Multivariable Terrain Methods

Angelo Lucia and Yang Feng
Dept. of Chemical Engineering, University of Rhode Island, Kingston, RI 02881

A novel geometric terrain methodology was recently presented for finding all physically relevant solutions and singular points to chemical process simulation problems based on intelligently moving along special integral curves of the gradient vector field to guide both uphill and downhill movement. The terrain methodology is extended here to include multivariable problems and integral curve bifurcations. The integral curves of interest are those associated with valleys and ridges and characterized as a collection of extrema in the norm of the gradient over a set of level curves. It is shown that integral curves can undergo tangent, pitchfork, and other types of bifurcations and that terrain methods are clearly superior to differential arc homotopy-continuation methods on problems that exhibit parametric disconnectedness. Several examples, including a pair of CSTR problems, a retrograde flash calculation, and the task of finding all azeotropes for a heterogeneous ternary mixture, are used to show that terrain methods represent a reliable, efficient, and global way of solving multivariable process engineering simulation and optimization problems. Geometric illustrations are used whenever possible to clarify underlying ideas.

Introduction

In a recent article, Lucia and Feng (2002) introduced a completely different, novel, and general approach to finding all physically meaningful solutions and singular points to mathematical models of physical systems by intelligently moving up and down the landscape of the least-squares function. The theoretical foundation for these global terrain methods rests on the fundamental geometric observations that:

(1) Solutions and singular points are smoothly connected when model functions are smooth;

(2) Valleys, ridges, ledges, and so on provide a natural characterization of this connectedness;

(3) Valleys and ridges can, in turn, be characterized as a collection of constrained extrema over a set of level curves;

(4) There is an equivalent characterization of valleys and ridges as solutions to generalized, constrained eigenvalue-eigenvector problems; and

(5) The natural flow of the Newton-like vector field tends to be along these distinct features of the landscape.

The overall algorithmic framework is based on the reliable and efficient computation of solutions and singular points, a terrain-following algorithm to efficiently move from one stationary point to another or to a boundary of the feasible region, and the concept of limited connectedness to decide when the method has finished. The resulting terrain methods were tested on a variety of two-dimensional (2-D) chemical process examples. These examples included a melt transition temperature calculation for a polymer blend, the determination of various minimum energy states in disk to vesicle transitions, finding the roots to the SAFT equation of state, and determining the conversion and temperature in an adiabatic CSTR. The associated numerical results clearly show that terrain methods are a reliable and efficient means of finding all physically relevant solutions and singular points for small dimensional problems. Figure 1 illustrates the connectedness of stationary points, and the path generated by our terrain-following algorithm, for the nonlinear CSTR example given in Shacham (1985).

The main objectives of this article are: (1) to extend the geometric theory of Lucia and Feng to include multivariable problems and integral path bifurcations; (2) to demonstrate that terrain methods provide a high degree of problem-solving reliability and efficiency when used to solve larger dimensional problems. Also, a simple illustration is given that clearly shows that, unlike continuation methods, terrain methods can solve problems that are disconnected in parameter space.

The fundamental geometric theory associated with terrain methods is briefly reviewed. Extensions of this theory to in-

Correspondence concerning this article should be addressed to A. Lucia.

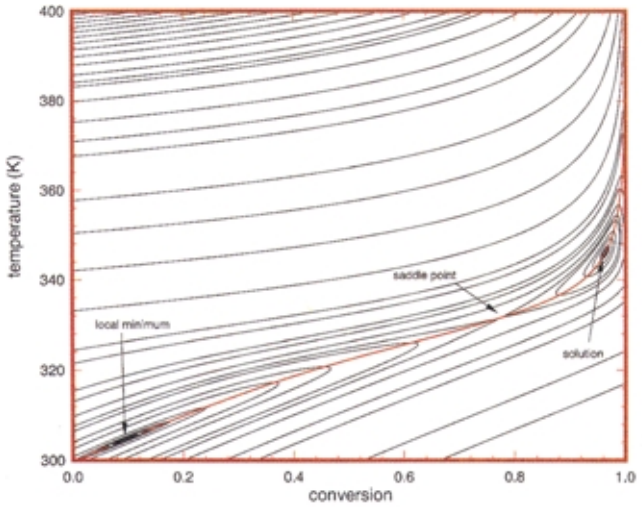

Figure 1. Connectedness of stationary points for a nonlinear CSTR.

tegral curve bifurcations are then presented along with geometric and computational illustrations for the 3-D steady-state Lorenz (1963) equations. The theoretical material is followed by numerical results for several multivariable chemical process examples, including a nonlinear CSTR, a retrograde flash calculation for a binary mixture, the task of finding all azeotropes for a heterogeneous ternary mixture, and a reactor model with solutions that lie on an isola. Finally, conclusions regarding the reliability, efficiency, and robustness of terrain methods are made.

Mathematical Background

Terrain methods are founded on two straightforward geometric concepts: (1) all stationary points are connected along (curved) valleys and ridges; (2) this connectedness is limited to neighboring stationary points.

Connectedness

Valleys, ridges, ledge, and so on of the least-squares landscape are specific integral curves of the gradient vector field and can be defined as a collection of solutions, say $\mathfrak{V}$, to a set of general nonlinearly, constrained optimization problems

$$\mathfrak{V} = \{\text{opt } g^T g \text{ such that } F^T F = L, \quad \text{for all} \quad L \in \mathfrak{L}\} \quad (1)$$

where F is a twice continuously differentiable vector function of the unknown variables Z, defined on R^n or C^n, $g = 2J^T F$, J is the Jacobian matrix of F, L is any given value (or level) of the least-squares objective function $F^T F$, and $\mathfrak{L}$ is some collection of level curves. Equation 1 forms the backbone for our terrain methodology and is interpreted as follows. For any given level curve, we find the point on L that corresponds to a local optimum in the gradient norm. The collection of these optima for all levels gives all (or part) of a valley, ridge, or ledge. We note here that it is not necessary to know the set of level curves $\mathfrak{L}$ *a priori*. $\mathfrak{L}$ is actually a computational byproduct of the terrain-following approach.

Multiplying $g^T g$ by one-quarter and the level constraint by one-half does not change the stationary points of the constrained optimization problems in Eq. 1, but allows us to conveniently write the associated necessary conditions for optimality in the form

$$F^T F - L = 0 \quad (2)$$

$$\mathfrak{K} g - \lambda g = 0 \quad (3)$$

where $\mathfrak{K}$ is the second derivative matrix of the least-squares function $\mathfrak{K} = J^T J + \Sigma F_i H_i$, and where H_i is the second derivative matrix of the component function F_i, and λ is a Lagrange multiplier associated with the level constraint. Note that Eq. 3 clearly shows that each point in $\mathfrak{V}$ (that is, point in a valley, on a ridge, and so on) represents a solution to a constrained eigenvalue-eigenvector problem and that the associated Lagrange multiplier λ is, in fact, the eigenvalue associated with the eigenvector g.

Limited Connectedness

Connectedness between stationary points is generally finite. That is, any given stationary point is usually connected to a small set of neighboring stationary points. In fact, Lucia and Feng (2002) present heuristic arguments and numerical examples that clearly support the idea that stationary points are really only connected to neighboring stationary points along specific eigendirections. These heuristics are based on the "attraction" that exists between neighboring stationary points and the resulting distortion of the level curves that give rise to valleys, ridges, as well as the general curvature of the least-squares landscape. For example, the direction of the attraction between the local minimum, saddle point, and solution shown in Figure 1 correspond to the eigenvectors associated with the smallest positive eigenvalue at either the local minimum or solution and the largest negative eigenvalue at the saddle point. Additional details of this will be presented in the section on Numerical Examples. Knowledge of this connectedness makes it possible to systematically move from one stationary point to another in an organized manner and determine when the terrain methodology has finished.

Basic Algorithmic Implementations

The foregoing ideas are readily implemented either in an algebraic or differential equations setting as described in Lucia and Feng (2002). In particular, Lucia and Feng provide considerable details for an algebraic implementation of a terrain method in which downhill movement takes place using an algebraic equation solver like a trust region method and uphill exploration is conducted using a predictor-corrector method. Predictor steps can be uphill Newton or approximate second-order Newton steps, while corrector steps, which return iterates to a valley or ridge if they have wandered too far, result from the intermittent solution of the optimization problem defined in Eq. 1 using successive quadratic programming. Rapid convergence to singular or saddle points is accomplished using quadratic acceleration. Both uphill and downhill movement results in the determination of a stationary point or a collision with a boundary of the feasible region. Because stationary points are really only connected to neigh-

boring stationary points along specific eigendirections, only a few eigenvalues and eigenvectors are needed to initiate movement to find the next stationary point. Moreover, this limited connectedness is easily implemented using linked-list concepts and can be used to determine when the algorithm has finished. See Appendix A for a brief description of the terrain methodology or Lucia and Feng (2002) for all of the associated details of this section.

Integral Curve Bifurcations

Integral curves can exhibit the same rich bifurcation structure as other types of nonlinear problems and are important because they can occur in chemical process engineering examples such as the task of finding all azeotropes.

Tangent, pitchfork, and other bifurcations in integral curves can occur at points where g^Tg is simultaneously a local minimum and a local maximum on some level curve of F^TF. These bifurcation points can be characterized by the singularity of the projected Hessian matrix of the Lagrangian function associated with Eq. 1, where the Hessian matrix of the Lagrangian function is given by

$$\mathfrak{M} = \mathfrak{H}^T\mathfrak{H} + \Sigma g_i G_i - \lambda \mathfrak{H} \tag{4}$$

where G_i is the second derivative matrix of the ith gradient function, g_i.

Tangent bifurcations in integral curves are similar to the more familiar bifurcations that occur in parametric S-shaped solution curves and characterized by the merging or splitting of a single minimum and single maximum on some level curve of F^TF. This type of bifurcation is easily identified in a numerical setting and corresponds to points where the integral curve becomes tangent to a level curve—something that is very easy to measure numerically. These bifurcations can also exhibit hysteresis. The relevance of tangent bifurcations in integral curves will be illustrated in the Numerical Results section using the problem of finding all azeotropes for isopropyl alcohol and water at 6.66612×10^4 Pa.

Pitchfork bifurcations can also occur in integral curves. This type of bifurcation is usually the result of competition between neighboring stationary points and corresponds to a point where two minima and one maximum in g^Tg occur simultaneously on some level curve of F^TF. Pitchfork bifurcations are also easily determined in a numerical setting since they correspond to a point where the projection of $\mathfrak{M}$ onto the tangent subspace of the level constraint has a zero eigenvalue.

Simple Illustrative Example

The Lorenz (1963) equations are a system of three first-order, ordinary differential equations, two of which contain quadratic nonlinearity, that have been studied extensively from a dynamical perspective to illustrate such things as chaotic dynamics, homoclinic explosions, Poincare return maps, and so on. A good overview and easily readable exposé on the Lorenz equations can be found in Sparrow (1982).

Here, we study the steady-state form of the Lorenz equations because:

(1) Like many chemical engineering examples, they contain a subset of linear equations;

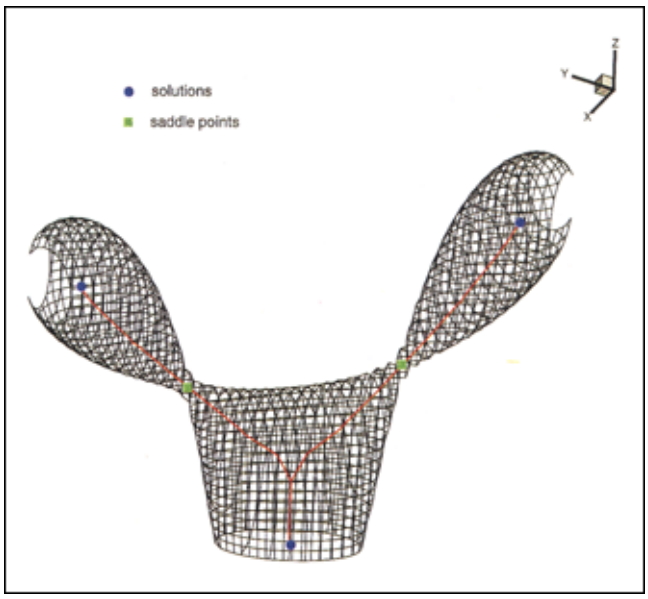

Figure 2. Level sets of F^TF for steady-state Lorenz equations.

(2) There is a pitchfork bifurcation in the integral curve connecting one of the solutions and a pair of saddle points; and

(3) They provide a clear 3-D illustration of the geometrical ideas that lie at the heart of the terrain methods.

The Lorenz equations are

$$dx/dt = \sigma(y - x) \tag{5}$$

$$dy/dt = rx - y - xz \tag{6}$$

$$dz/dt = xy - bz \tag{7}$$

where σ, r, and b are parameters. The parameter r is usually used as a bifurcation parameter and typical values for all parameters used in many dynamical studies are $\sigma = 10$, $b = 8/3$ and $r\epsilon(1, 30)$.

Figure 2 shows the level curves for the least-squares function for the 3-D vector function F, given by the steady-state form of Eqs. 5 through 7 for $\sigma = 10$, $b = 8/3$, and $r = 13$. The corresponding solutions and saddle points are given in Table 1. Before proceeding any further, it is important to explain how to interpret Figure 2. First, it is important to understand that *Figure 2 is a solid object*, where the outer cage-like surface represents the largest level of F^TF which, for the given parameter set, has a value of $F^TF = 150$. Thus, as one

Table 1. Solutions and Saddle Points for the Steady-State Lorenz Equations*

Stationary Point	Numerical Value
Solution	(0, 0, 0)
Saddle point	(2.49217, 2.89654, 6.963891)
Solution	$([b(r-1)]^{1/2}, [b(r-1)]^{1/2}, r-1)$
Saddle point	(−2.49217, −2.89654, 6.963891)
Solution	$(-[b(r-1)]^{1/2}, -[b(r-1)]^{1/2}, r-1)$

*$\sigma = 10$, $b = 8/3$, $r = 13$.

moves inward and orthogonal to the surface, the value of the level decreases. That is, each specific level encloses all lower levels. Second, the darker circular or elliptical regions of Figure 2 simply represent the intersection of the level surfaces with boundaries of the feasible region, which were set to $-7 \le x,\ y \le 7$ and $0 \le z \le 15$ for this illustration. Third, the two upper lobes of the figure are twisted about the torso. The lobe on the right comes out of the page, while the lobe on the left goes into the page. Finally, note that all solutions and saddle points are contained within this cage-like structure.

Unfortunately, it is difficult to see the smooth connectedness that exists between the solutions and saddle points in Figure 2, even though we have indicated this connectedness by the red integral curves generated by the terrain algorithm. It is equally difficult to see that these integral curves actually connect extrema in $g^T g$. Therefore, to provide a more convincing illustration of connectedness we show a 3-D relief map of the level curves of $F^T F$ in Figure 3. This relief map is simply a cut-away and clearly shows the connectedness of the stationary point of the steady-state Lorenz equations. Note that the integral curves connecting the solutions and saddle points pass directly through the collection of minima of $g^T g$ on the set of level curves shown in Figure 3. Also, note that Figures 2 and 3 both show that a pitchfork bifurcation in the integral curve occurs just a short distance from the solution at zero. The actual bifurcation point is $Z_b = (x, y, z) = (0, 0, 3.287)$ represents a point of transition between a pair of local minima and a local maximum in $g^T g$ and corresponds to a singularity of the projected Hessian matrix of the Lagrangian function defined in Eq. 4. The value of $F^T F$ at this bifurcation point is 19.207. This bifurcation is a pitchfork bifurcation because the extension of the vertical portion of the integral curve in the z-direction is a collection of maxima in $g^T g$ on the appropriate levels of $F^T F$.

Numerical Examples

In this section, we present four multivariable examples from chemical engineering to illustrate the reliability and efficiency of our terrain methodology. These examples include: (1) a nonlinear CSTR; (2) a retrograde flash calculation for a binary mixture; (3) the task of finding all azeotropes for a heterogeneous ternary mixture; (4) a nonadiabatic reactor with some solutions that lie on an isola. All calculations described in this section were performed in double precision arithmetic on a PC equipped with a Pentium III processor using a Lahey F77/90L-EM32 compiler. In all examples, the initial trust region radius R was 10. Convergence to a solution or stationary point was assumed when either $\|F\|$ or $\|g\| \le 10^{-8}$, quadratic acceleration was invoked when the parameter $\zeta \le 10^{-6}$ (see Eq. A2), and corrector steps were used whenever Eq. A4 was satisfied for $\theta \ge 5$.

Nonlinear Continuous Stirred Tank Reactor (CSTR)

Shacham (1985) gives a simple mathematical model for a nonlinear CSTR with lumped parameters. The equations for this reactor are the mass balance

$$F_1(x, T) = 120x - 75k(1 - x) = 0 \qquad (8)$$

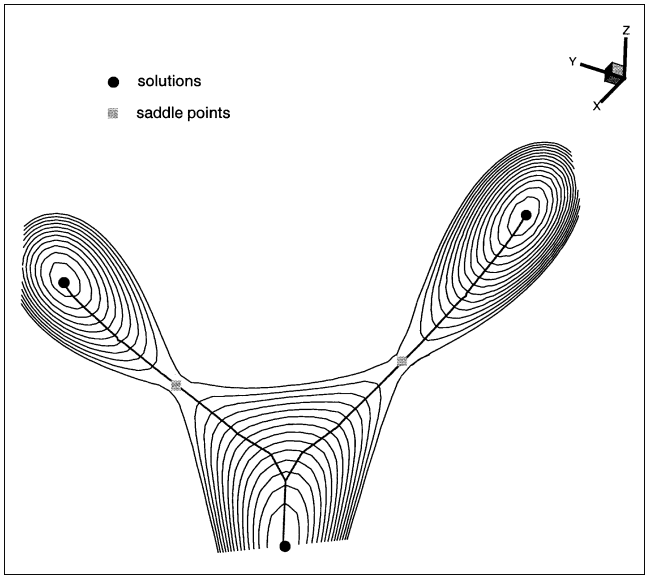

Figure 3. Relief map for the level curves for the Lorenz equations.

and the energy balance

$$F_2(x, T) = x(873 - T) - 11(T - 300) \qquad (9)$$

where x denotes the conversion, T is the reactor temperature in degrees Kelvin and k, the reaction rate constant, is given by $k = 0.12 \exp[12581(T - 298)/(298T)]$. The bounds on conversion and temperature for this example were $x \varepsilon [0, 1]$ and $T \varepsilon [298, 450]$.

There are three stationary points for this problem, a local minimum at low conversion and low temperature, a saddle point and a steady-state solution at high temperature and high conversion. These stationary points are given in Table 2. The corresponding eigenvalues and eigenvectors of the Hessian matrix of the least-squares function for all stationary points are listed in Table 3. Clearly, the negative eigenvalue at the saddle point indicates unstable reactor operation.

The numerical calculations for this example proceeded in the following way. From a starting point of $(x_0, T_0) = (0.6, 328.2)$, where $F_1 = -103.164$, $F_2 = 16.6800$, and $F^T F = 1.09210 \times 10^4$, our terrain algorithm found the local minimum that are shown in Table 5 in 20 iterations (or function and gradient calls) to an accuracy of $\|g\| = 1 \times 10^{-8}$. Of these 20 iterations, 16 were the downhill movement using a trust region strategy involving combined Newton and Cauchy steps and four were quadratic acceleration steps to the local minimum. At this point, the set of solutions is $\mathfrak{S} = \{s_1\}$, where $s_1 = (0.094594, 304.6772)$ and the set of connections is $\mathfrak{C} = \{c_1\}$, where $c_1 = (0.019607, 0.999807)$ is the eigenvector associated with the smallest eigenvalue of $\mathfrak{H}$ at the local minimum, $\lambda = 2.10713$.

Uphill movement from s_1 begins in the $+c_1$ direction. Eigen-perturbation gives the point $Z = (0.151915, 307.600)$ and terrain following finds the saddle point $s_2 = (0.774548, 331.5069)$ in 33 iterations to an accuracy of $\|g\| = 1 \times 10^{-8}$. 24 of these 33 iterations were predictor steps, one was a correc-

Table 2. Stationary Points for a Nonlinear CSTR Problem

Stationary Point	Numerical Value (x, T)
Local minimum	(0.094594, 304.6772)
Saddle point	(0.774548, 331.5069)
Solution	(0.963868, 346.1637)

tor step that returned iterates to the valley, and four were quadratic acceleration iterations used to converge to the saddle point to the accuracy shown. It is important to note that the corrector step required to solve the equality constrained optimization problem defined in Eq. 1 used 5 function and gradient calls (or iterations). Uphill movement from the local minimum to the saddle point required a total of 55 function and gradient calls. During the course of uphill exploration, the smallest eigenvalue and associated eigenvector at the local minimum smoothly transition to $\lambda = -92.7333$ and $c_2 = (-0.026170, -0.999657)$. Uphill movement in the $-c_1$ direction collides with the lower bound in reactor temperature in two iterations. Thus, $\mathcal{S} = \{s_1, s_2\}$ and $\mathcal{C} = \{c_2\}$ where $c_2 = (-0.026170, -0.999657)$.

Downhill movement from the saddle point s_2 takes place in the $-c_2$ direction, only because the saddle was approached from the $+c_2$ direction. Eigen-perturbation gives the point $Z = (0.883151, 335.655)$ and subsequent use of our trust region strategy finds the solution $s_3 = (0.963868, 346.1637)$ in 12 iterations to an accuracy of $\|F\| = 1 \times 10^{-8}$. At this point, $\mathcal{S} = \{s_1, s_2, s_3\}$, c_2 is deleted from the set of connections and $\mathcal{C} = \{c_3\}$ where $c_3 = (-0.004124, -0.999991)$. Uphill movement from the solution s_3 which takes place in the direction $-c_3$ because the solution was approached from the $+c_3$ direction, encounters an upper bound in conversion in 184 function and gradient calls from the point $Z = (0.964280, 346.264)$. Most of this effort was used in calculating corrector steps because the constrained optimization problems defined in Eq. 1 become extremely poorly scaled as iterates climb out of the valley toward the upper righthand side corner of the feasible region. Thus, line searching and between 10 and 20 iterations are required to solve the associated constrained optimization problems. In all, 273 function and gradient calls were required to find all stationary points for this example.

Retrograde Flash Calculations

Consider a fixed vapor flow rate-temperature (VT) flash for a binary mixture of carbon dioxide and n-butane where the vapor to feed ratio is $V/F = 0.90$ and the flash temperature is $T = 344.26$ K. See Figure 4 for an illustration. For

Table 3. Eigenvalues and Eigenvectors of $\mathcal{H}$ at All Stationary Points in Table 2

Stationary Point	Eigenvalues	Eigenvectors
Local minimum	2.10713	(0.019607, 0.999807)
	3.43486×10^5	$(-0.999807, 0.019607)$
Saddle point	-92.7333	$(-0.026170, -0.999657)$
	8.74274×10^5	$(-0.999657, 0.026170)$
Solution	98.27692	$(-0.004124, -0.999991)$
	1.13078×10^7	$(-0.999991, 0.004124)$

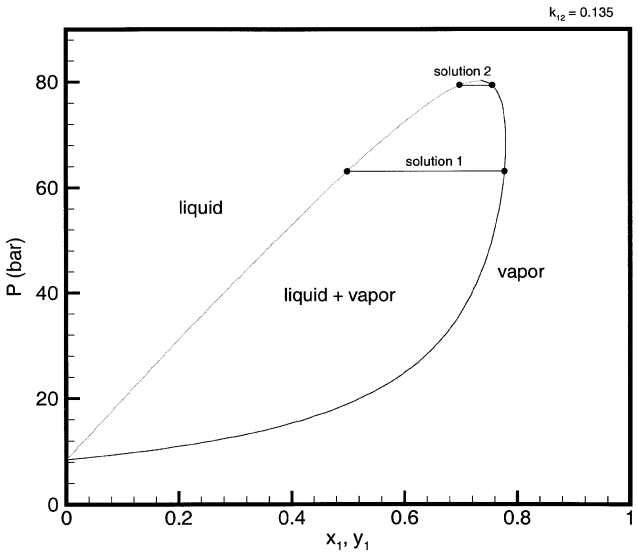

Figure 4. Isothermal VLE for CO_2-n-C_4H_{10} at 344.26 K using the SRK EOS.

the specified conditions, carbon dioxide is a supercritical component and this gives rise to the retrograde behavior of the mixture. The model equations are the n_c phase equilibrium equations

$$K_i l_i / (\Sigma l_j) - v_i / (\Sigma v_j) = 0 \tag{10}$$

the n_c component mass balance equations

$$f_i - l_i - v_i = 0 \tag{11}$$

and the specification equation

$$(\Sigma v_j) / (\Sigma f_j) - V/F = 0 \tag{12}$$

where f_i, l_i, and v_i are the feed, liquid and vapor molar flows, respectively. The component equilibrium ratios K_i are defined by the rule $K_i = \phi_i^L / \phi_i^V$, where ϕ_i^L and ϕ_i^V are the fugacity coefficients for the ith component in the liquid and vapor phases, respectively. The fugacity coefficients are generally nonlinear functions of temperature, pressure, and composition and defined through the compressibility factor using an equation of state (EOS). The Soave-Redlich-Kwong (SRK) EOS with classical mixing and combining rules with a binary interaction parameter $k_{12} = 0.135$ was used to model both phases. The SRK equation, classical mixing rules, and critical properties data we used can be found in Table 1.11 and Appendix D, respectively, in Walas (1985). The ten unknown variables in this problem are the complex-valued l_i's, v_i's, and the pressure p, and the bounds on the variables were $l_i, v_i \in [0, f_i]$ and $p \in [5. \times 10^6, 1. \times 10^7]$.

Table 4 gives the stationary points for one mole of feed containing 75 mol. % carbon dioxide and 25 mol. % n-butane for specifications of $V/F = 0.9$ and $T = 344.26$ K. In this table, the isolated stationary points are reported using the mole fraction of carbon dioxide in each phase and the pres-

Table 4. Stationary Points for a Retrograde Flash Problem

Stationary Point	Numerical Value (x_{CO_2}, y_{CO_2}, p)
Solution 1	(0.49944, 0.77784, 6.31084×10^6)
Saddle point	(0.54923, 0.77246, 6.93661×10^6)
Solution 2	(0.69735, 0.75585, 7.94380×10^6)

sure in units of Pa. All eigenvalues and eigenvectors of $\mathcal{H}$ at the stationary points in Table 4 are shown in Table 5.

From a starting point of $Z_0=(x_1,\ y_1,\ p)=(0.8,\ 0.5,\ 5.\times10^6)$, our complex domain trust region method finds the solution $s_1=(0.49944,\ 0.77784,\ 6.31084\times10^6)$ in 5 iterations. Thus, the set of solutions is set to $\mathcal{S}=\{s_1\}$. From Table 5, it is easily seen that the proper eigen-connection for locating the next stationary point using uphill movement is the eigenvector associated with the smallest eigenvalue $\lambda=6.2543\times10^{-5}$. Consequently, $c_1=(0.00069,\ -0.00064,\ -0.00668,\ 0.00536,\ 0.99996)$ and $\mathcal{C}=\{c_1\}$. Uphill movement in the $+c_1$ direction encounters a bound while uphill movement in the $-c_1$ direction finds the saddle point, $s_2=(0.54923,\ 0.77246,\ 6.93661\times10^6)$ in 35 iterations. Thus, $\mathcal{S}=\{s_1,\ s_2\}$, c_1 is deleted from the set of connections and $\mathcal{C}=\{c_2\}$, where $c_2=(-0.00105,\ -0.00105,\ -0.00946,\ 0.00946,\ -0.99991)$ is the eigenvector associated with the negative eigenvalue $\lambda=-0.9462\times10^{-5}$. Downhill movement from the saddle point in the $+c_2$ direction is not attempted since that is the approximate direction connecting s_1 and s_2. Downhill movement from the saddle in the direction $-c_2$ quickly locates the high pressure solution $s_3=(0.69735,\ 0.75585,\ 7.94380\times10^6)$ in 9 iterations using a complex domain trust region method. Thus, $\mathcal{S}=\{s_1,\ s_2,\ s_3\}$, c_2 is deleted from the set of connections and $\mathcal{C}=\{c_3\}$, where $c_3=(-0.00357,\ 0.00357,\ 0.03212,\ -0.03225,\ 0.99895)$ is the eigenvector associated with the smallest eigenvalue of $\mathcal{H}$ at the high-pressure solution. Since the direction $-c_3$ connects solutions s_2 and s_3, uphill movement from the high pressure solution proceeds in the $+c_3$ direction and encounters an upper bound in pressure in 10 iterations. Since c_3 is deleted from $\mathcal{C}$ and $\mathcal{C}$ is null, our terrain method terminates with $\mathcal{S}=\{s_1,\ s_2,\ s_3\}$. A total of 191 function and gradient calls and 3.41 s of computer time were required to solve this problem.

Finding All Azeotropes

Azeotropes can place limits on the separation of components in a mixture. Thus, knowledge of all azeotropes of homogeneous and heterogeneous multicomponent mixtures is useful in the analysis and design of separation and reactive separation processes. For a mixture of n_c components, the relevant model equations are the n_c phase equilibrium equations

$$K_i-1=\gamma_i f_i^0/\phi_i p-1=0 \tag{13}$$

and the summation equation

$$\Sigma x_i-1=0 \tag{14}$$

where γ_i, f_i^0, and ϕ_i are the liquid-phase activity coefficient, standard state fugacity, and the vapor phase fugacity coefficient, respectively, for the ith component. Since we implicitly use the condition that $x_i=y_i$, it follows that $\gamma_i=\gamma_i(x,T)$, $f_i^0(T)$ and $\phi_i=\phi_i(x,T)$, and, therefore, $K_i=K_i(x,T)$, where x denotes composition and T is the temperature. Thus, there are n_c+1 equations and n_c+1 unknowns (the x_i's and the azeotropic temperature T) in the homogeneous case. For so-called type I heterogeneous mixtures, the overall liquid composition satisfies Eqs. 13 and 14, but also splits into two or more liquid phases and, therefore, must satisfy additional equilibrium and summation equations.

Consider a mixture of *n*-propyl bromide (NPB), isopropyl alcohol (IPA), and water (H_2O) at 6.66612×10^4 Pa where it is assumed that the vapor phase behaves as an ideal gas and the liquid phase is nonideal. Using the Riedel vapor pressure constants and NonRandom Two Liquid (NRTL) binary interaction parameters given in Appendix B, this mixture exhibits two homogeneous and two heterogeneous azeotropes, as shown in Table 6. Note there are three binary and one minimum boiling ternary azeotropes. The form of the NRTL equation that we used is given on p. 183 in Walas (1985).

This example is interesting because there are a relatively large number of stationary points for the least-squares function despite the fact that there are only four equations. There are seven solutions—the three pure component vertices and the four azeotropes shown in Table 6. There are also a num-

Table 5. Eigenvalues and Eigenvectors of $\mathcal{H}$ at All Stationary Points in Table 4

Stationary Point	Eigenvalues	Eigenvectors
Solution 1	6.2543×10^{-5}	(0.00069, −0.00064, −0.00668,0.00536,0.99996)
	2.8114×10^{2}	(0.01974, −0.03203, −0.52978,0.84727,0.00818)
	9.2902×10^{3}	(−0.72131,0.64885, −0.22184, −0.09738,0.00004)
	1.8643×10^{4}	(0.37270,0.65601,0.54912,0.35947, −0.001906)
	8.4727×10^{4}	(0.58345,0.38421, −0.60709, −0.37869,0.00218)
Saddle point	-0.9462×10^{-5}	(−0.00105, −0.00105, −0.00946,0.00946, −0.99991)
	2.4181×10^{2}	(0.29021, −0.05988, −0.56715,0.82082,0.01323)
	8.7124×10^{3}	(0.69863, −0.69602,0.16174,0.36272,0.00028)
	1.5775×10^{4}	(−0.37610, −0.54153, −0.60604, −0.44498,0.00016)
	9.4447×10^{4}	(−0.60797, −0.46767,0.53368,0.35615, −0.00183)
Solution 2	5.3846×10^{-7}	(−0.00357,0.00357,0.03212, −0.03225,0.99895)
	3.6238	(0.07589, −0.07928, −0.69023,0.71375,0.04579)
	8.2036×10^{3}	(0.70568, −0.69951,0.09463, −0.06122,0.00001)
	1.4348×10^{4}	(−0.43998, −0.47105, −0.54681, −0.53435,0.00044)
	1.6161×10^{5}	(0.55015,0.53151, −0.46325, −0.44748,0.00052)

Table 6. Azeotropic Compositions/Temperatures for NPB/IPA/H_2O at 6.66612×10^4 Pa*

Composition	Temperature (K)	Type
(0.8008, 0., 0.1992)	318.630	heterogeneous
(0.8506, 0.1494, 0.)	320.316	homogeneous
(0., 0.6262, 0.3738)	345.628	homogeneous
(0.7842, 0.0403, 0.1755)	318.520	heterogeneous

*Component order is NPB = 1, IPA = 2, H_2O = 3.

ber of saddle points that lie between these solutions. Moreover, the valleys connecting these stationary points are quite curved and some contain tangent bifurcations.

To clearly illustrate that tangent bifurcations can occur in integral curves, consider the homogeneous binary mixture IPA and H_2O at 6.66612×10^4 Pa, which lies in a subspace of the equations defining all azeotropes for the ternary mixture. Figure 5 shows the path connecting the pure IPA vertex, the pure water vertex, the IPA-H_2O azeotropic solution, and the two saddle points between the pure component vertices and the binary azeotropic solution that was generated by our terrain method. Note that there is a pair of tangent bifurcations in the integral curves near the pure water vertex and hysteresis in the path, as indicated by the small set of level curves common to both branches of the path. The numerical values of the bifurcation points are $Z_b = (x_{\mathrm{IPA}},\ T) = (0.01126,$ 348.155 K) and $Z_b = (x_{\mathrm{IPA}}, T) = (0.0321,$ 351.487 K) at which the values of F^TF are 0.18086 and 0.14514, respectively. However, it is important to note that these tangent bifurcations do not affect the connectedness between the stationary points in this problem and our terrain method easily finds all stationary points irrespective of which stationary point is found first!

Figure 6 gives a complete summary of the numerical results of our terrain methodology applied to the problem of finding all azeotropes for the ternary heterogeneous mixture of NPB/IPA/H_2O at 6.66612×10^4 Pa. Note that 13 stationary points were located. More importantly, the integral curves that connect all azeotropes were correctly determined and all four azeotropes shown in Table 6 were successfully calculated using 472 function and gradient calls and 2.58 s of computer time.

Nonadiabatic CSTR with Isola

Consider the nonadiabatic CSTR studied by Balakotaiah and Luss (1981, 1983) and one mathematical model given by the mass and energy balance equations

$$F_1(x,T) = x(1+\theta k) - \theta k = 0 \tag{15}$$

$$F_2(x,T) = \left[\ \rho C_p(T-T_0) + (UA/F)(T_c - T)\right] / C_{A0} - x\Delta H_R = 0 \tag{16}$$

where x denotes the unknown conversion of species A, the residence time θ is given by $\theta = V/F$, V is the reactor volume, and F is the volumetric flow rate through the reactor. The rate constant k is defined by $k = k_0\exp(-E/RT)$ where k_0 is the pre-exponential factor, E is the activation energy, R is the gas constant, and T is the unknown temperature of the reactor. Also, ρ is the fluid density, C_p is the specific heat, U is an overall heat-transfer coefficient, A is the heat-transfer area, and ΔH_R is the heat of reaction. C_{A0} is the feed concentration for component A and T_0 and T_c represent the feed and cooling water temperatures, respectively. Note that the second term in the middle portion of Eq. 16 allows for both adiabatic and nonadiabatic operation of the reactor and it is nonadiabatic operation that gives rise to isola. The bounds on conversion and reactor temperature were $x\varepsilon[0,\ 1]$ and $T\varepsilon[298,\ 450]$. The data for this problem is given in Table 7.

The parametric behavior of conversion as a function of the Damköhler number $Da = \theta k_0$ is shown in Figure 7 and clearly indicates the presence of two disconnected solution branches

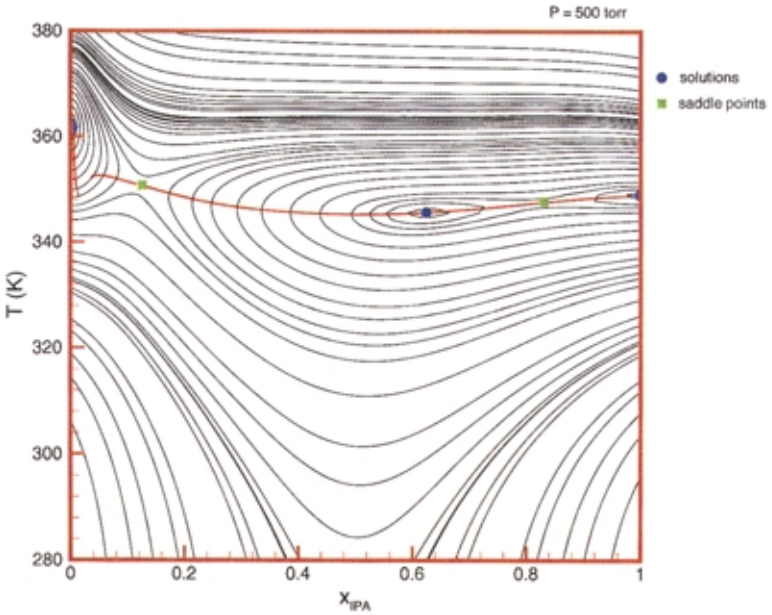

Figure 5. Terrain-following paths for the IPA-H_2O all azeotropes function.

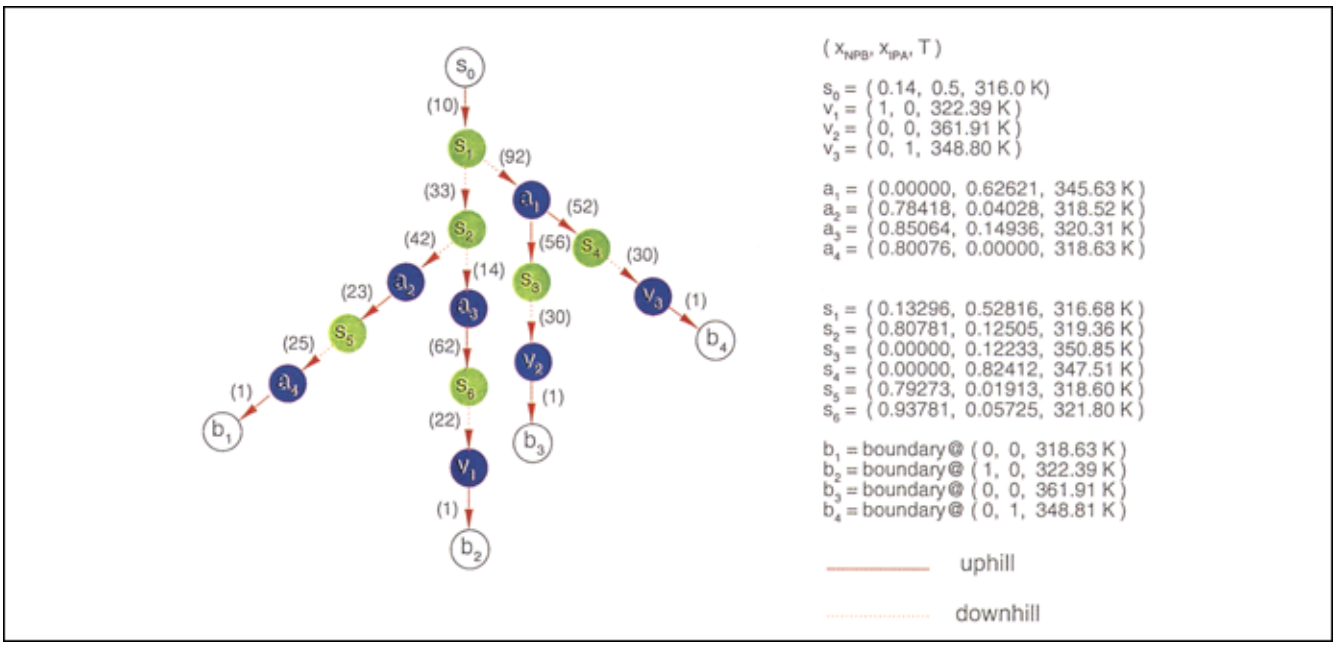

Figure 6. Computational tree for all azeotropes calculation.

—an isola and a monotonic solution curve. The isola exists for Da ε[0.011,0.077]. Solutions, singular points, principal eigenvalues, and normalized eigenvectors for $Da = 0.0497$ are given in Table 8 in the order in which they were determined by our terrain method from a starting point of $(x_0, T_0) = (0.20, 361.0)$ at which $F_1 = -3.10546$ and $F_2 = 1.049889 \times 10^5$ J/mol. This starting point is particularly poor because it is a combination of low conversion and high temperature. Figure 8 shows a complete summary of the numerical results for this example in the form of a computational tree. Several points are worth noting. First, only 104 total function and gradient calls were required to find all five stationary points shown in Table 8 from the given starting point. Uphill and downhill movement exactly follows the terrain methodology in that eigen-initiation corresponding to the smallest positive and largest negative eigenvalue, which correctly identifies the connectedness between solutions and saddle points. Moreover, the small associated computational overhead is quite encouraging. Secondly, our algorithm used no corrector steps throughout the computations because it correctly determined that the valley that connects the stationary points is essentially a straight line in this example. This is easily seen since the principal, normalized eigenvector shown in Table 8 remains unchanged from stationary point to stationary point. Finally, it is rather easy to see from Figure 7 that simple arc-length homotopy-continuation methods cannot find all solutions to this problem from a single starting point. To further support this last comment, we performed a continuation of the complex domain solutions that exist to the left and right of the isola in Figure 7. For example, for $Da = 0.0117710$, there are complex-valued solutions given by $(X, T) = (0.77378 +/- 0.03375\text{i}, 383.748 +/- 3.73976\text{i})$ and, as Da is decreased, the imaginary parts of the conversion and temperature increase in magnitude. Similarly, for $Da = 0.0771207$,

Table 7. Physical Properties and Other Specifications for Nonadiabatic CSTR

Quantity	Numerical Value
ρ (mol/m^3)	1×10^6
C_p (J/mol K)	4.18399
k_0 (s^{-1})	4.48×10^6
E (J/mol)	6.275386×10^4
ΔH_R (J/mol)	-1.763841×10^5
C_{A0} (mol/m^3)	3×10^3
T_0 (K)	298.
R (J/mol K)	8.313598
UA (J/s K)	2.24580×10^{11}
T_c (K)	298.

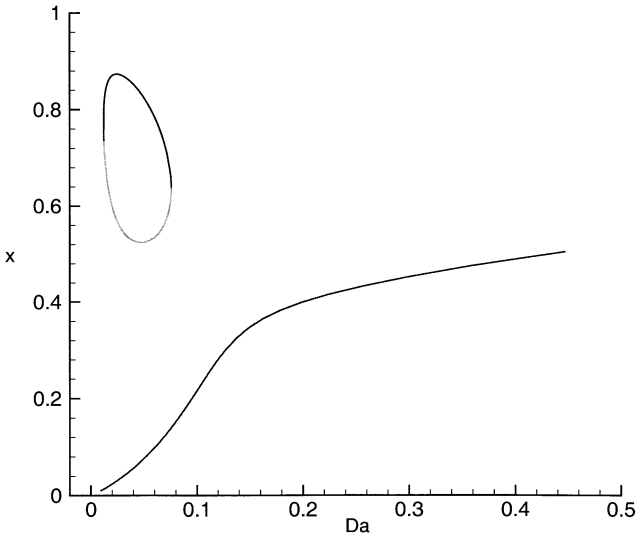

Figure 7. Nonadiabatic CSTR.

Table 8. Stationary Points and Eigeninformation for Nonadiabatic CSTR ($Da = 0.0497$)

Stationary Point	Numerical Value (x, T)	Eigenvalue	Eigenvector (x, T)
Saddle point	(0.71582, 354.709)	-1.584×10^{-4}	(0.01262, 0.99992)
Solution	(0.52459, 339.560)	6.070×10^{-5}	(0.01262, 0.99992)
Solution	(0.82659, 363.484)	6.534×10^{-4}	(0.01262, 0.99992)
Saddle point	(0.30564, 322.214)	-3.759×10^{-5}	(0.01262, 0.99992)
Solution	(0.07573, 303.999)	5.582×10^{-5}	(0.01262, 0.99992)

there are complex-valued solutions $(X, T) = (0.63446 +/- 0.04297i, 339.674 +/- 2.82290i)$, and, as Da increases, the imaginary parts of the solutions only grow in magnitude. These results clearly show that the complex-valued solutions are not parametrically connected to the unique real solutions at any values of Da and, thus, that continuation cannot find all solutions to this problem from a single starting point.

Conclusions

The terrain methodology of Lucia and Feng was used to solve a nontrivial set of multivariable, chemical process engineering examples, some of which exhibit integral curve bifurcations. The numerical results for the examples studied are summarized in Table 9. In our opinion, these results are extremely encouraging because they show that terrain methods make it possible to reliably find all solutions and singular points using only a modest number of total function and gradient evaluations, and a small amount of computer time. It is important for the reader to understand that the examples discussed at length in this article were chosen because they illustrate specific numerical characteristics (such as local minima, tangent and pitchfork bifurcations, isola, and so on) relevant to terrain methods and global equation solving. We have solved many other physical problems (such as distillation problems, phase split calculations, Gibbs energy minimization, potential energy minimization in molecular modeling), as well as many of the problems available through the Web (Shacham et al., 2002). These additional results, many of which will appear in future publication commitments of the principal author, only reinforce the general conclusion that terrain methods are a very reliable, robust, and efficient global equation-solving methodology.

We remark that the first example discussed in detail (that is, the nonlinear CSTR) is actually problem 31 from the Web library and is considered by Shacham et al. (2002) to be of a higher degree of difficulty. However, our terrain methodology has absolutely no difficulties whatsoever in finding the local minimum, saddle, and global minimum shown in Table

Table 9. A Summary of Numerical Results for the Terrain Methodology

Problem	Function & Gradient Calls	Computer Time (s)*
Nonlinear CSTR	273	0.06
Retrograde flash	191	3.41
All azeotropes	472	2.58
Reactor with isola	104	0.05

*Time on a Pentium III computer using a Lahey F77/90L-EM32 compiler.

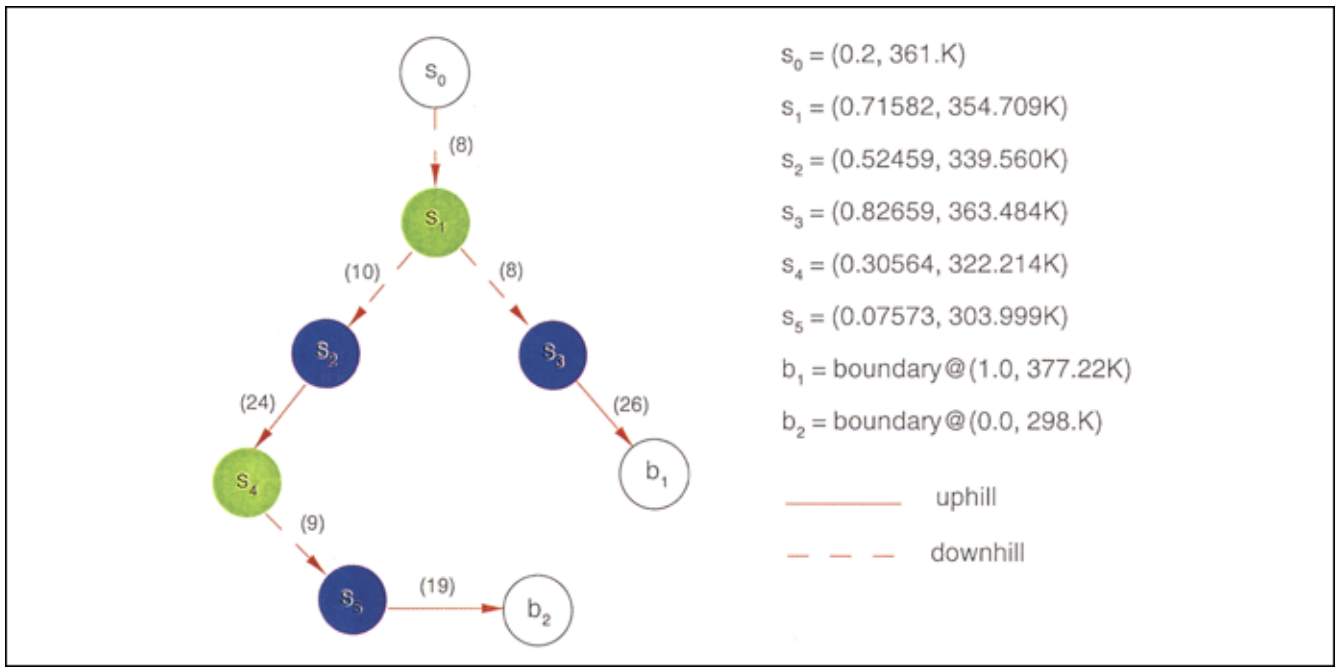

Figure 8. Computational tree for nonadiabatic CSTR.

Table 10. A Summary of Numerical Results for Problems from Polymath Library*

Problem†	Stationary Points	Function & Gradient Calls‡	Computer Time (s)‡
31	5	190	0.06
32	2	71	0.01
41	3	55	0.01
44	5	402	0.06
49	9	461	0.11
53	1	19	0.01

*www.polymath-software.com/library.
†Problem No. from Web library.
‡Averaged over all starting points given in Web library.

2. In particular, for the four initial guesses given in Shacham et al., $(x_0, T_0) = \{(1, 400), (0, 300), (0.5, 320), (0, 350)\}$, our terrain method finds all three stationary points in 213, 183, 188, and 178 total function and gradient evaluations, respectively. Shacham et al. (2002) state that "most methods won't converge from initial guesses 2 and 3 because of singularities." However, it is precisely because terrain methods correctly handle, and, in fact, exploit, singularities that our methodology easily finds all stationary points. All stationary points to other problems from this Web site are also easily found by our terrain methodology. Some of these numerical results are summarized in Table 10.

Finally, it is important to comment on the numerical characteristics of the optimization problems used to define valleys and ridges. The nonlinear optimization problems defined in Eq. 1 can be quite difficult to solve because they are inherently poorly scaled. We have tried a number of readily available, off-the-shelf optimization codes, including the successive quadratic programming (SQP) routine VF13AD from the Harwell Subroutine Library, and found them to be unsatisfactory. Sometimes, these codes converge and sometimes they do not converge. However, they often show quite erratic behavior during early iterations, which is a clear sign of potential numerical difficulties, and typically take 50 or more function and gradient calls when they do converge. We presently use our own in-house SQP methodology (Lucia and Xu, 1990) for solving the problems defined by Eq. 1, which always converges smoothly and usually in less than 20 function and gradient calls.

Acknowledgment

The authors would like to thank the National Science Foundation for support of this work under Grant No. CTS-0113091.

Literature Cited

Balakotaiah, V., and D. Luss, "Analysis of Multiplicity Patterns of a CSTR," *Chem. Eng. Commun.*, **13**, 111 (1981).
Balakotaiah, V., and D. Luss, "Multiplicity Features of Reacting Systems—Dependence of the Steady-States of a CSTR on the Residence Time," *Chem. Eng. Sci.*, **38**, 1709 (1983).
Lorenz, E. N., "Deterministic Nonperiodic Flows," *J. Atmos. Sci.*, **20**, 130 (1963).
Lucia, A., "Complex Domain Chemical Process Simulation in Theory and in Practice," *I & EC Res*., **39**, 1713 (2000).
Lucia, A., and Y. Feng, "Global Terrain Methods," *Comput*. & *Chem*. *Eng*., **26**, 529 (2002).
Lucia, A., and J. Xu, "Chemical Process Optimization Using Newton-like Methods," *Comput. Chem. Eng.*, **14**, 119 (1990).
Shacham, M., "Comparing Software for the Solution of Systems of Nonlinear Algebraic Equations Arising in Chemical Engineering," *Comput. & Chem. Eng.*, **9**, 103 (1985).
Shacham, M., N. Brauner, and M. B. Cutlip, "A Web-Based Library for Testing Performance of Numerical Software for Solving Nonlinear Algebraic Equations," *Comput*. & *Chem*. *Eng*., **26**, 547 (2002).
Sparrow, C., *The Lorenz Equations: Bifurcations, Chaos, and Strange Attractors*, Springer-Verlag, New York (1982).
Walas, S. M., *Phase Equilibria in Chemical Engineering*, Butterworth, Stoneham, MA (1985).

Appendix A: A Brief Description of the Terrain Methodology

Terrain methods assume that the stationary points of any objective function are connected along valleys and ridges. Successful movement from one stationary point to another proceeds by following an appropriate valley or ridge in the landscape of the objective function, either uphill or downhill, using predictor-corrector calculations and equation solving, respectively. Valleys and ridges are defined using principal eigenvalues and eigenvectors and termination occurs once all of the important eigen-directions have been exhausted.

Downhill movement is calculated using a trust region strategy in conjunction with a quadratic acceleration technique. Downhill steps are defined by

$$\Delta = \beta\Delta_N + (\beta - 1)g \tag{A1}$$

where Δ_N and g are the Newton and gradient directions, respectively, and where $\beta\varepsilon[0,1]$ is determined from the following simple rules. If $\|\Delta_N\| \le R$, then $\beta = 1$, where R is the trust region radius. If, on the other hand, $\|\Delta_N\| > R$ and $\|g\| \ge R$, then $\beta = 0$. Otherwise, β is the unique value in Eq. A1 on [0,1] that satisfies $\|\Delta\| = \mathrm{R}$. Furthermore, if during downhill movement, $\|g\|/\|\Delta_N\| \le \zeta$, then the equation

$$\Delta = -\,\mathcal{H}^{-1}g \tag{A2}$$

is used to calculate the step Δ, where $\mathcal{H}$ is the second derivative matrix of the least-squares function, $\mathcal{H} = J^TJ + \Sigma F_iH_i$, and where H_i is the second derivative matrix of the component function F_i. Note that Eq. A2 defines a second-order Newton (or quadratic acceleration) step. Downhill movement is terminated when either $\|F\| \le \epsilon$ or $\|g\| \le \epsilon$, where ϵ is a convergence tolerance and can result in convergence to either a minimum or saddle point of F^TF. Except for the initial starting point, which is arbitrary, downhill movement is always initiated in the eigen-direction associated with the largest negative eigenvalues of the Hessian matrix of the objective function.

Uphill movement is always initiated in the eigen-direction associated with the smallest positive eigenvalues of the Hessian matrix of the objective function followed by predictor-corrector calculations to get "near" a saddle (or singular) point. Predictor steps are simply controlled uphill Newton steps

$$\Delta_N = \alpha J^{-1}F \tag{A3}$$

where the step size $\alpha\varepsilon(0,1]$. These uphill Newton steps tend to follow valleys reasonably well, but do drift some. Therefore, corrector steps are used intermittently to return iterates

to the current valley and are invoked when

$$\theta = 57.295\arccos\left[\left(\Delta_N^T c\right)/\left(\|\Delta_N\|\,\|c\|\right)\right] \geq \Theta \qquad \text{(A4)}$$

where c is the current estimate of the eigenvector associated with the smallest positive eigenvalue of $\mathcal{H}$ and Θ is 5°. Corrector steps are defined by solving the optimality conditions associated with the definition of valleys and ridges method given by Eq. 1

$$F^T F - L = 0 \qquad \text{(A5)}$$

$$\mathcal{H}g - \lambda g = 0 \qquad \text{(A6)}$$

where L is the value of the level set at the last predictor iterate. There are many ways of solving Eqs. A5 and A6. We use the Successive Quadratic Programming (SQP) method given in Lucia and Xu (1990). Note that the solution to Eqs. A5 and A6 gives new estimates of both the eigenvalue and eigenvector of $\mathcal{H}$ that define the current valley. Uphill movement using predictor-corrector calculations continues until the objective function goes through a maximum value, or successive uphill Newton steps reverse direction or $\|g\|/\|\Delta_N\| \leq \zeta$. Once near a saddle point, quadratic acceleration (Eq. A2) is used to converge to the saddle (or next stationary point).

It is important for the reader to understand that the initial starting point does not matter. Irrespective of the stationary point found first (that is, saddle point, local minimum, or global minimum), movement from one stationary point to the next or to a boundary of the feasible region follows this same logic and the same set of stationary points will be located.

Termination of the terrain methodology is based on the concept of limited connectedness or that stationary points are really only connected to neighboring stationary points along specific eigendirections. We assume that the number of important connections between stationary points along valleys and ridges is limited to four or less and is related to dominant geometric distortions (that is, $+/-$ the smallest positive eigendirection and $+/-$ the most negative eigendirection) caused by the strongest "attractions" between neighboring stationary points. This makes it possible to dynamically catalog connections in a set $\mathcal{C}$ and to conclude that all connections have been explored for all stationary points when $\mathcal{C}$ is empty.

Appendix B. Physical Properties Data for NPB/IPA/H_2O

The Riedel vapor pressure equation for component j is given by

$$\ln P_j^{\text{sat}} = C_{1j} + C_{2j}/\left(T + C_{3j}\right) + C_{4j}T + C_{5j}\ln T + C_{6j}T^{C7j} \qquad \text{(B1)}$$

where T is in K and is P_j^{sat} in Pa. See Table B1 for Riedel constants for NPB/IPA/H_2O).

The NRTL binary interaction parameters $\tau_{i,j}$ and $G_{i,j}$ are defined by

$$\tau_{i,j} = a_{i,j} + b_{i,j}/RT \qquad \text{(B2)}$$

$$G_{i,j} = \exp(\alpha_{i,j}\tau_{i,j}) \qquad \text{(B3)}$$

where T is K.

In this work all $a_{i,j} = 0$, all $\alpha_{i,j} = 0.3$ and the $b_{i,j}$'s were generated using the UNIFAC group contribution method available in Apsen Plus version 10.2. The values of the b_{ij}s are given in Table B2.

Table B1. Riedel Constants for NPB/IPA/H_2O

Constant	NPB	IPA	H_2O
C_1	56.1709	82.9941	−31.3974
C_2	−5395.438	−7718.572	−2046.366
C_3	0.0	0.0	−75.40224
C_4	0.0	0.0	−0.01205428
C_5	−4.90746	−8.50595	9.165751
C_6	1.462525×10^{-15}	1.933595×10^{-15}	4.879195×10^{-16}
C_7	6.0	6.0	6.0

Table B2. NRTL Binary Interaction Parameters for NPB/IPA/H_2O

$b_{i,j}$ (K)	NPB	IPA	H_2O
NPB	0	737.657	763.178
IPA	−2.443	0	13.035
H_2O	2042.147	725.646	0

Manuscript received Sept. 11, 2002, and revision received Mar. 20, 2003.