then \(\Lambda(M_1) < \Lambda(M_2)\).
(c) $\Lambda(X)$ is a continuous function on the set of nonnegative matrices, where the matrices are thought of as points in $k^2$-space.

Proof. Since $Mv_M = \Lambda(M)v_M$ one has for $1 \leq i \leq k$,

$$0 < m_{ii}v_M(i) < \sum_{1 \leq j \leq k} m_{ij}v_M(j) = \Lambda(M)v_M(i),$$

since each $m_{ij}v_M(j) > 0$. Thus $m_{ii} < \Lambda(M)$. This proves (a).

For (b) suppose $0 \leq M_1 \leq M_2$. Note that for $x > 0$ and $1 \leq i \leq k$ one has

$$\frac{(M_1x)_i}{x_i} \leq \frac{(M_2x)_i}{x_i}.$$

By Proposition 2(b) it follows that $\Lambda(M_1) \leq \Lambda(M_2)$. If in addition every row sum of $M_1$ is less than the corresponding row sum of $M_2$ then for $x > 0$ and $1 \leq i \leq k$

$$\frac{(M_1x)_i}{x_i} < \frac{(M_2x)_i}{x_i}.$$

With this inequality Proposition 2(b) shows that $\Lambda(M_1) < \Lambda(M_2)$.

Since a square matrix $M$ and its transpose $M^t$ have the same characteristic polynomial, and hence the same eigenvalues, one can replace the condition on row sums by one on column sums, finishing the proof of (b).

To see that (c) is true note that the largest real eigenvalue of $M$ is simple, so the graph of the characteristic polynomial $p_M(\lambda)$ crosses the axis at $\lambda = \Lambda(M)$ with a non-zero slope. Consequently this crossing point changes continuously with changes in the entries of $M$. (Note: A special case of item (c) is stated on p. 2103 of Lalley [10], for certain Jacobian matrices denoted $J_z$ evaluated along certain curves.) □

2. Well-conditioned systems

The next definition gives a version of essentially well-known conditions which ensure that a system $y = G(x,y)$ has power series solutions $y_i = T_i(x)$ of the type encountered in generating functions for classes of trees. (See Drmota [5], [6].)

**Definition 4.** A system $y = G(x,y)$ is well-conditioned if it satisfies

(a) each $G_i(x,y)$ is a power series with nonnegative coefficients
(b) $G(x,y)$ is holomorphic in a neighborhood of the origin
(c) $G(0,y) = 0$
(d) for all $i$, $G_i(x, 0) \neq 0$

(e) $\det \left( I - J_G(0, 0) \right) \neq 0$ where $J_G$ is the Jacobian matrix $\left( \frac{\partial G_i}{\partial y_j} \right)$

(f) the system is irreducible

(g) for some $i, j, k$, $\frac{\partial^2 G_i(x, y)}{\partial y_j \partial y_k} \neq 0$ (so the system is nonlinear in $y$).

**Remark 5.** One can replace condition (b) by (b'): $G(x, y)$ converges at some positive $(a, b)$.

2.1. Solutions of Well-Conditioned Systems.

**Proposition 6.** If $y = G(x, y)$ is a well-conditioned system then the following hold:

(i) There is a unique vector $T(x)$ of formal power series $T_i(x)$ with nonnegative coefficients such that one has the formal identity

$$T(x) = G(x, T(x)).$$

(ii) Equation (6) gives a recursive procedure to find the coefficients of the $T_i(x)$.

(iii) Equation (6) holds for $x \in [0, \infty]$.

(iv) All $T_i(x)$ have the same radius of convergence $\rho \in (0, \infty)$ and all $T_i(x)$ converge at $\rho$, that is, $\tau_i := T_i(\rho) < \infty$.

(v) Each $T_i(x)$ has a singularity at $x = \rho$.

(vi) If $(\rho, \tau)$ is in the interior of the domain of $G(x, y)$ then

$$\det \left( I - J_G(\rho, \tau) \right) = 0.$$

**Proof.** By conditions (b) and (e) of Definition 4 the implicit function theorem guarantees a unique solution $y = T(x)$ to the system in a neighborhood of the origin. This solution is holomorphic in a neighborhood of the origin, so each $T_i(x)$ has a radius of convergence $> 0$. Let $\mu$ be the mapping on $m$-tuples of nonnegative power series defined by $\mu(A(x, y)) = G(x, A(x, y))$. In view of Conditions (c) and (d) of Definition 4, $\mu^n(x, 0)$ gives a sequence of $m$-tuples of nonnegative power series that converges to an $m$-tuple of nonnegative power series that provides a formal solution to the system, so it must be $T(x)$. Then by Proposition 1 the equation

$$T(x) = G(x, T(x))$$

This means the non-negative matrix $J_G$ is irreducible.
holds for $x \in [0, \infty]$. Since the system is irreducible, if $x > 0$ is such that some $T_i(x) = \infty$ then all $T_i(x) = \infty$. This means all $T_i(x)$ have the same radius of convergence $\rho$ which, as noted above, must be $> 0$. From the nonlinearity condition (g) of Definition 4 it follows that $\rho < \infty$ and all $\tau_i = T_i(\rho) < \infty$. Each $T_i(x)$ has a singularity at $x = \rho$ by Pringsheim’s Theorem, which gives (v). Thus if $(\rho, \tau)$ is in the interior of the domain of $G$ then the Implicit Function Theorem guarantees that the determinant of the Jacobian of the system must vanish, giving (vi).

\[ \square \]

The sequence $T(x)$ of power series described in Proposition 6 is the standard solution of the system, and the point $(\rho, \tau)$ is the extreme point (of the standard solution, or of the system). From (3) one has $T(0) = 0$, so the standard solution goes through the origin. The set

$$\text{Dom}^+(G) := \{(a, b) : a, b_1, \ldots, b_m > 0 \text{ and } G_i(a, b) < \infty, 1 \leq i \leq m\}$$

is the positive domain of $G$. For $(a, b) \in \text{Dom}^+(G)$ let

$$\Lambda(a, b) := \Lambda(J_G(a, b)),$$

the largest real eigenvalue of the Jacobian matrix $J_G(a, b)$.

2.2. Characteristic Systems, Characteristic Points. Flajolet and Sedgewick [7] VII.6 define the characteristic system of (4) to be

\[
\begin{align*}
y_1 &= G_1(x, y_1, \ldots, y_m) \\
\vdots \\
y_m &= G_m(x, y_1, \ldots, y_m) \\
0 &= \det(I - J_G(x, y)).
\end{align*}
\]

(7)

Let the positive solutions $(a, b) \in \mathbb{R}^{m+1}$ to this system be called the characteristic points of the system.\footnote{Flajolet and Sedgewick ([7] Chapter VII p. 468) only consider characteristic points in the interior of $\text{Dom}^+(G)$.} Requiring that $(\rho, \tau)$ be a characteristic point in the interior of the domain of $G(x, y)$ has been crucial to proofs that $x = \rho$ is a square-root singularity of the $T_i(x)$, leading to the asymptotics $t_i(n) \sim C_i \rho^{-n} n^{-3/2}$ for the non-zero coefficients.
There is considerable interest in finding practical computational means of estimating $\rho$—for then, at least in the case that $(\rho, \tau)$ is in the interior of the domain of $\mathbf{G}$, one has a good estimate for the growth of the non-zero coefficients of the $T_i(x)$, namely up to a constant it is $\rho^{-\eta n^{-3/2}}$. (To refine this further one needs estimates of the constants $C_i$.)

For the case that the $G_i(x, y)$ are polynomials we know that $(\rho, \tau)$ will be among the characteristic points and in the interior of the domain of $\mathbf{G}$. However until now, even in the polynomial case, no general attempt has been made to characterize $(\rho, \tau)$ among the characteristic points of the system—with one exception, namely the 1-equation systems.

3. A Collection of Basic Examples

The following examples explore the behavior of characteristic points of well-conditioned systems.

3.1. 1-equation systems. For a 1-equation system $y = G(x, y)$ there can be at most one characteristic point, and if such exists it must be $(\rho, \tau)$—see Proposition 29. Thus if there is a characteristic point then one has, in theory, a way to determine $(\rho, \tau)$, namely find the positive solution to the characteristic system. If there is no characteristic point then, at present, there is no simple alternative method to determine $(\rho, \tau)$.

In general $(\rho, \tau)$ lies either in the interior of the domain of $G$ or on the boundary of this domain—in either case $x = \rho$ is a singularity of the standard solution $T(x)$. If $(\rho, \tau)$ is in the interior of the domain of $G$ then $x = \rho$ is a square-root singularity of $T(x)$. The possibilities for the nature of this singularity when $(\rho, \tau)$ is on the boundary of the domain of $G$ have not been fully classified.

3.2. Examples for 1-equation systems. For 1-equation systems the following two examples show the three kinds of behavior described above, namely: (i) there is a characteristic point which is an interior point and thus equal to $(\rho, \tau)$, (ii) there is a characteristic point which is a boundary point and thus equal to $(\rho, \tau)$, and (iii) there is no characteristic point.

Each example starts with an equation $y = G(x, y)$ where the characteristic point $(\rho, \tau)$ is in the interior of the domain of $G(x, y)$. Then the example is modified to
give a system \( y = G^\ast(x, y) \) with \((\rho^\ast, \tau^\ast)\) on the boundary of the domain of \(G^\ast(x, y)\). \((\rho^\ast, \tau^\ast)\) is a characteristic point in Example 7 but not in Example 8.

**Example 7.** Let \( G(x, y) = x(1 + y^2) \). For the characteristic system

\[
\begin{align*}
y &= x(1 + y^2) \\
1 &= 2xy
\end{align*}
\]

of \( y = G(x, y) \) one has the characteristic point \((1/2, 1)\), an interior point of the domain of \(G(x, y)\), so for the standard solution \( y = S(x) \) of \( y = G(x, y) \) one has \((\rho, \tau) = (1/2, 1)\). The established theory for such a system (see [7], Chapter VII) shows that \( S(x) \) has a square-root singularity at \( x = \rho \).

Next let \( G^\ast(x, y) = S(x)(1 + y^2)/2 \). For the characteristic system

\[
\begin{align*}
y &= S(x)(1 + y^2)/2 \\
1 &= S(x)y
\end{align*}
\]

once again the characteristic point is \((1/2, 1)\), but now it is a boundary point of the domain of \(G^\ast(x, y)\). An examination of the standard solution (see Proposition [11]) of \( y = G^\ast(x, y) \), namely \( y = T(x) = S(S(x)/2) \), shows that it has a fourth-root singularity at \( x = 1/2 \).

**Example 8.** Let \( G(x, y) = x(1 + 2y + 2y^2) \). The characteristic system

\[
\begin{align*}
y &= x(1 + 2y + 2y^2) \\
1 &= 2x(1 + 2y)
\end{align*}
\]

of \( y = G(x, y) \) has the characteristic point

\[
\left( \frac{\sqrt{2} - 1}{2}, \frac{\sqrt{2}}{2} \right),
\]

an interior point of the domain of \(G(x, y)\), so for the standard solution \( y = S(x) \) of \( y = G(x, y) \) one has \( \rho = (\sqrt{2} - 1)/2 \) and \( \tau = \sqrt{2}/2 \). \( S(x) \) has a square-root singularity at \( x = \rho \).
Next let \( G^*(x, y) = x\left(1 + S(x) + y + 2y^2\right) \). The standard solution of \( y = G^*(x, y) \) is again \( y = S(x) \), so \((\rho^*, \tau^*) = (\rho, \tau)\). The characteristic system

\[
\begin{align*}
    y &= x(1 + S(x) + y + 2y^2) \\
    1 &= x(1 + 4y)
\end{align*}
\]

of \( y = G^*(x, y) \) has no characteristic point since the only candidate is \((\rho, \tau)\) and

\[
\rho(1 + 4\tau) = \left(\frac{1}{2}\right)(\sqrt{2} - 1)(1 + 2\sqrt{2}) \neq 1.
\]

\((\rho, \tau)\) is a boundary point of the domain of \( G^*(x, y) \) whose location is not detected by the method of characteristic points.

**Remark 9.** On p. 83 of their 1989 paper \[13\] Meir and Moon offer an interesting example of a 1-equation system without a characteristic point, namely \( y = A(x)e^y \) where \( A(x) = (1/6) \sum_n x^n/n^2 \). The characteristic system is

\[
y = A(x)e^y, \quad 1 = A(x)e^y,
\]

so a characteristic point \((a, b)\) must have \( b = 1, A(a) = 1/e \). But \( 1/e \) is not in the range of \( A(x) \), so there is no characteristic point. One can nonetheless easily find \((\rho, \tau)\) in this case since \((\rho, \tau)\) must lie on the boundary of the domain of \( A(x)e^y \). Thus \( \rho = 1 \), and then \( \tau = A(1)e^\tau = (\pi^2/36)e^\tau \), so \( \tau \approx 0.41529 \).

The paper goes on to claim that by differential equation methods one can show that the standard solution \( y = S(x) \) has coefficient asymptotics \( s(n) \sim C/n \). However this cannot be true since such a solution would diverge at its radius of convergence \( \rho = 1 \) (see \[2\]), whereas the given equation \( y = A(x)e^y \) is nonlinear in \( y \), so the solution must converge at \( \rho \).

### 3.3. 1-equation framework

This subsection gives a framework for 1-equation examples which will be useful for building the 2-equation examples in §3.4.

**Proposition 10.** Let \( A(x) \) be the standard solution of the recursive quadratic equation

\[
y = x(1 + ay + by^2)
\]

where \( a \geq 0 \) and \( b > 0 \). Then the following hold:
(a) 

\[ A(x) = \frac{1}{2bx} \left( (1 - ax) - \sqrt{(1 - ax)^2 - 4bx^2} \right). \]

(b) \( A(x) \) has non-negative coefficients.

(c) A sufficient condition for \( A(x) \) to have integer coefficients is that \( a \) and \( b \) are integers.

(d) \( A(x) \) has a positive radius of convergence \( \rho_A \) given by

\[ \rho_A = \frac{1}{a + 2\sqrt{b}}. \]

(e) \( \tau_A := A(\rho_A) \) is finite and is given by

\[ \tau_A = \frac{1}{\sqrt{b}}. \]

(f) \( \rho_A \) is a square-root branch point of the algebraic curve defined by (8).

(g) \( (\rho_A, \tau_A) \) is the unique characteristic point of (8), that is, it is the unique positive solution \((x, y)\) to

\[ \begin{align*}
  y &= x(1 + ay + by^2) \\
  1 &= x(a + 2by).
\end{align*} \]

Proof. Writing (8) in the form \( bx^2 + (ax - 1)y + x = 0 \), there are two solutions:

\[ A(x) = \frac{1}{2bx} \left( (1 - ax) \pm \sqrt{(1 - ax)^2 - 4bx^2} \right). \]

The standard solution goes through the origin, which is (9). (The other solution blows up at the origin.)

(b) and (c) are standard facts about recursively defined functions.

The radius of convergence \( \rho_A \) is found by setting the discriminant equal to 0, that is,

\[ (1 - ax)^2 - 4bx^2 = 0, \]

and choosing the least positive solution. This gives (10).

Then from (9) and (10) one calculates \( \tau_A := A(\rho_A) \):

\[ \tau_A = \frac{1 - a\rho_A}{2b\rho_A} = \frac{1}{\sqrt{b}}. \]
Equation (8) is a polynomial equation, so it defines an algebraic curve $A(x)$. For such curves the only singularities are poles and branch points, and $\rho_A$ is clearly of the latter kind.

(g) is a straightforward calculation. \qed

Let $S(x)$ be the standard solution to the quadratic equation

\begin{equation}
    y = x(1 + cy + dy^2)
\end{equation}

where $c \geq 0$ and $d > 0$. Then, from the previous result,

\begin{align*}
    S(x) &= \frac{1}{2dx} \left( (1 - cx) - \sqrt{(1 - cx)^2 - 4dx^2} \right) \\
    (\rho_S, \tau_S) &= \left( \frac{1}{c + 2\sqrt{d}}, \frac{1}{\sqrt{d}} \right).
\end{align*}

$(\rho_S, \tau_S)$ is the unique positive solution $(x, y)$ to

\begin{align*}
    y &= x(1 + cy + dy^2) \\
    1 &= x(c + 2dy),
\end{align*}

and $\rho_S$ is a square-root branch-point singularity of the algebraic curve defined by (11).

**Proposition 11.** Given $a, c \geq 0$ and $b, d > 0$ let $A(x)$ be the standard solution of

\begin{equation}
    y = x(1 + ay + by^2)
\end{equation}

and let $S(x)$ be the standard solution of

\begin{equation}
    y = x(1 + cy + dy^2).
\end{equation}

Let $T(x)$ be the standard solution of the recursive quadratic equation

\begin{equation}
    y = A(x)(1 + cy + dy^2).
\end{equation}

Then the following hold:

(a) $T(x) = S(A(x))$.

(b) $T(x) = \frac{1}{2dA(x)} \left( (1 - cA(x)) - \sqrt{(1 - cA(x))^2 - 4dA(x)^2} \right)$.

(c) $T(x)$ has non-negative coefficients.

(d) A sufficient condition for $T(x)$ to have integer coefficients is that $a, b, c, d$ are integers.
(e) If \( \sqrt{b} = c + 2\sqrt{d} \) then

\[
(\rho_T, \tau_T) = (\rho_A, \tau_S) = \left( \frac{1}{a + 2\sqrt{b}}, \frac{1}{\sqrt{d}} \right),
\]

and \( T(x) \) has a fourth-root singularity at \( \rho_T \).

**Proof.** Since \( S(x) \) solves (11),

\[
S(x) = x \left( 1 + c S(x) + d S(x)^2 \right).
\]

Substituting \( A(x) \) for \( x \) gives

\[
S(A(x)) = A(x) \left( 1 + c S(A(x)) + d S(A(x))^2 \right).
\]

Thus \( T(x) := S(A(x)) \) is the standard solution to \( y = A(x)(1 + cy + dy^2) \).

Since (8) defines an algebraic curve with a square-root branching point at \( \rho_A \), and likewise (11) defines an algebraic curve with a square-root branching point at \( \rho_S \), the composition \( T(x) = S(A(x)) \) defines part of an algebraic curve with a fourth-root branching point \( \rho_A \).

Given the premises of Proposition 11, the restriction \( \sqrt{b} = c + 2\sqrt{d} \) is called the critical composition condition (CCC); this is the condition needed for \( T(x) = S(A(x)) \) to be a critical composition (as defined by Flajolet and Sedgewick [7], p. 411).

3.4. Multi-equation systems. For systems \( y = G(x, y) \) of two or more equations one again has \( (\rho, \tau) \) in the interior or on the boundary of the domain of \( G \). However in general the situation is more involved than in the case of a single equation. Unlike the one-equation case, it is possible that there are multiple characteristic points, in which case they form an antichain (see Lemma 26). If \( (\rho, \tau) \) is a characteristic point then (see Proposition 31) among the characteristic points \( (a, b) \) it is the one with the largest value of \( a \).

The goal of this paper is to show that there is a simple additional constraint that one can put on characteristic points so that they are just as useful in the multi-equation case as in the single equation case. Namely (see §5) the property

1 is the largest positive eigenvalue of \( J_G(a, b) \)

gives a definitive test for a characteristic point \( (a, b) \) to be \( (\rho, \tau) \). A characteristic point with this property will be called an eigenpoint.
Before considering eigenpoints, a family of 2-equation systems is constructed, based on the results of §3.3. These systems yield the main examples in this paper, namely Examples 16 and 17.

**Proposition 12.** Suppose

\[ a, c_1 \geq 0, \quad b, c_2, \delta > 0, \quad \sqrt{b} = c + 2\sqrt{\delta}, \quad c = c_1 + c_2. \]

Let \( A(x) \) be the standard solution of

\[ y = x(1 + ay + by^2), \]

let \( S(x) \) be the standard solution of

\[ y = x(1 + cy + \delta y^2), \]

and let \( T(x) \) be the standard solution of

\[ y = A(x)(1 + cy + \delta y^2). \]

Then the following hold:

(a) The quadratic system

\[
(SYS): \quad \begin{cases}
y_1 = A(x)(1 + c_1 T(x) + c_2 y_2 + \delta y_1^2) \\
y_2 = A(x)(1 + c_1 T(x) + c_2 y_1 + \delta y_2^2)
\end{cases}
\]

is well-conditioned, and the standard solution is \( y_1 = y_2 = T(x) \).

(b) The extreme point \((\rho, \tau, \tau)\) of \((SYS)\) is given by

\[
(\rho, \tau, \tau) = \left( \frac{1}{a + 2\sqrt{b}} , \frac{1}{\sqrt{\delta}} , \frac{1}{\sqrt{\delta}} \right).
\]

It is on the boundary of the domain of \((SYS)\).

(c) \( T(x) = S(A(x)) \) has a fourth-root singularity at \( x = \rho \).

(d) A positive point \((x, y, y)\) is a characteristic point of \((SYS)\) iff either

\[
(*) \quad \begin{cases}
1 = A(x) \left( c_2 + 2\sqrt{\delta(1 + c_1 T(x))} \right) \\
y = \frac{1 - c_2 A(x)}{2\delta A(x)}
\end{cases}
\]
or

\[
\begin{align*}
\text{(**) } & \left\{ \begin{array}{l}
1 = A(x)\left( c_2 + 2\sqrt{c_2^2 + d(1 + c_1 T(x))} \right) \\
y = \frac{1 + c_2 A(x)}{2dA(x)}.
\end{array} \right.
\end{align*}
\]

(e) If \( c_1 = 0 \) then there are exactly two characteristic points of the form \((x, y, y)\): the first is \((\rho, \tau, \tau)\), a boundary characteristic point obtained from (\(\star\)), and the second is the unique positive solution to (\(\star\))\(^a\), an interior characteristic point. This is the only case where (\(\star\)) contributes a characteristic point, namely \((\rho, \tau, \tau)\), and this is the only case where \((\rho, \tau, \tau)\) is a characteristic point.

(f) If \( 0 < c_1 = 2c_2 \) then there is a unique characteristic point of the form \((x, y, y)\): it is the unique positive solution to (\(\star\))\(^a\) and it is a boundary point different from \((\rho, \tau, \tau)\).

(g) If \( 0 < c_1 < 2c_2 \) then there is a unique characteristic point of the form \((x, y, y)\): it is the unique positive solution to (\(\star\))\(^a\) and it is an interior point that is different from \((\rho, \tau, \tau)\).

(h) If \( 2c_2 < c_1 \) then there are no characteristic points of the form \((x, y, y)\), so again \((\rho, \tau, \tau)\) is not a characteristic point.

(i) The second characteristic point in (e) and the unique characteristic points in (f) and (g) are given explicitly by

\[
\begin{align*}
x &= \frac{c + \sqrt{c^2 + f}}{ac + 2c^2 + f + b + (a + 2c)\sqrt{c^2 + f}} \\
y &= \frac{c + c_2 + \sqrt{c^2 + f}}{2d}
\end{align*}
\]

where

\[
f = -6c_1c_2 + 3c_2^2 + 4d.
\]

**Proof.** The system is irreducible since \( c_2 > 0 \). \( T(x) \) satisfies \([13]\), consequently \( y_1 = y_2 = T(x) \) is a solution of \((SYS)\). This must be the standard solution of \((SYS)\), so the extreme point of \((SYS)\) is the stated \((\rho, \tau, \tau)\) (see \([3.3]\)), a point which is not in the interior of the domain of the system \((SYS)\) since \(A(x)\) has a singularity at \(\rho_A\), and \(\rho_A = \rho\) by \((CCC)\). This takes care of items (a)–(c).
The characteristic system of \((SYS)\) is
\[
(CSYS) : \begin{cases}
y_1 &= A(x)(1 + c_1 T(x) + c_2 y_2 + \vartheta y_1^2) \\
y_2 &= A(x)(1 + c_1 T(x) + c_2 y_1 + \vartheta y_2^2) \\
0 &= (1 - 2\vartheta A(x)y_1)(1 - 2\vartheta A(x)y_2) - c_2^2 A(x)^2;
\end{cases}
\]
so a positive point \((x, y, y)\) is a characteristic point of \(\Sigma\) iff \((x, y)\) satisfies
\[
(CSYS') : \begin{cases}
y &= A(x)(1 + c_1 T(x) + c_2 y + \vartheta y^2) \\
0 &= (1 - 2\vartheta A(x)y)^2 - c_2^2 A(x)^2.
\end{cases}
\]
This gives the conditions \((\ast)\) and \((\ast\ast)\) of item (d) as follows.

\((CSYS')\) yields
\[
1 = (2\vartheta \pm c_2) A(x), \tag{14}
\]
and then
\[
(2\vartheta \pm c_2)y = 1 + c_1 T(x) + c_2 y + \vartheta y^2, \tag{15}
\]
so
\[
\vartheta y^2 + (-1 \pm 1)c_2y = 1 + c_1 T(x).
\]
From \((14)\) and \((15)\) there are two cases to consider:

Case 1:
\[
1 = (2\vartheta + c_2) A(x) \quad \text{and} \quad \vartheta y^2 = 1 + c_1 T(x).
\]

Case 2:
\[
1 = (2\vartheta - c_2) A(x) \quad \text{and} \quad \vartheta y^2 - 2c_2 y = 1 + c_1 T(x).
\]
In each case, solving for \(y\) in the two equations and equating the results gives an equation to solve for \(x\):

For Case 1:
\[
y = \frac{1 - c_2 A(x)}{2\vartheta A(x)} = \sqrt{\frac{1 + c_1 T(x)}{\vartheta}} \tag{16}
\]
For Case 2:
\[
y = \frac{1 + c_2 A(x)}{2\vartheta A(x)} = \frac{1}{2\vartheta} \left(2c_2 + \sqrt{4c_2^2 + 4\vartheta(1 + c_1 T(x))}\right) \tag{17},
\]
Note that one cannot use the negative sign before the square root when solving for \( y \) in Case 2 since that would make \( y \) negative.

**CASE 1:**
To find candidates for \( x \) in the first case one needs to solve (16), which implies:

\[
1 = A(x) \left( c_2 + 2 \sqrt{d \left( 1 + c_1 T(x) \right)} \right)
\]

By the monotonicity of the right side, there is at most one positive \( x \) solving this equation. If such exists, then one finds \( y \) by (16):

\[
y := \frac{1 - c_2 A(x)}{2d A(x)} \left( \sqrt{\frac{1 + c_1 T(x)}{d}} > 0 \right).
\]

Thus if (18) has a positive solution for \( x \) then defining \( y \) by (19) gives the positive point \((x, y, y)\).

Conversely, given a positive point \((x, y, y)\) satisfying (18) and (19), the above steps are reversible to show that \((x, y)\) satisfies \((CSYS')\), so \((x, y, y)\) is a characteristic point of \((SYS)\).

**CASE 2:**
To find candidates for \( x \) in the second case one needs to solve (17), which implies:

\[
1 = A(x) \left( c_2 + 2 \sqrt{c_2^2 + d \left( 1 + c_1 T(x) \right)} \right)
\]

By the monotonicity of the right side, there is at most one positive \( x \) for this case. If such exists, then let

\[
y := \frac{1 + c_2 A(x)}{2d A(x)} > 0.
\]

Thus \((x, y, y)\) is a positive point.

Conversely, given a positive point \((x, y, y)\) satisfying (20) and (21), the above steps are reversible to show that \((x, y)\) satisfies \((CSYS')\), so \((x, y, y)\) is a characteristic point of \((SYS)\). This finishes the proof of (d).

For items (e)–(h) note that \((\rho, \tau, \tau)\) is a characteristic point of \((SYS)\) iff the second equation of \((CSYS')\) holds at this point, that is, iff

\[
0 = \left( 1 - 2d A(\rho) \tau \right)^2 - c_2^2 A(\rho)^2.
\]
Since $\tau = \tau_S$ and $A(\rho) = \tau_A = \rho_S$, \[(23) \quad 0 = (1 - 2d\rho_S\tau_S)^2 - c_2^2 \rho_S^2.\]

Now from (12) we have $1 = \rho_S(\epsilon + 2d\tau_S)$, so (23) holds iff

$$0 = \epsilon^2 \rho_S^2 - c_2^2 \rho_S^2,$$

and this holds iff $0 = \epsilon^2 - c_2^2$ iff $\epsilon_1 = 0$.

Let

$$F_1(x) := A(x)\left(c_2 + 2\sqrt{d(1 + c_1T(x))}\right)$$

$$F_2(x) := A(x)\left(c_2 + 2\sqrt{c_2^2 + d(1 + c_1T(x))}\right).$$

Then the first equations from ($\star$) and ($\star\star$) are

$$1 = F_1(x) \quad (24)$$

$$1 = F_2(x). \quad (25)$$

The functions $F_1(x)$ and $F_2(x)$ have radius of convergence $= \rho$, they are strictly increasing on $[0, \rho]$, and $F_1(0) = F_2(0) = 0$. Thus (24) has a positive solution for $x$ iff $1 \leq F_1(\rho)$, and (25) has a positive solution iff $1 \leq F_2(\rho)$.

There are two possibilities for a positive solution $x$ of $1 = F_i(x)$, for $i = 1, 2$:

$$x = \rho \text{ is a solution iff } 1 = F_i(\rho) \quad (26)$$

$$\text{some } x < \rho \text{ is a solution iff } 1 < F_i(\rho). \quad (27)$$

Now

$$F_1(\rho) = A(\rho)\left(c_2 + 2\sqrt{d(1 + c_1T(\rho))}\right)$$

$$= \frac{1}{\epsilon^2 + 2\sqrt{d}}\left(c_2 + 2\sqrt{d + c_1\sqrt{d}}\right),$$

so

$$1 = F_1(\rho) \quad \text{iff} \quad 1 = \frac{1}{\epsilon^2 + 2\sqrt{d}}\left(c_2 + 2\sqrt{d + c_1\sqrt{d}}\right)$$

$$\text{iff } \epsilon_1 = 0.$$
Likewise

\[ 1 < F_1(\rho) \iff \epsilon_1^2 < 0, \]

which is impossible.

Also,

\[
1 = F_2(\rho) \iff 1 = \frac{1}{\epsilon + 2\sqrt{\partial}} \left( \epsilon_2 + 2\sqrt{\epsilon_2^2 + \partial + \epsilon_1\sqrt{\partial}} \right)
\]

iff \( \epsilon_1 = 2\epsilon_2 \),

and

\[ 1 < F_2(\rho) \iff \epsilon_1 < 2\epsilon_2. \]

From the above one has the following cases, noting that \( F_1(x) < F_2(x) \) on \((0, \rho] \):

- \( \epsilon_1 = 0 \iff 1 = F_1(\rho) < F_2(\rho) \)
- \( 0 < \epsilon_1 = 2\epsilon_2 \iff F_1(\rho) < 1 = F_2(\rho) \)
- \( 0 < \epsilon_1 < 2\epsilon_2 \iff F_1(\rho) < 1 < F_2(\rho) \)
- \( 0 < 2\epsilon_2 < \epsilon_1 \iff F_1(\rho) < F_2(\rho) < 1. \)

Finally now consider (i). With the assumptions of (e) we have \( \epsilon_1 = 0 \). Condition (\( \star \)) gives the characteristic point

\[
(\rho, \tau, \tau) = \left( \frac{1}{a + 2\sqrt{\partial}}, \frac{1}{\sqrt{\partial}}, \frac{1}{\sqrt{\partial}} \right).
\]

Consider (\( \star \)). The first equation of (\( \star \)) is independent of \( T(x) \), giving

\[
A(x) = \frac{1}{\epsilon_2 + 2\sqrt{\epsilon_2^2 + \partial}}.
\]

Solving (9) for \( x \), using (28), and solving the second equation of (\( \star \)) for \( y \) using (28) for \( y \) and rearranging gives the desired expressions for \( x \) and \( y \) in the special case \( \epsilon_1 = 0 \).

Assume \( 0 < \epsilon_1 \leq 2\epsilon_2 \). Substitute \( y_1 = T(x) \), \( y_2 = T(x) \) into \( (SYS) \). Both equations of \( (SYS) \) give the same equation in \( A(x) \) and \( T(x) \), namely:

\[
0 = -T(x) + A(x) \left( 1 + (\epsilon_1 + \epsilon_2)T(x) + \partial T(x)^2 \right).
\]
From the first equation of (⋆⋆) one also has

\[(30) \quad 1 = A(x) \left( e_2 + 2 \sqrt{e_2^2 + d\left(1 + e_1 T(x)\right)}\right).\]

Solving (30) for \(T(x)\) gives a rational expression in \(A(x)\) which has denominator a multiple of \(A(x)^2\) and has numerator a quadratic polynomial in \(A(x)\). Substitute this in for \(T(x)\) in (29). Simplifying this gives a quartic polynomial in \(A(x)\) equal to 0. This polynomial can be solved explicitly for \(A(x)\). Let

\[
\begin{align*}
\mathfrak{f} &= -6c_1c_2 + 3c_2^2 + 4d \\
\mathfrak{g} &= 2c_1c_2 + 3c_2^2 + 4d
\end{align*}
\]

If \(\mathfrak{f} \neq 0\) then solving the polynomial gives

\[
\begin{align*}
A_1 &= \frac{-c + \sqrt{(c_1 + 2c_2)^2 + 4d}}{2c_1c_2 + 3c_2^2 + 4d} = \frac{-c + \sqrt{c^2 + \mathfrak{g}}}{\mathfrak{g}} \\
A_2 &= \frac{-c - \sqrt{(c_1 + 2c_2)^2 + 4d}}{2c_1c_2 + 3c_2^2 + 4d} = \frac{-c - \sqrt{c^2 + \mathfrak{g}}}{\mathfrak{g}} \\
A_3 &= \frac{-c + \sqrt{(c_1 - 2c_2)^2 + 4d}}{-6c_1c_2 + 3c_2^2 + 4d} = \frac{-c + \sqrt{c^2 + \mathfrak{f}}}{\mathfrak{f}} \\
A_4 &= \frac{-c - \sqrt{(c_1 - 2c_2)^2 + 4d}}{-6c_1c_2 + 3c_2^2 + 4d} = \frac{-c - \sqrt{c^2 + \mathfrak{f}}}{\mathfrak{f}}.
\end{align*}
\]

If \(\mathfrak{f} = 0\) then the quartic polynomial simplifies to the cubic polynomial

\[1 - 4cA(x) + 4(c_1^2 + c_2^2)A(x)^2 + 16cc_1c_2A(x)^3\]

Solving for \(A(x)\) gives \(A_1\) and \(A_2\) as before and the third root is

\[A_5 = \frac{1}{2c}\]

All that remains to do is to identify which \(A_i\) gives the characteristic point and then solve for \(x\) from (9) and for \(y\) from the second equation of (⋆⋆) and rearrange to obtain the explicit formulae.

The identification of the correct \(A_i\) is given in the following Lemmas.

\[\square\]

In order to compare the \(A_i\) the following elementary lemma will be handy.
Lemma 13. Suppose $c > 0$, $-c^2 \leq t \leq u$, $u \neq 0$, $t \neq 0$. Then

$$\frac{-c + \sqrt{c^2 + t}}{t} \geq \frac{-c + \sqrt{c^2 + u}}{u}$$

Proof. Simply calculate keeping track of the signs of $t$ and $u$:

$$\frac{-c + \sqrt{c^2 + t}}{t} \geq \frac{-c + \sqrt{c^2 + u}}{u}$$

$$\Rightarrow u\sqrt{c^2 + t} - t\sqrt{c^2 + u} \begin{cases} \geq c(u - t) & \text{if } t, u < 0 \text{ or } t, u > 0 \\ \leq (u - t) & \text{if } t < 0 < u \end{cases}$$

$$\Rightarrow u^2 + t^2 u + 2c^2 tu \begin{cases} \geq 2tu\sqrt{(c^2 + t)(c^2 + u)} & \text{if } t, u < 0 \text{ or } t, u > 0 \\ \leq (u - t)^2 & \text{if } t < 0 < u \end{cases}$$

$$\Rightarrow (u - t)^2 \geq 0$$

The second and fourth implications are by squaring both sides and rearranging; in both cases both sides of the previous step are positive regardless of the signs of $t$ and $u$. \qed

Lemma 14. With notation as in the proof of Proposition 12, if $f = 0$ then $A_5$ gives the characteristic point while if $f \neq 0$ then $A_3$ gives the characteristic point.

Proof. First determine when each $A_i$ is positive. $A_2$ is negative, and $A_1$ and $A_5$ are positive since $c_1 \geq 0$, $c_2 > 0$, and $d > 0$.

$A_3$ is no smaller than $A_1$ and hence positive by Lemma 13. $A_4$ has negative numerator and so is positive provided $f < 0$.

Comparing the $A_i$,

$$A_2 \leq 0 \leq A_1 \leq A_3 \leq A_4 \text{ when } f = 0$$

Also, $A_4$ can be discarded even when $f < 0$ as it is greater than the radius, which due to (CCC) is $1/(c + 2\sqrt{d})$. This is because $f + c^2 \geq 0$ so

$$f + c^2 + 2c\sqrt{d} \geq 0 \geq -(c + 2\sqrt{d})\sqrt{c^2 + f}$$

$$\Rightarrow \frac{1}{c + 2\sqrt{d}} \leq \frac{-c - \sqrt{c^2 + f}}{f}$$
Another source of spurious solutions is the fact that \((29)\) is quadratic in \(T\). Solving \((29)\) for \(T\) we are interested in the lower branch — the negative sign for the square root. So we must have

\[
-1 + 2c A_i + (-2c_1^2 - 2c_1 c_2 + 3c_2^2 + 4\delta)A_i^2 \geq 0
\]

As a polynomial in \(A_i\) this is negative at \(A_i = 0\) and has roots

\[
\text{Crit}_\pm = \frac{-c_1 \pm \sqrt{(2c_2 - c_1)(2c_2 + c_1) + 4\delta}}{-2c_1^2 - 2c_1 c_2 + 3c_2^2 + 4\delta} = \frac{-c \pm \sqrt{c^2 + h}}{h}
\]

where

\[
h = -2c_1^2 - 2c_1 c_2 + 3c_2^2 + 4\delta
\]

When \(h = 0\) \((31)\) becomes

\[
A_i \geq \frac{1}{2c}
\]

Suppose \(h = 0\). Then \(f = 2c_1(c_1 - 2c_2) \leq 0\) and \(g = 2c_1(c_1 + 2c_2)\). Similarly to Lemma 13 for any \(-c^2 \leq t, t \neq 0,

\[
\frac{1}{2c} \begin{cases} 
\geq & \frac{-c + \sqrt{c^2 + t}}{t} \quad \text{if } t > 0 \\
\leq & \frac{-c + \sqrt{c^2 + t}}{t} \quad \text{if } t < 0 
\end{cases}
\]

since

\[(32)\] holds

\[
\Leftrightarrow t + 2c^2 \geq 2c\sqrt{c^2 + t}
\]

\[
\Leftrightarrow t^2 \geq 0 \text{ by squaring both sides}
\]

Applying the above with \(f\) and \(g\) we see that \(A_1\) does not satisfy the requirement but \(A_3\) does. Also, when \(f = 0\), \(A_5\) trivially satisfies the requirement and so by the
same calculation $A_1 \leq A_5$. Thus $A_5$ gives the characteristic point if $f = 0$ and $A_3$ otherwise.

Now suppose $\mathfrak{h} \neq 0$. Note that since $c_1 \leq 2c_2$ we have $\mathfrak{h} \geq f$, with equality when $c_1 = 2c_2$. Thus by Lemma 13,

$$A_2 \leq 0 \leq A_1 \leq \text{Crit}_+ \leq A_3 \ (\leq A_4 \text{ when } f < 0)$$

When $\mathfrak{h} > 0$, $\text{Crit}_- < 0$ and so the positive part of (31) is below $\text{Crit}_-$ and above $\text{Crit}_+$. Also by (32) $\text{Crit}_+ < A_5$, so if $f = 0$ then $A_5$ must give the characteristic point. On the other hand if $f \neq 0$ then $A_4$ is out of the picture as it is either negative or larger than the radius and so $A_3$ must give the characteristic point.

Suppose $\mathfrak{h} < 0$. Then $f < 0$. Similarly to Lemma 13

$$\text{Crit}_- \geq A_4$$

since

$$\frac{-c - \sqrt{c^2 + \mathfrak{h}}}{\mathfrak{h}} \geq \frac{-c - \sqrt{c^2 + f}}{f}$$

$$\Leftrightarrow (\mathfrak{h} - f)c \geq -\mathfrak{h}\sqrt{c^2 + f} + f\sqrt{c^2 + \mathfrak{h}}.$$

The left hand side of the last line is nonnegative, so either the right hand is negative and we are done or we can square both sides and continue the chain of equivalences:

$$\Rightarrow -c - \sqrt{c^2 + \mathfrak{h}} \geq -f\sqrt{(c^2 + f)(c^2 + \mathfrak{h})}$$

which is true since the left hand side is nonnegative and the right hand side is negative.

Thus in this case

$$A_2 \leq 0 \leq A_1 \leq \text{Crit}_+ \leq A_3 \leq A_4 \leq \text{Crit}_-$$

and the region where (31) is positive is the region between $\text{Crit}_+$ and $\text{Crit}_-$. $A_4$ is larger than the radius and so again $A_3$ must give the critical point.

3.5. **Examples of 2-equation systems.** Now we look at three well-conditioned examples that show some of the varied behavior of characteristic points when one has more than one equation in the system. In the first example there are two characteristic points, both in the interior of the domain of $G(x,y)$ and one of them is
(ρ, τ). In the second example one has a characteristic point in the interior of the domain of $G(x, y)$ and $(ρ, τ)$ is a characteristic point on the boundary of the domain. In the third example one has a characteristic point in the interior of the domain of $G(x, y)$ but $(ρ, τ)$ is not a characteristic point. In the second and third examples, ρ is not a square-root singularity of the solutions. Such examples show the need for a more subtle use of characteristic points in the pursuit of information on $(ρ, τ)$ for multi-equation systems.

**Example 15.** For the system of two equations

\begin{align*}
y_1 &= x \cdot (1 + y_2 + 2y_1^2) \\
y_2 &= x \cdot (1 + y_1 + 2y_2^2)
\end{align*}

add

\[(1 - 4xy_1)(1 - 4xy_2) - x^2 = 0\]

to obtain the characteristic system. This is a polynomial system, so all characteristic points will be in the interior of the domain; and since $(ρ, τ_1, τ_2)$ is also in the interior it must be a characteristic point. Let $(a, b, c)$ be a characteristic point. If $b \neq c$ then subtracting the second equation from the first gives

\[b - c = a(c - b + 2b^2 - 2c^2).\]

Dividing through by $a(b - c)$ gives

\[\frac{1}{a} = 2(b + c) - 1.\]

The third equation says

\[\left(\frac{1}{a} - 4b\right)\left(\frac{1}{a} - 4c\right) = 1;\]

substituting in the expression for $1/a$ from the previous equation gives

\[(2c - 2b - 1)(2b - 2c - 1) = 1.\]

Thus

\[1 - 4(b - c)^2 = 1,\]

which is impossible if $b \neq c$. Thus $b = c$, so the characteristic points are the positive triples $(a, b, b)$ satisfying

\[b = a(1 + b + 2b^2)\]
\[ a^2 = (1 - 4ab)^2. \]

From this the system has two characteristic points:

\[
\left( \frac{2\sqrt{2} - 1}{7}, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right) \approx (0.2612, 0.7071, 0.7071)
\]

\[
\left( \frac{2\sqrt{3} - 1}{11}, \frac{1 + \sqrt{3}}{2}, \frac{1 + \sqrt{3}}{2} \right) \approx (0.2240, 1.3660, 1.3660).
\]

Now we are left with determining which of the two characteristic points gives \((\rho, \tau_1, \tau_2)\).
By applying either Proposition 31 or Lemma 32 below, it is the first of these.

Next we apply the results of the previous section to create two examples showing that the behavior of the extreme point is relatively independent of whether or not there are characteristic points in the interior of the domain of \(G\).

**Example 16.** Let \(a = 0\), \(b = 9\), \(c_1 = 0\), \(c_2 = 1\), and \(d = 1\). These numbers satisfy (CCC). Following the hypotheses of Theorem 12, let \(A(x)\) be the standard solution to \(y = x(1 + 9y^2)\) and consider the system

\[
y_1 = A(x) \cdot (1 + y_2 + y_1^2)
\]

\[
y_2 = A(x) \cdot (1 + y_1 + y_2^2).
\]

Since \(c_1 = 0\) there are two characteristic points of the form \((a, b, b)\). The first is the extreme point

\[ (\rho, \tau_1, \tau_2) = (1/6, 1, 1) \]

which lies on the boundary of the domain, and the second is the interior point obtained from the formulas in Proposition 12 (i):

\[
\left( \frac{1 + 16\sqrt{2}}{146}, 1 + \sqrt{2}, 1 + \sqrt{2} \right).
\]

**Example 17.** Let \(a = 0\), \(b = 16\), \(c_1 = 1\), \(c_2 = 1\), and \(d = 1\). These numbers satisfy (CCC). Following the hypotheses of Theorem 12, let \(A(x)\) be the standard solution to \(y = x(1 + 16y^2)\), and let \(T(x)\) be the standard solution to \(y = A(x)(1 + 2y + y^2)\). Consider the system

\[
y_1 = A(x) \cdot (1 + T(x) + y_2 + y_1^2)
\]

\[
y_2 = A(x) \cdot (1 + T(x) + y_1 + y_2^2).
\]
Since $0 < c_1 < 2c_2$, the extreme point 
\[(\rho, \tau_1, \tau_2) = (1/8, 1, 1)\]
is not a characteristic point, but there is a characteristic point of the form $(a, b, b)$ in the interior of the domain of $G$ given by the formulas of Proposition 12:
\[(a, b, b) = \left(\frac{30 + 17\sqrt{5}}{545}, \frac{3 + \sqrt{5}}{2}, \frac{3 + \sqrt{5}}{2}\right)\].

4. Characteristic Points of Well-Conditioned Systems

From now on it is assumed, unless stated otherwise, that we are working with a well-conditioned system $\Sigma : y = G(x, y)$ of $m$ equations.

4.1. Making substitutions in an irreducible system. A careful analysis of the characteristic points of $\Sigma$ is easier if $J_G(x, y)$ is a positive matrix for positive points $(x, y)$; this is the case precisely when no of entry $J_G(x, y)$ is 0. Fortunately there is a simple substitution procedure to transform the original system $\Sigma$ into a well-conditioned system $\Sigma^*$ with

(i) exactly the same positive solutions $(a, b)$, and

(ii) exactly the same set $\mathcal{CP}$ of characteristic points,
even though the Jacobian $J_{G^*}(x, y)$ of the new system can be quite different from that of the original.

The simplest substitutions to consider are $n$-fold iterations $G^{(n)}$ of the transformation $G$ defined by:
\[
G^{(1)}(x, y) := G(x, y) \\
G^{(n+1)}(x, y) := G^{(n)}(x, G_1(x, y), \ldots, G_m(x, y)).
\]

This uniform substitution method is used in [7] (see p. 492) on aperiodic polynomial systems $\Sigma$. The following example shows that, in general (without the aperiodic hypothesis), this uniform method of substitution fails to produce a system $\Sigma^*$ such that the Jacobian matrix $J_{G^*}(x, y)$ has all entries non-zero. However a more general form of substitution will provide the desired $\Sigma^*$.

---

A well-conditioned system $y = G(x, y)$ is aperiodic if the coefficients of each $T_i(x)$ are eventually positive, $T(x)$ being the standard solution—see [7], p. 489.
Example 18. Consider the irreducible system $y = G(x,y)$ of 4 equations:

$$\Sigma = \begin{cases} 
    y_1 &= G_1(x,y_1,\ldots,y_4) := x(1 + y_2^2 + y_4^2) \\
    y_2 &= G_2(x,y_1,\ldots,y_4) := x(1 + y_1^2 + y_3^2) \\
    y_3 &= G_3(x,y_1,\ldots,y_4) := x(1 + y_4^2) \\
    y_4 &= G_4(x,y_1,\ldots,y_4) := x(1 + y_1^2). 
\end{cases}$$

Let $M = J_{G^{(n)}}$. Then it is easy to check that $M_{11} \neq 0$ iff $n$ is odd, and $M_{12} \neq 0$ iff $n$ is even. Thus for $n \geq 1$, $J_{G^{(n)}}(x,y)$ has entries which are 0.

One can transform $\Sigma$ into a system $\Sigma^*$ where the Jacobian of $G^*$ has all entries non-zero by doing selective substitutions. For example, in the first equation of $\Sigma$ replace one of the two $y_2$'s by $G_2(x,y)$, giving the system

$$\Sigma^* = \begin{cases} 
    y_1 &= x(1 + y_2G_2(x,y) + y_4^2) \\
    y_2 &= x(1 + y_1^2 + y_3^2) \\
    y_3 &= x(1 + y_4^2) \\
    y_4 &= x(1 + y_1^2). 
\end{cases}$$

The first equation in this system is such that the right hand side depends on all 4 of the $y_i$. Continuing in this manner one obtains a system

$$\Sigma^* = \begin{cases} 
    y_1 &= G_1^*(x,y_1,\ldots,y_4) \\
    \vdots \\
    y_4 &= G_4^*(x,y_1,\ldots,y_4) 
\end{cases}$$

in which every $G_i^*(x,y)$ depends on each of $y_1,\ldots,y_4$.

Clearly every positive solution $(a,b)$ of the original system $\Sigma$ is also a solution of $\Sigma^*$. We need to look at the question of whether or not there are any further positive solutions of $\Sigma^*$, and how the set of characteristic points of $\Sigma^*$ compares to that of $\Sigma$. One finds in general, with a little caution in choosing the substitutions (in order to maintain irreducibility), that one can create a $\Sigma^*$ that has exactly the same positive solutions, and the same characteristic points, as $\Sigma$. The need for a carefully selected sequence of substitutions $y_j \leftarrow G_j(x,y)$ is largely removed in the following by using fractional substitutions $y_j \leftarrow \alpha G_j(x,y) + (1 - \alpha)y_j$, $\alpha \in (0,1)$. 
Given a system $\Sigma : y = G(x, y)$, a minimal self-substitution transformation creates the system $\Sigma_\alpha : y = G^{(\alpha)}(x, y)$ by selecting $\alpha \in [0, 1]$ and a pair of indices $i, j$ (possibly the same) with $\partial G_i(x, y)/\partial y_j \neq 0$ and then substituting $\alpha G_j(x, y) + (1 - \alpha)y_j$ for a single occurrence of $y_j$ in $G_i$. Suppose $H(x, y_0; y)$ is the result of replacing the single occurrence of $y_j$ in $G_i$ by a new variable $\alpha y_0$. Then the system $\Sigma_\alpha$ is

$$\Sigma_\alpha : \begin{cases} 
    y_1 = G_1^{(\alpha)}(x, y) := G_1(x, y) \\
    \vdots \\
    y_i = G_i^{(\alpha)}(x, y) := H(x, \alpha G_j(x, y) + (1 - \alpha)y_j; y) \\
    \vdots \\
    y_m = G_m^{(\alpha)}(x, y) := G_m(x, y)
\end{cases}$$

More generally, a system $\Sigma^* : y = G^*(x, y)$ is a self-substitution transform of $\Sigma : y = G(x, y)$ if there is a sequence $\Sigma_0, \Sigma_1, \ldots, \Sigma_r$ of systems such that $\Sigma = \Sigma_0$, $\Sigma^* = \Sigma_r$, and for $0 \leq i < r$ the system $\Sigma_{i+1}$ is a minimal self-substitution transform of $\Sigma_i$.

**Lemma 19.** For $\Sigma_\alpha$ and $\Sigma^*$ as described above:

(a) $\Sigma = \Sigma_0$.
(b) If $\Sigma$ is irreducible and $\alpha \in [0, 1)$ then $\Sigma_\alpha$ is irreducible.
(c) Suppose $\Sigma$ is irreducible. Then $\Sigma^*$ is irreducible iff each step $\Sigma_i$ is irreducible.
(d) Suppose $\Sigma$ is well-conditioned and $\alpha \in [0, 1]$. Then $\Sigma_\alpha$ is well-conditioned iff it is irreducible.
(e) Suppose $\Sigma$ is well-conditioned. Then $\Sigma^*$ is well-conditioned iff it is irreducible iff each step $\Sigma_i$ is irreducible.

Proof. (a) is obvious. For (b) note that $\alpha \in [0, 1)$ implies that $G_i(x, y)$ is still dependent on $y_j$ after the using the substitution $y_j \leftarrow \alpha G_j(x, y) + (1 - \alpha)y_j$ on a single occurrence of $y_j$ in $G_i(x, y)$. (c) follows from noting that if any step $\Sigma_i$ is reducible, then so is every step after it. (d) follows from the definition of irreducible and (b). (e) follows from (d). □

**Lemma 20.** Suppose $\Sigma^* : y = G^*(x, y)$ is a self-substitution transform of a well-conditioned $\Sigma : y = G(x, y)$. Then the following hold:

(a) $G(x, y)$ and $G^*(x, y)$ have the same positive domain of convergence.
(b) $\Sigma^*$ and $\Sigma$ have the same positive solutions and the same characteristic points.
(c) If $\Sigma^*$ is well-conditioned then $\Sigma$ and $\Sigma^*$ have the same standard solution $T(x)$ and extreme point $(\rho, \tau)$.

(d) If $\Sigma^*$ is well-conditioned then the Jacobians $J_{G}(x, y)$ and $J_{G^*}(x, y)$ have all entries finite at the same positive points $(a, b)$ in the domain of $G$.

Proof. It suffices to prove this for the case that $\Sigma^* = \Sigma_\alpha$, a minimal self-substitution transform of $\Sigma$ as described above, namely substituting $\alpha G_j(x, y) + (1 - \alpha) y_j$ for a single occurrence of $y_j$ in $G_i(x, y)$. Let

$$H(x, y_0; y) = A(x, y)y_0 + B(x, y),$$

where $A(x, y)$ and $B(x, y)$ are power series with non-negative coefficients, and neither is 0, be such that

$$G_i(x, y) = A(x, y)y_j + B(x, y)$$
$$G_i^{(\alpha)}(x, y) = A(x, y)(\alpha G_j(x, y) + (1 - \alpha)y_j) + B(x, y).$$

For item (a), first suppose that $(a, b) \in \text{Dom}^+(G)$. Then $A(a, b)$ and $B(a, b)$ are finite, so $G_i^{(\alpha)}(a, b)$ is finite. This suffices to show $(a, b) \in \text{Dom}^+(G^{(\alpha)})$ since the other $G_j^{(\alpha)}(x, y)$ are the same as those in $\Sigma$. Conversely, suppose $(a, b) \in \text{Dom}^+(G^{(\alpha)})$. Again $A(a, b)$ and $B(a, b)$ are finite, so $G_i(a, b)$ is finite; and as before, the other $G_j(a, b)$ are finite. Thus $(a, b) \in \text{Dom}^+(G)$.

For item (b), if $i \neq j$ then clearly the two systems have the same positive solutions since $y_j = G_j(x, y)$ is in both systems.

If $i = j$ first note that every positive solution of $\Sigma$ is also a solution of $\Sigma_\alpha$. We have

$$G_i(x, y) = A(x, y)y_i + B(x, y)$$
$$G_i^{(\alpha)}(x, y) = A(x, y)(\alpha A(x, y)y_i + B(x, y)) + B(x, y)$$
$$= A(x, y)\left(\alpha A(x, y)y_i + B(x, y)\right) + (1 - \alpha)y_i + B(x, y)$$
$$= \alpha A(x, y)^2 y_i + \alpha A(x, y)B(x, y) + (1 - \alpha)A(x, y)y_i + B(x, y).$$

Let $(a, b)$ be a positive solution of $\Sigma_\alpha$. Then $(a, b)$ solves all equations $y_j = G_j(x, y)$ of $\Sigma$ where $j \neq i$ since these equations are also in $\Sigma_\alpha$. Now

$$b_i = G_i^{(\alpha)}(x, y)$$
$$= \alpha A(a, b)^2 b_i + \alpha A(a, b)B(a, b) + (1 - \alpha)A(a, b)b_i + B(a, b),$$
so
\[ (1 - \alpha A(a, b)^2 - (1 - \alpha)A(a, b))b_i = (1 + \alpha A(a, b))B(a, b). \]
Since \( 1 + \alpha A(a, b) \) is positive, one can cancel to obtain
\[ (1 - A(a, b))b_i = B(a, b), \]
and thus
\[ b_i = A(a, b)b_i + B(a, b), \]
which says that \((a, b)\) satisfies the \(i\)th equation of \(\Sigma\), and thus all the equations of \(\Sigma\). Consequently \(\Sigma\) and \(\Sigma_\alpha\) have the same positive solutions \((a, b)\).

To show both systems have the same characteristic points, the partial derivatives \(\partial G_i/\partial y_k\) and \(\partial G_i^{(\alpha)}/\partial y_k\), \(1 \leq k \leq m\), are compared:

\[
\frac{\partial G_i(x, y)}{\partial y_k} = \frac{\partial A(x, y)}{\partial y_k} \cdot y_j + A(x, y) \cdot \delta_{jk} + \frac{\partial B(x, y)}{\partial y_k},
\]
\[
\frac{\partial G_i^{(\alpha)}(x, y)}{\partial y_k} = \frac{\partial A(x, y)}{\partial y_k} \cdot (\alpha G_j(x, y) + (1 - \alpha)y_j) + A(x, y) \cdot \left(\alpha \frac{\partial G_j(x, y)}{\partial y_k} + (1 - \alpha)\delta_{jk}\right) + \frac{\partial B(x, y)}{\partial y_k}.
\]

so
\[
\frac{\partial G_i^{(\alpha)}(x, y)}{\partial y_k} = \frac{\partial G_i(x, y)}{\partial y_k} + \alpha \frac{\partial A(x, y)}{\partial y_k} \cdot (G_j(x, y) - y_j) + \alpha A(x, y) \cdot \left(\frac{\partial G_j(x, y)}{\partial y_k} - \delta_{jk}\right).
\]

At a positive solution \((a, b)\) to \(\Sigma\) (hence to \(\Sigma^*\)), this gives
\[
(33) \quad \frac{\partial G_i^{(\alpha)}(a, b)}{\partial y_k} = \frac{\partial G_i(a, b)}{\partial y_k} + \alpha A(a, b) \cdot \left(\frac{\partial G_j(a, b)}{\partial y_k} - \delta_{jk}\right),
\]
so
\[
\delta_{ik} - \frac{\partial G_i^{(\alpha)}(x, y)}{\partial y_k} = \delta_{ik} - \frac{\partial G_i(a, b)}{\partial y_k} + \alpha A(a, b) \cdot \left(\delta_{jk} - \frac{\partial G_j(a, b)}{\partial y_k}\right).
\]

Thus, since \((a, b)\) is positive, one obtains \(J_\alpha(a, b) := I - J_{G^{(\alpha)}}(a, b)\) from \(J(a, b) := I - J_G(a, b)\) by adding a non-negative multiple of the \(j\)th row to the \(i\)th row when \(i \neq j\), and by multiplying the \(i\)th row by a positive constant when \(i = j\). It follows
that \( \det(J(a, b)) = 0 \) iff \( \det(J_\alpha(a, b)) = 0 \). Combining this with the fact that \( \Sigma \) and \( \Sigma_\alpha \) have the same positive solutions shows that \( CP = CP_\alpha \), that is, they have the same characteristic points.

For the next claim, item (c), note that the composition of minimal self-transforms using \( \alpha \in [0, 1) \) at each step preserves the well-conditioned property by Lemma 19.

For a well-conditioned system \( \Sigma \), the standard solution is the unique sequence \( T(x) \) of non-negative power series with \( T(0) = 0 \) that solve the system. The standard solution of \( \Sigma \) is clearly a solution of \( \Sigma_\alpha \). Thus if \( \Sigma_\alpha \) is well-conditioned then it has the same standard solution, and hence the same extreme point, as \( \Sigma \), so (d) holds.

For the final item, let \((a, b)\) be a point in \( \text{Dom}^+(G) \), hence a point in \( \text{Dom}^+(G^{(\alpha)}) \). \( A(a, b) \) is finite by looking at the expression above for \( G_i(x, y) \). Then, since \( G_j^{(\alpha)}(x, y) = G_j(x, y) \) for \( j \neq i \), (33) shows that \( \frac{\partial G_i^{(\alpha)}(a, b)}{\partial y_k} \) is finite iff \( \frac{\partial G_i(a, b)}{\partial y_k} \) is finite, so one has item (e).

\[ \square \]

**Lemma 21.** A well-conditioned system \( \Sigma : y = G(x, y) \) can be transformed by a self-substitution into a well-conditioned system \( \Sigma^* : y = G^*(x, y) \) such that the Jacobian matrix \( J_{G^*}(x, y) \) has all entries non-zero. Indeed, given any \( n > 0 \), one can find a \( \Sigma^* \) such that all \( n \)th partials of the \( G^*_i \) with respect to the \( y_j \) are non-zero.

**Proof.** The goal is to show that there is a sequence \( \Sigma_0, \ldots, \Sigma_r \) of minimal self-substitution transforms that go from \( \Sigma \) to the desired \( \Sigma^* \), and such that each system \( \Sigma_i \) is well-conditioned. The following four cases give the key steps in the proof.

**CASE I:** Suppose some \( G_i \) is such that all \( n \)th partials are non-zero. If \( G_j \) is dependent on \( y_i \) (there is at least one such \( j \)) then substituting \( (1/2)G_i + (1/2)y_i \) for some occurrence of \( y_i \) in \( G_j \) gives a well-conditioned system \( \Sigma' \) such that for \( G'_i = G_i \) and \( G'_j \), all \( n \)th partials are non-zero. Continuing in this fashion one eventually has the desired system \( \Sigma^* \).

**CASE II:** Suppose \( \frac{\partial^{mn} G_i}{\partial y_i^m} \neq 0 \) for some \( i \). This means \( y_i^{-mn} \) divides some monomial of \( G_i \). Use the fact that for any \( j \neq i \) there is a dependency path from \( y_i \) to \( y_j \) to convert, via self-substitutions that preserve the well-conditioned property, a product of \( n \) of the \( y_i \) in this monomial into a power series which has \( y_j^n \) dividing one of its
monomials. By doing this for each \( j \neq i \) one obtains a well-conditioned \( G'_i \) with
\[
\frac{\partial^{mn} G'_i}{\partial y_1^n \cdots \partial y_m^n} \neq 0.
\]
\( \Sigma' \) is now in Case I.

CASE III: Suppose \( \frac{\partial^2 G_i}{\partial y_i^2} \neq 0 \) for some \( i \). Substituting \( G_i \) for a suitable occurrence of \( y_i \) in \( G_i \) gives a well-conditioned \( \Sigma' \) where \( \frac{\partial^3 G'_i}{\partial y_i^3} \neq 0 \). Continuing in this fashion leads to Case II.

CASE IV: Suppose \( \frac{\partial^2 G_i}{\partial y_j \partial y_k} \neq 0 \) for some \( i, j, k \). If \( j \neq i \) there is a dependency path from \( y_j \) to \( y_i \) which shows how to make self-substitutions (that preserve the well-conditioned property) leading to \( \frac{\partial^2 G_i}{\partial y_i \partial y_k} \neq 0 \). Likewise, if \( k \neq i \) there is a dependency path from \( y_k \) to \( y_i \) which shows how to make self-substitutions (with each minimal step being well-conditioned) leading to \( \frac{\partial^2 G_i}{\partial y_i^2} \neq 0 \), which is Case III.

Since \( \Sigma \) is non-linear in \( y \), for some \( i, j, k \) we have
\[
\frac{\partial^2 G_i}{\partial y_i \partial y_k} \neq 0.
\]
Thus starting with Case IV and working back to Case I we arrive at the desired \( \Sigma^* \).

\[\Box\]

Lemma 22. Let \( \Sigma : y = G(x, y) \) be a well-conditioned system and let \( \Sigma^* : y = G^*(x, y) \) be a self-substitution transform of \( \Sigma \). If \( (a, b) \) is a characteristic point of \( \Sigma \), hence of \( \Sigma^* \), then \( \Lambda(a, b) = 1 \) iff \( \Lambda^*(a, b) = 1 \).

Proof. Let \( (a, b) \) be a characteristic point of \( \Sigma \). It suffices to consider the case where \( \Sigma^* \) is obtained from \( \Sigma \) by a minimal self-substitution. Let \( G_i(x, y) \) depend on \( y_j \), and let \( H(x, y_0; y) \) be the result of replacing a single occurrence of \( y_j \) in \( G_i(x, y) \) by \( y_0 \).

Then let \( \Sigma_\alpha : y = G^{(\alpha)}(x, y), \alpha \in [0, 1], \) be the minimal self-substitution transform of \( \Sigma \) obtained by applying the substitution \( y_0 \leftarrow \alpha G_j(x, y) + (1 - \alpha)y_j \) to \( H(x, y_0; y) \) to obtain
\[
G^{(\alpha)}_i(x, y) = H(x, \alpha G_j(x, y) + (1 - \alpha)y_j; y).
\]

Let \( \Lambda_\alpha := \Lambda_\alpha(a, b) \), the largest real eigenvalue of \( J_{G^{(\alpha)}}(a, b) \).
The only information that we need from the above construction of the $G_i^{(a)}$ is that the function $\alpha \mapsto J_{G^{(a)}}(a, b)$ is continuous on $[0, 1]$, and each $J_{G^{(a)}}(a, b)$ has 1 being an eigenvalue. Since $\Lambda$ is continuous on non-negative matrices by Corollary 3, it follows that $\alpha \mapsto \Lambda_\alpha$ is continuous on $[0, 1]$. The goal is to show that one has $\Lambda_0 = 1$ iff $\Lambda_\alpha = 1$.

Since $(a, b)$ is a characteristic point of $\Sigma_0$ it is also a characteristic point of $\Sigma_\alpha$, by Lemma 20 for $\alpha \in [0, 1]$. Thus 1 is an eigenvalue of $J_{G^{(a)}}(a, b)$ for $\alpha \in [0, 1]$. Suppose $\Lambda_0 = 1$. Suppose there is a $\beta \in (0, 1]$ with $\Lambda_\beta > 1$. From the continuity of $\Lambda_\alpha$ there is a $\gamma \in [0, \beta)$ such that: $\Lambda_\gamma = 1$, and $\Lambda_\alpha > 1$ for $\alpha \in (\gamma, \beta]$.

Let $p_\alpha(x)$ be the characteristic polynomial of $J_{G^{(a)}}(a, b)$. From

$$p_\alpha(1) = p_\alpha(\Lambda_\alpha) = 0$$

one has, for each $\alpha \in (\gamma, \beta)$, a $c_\alpha \in (1, \Lambda_\alpha)$ such that

$$\frac{dp_\alpha}{dx}(c_\alpha) = 0.$$ 

Since $\Lambda_\alpha$ is continuous on $[0, 1]$,

$$\lim_{\alpha \to \gamma^+} \Lambda_\alpha = \Lambda_\gamma = 1.$$ 

This implies

$$\lim_{\alpha \to \gamma^+} c_\alpha = 1,$$

and thus

$$\frac{dp_\gamma}{dx}(1) = \lim_{\alpha \to \gamma^+} \frac{dp_\alpha}{dx}(c_\alpha) = 0.$$ 

But from the Perron-Frobenius theory (see Proposition 2) we know that $\Lambda_\gamma = 1$ implies that 1 is a simple root of $p_\gamma(x)$, giving a contradiction. Thus $\Lambda_0 = 1$ implies $\Lambda_\alpha = 1$.

A similar proof gives the converse, that if $\Lambda_\alpha = 1$ then $\Lambda_0 = 1$, proving the lemma. 

\[\square\]

**Remark 23.** In view of the last two lemmas, given a well-conditioned system $\Sigma : \mathbf{y} = \mathbf{G}(x, \mathbf{y})$, when one wants to prove something about the positive solutions, the characteristic points, or whether or not $\Lambda(a, b) = 1$ at a characteristic point $(a, b)$, one can, given any $n > 0$, assume without loss of generality that all $n$th partials of each $G_i$ with respect to the $y_j$ are non-zero. In the rather scant literature on
nonlinear systems one finds a preference for working with aperiodic systems (see, e.g., [7]), no doubt because of the simplicity of using uniform substitutions to convert such a system into one where the Jacobian of $G$ has non-zero entries. With Lemmas 21 and 22, the need for the aperiodic hypothesis is avoided.

4.2. Basic Properties of $(\rho, \tau)$ and CP. Now we turn to the question of how to find information about the extreme point $(\rho, \tau)$ of a well-conditioned system $\Sigma$ without solving the system for the standard solution $T(x)$.

Lemma 24. Let $y = G(x, y)$ be a well-conditioned system with all entries of $J_G$ non-zero.

(a) One has the formal equality

$$T'(x) = G_x(x, T(x)) + J_G(x, T(x)) \cdot T'(x),$$

which also holds for $x \in [0, \infty]$.

(b) All $T'_i(\rho)$ are finite or all $T'_i(\rho) = \infty$.

(c) For all $i, j$ the following hold:

$$0 < \frac{\partial G_i}{\partial y_j}(\rho, \tau) \cdot \frac{\partial G_j}{\partial y_i}(\rho, \tau) \leq 1$$

$$0 < \frac{\partial G_i}{\partial y_j}(\rho, \tau) < \infty$$

$$0 < \frac{\partial G_i}{\partial y_i}(\rho, \tau) \leq 1.$$

Proof. Differentiating (6) gives (34), so $T'(x)$ is a solution to the irreducible system

$$u = G_x(x, T(x)) + J_G(x, T(x)) \cdot u,$$

implying (b). For $x \in (0, \rho)$, for each $i, j$, (34) implies

$$T'_i(x) > \frac{\partial G_i}{\partial y_j}(x, T(x)) \cdot T'_j(x),$$

and thus

$$1 > \frac{\partial G_i}{\partial y_j}(x, T(x)) \cdot \frac{\partial G_j}{\partial y_i}(x, T(x)) > 0,$$

giving the inequalities in (c) since the value of $\frac{\partial G_i}{\partial y_j}(\rho, \tau)$ is the limit of $\frac{\partial G_i}{\partial y_j}(x, T(x))$ as $x$ approaches $\rho$ from below. \qed
Lemma 25. Let \( y = G(x, y) \) be a well-conditioned system.

(a) If \((a, b) \in CP\) then \( \Lambda(a, b) \geq 1 \).

(b) \( 0 < \Lambda(a, T(a)) < 1 \), for \( 0 < a < \rho \).

Proof. For (a) note that \((a, b) \in CP\) implies that 1 is an eigenvalue of \( J_G(a, b) \), so \( \Lambda(a, b) \geq 1 \).

(b) Given \( 0 < a < \rho \), by the Perron-Frobenius theory of nonnegative matrices we know that there is a positive left eigenvector (a row vector) \( v \) belonging to \( \Lambda(a, T(a)) \).

By (33)

\[
 v \cdot T'(a) = v \cdot G_x(a, T(a)) + v \cdot J_G(a, T(a)) \cdot T'(a),
\]

so

\[
 v \cdot T'(a) = v \cdot G_x(a, T(a)) + \Lambda(a, T(a))v \cdot T'(a).
\]

Since \( v \cdot T'(a) > 0 \) and \( v \cdot G_x(a, T(a)) > 0 \) it follows that \( \Lambda(a, T(a)) < 1 \).

\[ \square \]

Lemma 26. Let \( y = G(x, y) \) be a well-conditioned system. Suppose \((a, b)\) and \((c, d)\) are characteristic points and \((a, b) \leq (c, d)\). Then \((a, b) = (c, d)\). Thus the set of characteristic points of the system form an antichain under the partial ordering \( \leq \).

Proof. For the proof assume, w.l.o.g., that all second partials of the \( G_i \) with respect to the \( y_j \) do not vanish. If \( b = d \) then \( G(a, b) = b = d = G(c, d) \), which forces \( a = c \) by the monotonicity of each \( G_i \).

Now assume \( b \neq d \). Since \( b \leq d \), all entries of \( d - b \) are non-negative. Using part of a Taylor series expansion,

\[
 G(c, d) \geq G(a, b) + J_G(a, b)(d - b) + \frac{1}{2} \begin{bmatrix} \frac{\partial^2 G_1(a, b)}{\partial y_1} (d_1 - b_1)^2 \\ \vdots \\ \frac{\partial^2 G_m(a, b)}{\partial y_m} (d_m - b_m)^2 \end{bmatrix}.
\]

Since \( G(a, b) = b \) and \( G(c, d) = d \),

\[
 d - b \geq J_G(a, b)(d - b) + \frac{1}{2} \begin{bmatrix} \frac{\partial^2 G_1(a, b)}{\partial y_1} (d_1 - b_1)^2 \\ \vdots \\ \frac{\partial^2 G_m(a, b)}{\partial y_m} (d_m - b_m)^2 \end{bmatrix}.
\]
Let $\lambda$ be the largest real eigenvalue of the positive matrix $J_G(a, b)$, and let $v$ be a positive left eigenvector belonging to $\lambda$. Then
\[
v(d - b) \geq v J_G(a, b)(d - b) + \frac{1}{2}v \begin{bmatrix}
\frac{\partial^2 G_1(a, b)}{\partial y_1^2}(d_1 - b_1)^2 \\
\vdots \\
\frac{\partial^2 G_m(a, b)}{\partial y_m^2}(d_m - b_m)^2
\end{bmatrix}
= \lambda v(d - b) + \frac{1}{2}v \begin{bmatrix}
\frac{\partial^2 G_1(a, b)}{\partial y_1^2}(d_1 - b_1)^2 \\
\vdots \\
\frac{\partial^2 G_m(a, b)}{\partial y_m^2}(d_m - b_m)^2
\end{bmatrix}
\]
so
\[
(1 - \lambda)v(d - b) \geq \frac{1}{2}v \begin{bmatrix}
\frac{\partial^2 G_1(a, b)}{\partial y_1^2}(d_1 - b_1)^2 \\
\vdots \\
\frac{\partial^2 G_m(a, b)}{\partial y_m^2}(d_m - b_m)^2
\end{bmatrix} > 0,
\]
and this forces $\lambda < 1$, contradicting Lemma 25 (a).

Lemma 27. Let $y = G(x, y)$ be a well-conditioned system.

(a) $(\rho, \tau)$ is in the domain of $J_G(x, y)$, that is, all entries of the matrix $J_G(\rho, \tau)$ are finite.

(b) If $(\rho, \tau)$ is in the interior of the domain of $G(x, y)$ then it is a characteristic point.

(c) $0 < \Lambda(\rho, \tau) \leq 1$.

(d) $\Lambda(\rho, \tau) = 1$ iff $1$ is an eigenvalue of $J_G(\rho, \tau)$ iff $(\rho, \tau) \in \mathcal{CP}$.

Proof. For item (a), first let $\Sigma^*$ be a well-conditioned self-substitution transform of $\Sigma$ with all entries in $J_G^*(x, y)$ non-zero. By Lemma 24, all entries of $J_G^*(\rho, \tau)$ are finite. Then Lemma 20 (e) shows that all entries of $J_G(\rho, \tau)$ are finite.

For the remainder of the proof we can assume that all entries in $J_G$ are non-zero. For part (b) one argues just as in the case of a single equation—if $(\rho, \tau)$ is an interior point but not a characteristic point then by the implicit function theorem there would be an analytic continuation of $T(x)$ at $\rho$, which is impossible.
For (c), since $\Lambda$ is a continuous nondecreasing function by Corollary 3 and since the limit of $J_G(x, T(x))$ as $x$ approaches $\rho$ from below is $J_G(\rho, \tau)$, it follows from Lemma 25 (b) that $\Lambda(\rho, \tau) \leq 1$.

For (d), clearly $\Lambda(\rho, \tau) = 1$ implies 1 is an eigenvalue of $J_G(\rho, \tau)$, and this in turn implies that $(\rho, \tau) \in \mathcal{CP}$. Now suppose that $(\rho, \tau) \in \mathcal{CP}$. Then 1 is an eigenvalue of $J_G(\rho, \tau)$, so $\Lambda(\rho, \tau) \geq 1$. Thus (c) gives $\Lambda(\rho, \tau) = 1$.

Lemma 28. Let $y = G(x, y)$ be a well-conditioned system. If $(a, b)$ is a characteristic point and $(a, b) \neq (\rho, \tau)$ then either

(a) $b_i > \tau_i$ for all $i$, or
(b) $a < \rho$ and $b_i > T_i(a)$ for all $i$, and some $b_j > \tau_j$.

Proof. Condition (e) in the definition of well-conditioned ensures that each $G_i(x, y)$ depends on $x$. W.l.o.g. assume that all second partials of each $G_i(x, y)$ with respect to the $y_j$ are non-zero. Suppose that (a) does not hold.

Claim 1: If some $b_i > \tau_i$ and some $b_j \leq \tau_j$ then $a < \rho$ and $T_i(a) < b_i$ for $1 \leq i \leq m$.

WLOG assume that

\[ b_1 \leq \tau_1, \ldots, b_k \leq \tau_k \]

and

\[ b_{k+1} > \tau_{k+1}, \ldots, b_m > \tau_m. \]

From the monotonicity and continuity of the $T_i$ on $[0, \rho]$ it follows that for $1 \leq i \leq k$ there exist unique $\xi_i \in (0, \rho]$ such that

\[ b_i = T_i(\xi_i). \]

WLOG assume that

\[ 0 < \xi_1 \leq \cdots \leq \xi_k \leq \rho. \]

For $i \in \{1, \ldots, k\}$

\[ T_i(\xi_1) \leq T_i(\xi_i) = b_i \]

and for $k+1 \leq i \leq m$

\[ T_i(\xi_1) \leq T_i(\rho) < b_i. \]

Now suppose $\xi_1 < a$. Then

\[ b_1 = T_1(\xi_1) \]
CHARACTERISTIC POINTS

\[ G_1(\xi_1, T_1(\xi_1), \ldots, T_m(\xi_1)) < G_1(a, b_1, \ldots, b_m) = b_1. \]

a contradiction. Thus

\[ 0 < a \leq \xi_1 \leq \cdots \leq \xi_k \leq \rho. \]

Using this one has, for \( 1 \leq i \leq k \):

\[
T_i(\xi_i) = b_i \\
= G_i(a, b_1, \ldots, b_m) \\
= G_i(a, T_1(\xi_1), \ldots, T_k(\xi_k), b_{k+1}, \ldots, b_m) \\
> G_i(a, T_1(a), \ldots, T_k(a), T_{k+1}(a), \ldots, T_m(a)) = T_i(a).
\]

Thus for \( 1 \leq i \leq k \),

\[ 0 < a < \xi_i \leq \rho \]

\[ T_i(a) < T_i(\xi_i) = b_i. \]

Furthermore, for \( k + 1 \leq i \leq m \),

\[ T_i(a) < T_i(\rho) < b_i. \]

Thus, in this case, for \( 1 \leq i \leq m \) one has \( T_i(a) < b_i \).

Claim 2: If \( b_i \leq T_i(\rho) \) for all \( i \) then \( a < \rho \) and \( b_i = T_i(a) \) for all \( i \).

Choose \( \xi_i \in (0, \rho) \) such that \( b_i = T_i(\xi_i) \). WLOG one can assume \( 0 < \xi_1 \leq \cdots \leq \xi_m \leq \rho \). If \( \xi_1 < a \) then

\[
b_1 = G_1(a, b_1, \ldots, b_m) \\
= G_1(a, T_1(\xi_1), \ldots, T_m(\xi_m)) \\
> G_1(\xi_1, T_1(\xi_1), \ldots, T_m(\xi_1)) \\
= T_1(\xi_1) = b_1,
\]

a contradiction. Thus \( a \leq \xi_1 \leq \cdots \leq \xi_m \leq \rho. \)

Next one has

\[
b_m = T_m(\xi_m) \\
= G_m(\xi_m, T_1(\xi_m), \ldots, T_m(\xi_m)) \\
\geq G_m(a, T_1(\xi_1), \ldots, T_m(\xi_m))
\]
\[ G_m(a, b_1, \ldots, b_m) = b_m, \]

so the \( \geq \) step must be an equality, and this implies \( \xi_m = a \). Thus all \( \xi_i = a \), and then for all \( i \) one has \( b_i = T_i(a) \). Since \( (a, b) = (a, T(a)) \) is assumed to be a different characteristic point from \((\rho, \tau)\), it follows that \( a < \rho \).

**Claim 3:** It is not the case that \( b_i \leq \tau_i \) for all \( i \).

Otherwise by Claim 2 we would have \((a, b) = (a, T(a))\) with \( 0 < a < \rho \), and then by Lemma 25 it would follow that \((a, b) \notin \mathcal{CP}\). But by assumption, \((a, b) \in \mathcal{CP}\).

**Proposition 29.** A well-conditioned 1-equation system \( y = G(x, y) \) has at most one characteristic point; if there is such a point it must be the extreme point \((\rho, \tau)\) of the standard solution \( T(x) \).

**Proof.** The characteristic system is

\[
\begin{align*}
y &= G(x, y) \\
1 &= G_y(x, y).
\end{align*}
\]

Suppose \((a, b) \in \mathcal{CP}\) is different from \((\rho, \tau)\). Then \( b > \tau \) by Lemma 28.

**CASE 1:** Suppose \( a > \rho \). Then \((\rho, \tau)\) is in the interior of \( \text{Dom}^+(G) \), so \((\rho, \tau) \in \mathcal{CP}\) by Lemma 27(b). But this violates the antichain condition of Lemma 26 for \( \mathcal{CP}\).

**CASE 2:** Suppose \( a \leq \rho \). Then \( b = G(a, b) \) and \( T(a) = G(a, T(a)) \) leads to \( 1 = G_y(a, \xi) \) for some \( T(a) < \xi < b \). But \( G_y(a, b) = 1 \) since \((a, b) \in \mathcal{CP}\), so again we have a contradiction by the strict monotonicity of \( G_y(x, y) \) in \( \text{Dom}^+(G) \).

Thus the only possible \((a, b) \in \mathcal{CP}\) is \((\rho, \tau)\). \( \square \)

**Remark 30.** Meir and Moon [13] prove that well-conditioned 1-equation systems have at most one characteristic point in the interior of \( \text{Dom}^+(G) \); and if such a point exists then it must be \((\rho, \tau)\). See also Flajolet and Sedgewick [7], Chapter VII §4.
Proposition 31. Suppose \((\rho, \tau)\) is a characteristic point of a well-conditioned system \(y = G(x, y)\). Then:

(a) \(\rho\) is the largest first coordinate of any characteristic point, that is
\[
\rho = \max \{a : (a, b) \in CP\},
\]

(b) \((\rho, \tau)\) is the only characteristic point whose first coordinate is \(\rho\).

Proof. Use Lemmas 26 and 28. \(\square\)

5. Eigenpoints

The results developed so far do not give a practical way of locating \((\rho, \tau)\) for well-conditioned systems with more than one equation. Even if one is successful in finding all the characteristic points, no means has yet been formulated to determine if \((\rho, \tau)\) is among them. In this section special characteristic points called eigenpoints are shown to provide the correct analog of characteristic points when moving from 1-equation systems to multi-equation systems.

Lemma 32. Suppose \((a, b)\) is a characteristic point of the well-conditioned system \(y = G(x, y)\). Then \(\Lambda(a, b) = 1\) iff \((a, b) = (\rho, \tau)\).

Proof. We can assume that no partial \(\partial G_i/\partial y_j\) is zero. The direction \((\Leftarrow)\) follows from Lemma 27 (d). To prove the direction \((\Rightarrow)\) assume \((a, b) \neq (\rho, \tau)\). By Lemma 28 one has two cases to consider:

(I) \(a \geq \rho\) and for all \(i, b_i > \tau_i\)
(II) \(a < \rho\) and for all \(i, b_i > T_i(a)\).

For (I), \((\rho, \tau)\) is in the interior of the domain of \(G\), so by Lemma 27 (b) it is a characteristic point. However this contradicts Lemma 26 which says the characteristic points form an antichain.

For (II), from the equations
\[
G(a, b) - b = 0
\]
\[
G(a, T(a)) - T(a) = 0
\]
one can apply a multivariate version of the mean value theorem to derive:
\[
(35) \quad \left(\frac{\partial G_i}{\partial y_j}(a, v_{ij})\right) (b - T(a)) = b - T(a)
\]
with \( v_{ij} = (v_{ij}(1), \ldots, v_{ij}(m)) \) satisfying

\[
\begin{cases}
  v_{ij}(r) = T_j(a) & \text{if } r > j \\
  T_i(a) < v_{ij}(r) < b_i & \text{if } r = j \\
  v_{ij}(r) = b_j & \text{if } r < j.
\end{cases}
\]

Clearly (35) shows that \( \lambda = 1 \) is an eigenvalue of \( \left( \frac{\partial G_i}{\partial y_j}(a, v_{ij}) \right) \), and from the properties of the \( v_{ij} \) we see that for all \( i, j \)

\[
\frac{\partial G_i}{\partial y_j}(a, v_{ij}) < \frac{\partial G_i}{\partial y_j}(a, b)
\]

since each \( \frac{\partial G_i}{\partial y_j} \) depends on all the variables \( x, y_1, \ldots, y_m \).

From these remarks and the monotonicity of \( \Lambda \) one has

\[
1 \leq \Lambda \left( \frac{\partial G_i}{\partial y_j}(a, v_{ij}) \right) < \Lambda(a, b),
\]

showing that \( (a, b) \neq (\rho, \tau) \) implies \( \Lambda(a, b) > 1 \).

\[\Box\]

**Definition 33.** A characteristic point \((a, b)\) is an eigenpoint if \( \Lambda(a, b) = 1 \).

The following theorem summarizes the key results proved in this section for well-conditioned systems.

**Theorem 34.** Let \( \Sigma : y = G(x, y) \) be a well-conditioned system. Then the following hold:

(a) \((\rho, \tau) \in \text{Dom}^+(G)\)

(b) If \((\rho, \tau)\) is in the interior of \( \text{Dom}^+(G) \) then it is an eigenpoint.

(c) The system \( \Sigma \) has at most one eigenpoint.

(d) If there is an eigenpoint of \( \Sigma \) then it must be \((\rho, \tau)\).

(e) If there is no eigenpoint of \( \Sigma \) then \((\rho, \tau)\) lies on the boundary of \( \text{Dom}^+(G) \) and one has \( \Lambda(\rho, \tau) < 1 \).

This result can be superior to Proposition 31 for computing purposes since the latter requires that one know all characteristic points of \( \Sigma \) before being able to isolate the one candidate for \((\rho, \tau)\). Theorem 34 says that if one can find a characteristic
point \((a, b)\) with \(J_G(a, b)\) having largest positive eigenvalue 1, it is \((\rho, \tau)\). As with the 1-equation case, if there are no eigenpoints of \(\Sigma\), then new methods are needed.

**Example 35.** The well-conditioned polynomial system

\[
\begin{align*}
y_1 &= G_1(x, y_1, y_2) := x(1 + 2y_1^3 + 2x^3y_1y_2) \\
y_2 &= G_2(x, y_1, y_2) := x(1 + x^3y_2 + 2y_1^3y_2^2)
\end{align*}
\]

has four characteristic points which, to 6 places of accuracy are:

- \((0.1818598, 1.556545, 0.3647603)\)
- \((0.2640956, 1.210710, 0.5353688)\)
- \((0.3867644, 0.6661246, 3.834789)\)
- \((0.4153198, 0.6217456, 0.4743552)\)

One sees that these four points form an antichain, as required by Lemma 26. The extreme point \((\rho, \tau_1, \tau_2)\) of a polynomial system is a characteristic point. By Proposition 31 it must be the last one since it has the largest \(x\)-value, assuming one has found all characteristic roots of this system. If one is not sure that there are only four characteristic points then, by Theorem 34, it suffices to verify that the indicated characteristic point is an eigenpoint.

### 6. Drmota’s Theorem Revisited

In 1993 Lalley [10] proved that the solutions \(y_i = T_i(x)\) to a well-conditioned polynomial system \(y = G(x, y)\) would have a square-root singularity at \(\rho\), and thus one had the familiar Pólya asymptotics for the coefficients. In 1997 [5], and again in 2009 [6], Drmota presented the first sweepingly general theorem concerning the asymptotic behavior of the coefficients of solutions of a well-conditioned system, namely the coefficients will again satisfy the same law that Pólya found to be true for several classes of trees (see [16]). However, as explained in Footnote 2 the hypotheses that Drmota has for the characteristic points of the system seem to be incorrect in the first publication, and vague in the second. To prove the theorem one needs to be able to show that \((\rho, \tau)\) is in the interior of the domain of \(G(x, y)\). The

---

Footnote 2: Having a polynomial system is a very strong condition since it immediately tells you that \(\rho\) is a branch point, which leads to a Puiseux expansion; it is only a matter of determining the order of the branch point (which is nonetheless a nontrivial task).
following subsection gives a clear statement of the hypotheses needed, along with a slightly different proof of the key induction step for the proof.

6.1. **Drmota’s Theorem.** The following version is somewhat simpler than that presented by Drmota since there are no parameters.

**Theorem 36.** Let $\Sigma : y = G(x, y)$ be a well-conditioned system with standard solution $T(x)$. Suppose $\Sigma$ has an eigenpoint in the interior of $\text{Dom}^+(G)$. Then each $T_i(x)$ is the standard solution to a well-conditioned 1-equation system $y_i = \hat{G}_i(x, y_i)$. Thus each $T_i(x)$ has a square-root singularity at $\rho$, and the familiar Pólya asymptotics (see, e.g., [2]) hold for the non-zero coefficients.

**Proof.** One only needs to consider the case that the system has at least two equations, and one can assume all second partials of the $G_i$ with respect to the $y_j$ are non-zero. The following shows that eliminating the first equation (and $y_1$) yields a well-conditioned system with one less equation which has the standard solution $(T_2(x), \ldots, T_m(x))$ and an eigenpoint in the interior of the domain of the system.

By the Implicit Function Theorem one can solve the first equation

$$y_1 = G_1(x, y)$$

for $y_1$, say

$$y_1 = H_1(x, y_2, \ldots, y_m),$$

where $H_1$ is holomorphic in a neighborhood of the origin, that is, $H_1(0, 0) = 0$ and

$$H_1(x, y_2, \ldots, y_m) = G_1(x, H_1(x, y_2, \ldots, y_m), y_2, \ldots, y_m)$$

in a neighborhood of the origin.

Since the $T_i(x)$ take small values near the origin (as they are continuous functions that vanish at $x = 0$), it follows that

$$H_1(x, T_2(x), \ldots, T_m(x)) = G_1(x, H_1(x, T_2(x), \ldots, T_m(x)), T_2(x), \ldots, T_m(x))$$

holds in a neighborhood of the origin. Also one has

$$T_1(x) = G_1(x, T_1(x), T_2(x), \ldots, T_m(x))$$

holding in a neighborhood of the origin, so by the uniqueness of solutions in such a neighborhood, we must have

$$T_1(x) = H_1(x, T_2(x), \ldots, T_m(x))$$
in a neighborhood of the origin. By Proposition 1, this equation actually holds globally for \( |x| \leq \rho \); in particular \( H_1 \) converges at \((\rho, \tau_2, \ldots, \tau_m)\). By Corollary 3(a) the Jacobian \(1 - \frac{\partial G_1}{\partial y_1}\) of the equation \(y_1 = G_1(x, y)\) does not vanish at \((\rho, \tau)\). Thus, by the Implicit Function Theorem, \(H_1\) is holomorphic at \((\rho, \tau_2, \ldots, \tau_m)\).

Now discarding the first equation and substituting \(H_1(x, y_2, \ldots, y_m)\) for \(y_1\) in the remaining equations gives a well-conditioned system of \(m - 1\) equations \(y_i = G^*_i(x, y_2, \ldots, y_m),\ 2 \leq i \leq m\), with standard solution \((x, T_2(x), \ldots, T_m(x))\) whose extreme point \((\rho, \tau_2, \ldots, \tau_m)\) is an eigenpoint, since it is a characteristic point of the system that is in the interior of \(\text{Dom}^+(G^*)\). Thus the elimination procedure can continue if \(G^*\) consists of more than one equation.

\[ \square \]

The extreme point of a well-conditioned polynomial system, such as Example 35, is always a characteristic point, and, as Lalley [10] proved, the coefficients of the solutions \(T_i(x)\) have the classical Pólya form \(C_i \rho^{-n} n^{-3/2}\). Drmota [5] extended Lalley’s result to well-conditioned power series systems with the extreme point in the interior of the domain of the system. A natural (and desirable) direction to consider for further research would be to drop the irreducible requirement. However, even in the polynomial case, this leads to substantial challenges. For example, consider the reducible polynomial system

\[
\begin{align*}
y_1 &= y_3 \cdot (1 + y_2 + y_1^2) \\
y_2 &= y_3 \cdot (1 + y_1 + y_2^2) \\
y_3 &= x \cdot (1 + 9y_3^2).
\end{align*}
\]

Let the third equation have the standard solution \(y_3 = A(x)\). One then sees that this example is really just an alternate presentation of Example 16 where the solutions for \(y_1\) and \(y_2\) have a fourth-root singularity at their radius of convergence.

6.2. A Wealth of Examples. In [2] we showed that single equation systems formed from a wide array of standard operators like Multiset, Cycle and Sequence led to square-root singularities and Pólya asymptotics for the coefficients. The arguments used there easily carry over to the setting of systems of equations since the conditions in that paper force the positive domain to be an open set, and this guarantees that \((\rho, \tau)\) is an interior point of the domain of the system. Thus, for example, consider
the collection of unlabeled rooted two-colored trees, say by the colors red and blue, defined by requiring that immediately below each red node that is not a leaf there must be a prime number of blue nodes and an even number of red nodes, and immediately below each blue node that is not a leaf all nodes have the same color. Then, by Theorem 36, the number of trees in this class has the Pólya asymptotics $C \rho^{-n} n^{-3/2}$.

7. Some Open Problems about Characteristic Points of Well-Conditioned Systems

Question 1. How can one locate $(\rho, \tau)$ if it is not a characteristic point?

Question 2. Is the set of characteristic points always finite?

As one has seen in the introductory examples, a system can have multiple characteristic points; the two equation polynomial system in Example 35 has four characteristic points. The following example shows that the set of real solutions to the characteristic system need not be finite. However Question 2 asks if the set of positive solutions is finite.

Example 37. For the characteristic system (belonging to a 2-equation system)

$$\begin{cases}
y_1 - x \cdot (1 + y_1 + y_2) &= 0 \\
y_2 - x \cdot (1 + y_2 + y_1) &= 0 \\
(x - 1) \cdot (x + xy_1 + xy_2 - 1) &= 0
\end{cases}$$

the real solutions include the infinite curve

$$\{(x, y_1, y_2) : x = 1, y_1y_2 = -1\}.$$

References

[1] Edward A. Bender, *Asymptotic methods in enumeration*. SIAM Rev. 16 (1974), 485–515.

[2] Jason P. Bell, Stanley N. Burris, and Karen A. Yeats, *Counting Rooted Trees: The Universal Law* $t(n) \sim C \cdot \rho^{-n} \cdot n^{-3/2}$. The Electron. J. Combin. 13 (2006), R63 [64pp.]

[3] E. Rodney Canfield, *Remarks on an asymptotic method in combinatorics*. J. Combin. Theory Ser. A 37 (1984), no. 3, 348352.

[4] A. Cayley, *On the theory of the analytical forms called trees*. Phil. Magazine 13 (1857), 172–176.
[5] Michael Drmota, *Systems of functional equations*. Random Structures and Algorithms **10** (1997), 103–124.

[6] Michael Drmota, *Random Trees*. Springer, 2009.

[7] Philippe Flajolet and Robert Sedgewick, *Analytic Combinatorics*. Cambridge University Press, 2009.

[8] F.R. Gantmacher, *Applications of the Theory of Matrices*. Interscience Publishers, Inc., New York, 1959.

[9] E. Hille, *Analytic function theory*, Blaisdell Publishing Company, Waltham, 1962, 2 Volumes.

[10] Stephen P. Lalley, *Finite range random walk on free groups and homogeneous trees*. The Annals of Probability **21**, No. 4 (1993), 2087–2130.

[11] A. Meir and J. W. Moon, *On the altitude of nodes in random trees*, Canadian Journal of Mathematics **30** (1978), 997–1015.

[12] A. Meir and J.W. Moon, *Some asymptotic results useful in enumeration problems*. Aequationes Math. **33** (1987), 260–268.

[13] A. Meir and J.W. Moon, *On an asymptotic method in enumeration*. J. Combin. Theory Ser. A **51** (1989), no. 1, 77 *Erratum*: J. Combin. Theory Ser. A **52** (1989), no. 1, 163.

[14] A.M. Odlyzko, *Asymptotic enumeration methods*. Handbook of Combinatorics, Vol. 1, 2, 1063–1229, Elsevier, Amsterdam, 1995.

[15] R. Otter, *The number of trees*. Annals of Mathematics **49** (1948), 583–599.

[16] G. Pólya and R.C. Read, *Combinatorial enumeration of groups, graphs and chemical compounds*. Springer Verlag, New York, 1987.

[17] Alan R. Woods, *Coloring rules for finite trees, probabilities of monadic second order sentences*. Random Structures Algorithms **10** (1997), 453–485.