INTERNATIONAL JOURNAL FOR NUMERICAL METHODS IN ENGINEERING, VOL. 37, 3939–3955 (1994)

THE MONOMIAL METHOD AND ASYMPTOTIC PROPERTIES OF ALGEBRAIC SYSTEMS

SCOTT A. BURNS*

University of Illinois at Urbana-Champaign, 104 S. Mathews #117, Urbana, IL 61801, U.S.A.

SUMMARY

The monomial method is a numerical method for solving systems of algebraic equations. It is related to Newton's method, but is based on an approximation that is monomial, rather than linear, in form. It has been shown previously that the monomial method has a number of properties not shared by Newton's method that are responsible for enhanced performance. This paper demonstrates that many of the performance characteristics of the monomial method can be explained in terms of asymptotic properties of algebraic systems, and the ability of the monomial method to exploit these properties. The monomial method recasts the algebraic system to have monomial asymptotes in all directions, so that the monomial approximation becomes asymptotically exact. This explains the very rapid movement toward a solution in the first iteration from distant starting points that has been observed with the monomial method. Problem types with 'good asymptotics' are found to be solved very effectively by the monomial method. Several specific engineering applications with good asymptotics are identified, and suggestions are made regarding other types of engineering applications that are likely to be handled effectively by the monomial method.

INTRODUCTION

The monomial method is a numerical method for solving systems of algebraic equations. It can be thought of conceptually as a monomial-based version of Newton's method. The monomial approximation becomes linear when the variables are transformed logarithmically, giving rise to a linear system to be solved at each iteration. Of the many favourable performance characteristics that have been associated with the monomial method, one of the most striking is that movement toward a solution in the first iteration is often remarkably better than that of Newton's method, particularly from starting points very distant from a solution. This paper shows that this rapid initial movement can be explained in terms of the asymptotic properties of algebraic systems. Certain types of engineering applications are shown to have favourable asymptotic properties that explain other observed performance characteristics of the monomial method, such as an improved conditioning of the linear system produced by the monomial method, in comparison to that of Newton's method.

PREVIOUS WORK

The monomial approximation used by the monomial method is based on a process called condensation, which appeared in the late 1960s and early 1970s as a non-linear programming

*Associate Professor of General Engineering.

CCC 0029–5981/94/223939–17$9.00

Received 27 July 1993
Revised 22 March 1994

tool.[1,2,3] Its use specifically to solve systems of non-linear equations was first suggested by Rijckaert and Martens.[4] Although the subject of their paper was the solution of non-linear programs, they solved a system of non-linear equations using condensation as part of their non-linear programming solution process. Several subsequent publications, mostly originating from the Colorado School of Mines between 1976 and 1984, further developed the idea of solving a system of algebraic equations using condensation, or a mathematically equivalent process. They are summarized in the author's previous paper.[5] Greening's work of 1982 showed that all of the monomial-based methods presented in the Colorado School of Mines works are mathematically equivalent to one another, and to the method of Rijckaert and Martens;[6] the name 'monomial method' is used by author to describe them collectively, as well as to describe the same method developed independently by the author.[7]

One of the Colorado School of Mines works is particularly relevant to the subject of this paper. Tulk[8] observed that the monomial method is 'indifferent' to the selection of starting point in the one-dimensional case for large values of the variable. He demonstrated that the monomial method reformulates the algebraic equation so that is has a monomial asymptote, giving rise to the observed indifference to starting point. A similar behaviour in the two-dimensional case has been reported by the author in previous works.[7,5]

This paper builds on these observations. In the one-dimensional case, the reformulated algebraic equation is shown to have dual asymptotes, which provides an indifference to very small, as well as very large, starting points. Then the multi-dimensional case is considered, and it is shown that monomial asymptotes exist in all directions of extreme starting points on the plane. A graphical construction is presented to aid in the identification of the monomial asymptotes in the two-dimensional case. Finally, the relevance of the monomial asymptotes are discussed in relation to specific engineering applications.

ONE-DIMENSIONAL CASE

Consider the case of a single *signomial*, or generalized polynomial,

$$\sum_{i=1}^{T} c_i x^{a_i} = 0 \tag{1}$$

with arbitrary real exponents and coefficients, but with a variable, x, that is restricted to be positive. The monomial method solution of this equation proceeds as follows. First, the terms are grouped according to their sign, positive terms in one group, $p(x)$, and negative terms in the other, $q(x)$, so that the signomial becomes $p(x) - q(x) = 0$. The negative terms are brought to the right-hand side and then are divided through to yield a rational form, $p(x)/q(x) = 1$. A monomial approximation is made separately to $p(x)$ and $q(x)$, using the process of condensation about the current operating point. Since the quotient of two monomials is again a monomial, the result is of the form $\beta x^{\alpha} = 1$. This monomial is solved to give the new operating point, $x = \beta^{-1/\alpha}$. The process repeats to convergence. Previous literature contains further explanation of the monomial method.[7,5]

It has been demonstrated that the monomial method can be interpreted as Newton's method applied to the reformulated problem, $g(z) = 0$, where $g(z) \equiv \ln[p(e^z)/q(e^z)]$ and $z = \ln(x)$.[5-10] This logarithmic/rational/exponential form has some unusual properties. Figure 1 presents a graph of 20 randomly generated $g(z)$ functions with six positive and six negative terms whose coefficients and exponents are randomly chosen as real numbers in the range 0·0–10·0. It is visually evident from this figure that $g(z)$ acts like a linear function of z for large positive and negative values of z. Since Newton's method is based on linearization, when the starting point is

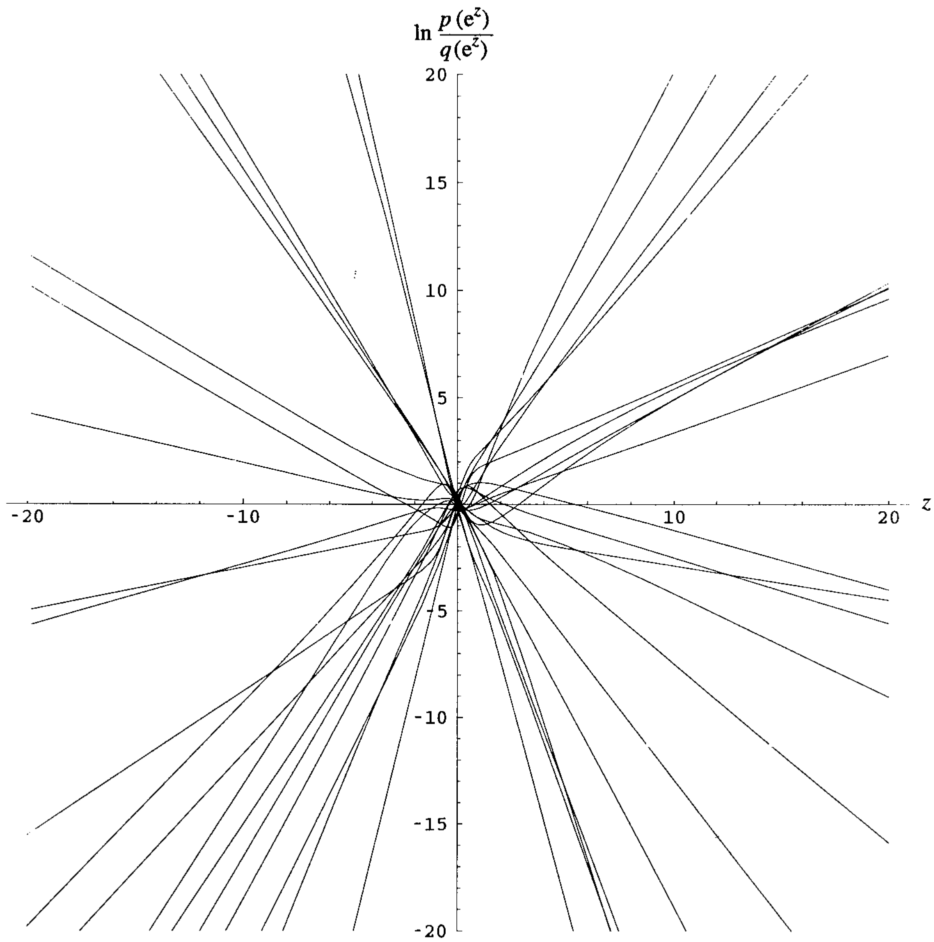


Figure 1. A plot of 20 randomly-generated generalized polynomial functions after being recast in logarithmic/rational/exponential form

large (positive or negative), the Newton approximation will match the linear asymptote of $g(z)$ nearly perfectly. The outcome of the first iteration will be essentially identical for all large z starting points of the same sign.

The linear asymptotes of $g(z)$ correspond to monomial asymptotes of $p(x)/q(x)$. The reason for $p(x)/q(x)$ becoming asymptotically monomial can be understood in terms of dominant terms in $p(x)$ and $q(x)$. Extreme positive and negative values of z correspond to very large and very small (approaching zero from the positive side) values of x, respectively, since $z = \ln(x)$. For these extreme values of x, one of the terms in $p(x)$ dominates, as does one of the terms in $q(x)$. For large x, the highest-order term dominates; for small x, the lowest. Therefore, the ratio $p(x)/q(x)$ approaches a ratio of monomials (equivalent to a single monomial) for extreme x, which is nearly perfectly approximated by the monomial method. This helps to explain the very rapid movement toward a solution in the first iteration that has been observed with the monomial method. when initiated from a starting point very distant from a solution.

As x grows large enough to make the highest-order term in each of $p(x)$ and $q(x)$ dominate, $p(x)/q(x) = 1$ approaches $(c^p_{max}/c^q_{max})\, x^{a^p_{max} - a^q_{max}} = 1$, where a^p_{max} is the largest positive exponent in $p(x)$, and c^p_{max} is its associated coefficient (*not* the largest coefficient in $p(x)$). Similarly, a^q_{max} and c^q_{max} correspond to the term in $q(x)$ with largest positive exponent. The solution of this equation (and the outcome of the first iteration of the monomial method when x is very large) is

$$x = (c^p_{max}/c^q_{max})^{-1/(a^p_{max} - a^q_{max})} \tag{2}$$

The only instance in which this would fail is when $a^p_{max} = a^q_{max}$. This will never happen since terms with identical exponents will not appear in $p(x)$ and $q(x)$ simultaneously—they will have been combined prior to separating the signomial into positive and negative parts. The solution for that case of very small x is

$$x = (c^p_{min}/c^q_{min})^{-1/(a^p_{min} - a^q_{min})} \tag{3}$$

where 'min' indicates the term with the lowest order (largest negative exponent). The two extreme starting point solutions, (2) and (3), will be referred to henceforth as the 'asymptotic solution set'.

The monomial method derives much of its performance strength from the fact that $p(x)/q(x) = 1$ is asymptotically monomial in both extremes. To demonstrate what effect this has on performance in comparison to Newton's method, a set of 10 000 randomly generated polynomial equations were solved by both Newton's method and the monomial method. The equations were created by expanding 10 000 instances of $(x - s_1)(x - s_2)\ldots(x - s_n) = 0$, where n is a random integer in the range 2–10, inclusive, and the s_i are randomly selected real values in the range 1·0–10·0. The 10 000 equations were each solved six times from a series of different starting points: $x = \{1, 10, 100, 10^3, 10^4, 10^5\}$. Figure 2 plots histograms of the outcomes of one iteration of the two methods applied to each of the 10 000 equations. The circle indicates the starting point. For starting points near the region where the solutions exist ($1{\cdot}0 \leqslant x \leqslant 10{\cdot}0$), the two methods give comparable results. But as the starting point grows, the monomial method distribution eventually becomes invariant to further increases in the starting point; this invariant distribution describes 10 000 upper asymptotic solutions, as described by equation (2). In contrast, the first iterate of Newton's method closely follows the starting point no matter how large it becomes. This causes Newton's method to require many more iterations to be performed before it reaches a solution.

Figure 3 plots histograms of the number of iterations required by each method to converge to one of the solutions to within a tolerance of 10^{-5}. Again the two methods are comparable when initiated in the region of the solutions, but Newton's method requires an increasing number of iterations as the starting point grows. An unusual periodic structure can be seen in the Newton histograms for very large starting points. Further investigation reveals that each lobe of the histogram corresponds to the randomly generated equations of a common order; the right-most lobe corresponds to all tenth-order equations, the next, all ninth-order, etc. The *speed* of convergence of Newton's method (not to be confused with the quadratic *rate* of convergence) is strongly correlated to the order of the equation. In the case of polynomial equations of order n, this correlation is easy to show. Start with Newton's method, $x_{new} = x - f(x)/f'(x)$ and evaluate

$$\lim_{x\to\infty} \frac{\Delta x}{x} = \lim_{x\to\infty} -\frac{f(x)}{xf'(x)} = -\frac{1}{n} \tag{4}$$

Therefore, the first iterate by Newton's method, when initiated from a point where the highest term dominates, will be approximately $(1 - 1/n)x$. The first iteration histograms of Figure 2 show what appears to be three peaks for the Newton method case as the starting point gets very large. Actually, there are nine closely spaced peaks that are masked by the low resolution of the

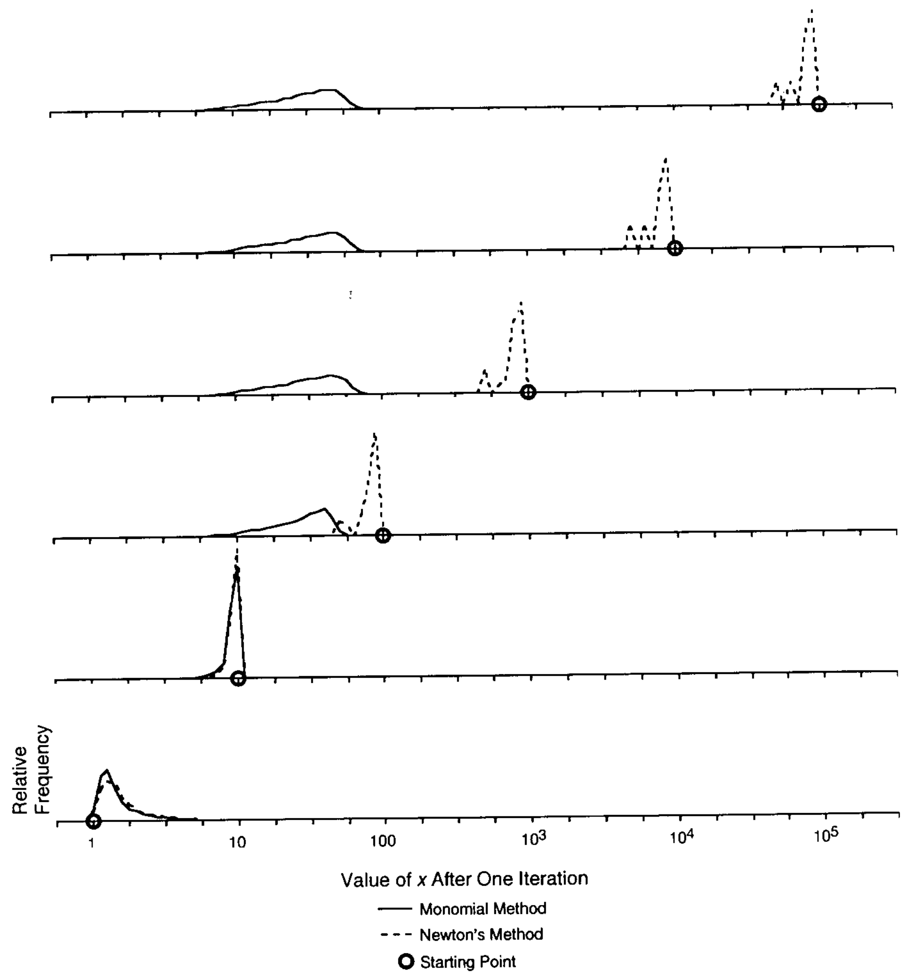


Figure 2. Histograms of first iterate locations for 10 000 randomly-generated polynomial equations solved by Newton's method (dashed) and the monomial method (solid), for six different starting points (circles)

histogram intervals selected for collecting the data. The nine peaks correspond to the nine different equation orders (second order–tenth order) that were randomly generated. The highest-order produce the slowest progress toward a solution by Newton's method. The monomial method does not suffer from this 'order locking'. The first iterates are independent of the equation order, and the vast majority of them fall within an order of magnitude of the solution, regardless of the starting point chosen.

MULTI-DIMENSIONAL CASE

The linear asymptotes of $g(z)$ (or equivalently, the monomial asymptotes of $p(x)/q(x)$) extend to the multi-dimensional case in which $P(X)$ and $Q(X)$ each have the form

$$\sum_{i=1}^{T} c_i \prod_{j=1}^{N} x_j^{a_{ij}}, c_i > 0, \quad x_j > 0 \tag{5}$$

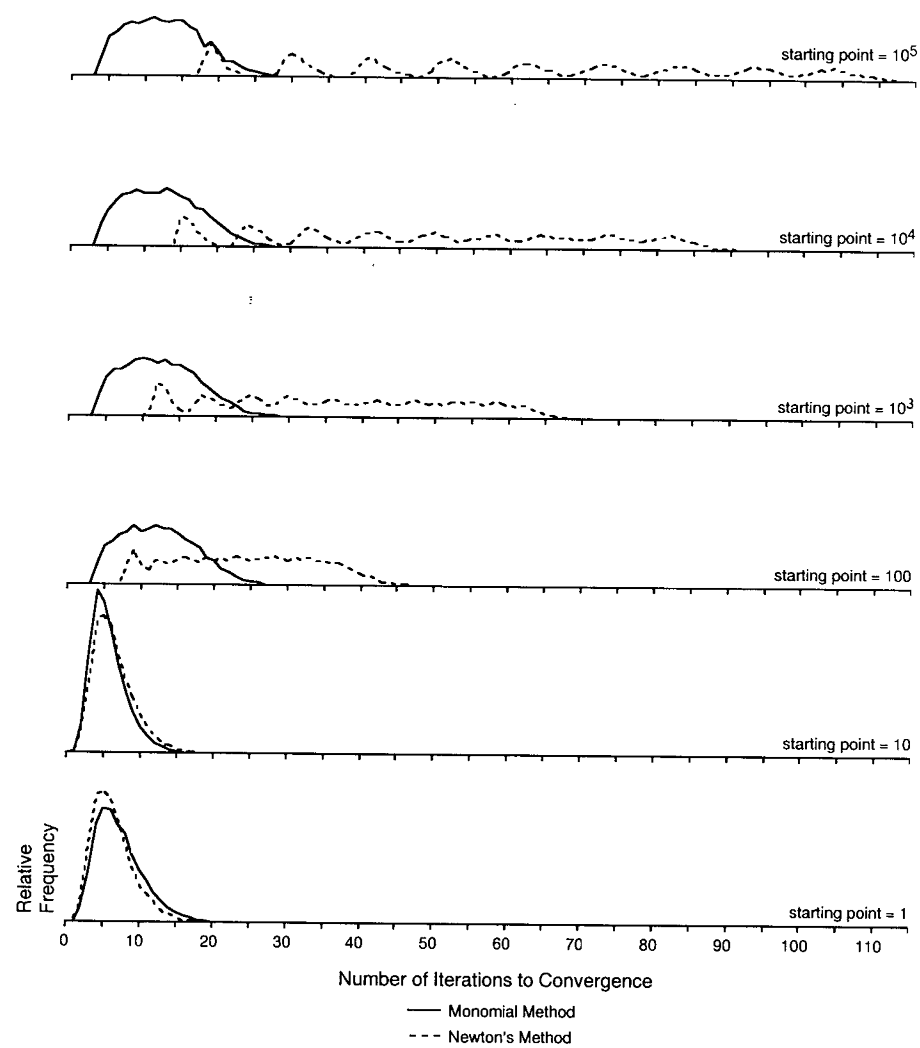


Figure 3. Histograms of iteration count at convergence for 10 000 randomly-generated polynomial equations solved by Newton's method (dashed) and the monomial method (solid), for six different starting points

Figure 4 is a surface plot of a randomly-generated, two-dimensional generalized polynomial, $P(x_1, x_2) - Q(x_1, x_2) = 0$, recast in the form $G(z_1, z_2) \equiv \ln[P(e^{z_1}, e^{z_2})/Q(e^{z_1}, e^{z_2})]$. Note that the linear asymptotes form planar 'facets' in the surface; each facet corresponds to a pair of dominant terms, one in $P(x_1, x_2)$ and one in $Q(x_1, x_2)$.

The location of the facets can be identified using a simple graphical construction. First, consider the terms in $P(x_1, x_2)$. Each term has a pair of exponents, $a^p_{i,1}$ and $a^p_{i,2}$, acting on x_1 and x_2, respectively. To create the graphical construction, we plot the exponent pair for the first term, $(a^p_{1,1}, a^p_{1,2})$, as a point in cartesian co-ordinates, and then draw a circle that passes through $(a^p_{1,1}, a^p_{1,2})$ and the origin as opposite extremes of the circle. We draw additional circles for each additional term in $P(x_1, x_2)$. Now consider the outer boundary of the union of all of the circles. As

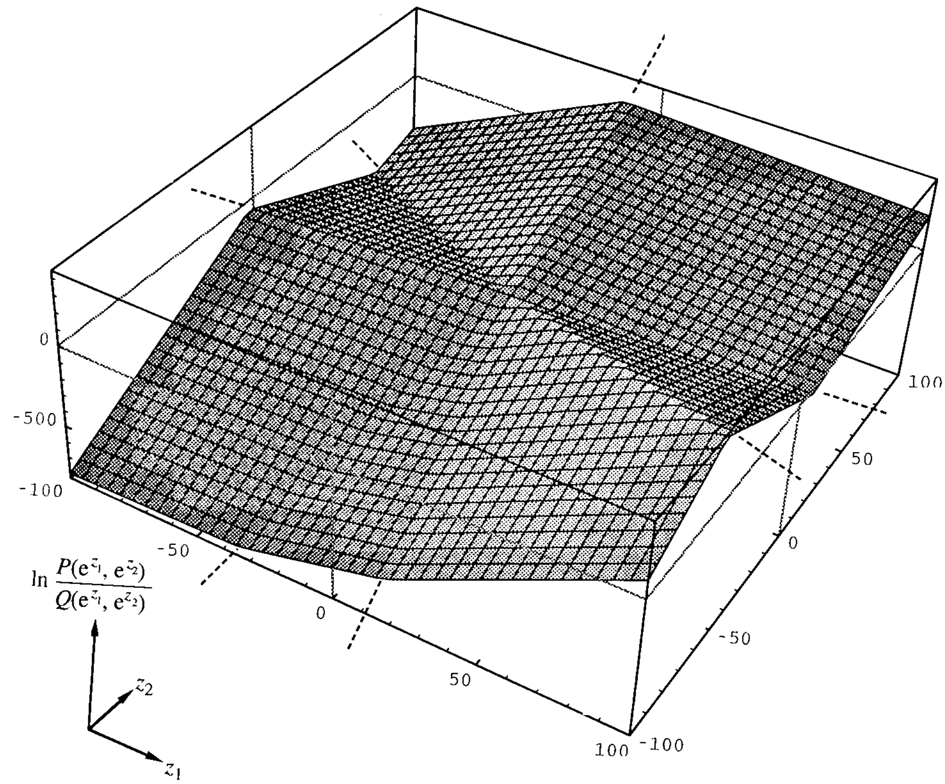


Figure 4. Two-dimensional generalized polynomial function (equation (6)) after being recast in logarithmic/rational/exponential form (dashed lines indicate boundaries between planar facets)

we traverse this outer boundary, we encounter several intersection points between circles. Each of these intersection points defines the orientation of an edge between two planar facets of $G(z_1, z_2)$. The edge is defined by a ray passing from the origin through the intersection point. Note that although the circles are plotted on the $a_1 - a_2$ plane, the orientation of the facet edges should be interpreted by replacing the $a_1 - a_2$ co-ordinate system with the z_1–z_2 co-ordinate system. The dominant term in each wedge-shaped region between adjacent edges of the construction corresponds to the circle formi ng the outer boundary in that region. A separate construction is then made for the terms of $Q(x_1, x_2)$. (See Appendix A for justification of this construction, and for additional construction aids.)

Figure 5 presents this construction for the positive terms of the surface plotted in Figure 4, which is based on the generalized polynomial:

$$P - Q = 760x_1^{-3.04}x_2^{-1.27} + 626x_1^{1.09}x_2^{-0.525} + 326x_1^{-6.72}x_2^{-0.0263} + 949x_1^{0.0993}x_2^{9.16} - (603x_1^{-7.15}x_2^{-7.66} + 520x_1^{-6.07}x_2^{-3.99} + 265x_1^{-7.44}x_2^{5.22} + 757x_1^{0.764}x_2^{6.37}) \quad (6)$$

The four heavy lines in Figure 5 identify the boundaries between regions of dominant terms. A similar construction is shown in Figure 6 for the negative terms that appear in $Q(x_1, x_2)$. Note that the latter figure indicates that term 2 is never dominant in $Q(x_1, x_2)$. When superimposed, the two constructions identify the ratio of terms (one each from P and Q) that govern the asymptotic

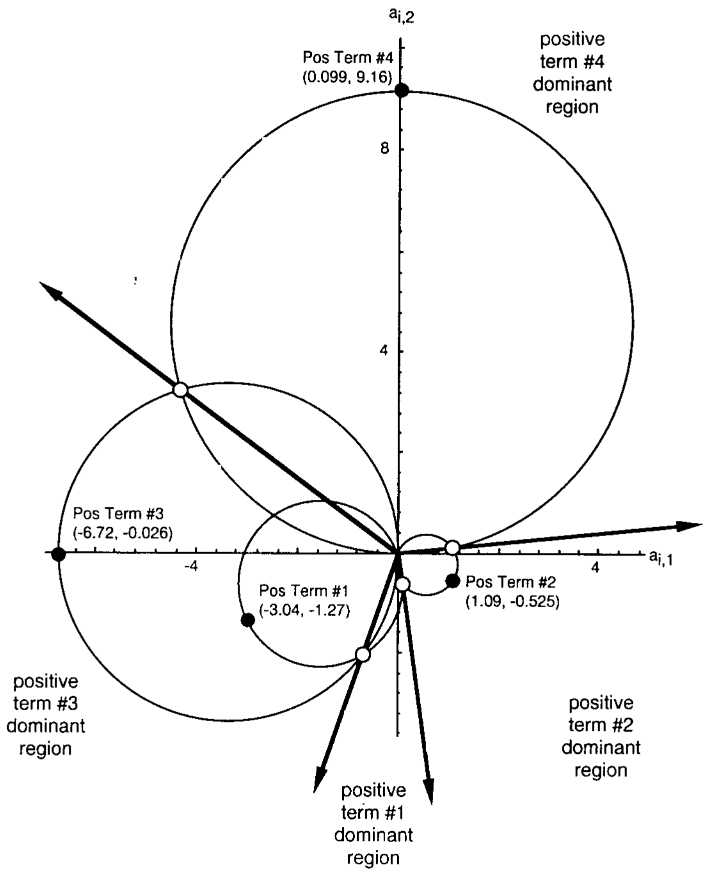


Figure 5. Regions of dominance for positive terms (P) in equation (6)

behaviour of $P(x_1, x_2)/Q(x_1, x_2)$ in each region. This figure is referred to as a *dominance diagram*, and is shown in Figure 7. The regions in the dominance diagram can be seen to match the facets of Figure 4.

The dominant terms in $P(x_1, x_2)$ and $Q(x_1, x_2)$ in a region of dominance define the monomial asymptote of $P(x_1, x_2)/Q(x_1, x_2)$ in the region. This monomial approximation is

$$(c^p_{\text{dom}}/c^q_{\text{dom}})x_1^{a^p_{\text{dom},1} - a^q_{\text{dom},1}} x_2^{a^p_{\text{dom},2} - a^q_{\text{dom},2}} \tag{7}$$

where c^p_{dom}, $a^p_{\text{dom},1}$ and $a^p_{\text{dom},2}$ are the coefficient and exponents belonging to the dominant term in P, and c^q_{dom}, $a^q_{\text{dom},1}$ and $a^q_{\text{dom},2}$ are the coefficient and exponents belonging to the dominant term in Q. Since the monomial method operates on the transformed equation $P(x_1, x_2)/Q(x_1, x_2) = 1$, an asymptotic solution is found by setting monomial (7) equal to unity

$$(c^p_{\text{dom}}/c^q_{\text{dom}})x_1^{a^p_{\text{dom},1} - a^q_{\text{dom},1}} x_2^{a^p_{\text{dom},2} - a^q_{\text{dom},2}} = 1 \tag{8}$$

for each of the two equations of the system, and then solving them simultaneously.

When a system of two non-linear equations are being solved simultaneously, a pair of dominance diagrams serve to identify sets of simple systems of equations of the form (8) whose solutions constitute the asymptotic solution set. By superimposing the two dominance diagrams,

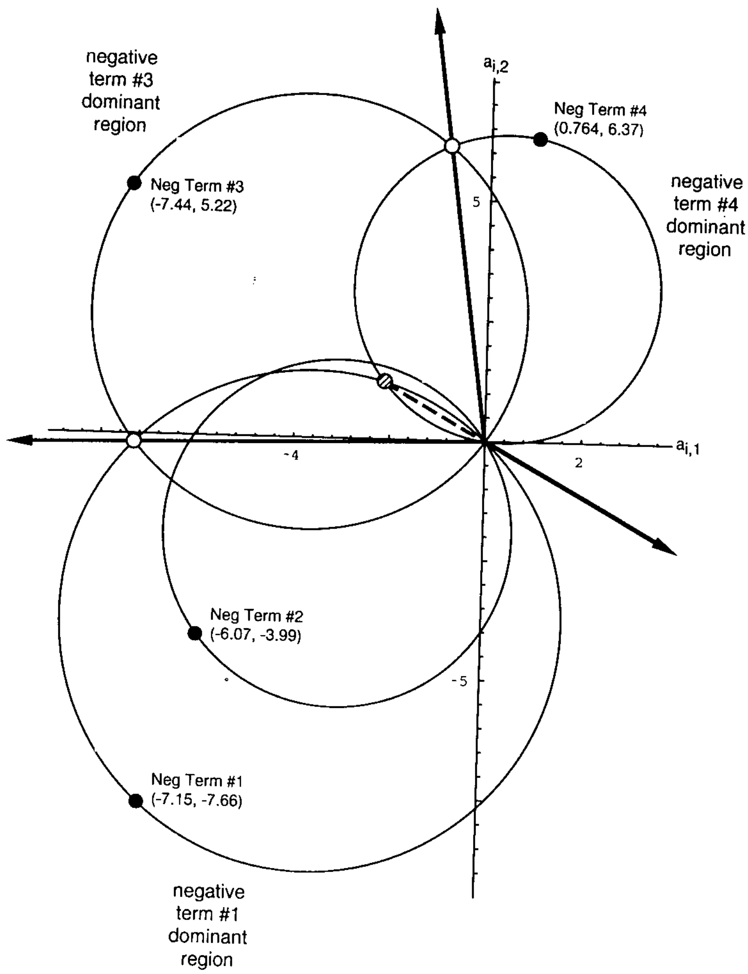


Figure 6. Regions of dominance for negative terms (Q) in equation (6)

a *system dominance diagram* is produced. Since the number of regions in any one dominance diagram cannot exceed the number of terms in the generalized polynomial, the number of regions in the system dominance diagram is at most equal to the total number of terms in the two equations. In many cases, the number of regions is far less than the total number of terms because of coincident edges of dominant regions or because of non-dominant terms.

To demonstrate how to construct the system dominance diagram and what useful information can be obtained from it, we turn to the system of equations describing the frame problem presented in section 7 of the original monomial method paper.[7]

$$11\,664x_1x_2^{-4} + 54x_1^2x_2^{-4} + 21{\cdot}6x_1^{-2} - 10{\cdot}8x_1^4x_2^{-4} - 4{\cdot}32 = 0 \tag{9}$$

$$6998{\cdot}4x_1^4x_2^{-7} + 18x_1^4x_2^{-6} + 8398x_2^{-3} - 12{\cdot}96x_1^4x_2^{-4} - 5{\cdot}184 = 0 \tag{10}$$

Figure 8 presents the dominance diagrams for each of the two equations in this system. Each region is identified by the linear function that dominates in that region. Figure 9 presents the

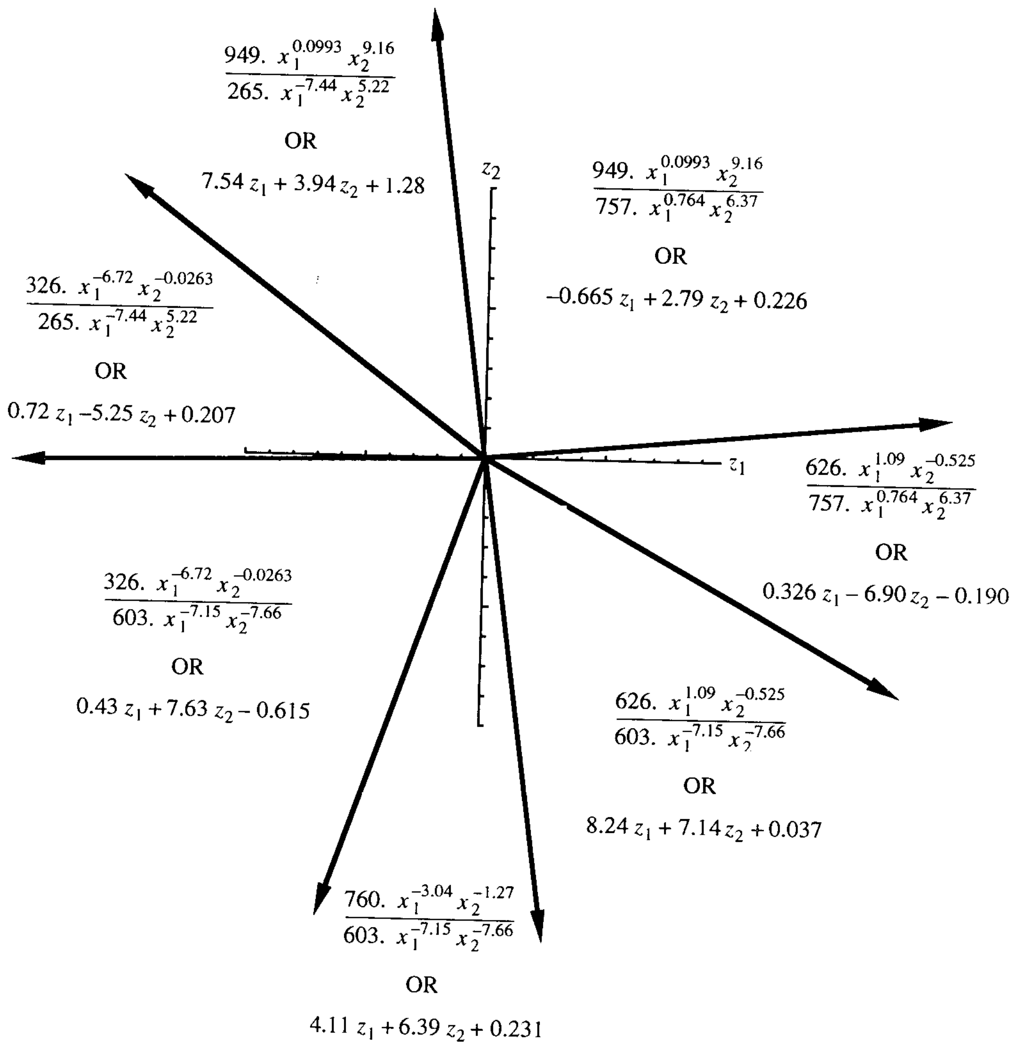
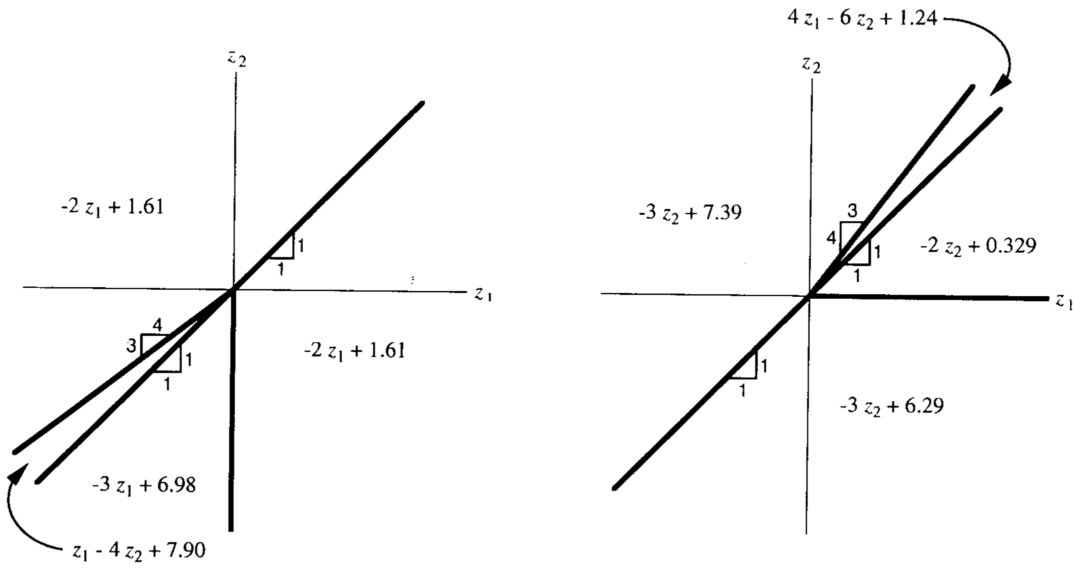


Figure 7. Dominance diagram for equation 6

system dominance diagram, which has six regions, and the six asymptotic solutions that are obtained by simultaneously solving the dominant linear equations corresponding to each region.

The system dominance diagram explains why the monomial method was found to perform so well when applied to the frame problem. First, the asymptotic solutions are all fairly near to the solutions of this problem, which are (2·92, 11·72), (6·45, 11·15), (9·12, 9·71) and (− 1·98, 11·74). This means that starting points distant from the region of the solutions *in any direction* will yield very good results in the first iteration of the monomial method.

The second observation relates to numerical conditioning. In general, each iteration of the monomial method produces a system of linear equations by transforming the system of approximating monomial equations using a logarithmic change of variables, $z_j = \ln(x_j)$. The system of linear equations can be difficult to solve accurately if it is ill-conditioned. It was observed



(a) Equation 9

(b) Equation 10

Figure 8. Dominance diagrams for the frame problem (equation (9) and (10))

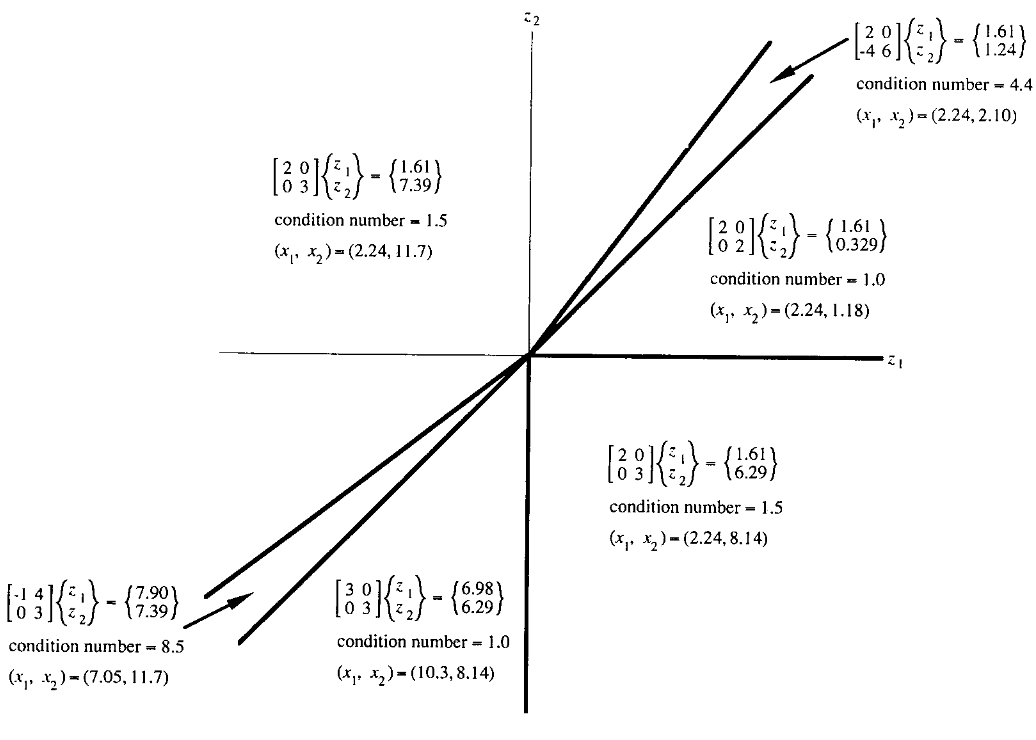


Figure 9. System dominance diagrams for the frame problem (Equations (9) and (10)), showing the asymptotic linear system, the condition number of the linear system, and the asymptotic solution

previously that the condition number of the linear system for the frame problem was exceptionally good for the monomial method, approaching the absolute best condition number (unity) in many regions.[7] The system dominance diagram explains why this is so. The coefficient matrix of the linear system in each region and the corresponding condition number are shown in Figure 9 Extreme starting points will give rise to linear systems approaching one of these condition numbers, depending on the direction of the distant starting point. For example, for starting points with very small x_1 and large x_2, the condition number will approach 1·5 (the upper-left quadrant of the system dominance diagram).

A final noteworthy observation is that the monomial system becomes uncoupled asymptotically in the four largest regions of the system dominance diagram. In other words, for extreme starting points, only one variable appears in each governing monomial in these four regions. This greatly simplifies the solution of the linear system.

The system dominance diagram also explains the poor behaviour experienced with certain problems solved by the monomial method.[5] Consider the problem $x_1^2 + x_2^2 - 2x_1 - 2x_2 + 1 = 0$ and $x_1^2 + x_2^2 - 4x_1 - 4x_2 + 7 = 0$, which describes the intersection of two circles. Since the exponent structure of the two equations is identical, the dominant system of equations in each region of the system dominance diagram will be singular. One can expect very poor performance from the monomial method for distant starting points.

For systems of equations with more than two variables, the dominance diagrams are more difficult to construct and visualize. Nevertheless, it is a simple matter to sample the regions of dominance for systems of any size by specifying a direction of a distant starting point. Appendix A explains how this can be accomplished and presents a few examples. The sampled direction may indicate that the monomial method is worth pursuing as a solution technique if the dominant system is uncoupled or if it has good conditioning. It also provides an asymptotic solution, which sometimes falls within the general vicinity of the solutions of the non-linear system. On the other hand, if the sampled direction indicates that the dominant system is defective (i.e. rank deficient or poorly conditioned), then the monomial method is probably not a good choice for solving the problem, at least from starting points in that direction.

Limited numerical experiments have shown that if a problem is asymptotically defective in several adjacent regions, then selecting a starting point on the boundary between defective regions sometimes yields good results. Along these boundaries, two terms are competing for dominance, and together they may give rise to an asymptotic system that is not defective. Appendix A presents an example of this.

ENGINEERING APPLICATIONS

The frame problem described previously is solved effectively by the monomial method because it has 'good asymptotics'. The asymptotic solution set falls very near to the true solutions, and the asymptotic systems of equations are mostly uncoupled and have excellent conditioning. However, one should not infer from this example that a well-structured asymptotic system implies that the monomial method will perform well. For example, it is possible that the asymptotic solution corresponding to one region of dominance may fall within another region of dominance, which in turn, may send it back to the first one. This will lead to undesirable cyclic behaviour and failure to converge. Many similar scenarios are possible, even though all individual asymptotic systems might be well-structured and easily solved.

Based on numerous experiments performed by the author and others, some generalizations about the performance of the monomial method applied to engineering problems can be observed. The monomial method tends to perform well for problems that have the following structure: a basic underlying monomial power law, which is a function of the independent variables, combined with a summation of intermediate quantities that are based on the power law. These problems tend have good asymptotics and have been found to be treated very effectively by the monomial method. Such problems include the following:

Structural frame design.[7] The basic monomial power laws include the beam bending stress relationship, $\sigma = Mc/I$, and the axial stress relationship, $\sigma = P/A$, where c, I, and A are each monomial functions of the independent variables (member dimensions). The stresses are then summed in each member and set equal to an allowable stress.

Chemical equilibrium.[6,11–14] Determining the chemical species distribution of a reaction at equilibrium using the kinetic approach involves a monomial power law describing the reaction rate equations, $\Pi_{i \in \text{products}} a_i / \Pi_{j \in \text{reactants}} a_j = K$, where the ratio of products of forward and reverse reaction rates are set equal to an equilibrium constant. In addition, mass balance and charge balance are expressed as a summation of species concentrations. The reaction rates in the monomial power law are monomial functions of the species concentrations, which are the independent variables.

Investment rate of return.[11,15] The monomial power law is the compound interest relationship, $PV = FV/(1+i)^n$, where the independent variable is taken as $x = (1+i)$. To determine the discounted cash flow rate of return, the present values of all cash flows are summed and set equal to zero, $\sum PV = 0$.

The important feature present in the above problems is that the independent variables appear directly in the monomial power law. Based on the success of the monomial method with these applications, one would expect success with other applications that have a similar mathematical structure. A large class of engineering applications share a common mathematical structure,[16]

$$e = b - Ax, \quad y = Ce, \quad A^{\mathrm{T}} y = f \tag{11}$$

In many cases, C is a diagonal matrix of monomial power law functions of the independent variables. For example, for the analysis of a truss structure, matrix A describes the connectivity of members, x is a vector of nodal displacements, b is a vector of imposed member elongations (usually zero) e is a vector of member elongations, y is a vector of element axial forces, and f is a vector of externally applied nodal loads. The C matrix is the constitutive relationship (Hooke's law), which relates member forces to member elongations through a monomial power law function of member properties: length, L, cross-sectional area, a, and modulus of elasticity, E. Normally, E, a, and L are fixed in value, and system (11) forms the truss *analysis* problem, which is linear. The truss *design* problem would consider one of the member properties, usually a in this case, to be the independent variable, making system (11) non-linear, and making the entries of the C matrix monomial functions of the independent variables. To complete the design problem, additional equations would be specified to impose performance conditions on the system, such as requiring the member stresses (y/a) to be equal to an allowable stress.

Another engineering application that has this mathematical structure is the design of an electrical resistive network, in which the independent variables are the individual resistance values, the monomial power law (matrix C) is Ohm's law, and system (11) relates connectivity of the circuit (A), nodal voltage potentials (x), voltage sources (b), voltage drops across each resistor (e), flow of current through each resistor (y), and current sources at each node (f). The design problem would entail determining the resistance values to produce a desired circuit behavior in terms of voltages and/or currents.

A final example of an engineering application with this structure is the design of a fluid piping network, such as a municipal water supply network. Here, the monomial power law can be taken as the Hazen–Williams formula, which is a monomial function of the independent variables (pipe diameter, length, and/or roughness) that relates head loss and fluid velocity. System (11) relates connectivity of the network (A), the head at each junction (x), head sources along a branch, such as pumps (b), the head loss across each pipe (e), the flow through each pipe (y), and the demand at each junction (f). The design problem would entail solving for pipe characteristics that provide enough flow at each junction to satisfy demand at some minimum pressure. A possible complicating factor in this application is that the Hazen–Williams formula is not differentiable at the point where the flow changes direction.

To treat the engineering design problem more generally, the solution of a system of equations is often insufficient. One must generally include inequality constraints and an objective function in the problem statement. In many cases the auxiliary performance equations, such as the allowable stress equations discussed previously, are actually unilateral; member stresses can be lower than, but not greater than, the allowable stress. This promotes the problem to the status of non-linear programming. The monomial method is still useful in this case. It has been shown to fit seamlessly within the framework of the generalized geometric programming (GGP) non-linear programming method.[10,17,18]

In contrast to the applications described above, some engineering applications have a mathematical structure not as well-suited to the monomial method, similar to the intersection of circles problem that gave rise to a singular asymptotic system. Problems in mechanism design are often of this form; the kinematics of the mechanism are governed by a set of member length constraints of the form $x_i^2 + y_i^2 = l_i^2$. These basic building blocks of the mechanism problem are not monomial in form, and often give rise to an asymptotic system that is defective. Unpublished experiments with mechanism design problems conducted by the author confirm that the monomial method has considerable difficulty when initiated from distant starting points.

CONCLUSIONS

Many of the performance characteristics of the monomial method can be explained in terms of asymptotic properties of algebraic systems and the ability of the monomial method to exploit these properties. The monomial method recasts the algebraic system to have monomial asymptotes in all directions, so that the monomial approximation becomes asymptotically exact. This explains the very rapid movement toward a solution in the first iteration from distant starting points that has been observed with the monomial method.

The suitability of the monomial method for solving a system of algebraic equations can be assessed, to some extent, by examining the structure of the monomial asymptotes of the reformulated system. In some cases, the monomial method uncouples the algebraic system asymptotically or provides a well-conditioned approximating linear system. In undesirable cases, the monomial method produces a defective asymptotic system that cannot be solved. As a rule of thumb, problems that are based on an underlying monomial power law tend to have good asymptotics, and have been found to be treated effectively by the monomial method.

ACKNOWLEDGEMENTS

This material is based on work supported by the National Science Foundation under PYI Grant No. DDM-8957410, and was conducted during an appointment at the University of Illinois Center for Advanced Study.

APPENDIX A

This section derives the conditions for which a particular term in a sum of positive terms will dominate for extreme values of the strictly positive variables. The one-dimensional case is very simple—the term with the largest exponent will dominate when the variable is very large, and the term with the smallest (most negative) exponent will dominate when the variable is very small (approaching zero).

A simple graphical construction facilitates identification of the dominant term in the two-dimensional case. Let the sum of positive terms (*posynomial*), representing either $P(x_1, x_2)$ or $Q(x_1, x_2)$, be given by

$$\sum_{i=1}^{T} c_i x_1^{a_{i,1}} x_2^{a_{i,2}} \tag{A1}$$

If the variables are transformed according to $z_1 = \ln(x_1)$ and $z_2 = \ln(x_2)$, then the posynomial takes the form

$$\sum_{i=1}^{T} c_i e^{z_1 a_{i,1} + z_2 a_{i,2}} \tag{A2}$$

Suppose that the vector z is described in polar co-ordinates, (r, θ), so that $z_1 = r\cos\theta$ and $z_2 = r\sin\theta$. The posynomial now takes the form

$$\sum_{i=1}^{T} c_i e^{r(a_{i,1}\cos\theta + a_{i,2}\sin\theta)} \tag{A3}$$

To identify which term, i, dominates in a given direction, θ, as z gets very large in that direction (i.e. as r gets large), it is sufficient to identify which term has the largest value of $a_{i,1}\cos\theta + a_{i,2}\sin\theta$. This follows from the observation that the coefficients, c_i, are eventually dwarfed by the exponential terms as r grows large, and so the term with the largest coefficient of r eventually dominates. Note that the function $a_{i,1}\cos\theta + a_{i,2}\sin\theta$ is a circle in polar co-ordinates on the a_1–a_2 plane with a diameter corresponding to the line segment between $(a_1, a_2) = (0, 0)$ and $(a_1, a_2) = (a_{i,1}, a_{i,2})$. This justifies the graphical construction described earlier for identifying the dominant term, since the circle appearing on the boundary of the union of the circles in a given direction from the origin has the largest value of $a_{i,1}\cos\theta + a_{i,2}\sin\theta$ in that direction.

As mentioned earlier, the outer boundary of the union of all circles contains a number of intersections of circles. These intersections are called *external* intersections. All other intersections of circles are called *internal* intersections. The partitioning of the z plane into wedge-shaped *regions of dominance* is done by drawing a series of rays from the origin through each external intersection. Each region of dominance defines an interval of θ for which one of the terms in the posynomial will eventually dominate for large enough r. The particular term that will dominate corresponds to the circle that forms the outer boundary of the union of circles in that region.

In some cases, an exterior intersection coincides with the origin, such as between circles one and four in Figure 6. In these cases, the orientation of the boundary ray is determined by extending a ray from other (internal) intersection of the two circles (the shaded dot at the intersection of circles one and four in Figure 6), through the origin. In general, the precise orientation of the boundary ray between dominant terms i and j can be determined by $\tan\theta = -(a_{i,1} - a_{j,1})/(a_{i,2} - a_{j,2})$.

For higher-dimensional problems, the graphical construction becomes more difficult to draw; in three dimensions, each region of dominance becomes an origin-based polyhedral cone.

Although it is more difficult to visualize the overall structure of the regions of dominance, they can be easily described (or sampled) mathematically. An N-dimensional posynomial, equation (5), has an associated exponent matrix, $\mathbf{E} = [a_{i,j}]$, with T rows and N columns. Consider a z vector, $z_j = \ln(x_j)$, oriented in some direction as defined by a vector of direction numbers or direction cosines, n (even the individual components of z will suffice here, $n_j = z_j$). The dominant term in that direction will correspond to the element of product $\mathbf{E}n$ with the largest value. Thus, the region of dominance for term t is the set of orientations, n, for which the tth element of $\mathbf{E}n$ is the largest (most positive) element.

As an example, consider the system of equations describing the intersection of two parabolas, $(x_1 - 1)^2 - x_2 + 1 = 0$ and $(x_2 - 1)^2 - x_1 + 1 = 0$, which has two solutions, $(1, 1)$ and $(2, 2)$. After expanding the equations and grouping the terms into posynomials, we have $P_1 = x_1^2 + 2$, $Q_1 = 2x_1 + x_2$, $P_2 = x_2^2 + 2$, and $Q_2 = 2x_2 + x_1$. The exponent matrices corresponding to these four posynomials are

$$\mathbf{E}_{P_1} = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}, \quad \mathbf{E}_{Q_1} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \mathbf{E}_{P_2} = \begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix}, \quad \mathbf{E}_{Q_2} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \tag{A4}$$

To determine the asymptotic behaviour in a particular direction, say along $z = (1, 2)$, we multiply each of the exponent matrices by a direction vector, $n = (1, 2)^{\mathrm{T}}$:

$$\mathbf{E}_{P_1} n = \begin{Bmatrix} 2 \\ 0 \end{Bmatrix}, \quad \mathbf{E}_{Q_1} n = \begin{Bmatrix} 1 \\ 2 \end{Bmatrix}, \quad \mathbf{E}_{P_2} n = \begin{Bmatrix} 4 \\ 0 \end{Bmatrix}, \quad \mathbf{E}_{Q_2} n = \begin{Bmatrix} 2 \\ 1 \end{Bmatrix} \tag{A5}$$

The element of the product $\mathbf{E}n$ with the largest value corresponds to the dominant monomial term in the posynomial, thus the asymptotic monomial system in the direction $n = (1, 2)$ is $x_1^2/x_2 = 1$ and $x_2^2/2x_2 = 1$. The asymptotic solution is $(x_1, x_2) = (1{\cdot}414, 2{\cdot}000)$ and the condition number of the linear system (after transforming into log space) is 2·618. Thus, the monomial method appears to be a good candidate for solving this system from distant starting points in this particular direction. (Note that distant points the direction $(1, 2)$ in z-space correspond to distant points in x-space along the curve $x_2 = x_1^2$.)

If we consider distant starting points in another direction, $n = (-1, 2)$, we find that the asymptotic linear system is rank deficient and thus singular. The monomial method is likely to give poor results from distant starting points in this direction. In fact, along any direction $n = (a, b)$ where $a < 0$ and $a < b$, the asymptotic linear system proves to be singular. This is also true for $b < 0$ and $a > b$. However, between these two adjacent regions of defective asymptotic behaviour, there is a single direction, $n = (-1, -1)$, where the linear system is not defective. Along this direction, the asymptotic monomial system degenerates into a linear system with solution $x_1 = x_2 = \frac{2}{3}$ and condition number 3·0. Distant starting points in this direction (even approximately in this direction) readily converge to a solution by the monomial method with none of the numerical difficulties encountered with other starting points in this third quadrant. The lesson to be learned here is that the asymptotic behaviour along boundaries between regions of dominance can exhibit behaviour quite different than that of all adjacent regions.

REFERENCES

1. R. J. Duffin, 'Linearizing geometric programs', *SIAM Review*, **12**, 211–227 (1970).
2. M. Avriel and A. Williams, 'Complementary geometric programming', *SIAM J. Appl. Math.*, **19**, 125–141 (1970).
3. U. Passy, 'Generalized weighted mean programming', *SIAM J. Appl. Math.*, **20**, 763–778 (1971).
4. M. J. Rijckaert and X. M. Martens, 'A condensation method for generalized geometric programming', *Math. Prog.*, **11**, 89–93 (1976).

5. S. A. Burns, 'The monomial method: extensions, variations, and performance issues', *Int. j. numer. methods eng.*, **37**, 2093–2107 (1994).
6. D. R. Greening, 'A proof of convergence for a condensation approach to the solution of systems of nonlinear equations', *M.S. Thesis*, Colorado School of Mines, T-2637, 1982.
7. S. A. Burns and A. Locascio, 'A monomial-based method for solving systems of nonlinear algebraic equations', *Int. j. numer. methods eng.*, **31**, 1295–1318 (1991).
8. R. K. Tulk, 'Transformed Newton's method', *Ph.D Thesis*, Colorado School of Mines, T-1884, 1976.
9. E. L. Peterson, 'The solution of posynomial equations', *Technical Report* # 010887-01, Department of Mathematics and Graduate Program in Operations Research, North Carolina State University, Raleigh, NC, 1987.
10. S. A. Burns, 'Generalized geometric programming with many equality constraints', *Int. j. numer. methods eng.*, **24**, 725–741 (1987).
11. David W. Baker, 'Application of the geometric inequality to the solution of systems of nonlinear equations', *Ph.D. Thesis*, Colorado School of Mines, T-2339, 1980.
12. Michael J. Cusick, 'Some insights into the fluidized-bed electrowinning of zinc from acidic zinc-sulfate solutions', *M.S. Thesis, Colorado School of Mines, T-*2551, 1981.
13. Thomas Wayne Wall, 'A numerical algorithm for the solution of chemical equilibrium problems', *M.S. Thesis*, Colorado School of Mines, T-2842, 1984.
14. Thomas Wayne Wall, Doran Greening and R. E. D. Woolsey, 'Solving complex chemical equilibria using a geometric programming based technique', *Oper. Res.*, **34**, 345–355, (1986).
15. Paul L. Bernard, 'An algorithm for the solution of investment problems having dual rates of return', *M.S. Thesis*, Colorado School of Mines, T-2711, 1982.
16. Gilbert Strang, *Introduction to Applied Mathematics*, Wellesley-Cambridge Press, Wellesley, MA, 1986, p. 90.
17. M. Avriel, R. Dembo and U. Passy, 'Solution of generalized geometric programs', *Int. j. numer methods eng.*, **9**, 149–168 (1975).
18. S. A. Burns, 'Application of the monomial method to shape optimization', *Adv. Des. Automat. ASME*, **DE 65-1**, 725–733 (1993).