A Dual Extremum Principle in Thermodynamics

Alexander Mitsos and Paul I. Barton
Dept. of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139

DOI 10.1002/aic.11230
Published online June 22, 2007 in Wiley InterScience (www.interscience.wiley.com).

Phase equilibria of multicomponent mixtures are considered and a reinterpretation of the Gibbs tangent plane stability criterion is proposed via Lagrangian duality. The starting point is the natural primal problem of minimizing the Gibbs free energy subject to material balance. The stable phase split is the solution of the corresponding dual problem, providing a necessary and sufficient dual extremum principle. Only in the absence of duality gap is the physical phase split also the solution of the primal problem. The only requirements are continuity of the Gibbs free energy and the trivial requirement that each species is present in the overall composition. The number of phases is permitted to be infinite, and does not need to be known a priori. No assumption is made on the presence of all species in all phases. Case studies are presented based on the NRTL and UNIQUAC activity coefficient model. © 2007 American Institute of Chemical Engineers *AIChE J,* 53: 2131–2147, 2007
Keywords: phase equilibrium, Gibbs tangent plane, duality, Lagrangian, extremum principle

Introduction

Extremum principles are fundamental concepts in thermodynamics and offer one of the alternative formulations of the second law of thermodynamics. The general statement of a (primal) extremum principle is that a state is stable if and only if among all admissible states it is an extremum (minimum or maximum) of an appropriately selected physical property. The set of admissible states is limited by considerations such as material balances. Which physical property is maximized or minimized depends on the constraints considered. For instance, at constant volume and internal energy, entropy tends to a maximum. At constant pressure and temperature, Gibbs free energy tends to a minimum. Throughout this article we consider the latter case.

The two simplest classes of chemical equilibria are reaction equilibrium in a single phase and phase equilibrium without reactions. The latter case is considered throughout this article for multicomponent mixtures ($n > 1$), where n is the number of species. Given an overall composition $\mathbf{x}^0$ (mole fractions) we are interested in obtaining a stable state comprising one or more phases at a fixed temperature T and pressure P. Calculating phase and reaction equilibria is a very challenging mathematical task and in state-of-the-art algorithms multiple nonconvex optimization problems are solved to guaranteed global optimality. An established stability criterion is construction of the Gibbs tangent plane,[1] described extensively later on. A state comprising one or many phases with equal chemical potential is stable if and only if its Gibbs tangent plane lies below the Gibbs surface for all mole fractions.

This article presents an alternative interpretation of the Gibbs tangent plane based on Lagrangian duality. Duality is a fundamental mathematical concept used in a variety of sub-

This article includes Supplementary Material available from the authors upon request or via the internet at http://www.interscience.wiley.com/jpages/0001–1541/suppmat/.

Correspondence concerning this article should be addressed to P. I. Barton at amitsos@alum.mit.edu, pib@mit.edu.

disciplines such as topology, functional analysis, algebra, and mathematical programming. Often it is simpler to prove properties of a primal problem by considering its dual problem. Moreover, in mathematical programming, duality is used extensively in algorithms. In convex programs satisfying a constraint qualification, the optimal solution values of the primal and dual programs are equal, while in nonconvex programs a duality gap typically exists. In this latter case, the optimal solution value of the dual is a strict lower bound to the optimal solution value of the primal.

We first discuss the geometric interpretation of duality in mathematical programming based on the treatment by Bertsekas[2] and present some mathematical background necessary for our proofs. Subsequently, we provide some thermodynamic background focusing on necessary and sufficient stability criteria for phase equilibrium and in particular on the Gibbs tangent plane. We present a more direct proof of this stability criterion than Baker et al.[1] to facilitate the transition to our main result. The main contributions of this article are the reinterpretation of the Gibbs tangent plane via Lagrangian duality and the proposition of a dual extremum principle: a state is stable if and only if it is a solution of the dual problem. Furthermore, our treatment provides a necessary and sufficient phase stability criterion that does not require any differentiability assumptions for the Gibbs free energy surface. We do not require a finite number of phases, nor that all species are present in all phases. Case studies are presented based on the NRTL and UNIQUAC activity coefficient models. Finally, we discuss potential generalizations and algorithmic implications of our results.

Mathematical Background

Convex hull

Given m points $\mathbf{x}^j \in \mathrm{R}^n$, their convex combinations are given by $\Sigma_{j=1}^m \kappa^j \mathbf{x}^j$ for scalars $\kappa^j \in [0,1]$ satisfying $\Sigma_{j=1}^m \kappa^j = 1$. By definition a set is convex if the convex combinations of any pair of its elements is also an element of the set. Consider any $D \subset \mathrm{R}^n$. The *convex hull* of D, denoted conv(D), is the set of all convex combinations of elements of D and it follows that it is the smallest convex superset of D. The following theorem, known as Caratheodory's theorem, states that any element of conv(D) can be represented as a convex combination of $n + 1$ elements of D.[2] Case studies are presented based on the NRTL and UNIQUAC activity coefficient model.

Theorem 1 (Caratheodory's Theorem). *Let $D \subset \mathrm{R}^n$ be a nonempty set and* conv(D) *its convex hull. Consider $\bar{\mathbf{x}} \in$* conv(D). *There exist $\kappa^j \in [0,1]$ and $\mathbf{x}^j \in D$, such that*

$$\bar{\mathbf{x}} = \sum_{j=1}^{n+1} \kappa^j \mathbf{x}^j \quad \textit{and} \quad \sum_{j=1}^{n+1} \kappa^j = 1.$$

Note that for the trivial case that D contains fewer than $n + 1$ elements in the above formulation some of the $\mathbf{x}^j$ will coincide.

Duality in mathematical programming

Duality in mathematical programming is a well established and widely used concept.[2] Besides the theoretical insight provided by duality, there exist efficient duality-based algorithms, e.g., in linear programming.[3] Here we briefly review basic concepts of duality with the focus on the geometric interpretation. Our presentation is based on the treatment by Bertsekas[2] with the major difference that we consider equality constrained problems.

Algebraic Treatment. The starting point is the *primal problem*

$$\begin{aligned} f^* = \inf_{\mathbf{x} \in X} & f(\mathbf{x}) \\ \mathbf{g}(\mathbf{x}) &= \mathbf{0}, \end{aligned} \tag{1}$$

with the host set $X \subset \mathrm{R}^n$, the objective function f: $X \to \mathrm{R}$ and the constraints $\mathbf{g}$: $X \to \mathrm{R}^m$. We assume that the feasible set $X_{\mathrm{feas}} = \{\mathbf{x} \in X: \mathbf{g}(\mathbf{x}) = \mathbf{0}\}$ is nonempty and that the optimal solution value f^* is bounded. Note that since we do not make any assumptions on the host set X the isolation of constraints $\mathbf{g}$ to be dualized is arbitrary.

The *Lagrangian function* L: $X \times \mathrm{R}^m \to \mathrm{R}$ is defined as

$$L(\mathbf{x}, \boldsymbol{\lambda}) = f(\mathbf{x}) + \sum_{j=1}^{m} \lambda_j g_j(\mathbf{x}),$$

where $\boldsymbol{\lambda} \in \mathrm{R}^m$ are called (duality) multipliers. The *dual function* is defined as

$$q(\boldsymbol{\lambda}) = \inf_{\mathbf{x} \in X} L(\mathbf{x}, \boldsymbol{\lambda})$$

and the *dual problem* as

$$q* = \sup_{\boldsymbol{\lambda} \in \mathrm{R}^m} q(\boldsymbol{\lambda}). \tag{2}$$

Note that because we consider equality constraints, there are no bounds on the multipliers $\boldsymbol{\lambda}$. The weak duality theorem relates the optimal objective value of the primal and dual problems.

Theorem 2 (Weak Duality Theorem). *The optimal solution value of the dual problem is a lower bound to the optimal solution value of the primal problem, i.e., $q^* \leq f^*$.*

If $q^* = f^*$ we say that there is no duality gap and that strong duality holds. A sufficient condition for strong duality is convexity of the primal problem along with a constraint qualification, e.g., a convex objective function and linear equality constraints. For nonconvex problems, typically $q^* < f^*$, in which case there is a duality gap.

Geometric Interpretation. An insightful interpretation of duality is the geometric interpretation in the space of constraint-cost pairs ($\mathbf{g}(\mathbf{x})$, $f(\mathbf{x})$).[2] In Figure 1 the horizontal axis ($w = 0$) is used for the constraint values and the vertical axis ($\mathbf{z} = \mathbf{0}$) for the objective value. The patterned area corresponds to the set of all possible constraint-cost pairs

$$S = \left\{ \left(\mathbf{g}(\mathbf{x}), f(\mathbf{x})\right) \middle| \, \mathbf{x} \in X \right\}.$$

The only feasible pairs are the ones that satisfy $\mathbf{g}(\mathbf{x}) = \mathbf{0}$ and therefore the primal problem (1) can be interpreted as the minimal common point of the vertical axis $\mathbf{z} = \mathbf{0}$ with the set S. Note that by our assumption of feasibility at least one common point exists.

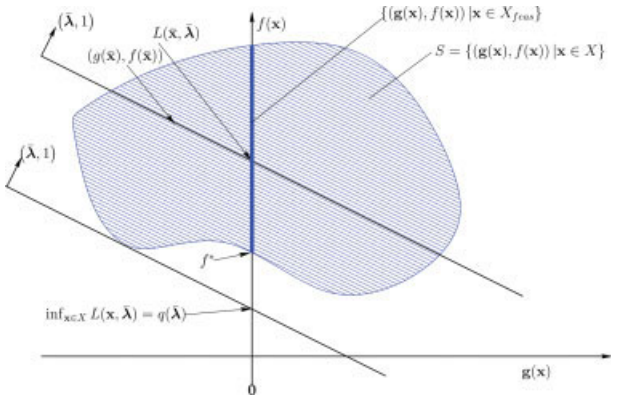

Figure 1. Geometrical interpretation of the Lagrangian and dual problems.

[Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

The geometric interpretation of the dual problem is somewhat more elaborate and requires some basic concepts from linear algebra. Let $\bar{\mathbf{z}} \in \mathrm{R}^m$ and $\bar{w} \in \mathrm{R}$. A nonvertical hyperplane H in R^{m+1} that passes through the point $(\bar{\mathbf{z}}, \bar{w})$ is described by a (scaled) normal vector $(\boldsymbol{\lambda}, 1)$ and a linear equation

$$H = \left\{(\mathbf{z}, w) | \ \mathbf{z} \in \mathrm{R}^m, \ w \in \mathrm{R}, \ \boldsymbol{\lambda}^{\mathrm{T}}\mathbf{z} + w = \boldsymbol{\lambda}^{\mathrm{T}}\bar{\mathbf{z}} + \bar{w}\right\}.$$

Associated with the hyperplane is the *positive halfspace*

$$H^+ = \left\{(\mathbf{z}, w) | \ \mathbf{z} \in \mathrm{R}^m, \ w \in \mathrm{R}, \ \boldsymbol{\lambda}^{\mathrm{T}}\mathbf{z} + w \geq \boldsymbol{\lambda}^{\mathrm{T}}\bar{\mathbf{z}} + \bar{w}\right\}.$$

When a set is contained in the positive halfspace of a hyperplane, this hyperplane is called a *supporting hyperplane* of the set.

As is shown in the following proposition, the dual problem is equivalent to finding a supporting hyperplane of S which among all supporting hyperplanes has the highest intercept with the vertical axis. This equivalence motivates the use of $\boldsymbol{\lambda}$ to form the normal vector.

Proposition 3. *The solution value q^* of the dual problem (2) is equal to the highest intercept of the supporting hyperplanes of S with the vertical axis. For each such hyperplane with (scaled) normal vector $(\boldsymbol{\lambda}^*, 1)$ the vector $\boldsymbol{\lambda}^* \in \mathrm{R}^m$ is a solution point of the dual.*

Having established geometric interpretations of the primal and dual problems, Figure 2 shows two important cases. On the left hand side Figure 2 a case is shown without duality gap $f^* = q^*$. The only common point of the maximum intercept supporting hyperplane with the set S is $(\mathbf{0}, f^*)$. On the right hand side of Figure 2 a case is shown with duality gap $f^* > q^*$. The maximum intercept supporting hyperplane has two common points with the set S none of which is $(\mathbf{0}, f^*)$.

As mentioned above, convexity plus a constraint qualification is a sufficient condition for strong duality. In the geometric interpretation it suffices to consider the extension of S upwards

$$\hat{S} = \left\{\mathbf{z} \in \mathrm{R}^m, w \in \mathrm{R} | \ \mathbf{z} = \mathbf{g}(\bar{\mathbf{x}}) \ \text{ and } \ w \geq f(\bar{\mathbf{x}}) \text{ for some } \bar{\mathbf{x}} \in X\right\}.$$

Note that in Figure 2 $\hat{S}$ is convex on the left hand side and nonconvex on the right hand side. In general, for instances where $\hat{S}$ is convex under a constraint qualification, the maximal intercept and the minimal common point coincide and strong duality holds.

Thermodynamic Background

Stability

A consequence of the second law of thermodynamics is that at constant temperature and pressure, a state is stable if and only if it minimizes the Gibbs free energy:

Theorem 4. *At a given temperature T, pressure P and overall number of moles n_i^0 for the species $i = 1, \ldots, n$, a state described by a collection of phases with index set J, with compositions $\mathbf{x}^j$ and with total mole numbers in each phase n_t^j such that $\sum_{j \in J} n_t^j x_t^j = n_i^0$ for all $i = 1, \ldots, n$, is stable if and only if among all possible states it has the lowest Gibbs free energy, i.e.,*

$$\sum_{j \in J} n_t^j G(\mathbf{x}^j) \leq \sum_{k \in K} n_t^k G(\mathbf{x}^k), \quad \forall (K, n_t^k, \mathbf{x}^k) : \sum_{k \in K} n_t^k x_i^k = n_i^0 \quad \forall i = 1, \ldots, n, \qquad (3)$$

where G is the intensive Gibbs free energy.

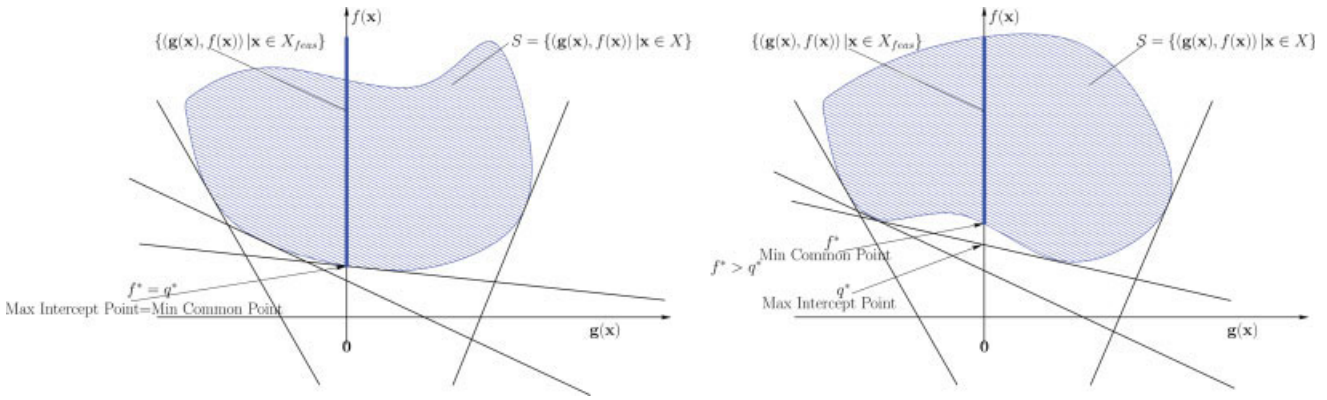

Figure 2. Illustration of primal (minimum common point) and dual (maximum intercept point) problems without (left) and with (right) duality gap.

[Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

Theorem 4 is quite theoretical and lacks a description of a practical method to establish stability. Moreover, Theorem 4 by itself does not indicate how many phases are possible.

Chemical potential and partial derivatives

A traditional stability criterion in phase equilibrium is based on chemical potentials. The chemical potential μ_i^j of species i in phase j is defined as the partial derivative of the extensive Gibbs free energy with respect to the mole number n_i^j of species i in phase j at constant mole numbers n_k^j for $k \in \{\{1,\ldots,n\}: k \neq i\}$. Equality of chemical potentials in all phases

$$\mu_i^j = \mu_i^{j^*}, \qquad \forall i = 1,\ldots,n, \quad j \in J,$$

where $j^* \in J$ is some reference phase, is a necessary condition for stability. It is well known, e.g. Ref. 1, that this is merely a necessary stability criterion, not a sufficient one.

To see the necessity note that if the above equality is violated, an infinitesimally small exchange of a species from a phase with higher potential to a phase with lower potential would lead to a decrease of Gibbs free energy. Indeed, without loss of generality assume for species $\bar{i}$ in phase $\bar{j}$ we have $\mu_{\bar{i}}^{\bar{j}} > \mu_{\bar{i}}^{j^*}$. Perturb now the state by transferring $\varepsilon > 0$ moles of species $\bar{i}$ from phase $\bar{j}$ to j^*. Assuming differentiability, for sufficiently small ε the first order term

$$\varepsilon\left(-G_{n_{\bar{i}}^{\bar{j}}} + G_{n_{\bar{i}}^{j^*}}\right) = \varepsilon\left(\mu_{\bar{i}}^{j^*} - \mu_{\bar{i}}^{\bar{j}}\right) < 0$$

dominates and the change in Gibbs free energy is negative, $\delta G < 0$.

In the remainder of this article we eliminate mole fraction x_n and consider partial derivatives of the Gibbs free energy at constant temperature and pressure with respect to the mole fraction x_j^i at constant mole fractions x_k^i for $k \in \{\{1,\ldots,n-1\}: k \neq i\}$. To simplify notation we denote this partial derivative G_{x_i}. The following well-known relations from multicomponent thermodynamics, e.g. equations (9–53) and (9–55) in Ref. 4, relate these partial derivatives with the chemical potentials

$$\mu_i^j(\mathbf{x}^j) = G(\mathbf{x}^j) + G_{x_i}(\mathbf{x}^j) - \sum_{k=1}^{n-1} x_k G_{x_k}(\mathbf{x}^j), \quad \forall i = 1,\ldots,n-1 \tag{4}$$

$$\mu_n^j(\mathbf{x}^j) = G(\mathbf{x}^j) - \sum_{k=1}^{n-1} x_k G_{x_k}(\mathbf{x}^j) \tag{5}$$

$$\mu_i^j(\mathbf{x}^j) = \mu_n^j + G_{x_i}(\mathbf{x}^j), \quad \forall i = 1,\ldots,n-1. \tag{6}$$

Gibbs phase rule

At constant temperature and pressure the well-known Gibbs phase rule states that the number of possible phases cannot exceed $n + 2$, where n is the number of species. This is true under linear independence of the Gibbs-Duhem equations.[5] Noll[6] shows that the Gibbs phase rule result holds only for strictly stable states, i.e., for states which satisfy (3) with a strict inequality.[5] However, if multiple states are stable, at least one of them has no more than $n + 2$ phases. Feinberg[5] gives an alternative proof of these results with fewer assumptions. In particular if the Gibbs free energy is continuous, every equivalence class has a state with no more than $n + 2$ phases. Mistura[7] derives the Gibbs phase rule by interpreting the Gibbs-Duhem equation as a differential equation, regarding all densities and fields as independent variables. Rabier and Griewank[8] derive the Gibbs phase rule along with its exceptions using convexification of a smooth function based on singularity theory. Gutierrez[9] obtains the Gibbs phase rule by associating a graph to each thermodynamic system.

While the Gibbs phase rule would make some of our proofs simpler, we do not restrict the discussion to at most $n + 2$ phases or even a finite number of phases. The solution of the dual problem furnishes one phase; other phases can be found from the solutions of a system of equations without specifying the number of phases a priori.

Gibbs tangent plane

Baker et al.[1] formalize and prove an important stability criterion for phase equilibria which is originally based on the work by Gibbs and is known as the Gibbs tangent plane. Here we review the concepts relevant to our interpretation of this criterion as a dual extremum principle. We present the results in the space of mole fractions x_i for $i = 1,\ldots,n-1$ after eliminating the mole fraction of species n, i.e., on the set

$$X = \left\{\mathbf{x} \in [0,1]^{n-1} : \sum_{i=1}^{n-1} x_i \leq 1\right\}.$$

Recall that we restrict the discussion to multicomponent mixtures, i.e., to $n > 1$. In this section we assume that the Gibbs free energy as a function of the mole fractions is differentiable on $\mathrm{int}(X)$ and continuous on X. These are standard assumptions in thermodynamics, for otherwise quantities such as chemical potentials would not be defined. Note that it would not be reasonable to require the existence of one-sided limits of the derivatives with respect to x_i approaching the boundary of X, since even for ideal mixtures $G_{x_i} \to -\infty$ for $x_i \to 0$. Moreover, similar to Baker et al.,[1] we require that each species is present in each phase. As a consequence of this assumption each species has to be present in the overall composition; the latter is a trivial requirement, since if a species is not present the treatment can be performed in a lower dimensional space. Later, in our main result, we eliminate the assumptions of differentiability and that each species is present in each phase.

The first step in the determination of phase equilibria is to find a postulated solution. To that effect the Gibbs free energy $G(\mathbf{x})$ is calculated as a function of the mole fractions on X and a plane $T(\mathbf{x})$ is constructed which has one or more common points with $G(\mathbf{x})$. The common points $\mathbf{x} \in X^t$, where

$$X^t = \left\{\mathbf{x} \in X |\ T(\mathbf{x}) = G(\mathbf{x})\right\},$$

correspond to the compositions $(x_i^j,\ i = 1,\ldots,n-1)$ of the predicted equilibrium phases $j \in J$. As the name of the criterion suggests, we are interested in a tangent plane and

therefore it is expected that locally around the common points $T(\mathbf{x})$ is a first-order approximation to the surface $G(\mathbf{x})$. Therefore, the first order partial derivatives with respect to the mole fractions are expected to be equal.

$$G_{x_i}(\mathbf{x}) = T_{x_i}(\mathbf{x}), \quad \forall i = 1, \ldots, n-1, \quad \forall \mathbf{x} \in X^t.$$

Recall the relation of these partial derivatives and the chemical potentials from Eqs. 4 and 5. For arbitrarily shaped surfaces, an uncountable number of points is conceivable, i.e., the cardinalities of X^t and J are potentially uncountable. As was discussed in the subsection dedicated to the Gibbs Phase rule, there are some limitations on the number of possible phases and thus of tangency points. In our treatment we will not make full use of these facts and permit infinitely many points of tangency; to keep the mathematical requirements to a minimum we restrict the discussion to the countable case.

The species balance $\Sigma_{j\in J} n_i^j = n_i^0$ for $i = 1, \ldots, n$ mandates that the overall composition lies in the "region bounded by the points of tangency of the surface and the tangent plane."[1] More precisely, the overall composition $\mathbf{x}^0$ must be in the convex hull of X^t, i.e., $\mathbf{x}^0 \in \text{conv}(X^t)$. For a countable number of phases, $\mathbf{x}^0$ must be a convex combination of the phase compositions, i.e., there exist scalars $\kappa^j \in [0,1]$ for $j \in J$ such that

$$\sum_{j\in J} \kappa^j x_i^j = x_i^0, \quad \forall i = 1, \ldots, n-1 \quad \text{and} \quad \sum_{j\in J} \kappa_j = 1.$$

These scalars correspond to the quotient of the total mole numbers in each phase with the total mole number in the system. The overall (intensive) Gibbs free energy is therefore given by $G = \sum_{j\in J} \kappa^j G^j$, where G^j is the (intensive) Gibbs free energy in phase j. For n or less phases the scalars κ^j are uniquely determined (assuming linearly independent phase compositions) while for more than n phases additional information is required, such as specifying some extensive properties.

The second step in the determination of phase equilibria is to establish the stability of the postulated solution and the Gibbs tangent plane is a very simple (conceptually) stability criterion. Assuming equality of chemical potentials, the points of tangency described above correspond to a stable solution of the phase equilibrium problem if and only if the tangent plane lies below the Gibbs free energy surface for all $\mathbf{x} \in X$. This result is quite intuitive since at constant temperature and pressure the stable state of the system is the one that minimizes the Gibbs free energy. If the tangent plane is not below the Gibbs free energy surface, then it is possible to find another tangent plane, for which the convex combination of tangency points leads to a lower overall Gibbs free energy.[1] Numerical implementations of this stability criterion are challenging and for a discussion the reader is referred to.[10–12] Theorem 5 contains the mathematical statement of the stability criterion.

Theorem 5 (Gibbs Tangent Plane). *Consider a system containing n species at a given temperature T and pressure P and an overall mole number $n_i^0 > 0$ for the species $i = 1, \ldots, n$. Denote $n_t^0 = \Sigma_{i=1}^n n_i^0$. Consider a state described by a collection of phases with index set J and composition $\mathbf{x}^j \in$ int(X) where*

$$X = \left\{ \mathbf{x} \in [0,1]^{n-1} : \sum_{i=1}^{n-1} x_i \le 1 \right\}$$

and with nonzero total mole number in each phase $n_t^j > 0$ such that

$$\sum_{j\in J} n_t^j x_i^j = n_i^0 \qquad \forall i = 1, \ldots, n-1.$$

Consider finally the associated tangent at $\mathbf{x}^{j^*}$ for some $j^* \in J$

$$T(\mathbf{x}) = G(\mathbf{x}^{j^*}) + \sum_{i=1}^{n-1} G_{x_i}(\mathbf{x}^{j^*})(x_i - x_i^{j^*}).$$

The state is stable if and only if for all $\mathbf{x} \in X$

$$T(\mathbf{x}) \le G(\mathbf{x}) \tag{7}$$

$$T(\mathbf{x}) = G(\mathbf{x}^j) + \sum_{i=1}^{n-1} G_{x_i}(\mathbf{x}^j)(x_i - x_i^j), \quad \forall j \in J. \tag{8}$$

In the Appendix we give a direct proof of Theorem 5.

Illustrative Example. To illustrate the application of the Gibbs tangent plane stability criterion consider a hypothetical binary mixture for which the Gibbs free energy is given by

$$G(x) = 0.01x + (x - 0.5)^4 - 0.2(x - 0.5)^2.$$

Figure 3 shows four cases. In the upper-left subfigure the overall composition is $x^0 = 0.9$; locally around x^0 the Gibbs free energy is convex; moreover, the tangent to the Gibbs free energy surface at x^0 lies below the Gibbs free energy surface and a single phase with the composition x^0 is stable. In the upper-right subfigure the overall composition is $x^0 = 0.5$; locally around x^0 the Gibbs free energy is concave, and therefore a single phase with this composition is unstable. In the lower-left subfigure the overall composition is $x^0 = 0.2$; locally around x^0 the Gibbs free energy is convex, but for values around $x = 0.68$ the tangent crosses above the surface; therefore a single phase with the composition $x^0 = 0.2$ is unstable. Finally, the lower-right subfigure shows the phase split for overall compositions $x^0 \in [0.1835, 0.8163]$ to two phases; the two-phase tangent lies below the Gibbs free energy for all x. Note that the necessary phase equilibrium criteria $\mu_1^1 = \mu_1^2$ and $\mu_2^1 = \mu_2^2$ are met since the chemical potentials are given by the intercept with the axis $x = 0$ and $x = 1$.

Dual Extremum Principle

We will now interpret the Gibbs tangent plane criterion via Lagrangian duality. We assume that the Gibbs free energy is continuous for all $\mathbf{x} \in X$, where

$$X = \left\{ \mathbf{x} \in [0,1]^{n-1} : \sum_{i=1}^{n-1} x_i \le 1 \right\}$$

and that $\mathbf{x}^0 \in$ int(X), or equivalently $n_i^0 > 0$, $\forall i = 1, \ldots, n$. Note again that the requirement of all species being present in the overall composition is trivial, since if a species is not present the treatment can be performed in a lower-dimensional

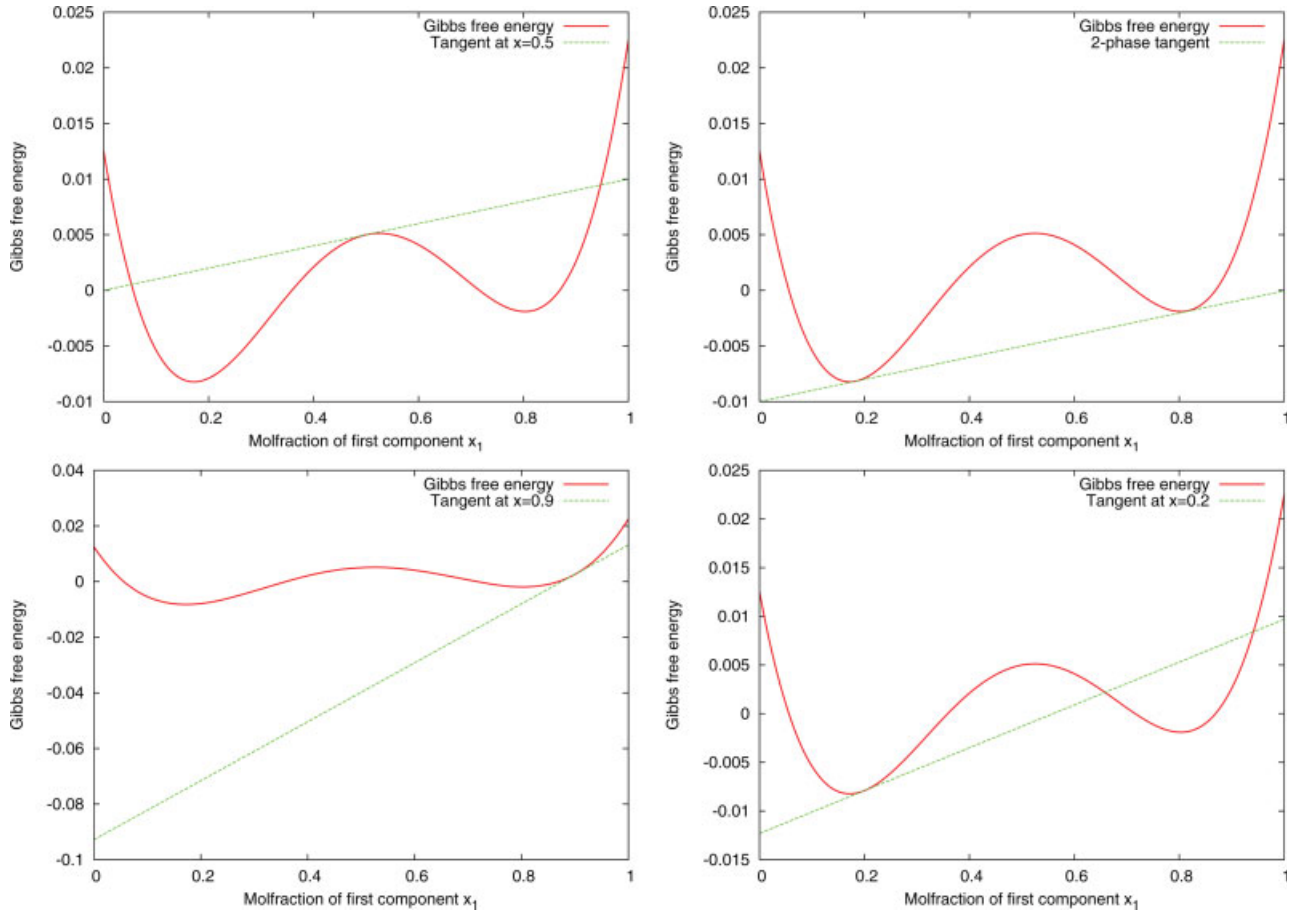

Figure 3. Illustration of Gibbs tangent stability criterion.

[Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

space. We do not make any assumptions on differentiability, nor do we require all species to be present in each phase. Furthermore, we allow an infinite number of phases; to keep the mathematical requirements to a minimum we limit the discussion to the countable case. Note that continuity of the Gibbs free energy ensures that this is not restrictive.[5]

Theorem 4 suggests that to find a single stable phase at constant temperature and pressure one has to minimize Gibbs free energy subject to a constant overall composition

$$\begin{aligned} G^p &= \min_{\mathbf{x}\in X} G(\mathbf{x}) \\ & x_i^0 - x_i = 0 \qquad \forall i = 1,\ldots,n-1. \end{aligned}$$

Note that the number of (linearly independent) constraints is equal to the number of variables and therefore this *primal problem* has a trivial solution. The Lagrangian function, dual function and dual problem, are respectively given by

$$\begin{aligned} L(\mathbf{x},\boldsymbol{\lambda}) &= G(\mathbf{x}) + \sum_{i=1}^{n-1} \lambda_i (x_i^0 - x_i) \\ q(\boldsymbol{\lambda}) &= \min_{\mathbf{x}\in X} \left(G(\mathbf{x}) + \sum_{i=1}^{n-1} \lambda_i (x_i^0 - x_i) \right) \\ G^d &= \max_{\boldsymbol{\lambda}\in \mathrm{R}^{n-1}} q(\boldsymbol{\lambda}). \end{aligned} \tag{9}$$

In the Appendix we establish that the above minima exist as asserted. We now present the main contribution of this article, i.e., the equivalency of stability and the solution of the dual problem. Note that we do not require X^* to have finite cardinality.

Theorem 6 (Dual Extremum Principle). *Consider a system of n species at a given temperature T, pressure P, an overall mole number $n_i^0 > 0$ for the species $i = 1,\ldots,n$ and corresponding overall composition $\mathbf{x}^0 \in \mathrm{int}(X)$. Take any solution $\boldsymbol{\lambda}^* \in \mathrm{R}^{n-1}$ of the dual problem (9) and any $\mathbf{x}^* \in X$, such that $L(\mathbf{x}^*, \boldsymbol{\lambda}^*) = q(\boldsymbol{\lambda}^*) = G^d$. The hyperplane*

$$T(\mathbf{x}) = G(\mathbf{x}^*) + \sum_{i=1}^{n-1} \lambda_i^* (x_i - x_i^*).$$

is a supporting hyperplane of G on X, i.e.,

$$T(\mathbf{x}) \le G(\mathbf{x}) \qquad \forall \mathbf{x} \in X.$$

Moreover

$$\mathbf{x}^0 \in \mathrm{conv}(X^*),$$

where X is the set of common points between the Gibbs surface and the hyperplane, i.e.,*

$$X^* = \{\mathbf{x} \in X : T(\mathbf{x}) = G(\mathbf{x})\}.$$

Finally, a state described by a collection of phases with index set J and compositions $\mathbf{x}^j \in X$ *and with nonzero total mole number in each phase* $n^j{}_t > 0$ *such that*

$$\mathbf{x}^0 \in \text{conv}(X^J) \quad \textit{where } X^J = \{\mathbf{x}^j \in X : j \in J\},$$

is stable if and only if $\mathbf{x}^j \in X^*$ *for all* $j \in J$, *or equivalently all* $L(\mathbf{x}^j, \boldsymbol{\lambda}^*) = G^d$.

In the Appendix we give a proof of Theorem 6.

A direct consequence of Theorem 6 is the following necessary (not sufficient) stability criterion.

Corollary 7. *Consider a system of n species at a given temperature T and pressure P and an overall mole number* $n_i^0 > 0$ *for the species* $i = 1, \ldots, n$. *A phase with composition* $\mathbf{x}^j \in X$ *can be part of a stable phase split only if a supporting hyperplane exists through* $\mathbf{x}^j \in X$, *i.e., there exists* $\boldsymbol{\lambda}^* \in \mathrm{R}^{n-1}$ *such that for all* $\mathbf{x} \in X$ *we have* $T(\mathbf{x}) \leq G(\mathbf{x})$ *where*

$$T(\mathbf{x}) = G(\mathbf{x}^j) + \sum_{i=1}^{n-1} \lambda_i (x_i - x_i^j).$$

At this point it is interesting to note the importance of the requirement $\mathbf{x}^0 \in \text{int}(X)$ or equivalently $n_i^0 > 0$ for all i. Take the simple example of an ideal binary mixture with zero reference Gibbs free energy

$$G(x) = x\ln(x) + (1-x)\ln(1-x)$$

and suppose that the overall composition is given by $x^0 = 0$. The unique stable state is a single phase with $x = x^0 = 0$. The optimal solution values of the primal and dual are equal $G^p = G^d = 0$, but the supremum of the dual is not attained. For any λ the optimal solution of the inner program is given by $\bar{x}(\lambda) = e^{\lambda}/(1+e^{\lambda})$ and therefore the dual function is not coercive: for $\lambda \to -\infty$ we obtain $\bar{x}(\lambda) \to 0 = x^0$ and $q(\lambda) \to 0 = G^d$. The only "supporting" hyperplane is vertical.

Numerical solution of the dual

While the dual problem is always convex[2] it is potentially very hard to solve. Efficient algorithms for computing stable phase splits are outside the scope of this article and we only remark that algorithms for min-max problems and linear (in the variables) semi-infinite programs (SIP) can be applied. In Algorithm 1 we sketch an algorithm for the dual problem based on the SIP algorithm by Blankenship and Falk[13] and the min-max algorithm by Falk and Hoffman.[14]

Algorithm 1, illustrated in Figure 4, is based on an upper bound (UBD) and a lower bound (LBD) to the optimal solution value, i.e., LBD $\leq G^d \leq$ UBD and finite termination is established when these bounds converge within a prescribed tolerance $\varepsilon > 0$, i.e., when UBD $-$ LBD $\leq \varepsilon$. The upper bounding problem is a linear program (LP) and the lower bounding problem a nonconvex nonlinear program (NLP) solved to global optimality. To obtain an upper bound a linear SIP equivalent to the dual program is considered and the infinite cardinality set X is replaced with a finite set X^d. This is a relaxation and therefore furnishes an upper bound. Because at each iteration the discretization is refined, convergence is achieved. The lower bound is obtained by calculating $q(\boldsymbol{\lambda}^*)$ for the solutions of the upper bounding problem.

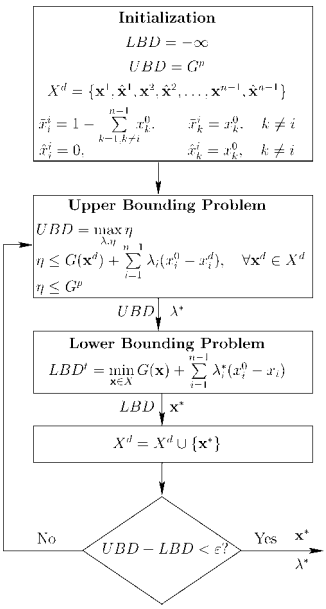

Figure 4. Algorithm for the solution of the dual problem.

Note that the lower bound may not increase monotonically, but converges. By the weak duality theorem (Theorem 2) the primal solution value can be used as an upper bound: $G^d \leq$ UBD $\leq G^p = G(\mathbf{x}^0)$.

Algorithm 1

Step 1: **Initialization**

Set LBD $= -\infty$, UBD $= G^p$ and $X^d = \{\bar{\mathbf{x}}^1, \hat{\mathbf{x}}^1, \bar{\mathbf{x}}^2, \hat{\mathbf{x}}^2, \ldots, \bar{\mathbf{x}}^{n-1}, \hat{\mathbf{x}}^{n-1}\}$ where

$$\bar{x}_i^i = 1 - \sum_{\substack{k=1 \\ k\neq i}}^{n-1} x_k^0, \quad \bar{x}_k^i = x_k^0, \quad \forall k \in \{\{1, \ldots, n-1\} : k \neq i\}$$

$$\hat{x}_i^i = 0, \quad \hat{x}_k^i = x_k^0, \quad \forall k \in \{\{1, \ldots, n-1\} : k \neq i\}.$$

Step 2: **Termination test**

If UBD $-$ LBD $\leq \varepsilon$ terminate.

Step 3: **Upper bounding problem**

Solve

$$\begin{aligned} \text{UBD} = \max_{\boldsymbol{\lambda},\eta} \ & \eta \\ \text{s.t.} \ & \eta \leq G(\mathbf{x}^d) + \sum_{i=1}^{n-1} \lambda_i (x_i^0 - x_i^d), \quad \forall \mathbf{x}^d \in X^d \\ & \eta \leq G^p \\ & \boldsymbol{\lambda} \in \mathrm{R}^{n-1} \end{aligned} \tag{10}$$

and set $\boldsymbol{\lambda}^*$ equal to an optimal solution point.

Step 4: **Lower Bounding Problem**

Solve

$$\begin{aligned} \text{LBD} = \min_{\mathbf{x}} \ & G(\mathbf{x}) + \sum_{i=1}^{n-1} \lambda_i^* (x_i^0 - x_i) \\ \text{s.t.} \ & \mathbf{x} \in X \end{aligned} \tag{11}$$

and set $\mathbf{x}^*$ equal to an optimal solution point.

Set $X^d = X^d \cup \{\mathbf{x}^*\}$.

Goto Step 2.

At finite termination $\mathbf{x}^*$ contains the (approximate) composition of one of the phases and $\boldsymbol{\lambda}^*$ the (approximate) normal of a maximal Gibbs supporting hyperplane. It is very important to note that the solution of the dual did not require an a priori specification of the number of phases. To find the number of phases all solutions of $G(\mathbf{x}) = G(\mathbf{x}^*) + \boldsymbol{\lambda}^{*\mathrm{T}}(\mathbf{x} - \mathbf{x}^*)$ can be found a posteriori. Note that algorithms for this purpose exist, e.g. Ref. 15.

Since the upper bounding problem (10) is a LP and its objective function is bounded above by $G(\mathbf{x}^0)$, an optimal solution exists.[3] The purpose of populating X^d in the initialization step is to enforce implicit bounds on $\boldsymbol{\lambda}$. For each $i = 1,\ldots,n-1$, the points $\bar{\mathbf{x}}^i$, $\hat{\mathbf{x}}^i$ satisfy $x_i^0 - \hat{x}_i^i > 0 > x_i^0 - \bar{x}_j^i$ and therefore enforce

$$\frac{\eta - G(\hat{\mathbf{x}}^i)}{x_i^0 - \hat{x}_i^i} \leq \lambda_i \leq \frac{\eta - G(\bar{\mathbf{x}}^i)}{x_i^0 - \bar{x}_i^i}.$$

While these bounds are not needed for the solution of the upper bounding problem, they are included to ensure that at each iteration sensible values for $\boldsymbol{\lambda}^*$ are obtained. These bounds guarantee the convergence of the algorithm as well as a well-behaved lower bounding problem at each iteration. Note that these bounds offer another motivation for the requirement $\mathbf{x}^0 \in \mathrm{int}(X)$. Note also that in a numerical implementation, caution needs to be used when calculating $G(\hat{\mathbf{x}}^i)$ and $G(\hat{\mathbf{x}}^i)$, since the points are by construction on the boundary of X. Figure 5 illustrates how these points are obtained for a ternary mixture. The convergence of Algorithm 1 is a direct consequence of the properties of the algorithm by Blankenship and Falk.[13] For the sake of completeness we give a formal convergence statement and a proof in the Appendix.

For state-of-the-art global optimization solvers the host set needs to be compact and all functions involved bounded.[16] Typically $G(\mathbf{x})$ also contains the ideal mixing term $x\ \ln(x)$ and while $x\ \ln(x)$ is bounded on X, $\ln(x)$ is not. Therefore, for numerical implementation the host set in (11) needs to be restricted to a closed subset of the interior of X, e.g., to

$$\widetilde{X} = \left\{ \mathbf{x} \in [\delta, 1-\delta]^{n-1} : \sum_{i=1}^{n-1} x_i \leq 1 - \delta \right\},$$

for some small positive $\delta > 0$, or $x\ \ln(x)$ has to be treated as a special intrinsic function.

Although our focus is not algorithmic, a brief comparison with existing methods and algorithms is warranted. Michelsen[10,11] introduced an efficient implementation of phase split calculations based on the Gibbs tangent plane criterion. It is essentially a two-stage approach. At the first stage a phase split is postulated and at the second stage the stability of this phase split is checked with the tangent plane criterion. If the stability criterion succeeds, the algorithm terminates, otherwise a new phase split is identified. McDonald and Floudas[12,17,18] then introduced concepts from global optimization that provide rigorous guarantees for the success of the phase stability criterion. In particular they used a primal/dual method by Visweswaran and Floudas[19–21] to globally solve the (primal) problem of tangent distance minimization. Similar to the aforementioned methods, our proposal is a two-stage process. In the first stage the normal of a hyperplane is obtained through a linear program. In the second stage the maximal interception of this hyperplane is obtained through the global solution of a nonlinear program, which is equivalent to the tangent distant minimization of the existing methods. Our proposal does not require differentiability and can use any global optimization algorithm. Finally, our method solves a dual optimization problem which is different than applying duality-based algorithms for the solution of a primal problem.

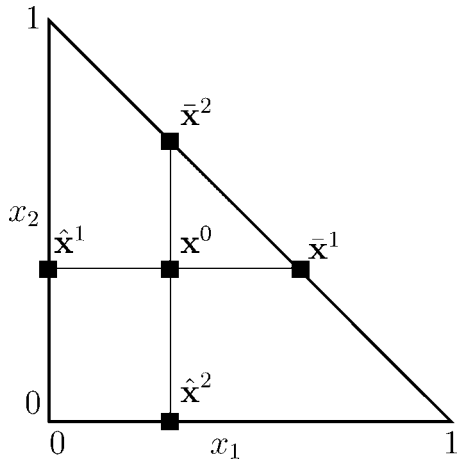

Figure 5. Illustration of initial points in set X^d for a ternary mixture.

Demonstration examples

We now demonstrate the dual extremum principle along with its numerical solution for a few literature examples. The upper bounding problem is solved with CPLEX and the lower bounding problem with BARON both through the GAMS[22] version 22.3 and default options, except for the optimality tolerance which we set to 10^{-7}. Such a tight tolerance may not be necessary for most applications, but here we are interested in high accuracy. We run the code on a 64-bit Xeon processor 3.2GHz running Linux 2.6.13. The GAMS code for all case studies as well as the main numerical results are given in the supplementary material. Note that the problems are sufficiently small that they can be run with the demo version of GAMS.

Binary NRTL Model. As an illustrative example we consider a binary mixture of butyl-acetate and water modeled with the NRTL activity coefficient model[4] with zero Gibbs free energy of reference.

$$G(x) = x\ln x + (1-x)\ln(1-x) + x(1-x)\left(\frac{\tau_{12}e^{\alpha_{12}\tau_{12}}}{(1-x)+xe^{-\alpha_{12}\tau_{12}}} + \frac{\tau_{21}e^{-\alpha_{21}\tau_{21}}}{x+(1-x)e^{-\alpha_{21}\tau_{21}}}\right)$$

with the values from Ref. 12: $\tau_{12} = 3.00498$, $\tau_{21} = 4.69071$, $\alpha_{12} = 0.391965$, and an overall composition $x^0 = 0.5$. As is shown in Figure 6 the tangent at x^0 is locally below the Gibbs surface, but at approximately 0.05 crosses over it; therefore a single phase is unstable.

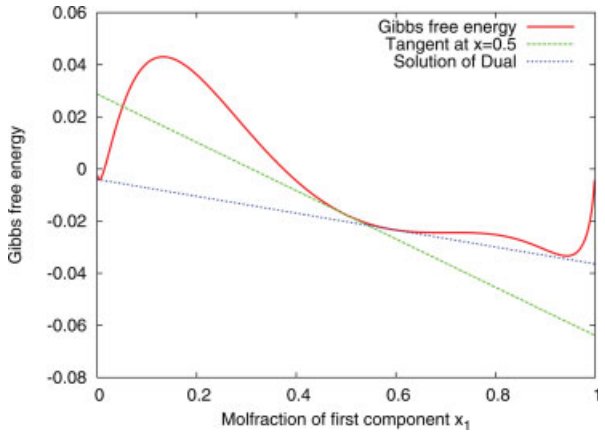

Figure 6. Illustration of binary NRTL example.

[Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

The primal problem is given by

$$\min_{x\in[0,1]} \; x\ln x + (1-x)\ln(1-x) + x(1-x) \times\left(\frac{\tau_{12}e^{-\alpha_{12}\tau_{12}}}{(1-x)+xe^{-\alpha_{12}\tau_{12}}} + \frac{\tau_{21}e^{-\alpha_{21}\tau_{21}}}{x+(1-x)e^{-\alpha_{21}\tau_{21}}}\right)$$
$$\text{s.t.} \quad x = 0.5,$$

with an optimal solution value of approximately -0.01758. Recall that this solution is not stable.

The Lagrangian and dual functions are given by

$$L(x,\lambda) = \lambda(x-0.5) + x\ln x + (1-x)\ln(1-x) + x(1-x) \times\left(\frac{\tau_{12}e^{-\alpha_{12}\tau_{12}}}{(1-x)+xe^{-\alpha_{12}\tau_{12}}} + \frac{\tau_{21}e^{-\alpha_{21}\tau_{21}}}{x+(1-x)e^{-\alpha_{21}\tau_{21}}}\right)$$
$$q(\lambda) = \min_{x\in[0,1]} \; \lambda(x-0.5) + x\ln x + (1-x)\ln(1-x) + x(1-x) \times\left(\frac{\tau_{12}e^{-\alpha_{12}\tau_{12}}}{(1-x)+xe^{-\alpha_{12}\tau_{12}}} + \frac{\tau_{21}e^{-\alpha_{21}\tau_{21}}}{x+(1-x)e^{-\alpha_{21}\tau_{21}}}\right).$$

The dual problem is given by

$$\max_{\lambda\in R} \min_{x\in[0,1]} \; \lambda(x-0.5) + x\ln x + (1-x)\ln(1-x) + x(1-x) \times\left(\frac{\tau_{12}e^{\alpha_{12}\tau_{12}}}{(1-x)+xe^{-\alpha_{12}\tau_{12}}} + \frac{\tau_{21}e^{-\alpha_{21}\tau_{21}}}{x+(1-x)e^{-\alpha_{21}\tau_{21}}}\right).$$

Since this problem involves only two variables it can be solved efficiently via Algorithm 1. Each subproblem is a univariate problem. For solution we restrict x to [0.00001, 0.99999]. After just five iterations and a total of approximately 0.8 s of CPU, the lower and upper bounds converge in the sixth significant digit (-0.0201983305 vs. -0.0201983083). The (approximate) solution point obtained is $x^* = 0.004557$ and $\lambda^* = -0.03244$. As is shown in Figure 6 the line $G(x^*) + \lambda^*(x - x^*)$ is the maximal Gibbs supporting hyperplane corresponding to the stable split in two phases. The composition of one of the two phases is given by $x^{\mathrm{I}} = x^*$; the other can be found by solving $G(x) = G(x^*) + \lambda^*(x - x^*)$ to be $x^{\mathrm{II}} \approx 0.5920$. The duality gap $G^p - G^d \approx 0.002628$ corresponds to approximately 15% of the primal objective value. After the second iteration, the solution points of the lower bounding problem alternate between points converging to x^{I} and x^{II}.

Ternary NRTL Model. We now consider the ternary toluene-water-aniline mixture modeled with the NRTL activity coefficient model. The excess Gibbs free energy of a multicomponent mixture is given by

$$G^E(\mathbf{x}) = \sum_{i=1}^{n} x_i \frac{\sum_{j=1}^{n} x_j\tau_{ji}e^{-\alpha_{ji}\tau_{ji}}}{\sum_{j=1}^{n}\tau_{ji}e^{-\alpha_{ji}\tau_{ji}}},$$

with the convention $\tau_{ii} = 0$ and $\alpha_{ij} = \alpha_{ji}$. For a ternary mixture after elimination of the molefraction of the third component ($x_3 = 1 - x_1 - x_2$) we obtain

$$\begin{aligned} G(\mathbf{x}) = {} & x_1\ln(x_1) + x_2\ln(x_2) + (1-x_1-x_2)\ln(1-x_1-x_2) \\ & + x_1\frac{\tau_{21}e^{-\alpha_{12}\tau_{21}}x_2 + \tau_{31}e^{-\alpha_{13}\tau_{31}}(1-x_1-x_2)}{x_1 + e^{\alpha_{12}\tau_{21}}x_2 + e^{-\alpha_{13}\tau_{31}}(1-x_1-x_2)} \\ & + x_2\frac{\tau_{12}e^{-\alpha_{12}\tau_{12}}x_1 + \tau_{32}e^{-\alpha_{23}\tau_{32}}(1-x_1-x_2)}{e^{-\alpha_{12}\tau_{12}}x_1 + x_2 + e^{-\alpha_{23}\tau_{32}}(1-x_1-x_2)} \\ & + (1-x_1-x_2)\frac{\tau_{13}e^{-\alpha_{13}\tau_{13}}x_1 + \tau_{23}e^{-\alpha_{23}\tau_{23}}x_2}{e^{\alpha_{13}\tau_{13}}x_1 + e^{-\alpha_{23}\tau_{23}}x_2 + 1 - x_1 - x_2}, \end{aligned}$$

taking again zero Gibbs free energy reference for all species. The binary interaction parameters are taken from Ref. 12: $\tau_{12} = 4.93035$, $\tau_{21} = 7.77063$, $\alpha_{12} = 0.2485$, $\tau_{13} = 1.59806$, $\tau_{31} = 0.03509$, $\alpha_{13} = 0.3000$, $\tau_{23} = 4.18462$, $\tau_{32} = 1.27932$, $\alpha_{23} = 0.3412$ and we consider an overall composition $x_1^0 = 0.3$ and $x_2^0 = 0.2$. The optimal solution value of the primal problem is $G^p = -.3247905329$. As is shown in Figure 7 this solution does not correspond to a stable phase split.

This problem involves four variables and each subproblem of Algorithm 1 contains two variables. For solution we restrict $\mathbf{x}$ to $[0.000001, 0.999999]^2$. After 13 iterations and a total of approximately 35 s of CPU, the lower and upper bounds converge in the sixth significant digit (-0.3529753313 vs. -0.3529746879). The (approximate) solution point obtained is $x_1^* = 0.34759$, $x_2^* = 0.07562$, $\lambda_1^* = 0.0917013$, and $\lambda_2^* = 0.46985$. The reason for the increased CPU requirement compared to the binary example is the increased cost of solving the lower bounding problem. We use a general-purpose solver for this global optimization problem, which cannot take advantage of specialized techniques. For instance, the term $x\ln(x)$ is convex, but the solver does not have this information. Since the upper bounding problem is a linear program and very fast, the algorithm performance can be improved by initializing the set X^d with more points in the molefraction space, e.g., the pure species.

As is shown in Figure 7 the plane $G(\mathbf{x}^*) + \lambda_1^*(x_1 - x_1^*) + \lambda_2^*(x_2 - x_2^*)$ is a supporting hyperplane corresponding to the stable split in two phases. The composition of one of the two phases is given by $\mathbf{x}^I = \mathbf{x}^*$; the other can be found by finding the second global minimizer of $G(\mathbf{x}) = G(\mathbf{x}^*) + \Sigma_{i=1}^2\lambda_i^*(x_i - x_i^*)$ to be $\mathbf{x}^{\mathrm{II}} \approx (0.00009, 0.9994950)$. The duality gap $G^p - G^d \approx 0.0282$ corresponds to approximately 9% of the primal objective value. During the iterations of the algorithm the solution points of the lower bounding problem alternate between a water rich and a toluene rich phase.

Binary UNIQUAC Model. We now consider a binary mixture of toluene and water modeled with the UNIQUAC

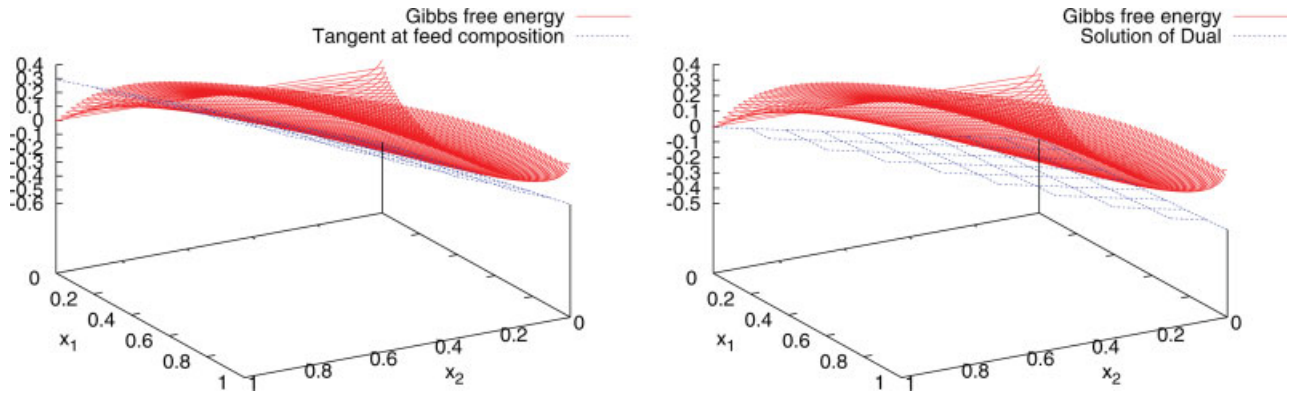

Figure 7. Illustration of ternary NRTL example.

[Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

activity coefficient model with zero Gibbs free energy of reference

$$G = \sum_{i=1}^{n} x_i \left(\left(1 - \frac{z}{2} q_i\right) \ln(\Phi_i) + \frac{z}{2} q_i \ln(\theta_i) q_i' \ln\left(\sum_{j=1}^{n} \theta_j' \tau_{ji} \right) \right),$$

with

$$\Phi_i = \frac{r_i x_i}{\sum_{j=1}^{n} r_j x_j}, \quad \theta_i = \frac{q_i x_i}{\sum_{j=1}^{n} q_j x_j}, \quad \theta_j' = \frac{q_j' x_j}{\sum_{k=1}^{n} q_k' x_k}$$

and the convention $\tau_{ii} = 1$. We use the parameter values from Ref. 12: $\tau_{12} = 0.09867$, $\tau_{21} = 0.59673$, $q_1 = 2.97$, $q_2 = 1.40$, $q_1' = 2.97$, $q_2' = 1.00$, $r_1 = 3.92$, $r_2 = 0.92$, and $z = 10$.[4] We use equimolar overall composition, i.e., $x^0 = 0.5$. As is shown in Figure 8 the tangent at x^0 is locally and globally above the Gibbs surface; therefore a single phase is unstable.

Since this problem only involves two variables it can be solved efficiently via Algorithm 1. Each subproblem is a univariate problem. For solution we restrict x to [0.00001, 0.99999]. After just five iterations and a total of approximately 1.3 s of CPU, the lower and upper bounds converge in the sixth significant digit (-0.0197587959 vs. -0.0197587955), giving a duality gap $G^p - G^d \approx 0.33$. The (approximate) solution point obtained is $x^* = 0.0008937706$ and $\lambda^* = -0.0377627074$. As is shown in Figure 8 the line $G(x^*) + \lambda^*(x - x^*)$ is the maximal Gibbs supporting hyperplane corresponding to the stable split in two phases. The composition of one of the two phases is given by $x^{\mathrm{I}} = x^*$; the other can be found by solving $G(x) = G(x^*) + \lambda^*(x - x^*)$ to be $x^{\mathrm{II}} \approx 0.9564$. After the second iteration, the solution points of the lower bounding problem alternate between points converging to x^{I} and x^{II}.

Conclusions and Future Work

We have established that the well-known stability criterion of the Gibbs tangent plane can be interpreted as the solution of the dual problem of a primal problem that minimizes Gibbs free energy subject to material balance. Since we constructed the dual problem based on the natural primal problem for Gibbs free energy minimization, this equivalency suggests a *dual extremum principle* for phase equilibrium, i.e., that the stable phase splits are the solutions of a dual problem as opposed to its primal problem. Only in the absence of duality gap does the solution coincide with the solution of the primal problem; this case can correspond to a single phase being stable (no guarantee of uniqueness). No assumptions on differentiability of the Gibbs free energy are required, which by itself is a significant advance. Moreover, the number of phases is not restricted to a finite number and the presence of all species in all phases is not required.

The observed result is expected to hold for different reference variables. For instance, we could consider constant volume and internal energy and maximize entropy. It is more challenging to examine whether this result can be generalized to other systems, including phase-equilibrium for reacting systems, or systems where a different thermophysical property model is used in different phases.[23] Furthermore, it would be interesting to consider the significance of a perturbed primal problem in which the temperature and pressure are considered as variables along with constraints fixing them to the nominal values; the dual program would be obtained by dualizing these constraints in addition to the

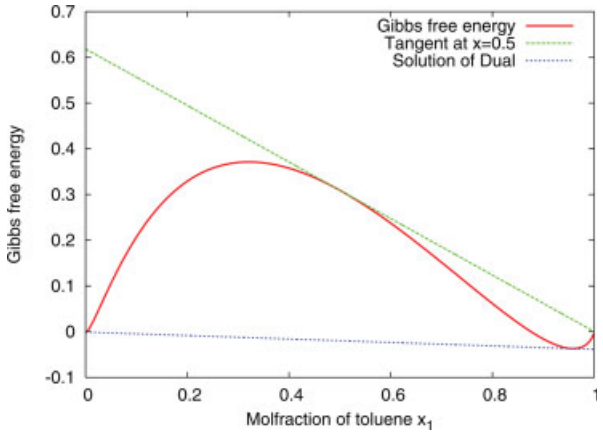

Figure 8. Illustration of binary UNIQUAC example.

[Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

overall composition. Finally, an extension to discrete systems would be very interesting for applications involving few molecules such as nanotechnology. In that case the mole fractions would be limited to rational numbers and the most likely dual would be obtained by incorporating this requirement into the host set, deviating from the typical treatment of discrete constraints in mathematical programming. Even more challenging would be to consider if dual extremum principles exist in other disciplines where extremum principles are applied such as mechanical, ecological or economical systems[24]; of particular promise are such systems that allow for the equivalent of different phases.

In addition to theoretical interest, the observations of this article may also have algorithmic implications. In general the dual optimization problem is convex, but believed to be NP-hard in general. The dual program considered here has a special structure and can be formulated as a linear (in the variables) SIP. It is therefore conceivable to formulate tractable algorithms. For instance, the computational complexity of Algorithm 1 is likely to be improved by adding more points to the discretization set in the upper bounding problem, or solving a convex relaxation of the lower bounding problem for most iterations. The main potential benefits compared to state-of-the-art algorithms is that the solution of the dual program does not require an a priori specification of the number of phases and (at least in principle) no differentiability assumptions. Moreover, since it is computationally inexpensive to obtain upper bounds to the optimal solution value of the dual problem, these results can be used to disprove the stability of postulated phase splits.

Acknowledgments

This material is based upon work supported by the Department of Energy Nuclear Energy Research Initiative under Award Number DE-FC07-06ID14751.

Literature Cited

1. Baker LE, Pierce AC, Luks KD. Gibbs energy analysis of phase equilibria. *Soc Petrol Eng J*. 1982;22:731–742.
2. Bertsekas DP. *Nonlinear Programming*, 2nd ed. Belmont, MA: Athena Scientific, 1999.
3. Bertsimas D, Tsitsiklis JN. *Introduction to Linear Optimization*. Belmont, MA: Athena Scientific, 1997.
4. Tester JW, Modell M. *Thermodynamics and Its Applications*, 3rd ed. New Jersey: Prentice Hall, 1997. Prentice Hall International Series in the Physical and Chemical Engineering Sciences.
5. Feinberg M. On Gibbs' phase rule. *Arch Rational Mech Anal*. 1979;70:219–234.
6. Noll W. On certain convex sets of measures and on phases of reacting mixtures. *Arch Rational Mech Anal*. 1970;38:1–12.
7. Mistura I. Solutions of the Gibbs-Duhem equation. *J Phys A Math Gen*. 1979;12:1547–1562.
8. Rabier PJ, Griewank A. Generic aspects of convexification with applications to thermodynamic equilibrium. *Arch Rational Mech Anal*. 1992;118:349–397.
9. Gutierrez G. Gibbs' phase rule revisited. *Theor Math Phys*. 1996;108:1222–1224.
10. Michelsen ML. The isothermal flash problem. 1. Stability. *Fluid Phase Equilibria*. 1982;9:1–19.
11. Michelsen ML. The isothermal flash problem. 2. Phase-split calculation. *Fluid Phase Equilibria*. 1982;9:21–40.
12. Mcdonald CM, Floudas CA. Global optimization for the phase-stability problem. *AIChE J*. 1995;41:1798–1814.
13. Blankenship JW, Falk JE. Infinitely constrained optimization problems. *J Optim Theory Appl*. 1976;19:261–281.
14. Falk JE, Hoffman K. A nonconvex max-min problem. *Naval Res Logistics*. 1977;24:441–450.
15. Maranas CD, Floudas CA. Finding all solutions of nonlinearly constrained systems of equations. *J Global Optim*. 1995;7:143–182.
16. Tawarmalani M, Sahinidis NV. *Convexification and Global Optimization in Continuous and Mixed-Integer Nonlinear Programming. Nonconvex Optimization and its Applications*. Boston: Kluwer Academic, 2002.
17. Mcdonald CM, Floudas CA. Decomposition based and branch-and-bound global optimization approaches for the phase-equilibrium problem. *J Global Optim*. 1994;5:205–251.
18. McDonald C, Floudas CA. Global optimization and analysis for the Gibbs free-energy function using the UNIFAC, Wilson, and ASOG equations. *Ind Eng Chem Res*. 1995;34:1674–1687.
19. Floudas CA, Visweswaran V. A Global optimization algorithm (GOP) for certain classes of nonconvex NLPs. 1. Theory. *Comput Chem Eng*. 1990;14:1397–1417.
20. Visweswaran V, Floudas CA. A global optimization algorithm (GOP) for certain classes of nonconvex NLPs. 2. Application of theory and test problems. *Comput Chem Eng*. 1990;14:1419–1434.
21. Floudas CA, Visweswaran V. Primal-relaxed dual global optimization approach. *J Optim Theory Appl*. 1993;78:187–225.
22. Brooke A, Kendrick D, Meeraus A. GAMS: *A User's Guide*. Redwood City, CA: The Scientific Press, 1988.
23. Smith JV, Missen RW, Smith WR. General optimality criteria for multiphase multireaction chemical-equilibrium. *AIChE J*. 1993;39:707–710.
24. Sieniutycz S, Farkas H, editors. *Variational and Extremum Principles in Macroscopic Systems*. Oxford: Elsevier, 2005.
25. Baier R, Gerdts M. Intensive course on set-valued numerical analysis and optimal control. http://www.math.uni-hamburg.de/home/gerdts/TEACHING/num_analysis_opt_control.pdf 2005.
26. Fosdick R, Patiño J. On the Gibbsian thermostatics of mixtures. *Arch Rational Mech Anal*. 1986;93:203–221.
27. Bank B, Guddat J, Klatte D, Kummer B, Tammer K. *Non-Linear Parametric Optimization*. Stuttgart: Birkhäuser Verlag, 1983.

Appendix: Mathematical Results and Proofs

Geometric interpretation of duality

For the sake of completeness we give a proof of Proposition 3.

Proposition 3. *The solution value q^** of the dual problem *(2) is equal to the highest intercept of the supporting hyperplanes of S with the vertical axis. For each such hyperplane with (scaled) normal vector ($\boldsymbol{\lambda}^*$, 1) the vector $\boldsymbol{\lambda}^* \in R^m$ is a solution point of the dual.*

Proof. Following the proof by Bertsekas[2] (for inequality constrained problems) we first prove some intermediate results.

1. The hyperplane with normal vector ($\bar{\boldsymbol{\lambda}}$, 1) that passes through a point ($\mathbf{g}(\mathbf{x})$, $f(\mathbf{x})$) for some $\mathbf{x} \in X$ has an intercept with the vertical axis $\mathbf{z} = \mathbf{0}$ at the level $L(\mathbf{x}, \bar{\boldsymbol{\lambda}})$.

The hyperplane is defined by

$$H = \{(\mathbf{z}, w) |\ \mathbf{z} \in \mathrm{R}^m, w \in \mathrm{R}, \bar{\boldsymbol{\lambda}}^{\mathrm{T}}\mathbf{z} + w = \bar{\boldsymbol{\lambda}}^{\mathrm{T}}\mathbf{g}(\mathbf{x}) + f(\mathbf{x})\}.$$

Setting $\mathbf{z} = \mathbf{0}$ in the defining equation we therefore obtain

$$w = \bar{\boldsymbol{\lambda}}^{\mathrm{T}}\mathbf{g}(\mathbf{x}) + f(\mathbf{x}).$$

Since $L(\mathbf{x}, \bar{\boldsymbol{\lambda}}) = \bar{\boldsymbol{\lambda}}^{\mathrm{T}}\mathbf{g}(\mathbf{x}) + f(\mathbf{x})$ the intercept with the vertical axis is at the level $L(\mathbf{x}, \bar{\boldsymbol{\lambda}})$.

2. Among all supporting hyperplanes of S with normal vector ($\bar{\boldsymbol{\lambda}}$, 1) the highest attained intercept with the vertical axis $\mathbf{z} = \mathbf{0}$ is $\inf_{\mathbf{x}\in X} L(\mathbf{x}, \bar{\boldsymbol{\lambda}})$.

A hyperplane with normal vector $(\bar{\lambda},1)$ that intercepts the vertical axis at a level w^* satisfies

$$\bar{\lambda}^{\mathrm{T}}\mathbf{z} + w = \bar{\lambda}^{\mathrm{T}}\mathbf{0} + w^* = w^*, \quad \forall (\mathbf{z}, w) \in H.$$

By assumption the hyperplane contains S in its positive halfspace and therefore

$$\bar{\lambda}^{\mathrm{T}}\mathbf{g}(\mathbf{x}) + f(\mathbf{x}) \geq w^*, \quad \forall \mathbf{x} \in X.$$

Since by definition $L(\mathbf{x},\bar{\lambda}) = \bar{\lambda}^{\mathrm{T}}\mathbf{g}(\mathbf{x}) + f(\mathbf{x})$ the maximum intercept is $\inf_{\mathbf{x}\in X} L(\mathbf{x}, \bar{\lambda})$.

Recall now that the dual problem is the maximization of the dual function $q(\bar{\lambda}) = \inf_{\mathbf{x}\in X} L(\mathbf{x},\bar{\lambda})$. Equivalently, a vector $\bar{\lambda}^*$ is sought that maximizes the intercept level. Note that this vector is not necessarily unique. □

General results

Existence of minima: A question which arises in optimization is whether for a given feasible set $X_{\mathrm{feas}} \subset \mathrm{R}^n$ and a given objective function f: $X_{\mathrm{feas}} \to \mathrm{R}$, the infimum $f^* = \inf_{\mathbf{x}\in X_{\mathrm{feas}}} f(\mathbf{x})$ is attained, i.e., if there exists $\mathbf{x}^* \in X_{\mathrm{feas}}$, such that $f(\mathbf{x}^*) = f^*$. Examples where this is not the case include $\inf_{x\in\mathrm{R}} e^x$ and $\inf_{x>0} x$. Typically, existence of a minimum is established by the application of Weierstrass' theorem:

Theorem 8 (Weierstrass' Theorem). *Let $X_{\mathrm{feas}} \subset \mathrm{R}^n$ be a nonempty closed set and f: $X_{\mathrm{feas}} \to \mathrm{R}$ be continuous on X_{feas}. If either of the following properties holds*

- *X_{feas} is bounded*
- *f is coercive, i.e., $f(\mathbf{x}) \to \infty$ for $\|\mathbf{x}\| \to \infty$*

then there exists $\mathbf{x}^* \in X_{\mathrm{feas}}$ such that $f(\mathbf{x}^*) = \inf_{\mathbf{x}\in X_{\mathrm{feas}}} f(\mathbf{x})$.

The reader is referred to Ref. 2 for a more general statement of Weierstrass' theorem. A closed feasible set is typically obtained by the intersection of a closed host set $X \subset \mathrm{R}^n$ with constraints defined by continuous functions $\mathbf{g}$: $X \to \mathrm{R}^m$:

$$X_{\mathrm{feas}} = \{\mathbf{x} \in X : \mathbf{g}(\mathbf{x}) \leq \mathbf{0}\}.$$

Equality constraints can be defined as pairs of inequality constraints. In R^n if a set is closed and bounded it is also compact. In the case that existence of a minimum is established we use min instead of inf.

In the proof of our main contribution we will make use of the following lemma.

Lemma 9 (Convex Hull of a Compact Set). *The convex hull of a compact set $D \subset \mathrm{R}^n$ is compact.*

Proof. The proof is based on Ref. 25. Recall that in R^n a compact set is equivalent to a closed and bounded set. For simplicity denote $C \equiv \mathrm{conv}(D)$. Clearly, since D is bounded, so is C. Consider now a convergent sequence of points $\{\mathbf{x}^m\}$, $\mathbf{x}^m \in C$, and denote its limit $\bar{\mathbf{x}}$. By the definition of the convex hull and Theorem 1 each point $\mathbf{x}^m$ can be represented as

$$\mathbf{x}^m = \sum_{j=1}^{n+1} \kappa^{j,m}\mathbf{x}^{j,m},$$

for some $\mathbf{x}^{j,m} \in D$ and $\kappa^{j,m} \in [0,1]$ satisfying $\sum_{j=1}^{n+1} \kappa^{j,m} = 1$. Since $[0,1]$ and D are compact the sequences $\{\kappa^{j,m}\}$ and $\{\mathbf{x}^{j,m}\}$ contain convergent subsequences. Without loss of generality consider common convergent subsequences $\{\kappa^{j,m_k}\}$ and $\{\mathbf{x}^{j,m_k}\}$, such that

$$\text{for } k \to +\infty \text{ we have: } \quad \kappa^{j,m_k} \to \kappa^j \quad \text{and} \quad \mathbf{x}^{j,m_k} \to \mathbf{x}^j, \qquad j = 1, \ldots, n+1.$$

Since $[0,1]$ and D are closed we also have $\kappa^j \in [0,1]$ and $\mathbf{x}^j \in D$. Therefore the sequence of points $\{\mathbf{x}^{m_k}\} = \{\sum_{j=1}^{n+1} \kappa^{j,m_k} \mathbf{x}^{j,m_k}\}$ converges to $\sum_{j=1}^{n+1} \kappa^j \mathbf{x}^j \in C$. Since $\sum_{j=1}^{n+1} \kappa^{j,m_k} = 1$, also $\sum_{j=1}^{n+1} \kappa^j = 1$. Since $\{\mathbf{x}^m\}$ converges to $\bar{\mathbf{x}}$, so does $\{\mathbf{x}^{m_k}\}$. Therefore $\bar{\mathbf{x}}$ is a convex combination of the points $\mathbf{x}^j \in D$, and thus $\bar{\mathbf{x}} \in C$, which proves that C is closed. □

Separating hyperplanes: A well-known result for convex sets is the separating hyperplane theorem, e.g. Ref. 3:

Theorem 10 (Separating Hyperplane Theorem). *Let $S \subset \mathrm{R}^n$ be a nonempty convex, closed set and $\bar{\mathbf{x}} \in \mathrm{R}^n$ be a point that does not belong to this set ($\bar{\mathbf{x}} \notin S$). There exists a $\mathbf{c} \in \mathrm{R}^n$, such that*

$$\mathbf{c}^{\mathrm{T}}\bar{\mathbf{x}} < \mathbf{c}^{\mathrm{T}}\mathbf{x}, \quad \forall \mathbf{x} \in S.$$

The following corollary of Theorem 10 is used in the proof of our main result.

Corollary 11 (Separating Hyperplane Theorem Variant). *Let $S \subset \mathrm{R}^n$ be a nonempty convex, compact set and $\bar{\mathbf{x}} \in \mathrm{R}^n$ be a point that does not belong to this set ($\bar{\mathbf{x}} \notin S$). There exist $\mathbf{c}$, $\hat{\mathbf{x}} \in \mathrm{R}^n$, such that*

$$\mathbf{c}^{\mathrm{T}}\hat{\mathbf{x}} < \mathbf{c}^{\mathrm{T}}\mathbf{x}, \quad \forall \mathbf{x} \in S$$
$$\mathbf{c}^{\mathrm{T}}\hat{\mathbf{x}} > \mathbf{c}^{\mathrm{T}}\bar{\mathbf{x}}.$$

Proof. By Theorem 10 there exists a $\mathbf{c}$ such that

$$\mathbf{c}^{\mathrm{T}}\bar{\mathbf{x}} < \mathbf{c}^{\mathrm{T}}\mathbf{x} \qquad \forall \mathbf{x} \in S.$$

As a direct consequence of Theorem 8 there exists $\mathbf{x}^s \in S$ such that

$$\mathbf{c}^{\mathrm{T}}\mathbf{x}^s \leq \mathbf{c}^{\mathrm{T}}\mathbf{x} \qquad \forall \mathbf{x} \in S.$$

Since $\mathbf{x}^s \in S$

$$\mathbf{c}^{\mathrm{T}}\mathbf{x}^s > \mathbf{c}^{\mathrm{T}}\bar{\mathbf{x}}.$$

Take $\hat{\mathbf{x}} = \frac{1}{2}(\mathbf{x}^s + \bar{\mathbf{x}})$ and note that

$$\mathbf{c}^{\mathrm{T}}\hat{\mathbf{x}} = \frac{1}{2}\mathbf{c}^{\mathrm{T}}\mathbf{x}^s + \frac{1}{2}\mathbf{c}^{\mathrm{T}}\bar{\mathbf{x}}.$$

Therefore

$$\mathbf{c}^{\mathrm{T}}\hat{\mathbf{x}} > \frac{1}{2}\mathbf{c}^{\mathrm{T}}\bar{\mathbf{x}} + \frac{1}{2}\mathbf{c}^{\mathrm{T}}\bar{\mathbf{x}} = \mathbf{c}^{\mathrm{T}}\bar{\mathbf{x}}.$$

Similarly

$$\mathbf{c}^{\mathrm{T}}\hat{\mathbf{x}} < \frac{1}{2}\mathbf{c}^{\mathrm{T}}\mathbf{x}^s + \frac{1}{2}\mathbf{c}^{\mathrm{T}}\mathbf{x}^s$$

and therefore

$$\mathbf{c}^{\mathrm{T}}\hat{\mathbf{x}} < \mathbf{c}^{\mathrm{T}}\mathbf{x}, \quad \forall \mathbf{x} \in S. \qquad \square$$

Note that boundedness is only assumed for simplicity. Corollary 11 is illustrated graphically in Figure A1.

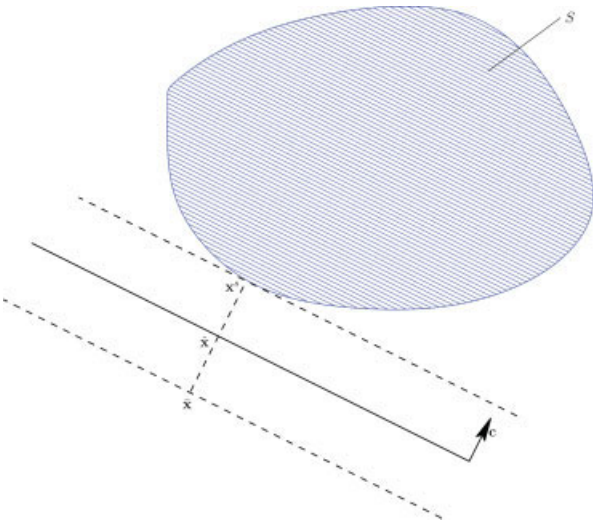

Figure A1. Illustration of separating hyperplanes, Theorem 10 and Corollary 11.

[Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

Tangents and Secants: The following lemma will be the basis of our direct proof for the Gibbs tangent plane under differentiability.

Lemma 12 (Tangent Lemma). *Let* $Z \subset \mathrm{R}$ *be an interval and* z^0 *be an interior point of this interval,* $z^0 \in \mathrm{int}(Z)$. *Consider* f: $Z \to \mathrm{R}$, *differentiable on* int(Z). *Denote the derivative* f': int(Z) $\to$ R. *Consider also the tangent* t^0 *at the point* z^0 *given by*

$$t^0(z) = f(z^0) + f'(z^0)(z - z^0).$$

If there exists a point $z^{\mathrm{I}} \in Z$ *such that* $z^{\mathrm{I}} < z^0$ *and* $t^0(z^{\mathrm{I}}) > f(z^{\mathrm{I}})$ *then there exists* $h > 0$, *such that all points* $Z^{II} \in (z^0, z^0 + h)$ *satisfy*

$$f(z^I) + \frac{f(z^{\mathrm{II}}) - f(z^{\mathrm{I}})}{z^{\mathrm{II}} - z^{\mathrm{I}}}(z^0 - z^{\mathrm{I}}) < f(z^0).$$

The proof of Lemma 12 is illustrated in Figure A2. Note that $z^{\mathrm{I}} < z^0$ was used for simplicity and the case $z^{\mathrm{I}} > z^0$ is analogous.

Proof. Consider the secant $s^{\mathrm{I},0}$ through $(z^{\mathrm{I}}, f(z^{\mathrm{I}}))$ and $(z^0, f(z^0))$, given by

$$s^{I,0}(z) = f(z^{\mathrm{I}}) + \frac{f(z^0) - f(z^{\mathrm{I}})}{z^0 - z^{\mathrm{I}}}(z - z^{\mathrm{I}}).$$

Consider also points $z^h = z^0 + h$, for $h > 0$. By the differentiability of f we have

$$\lim_{h \to 0} \frac{f(z^h) - t^0(z^h)}{h} = 0.$$

On the other hand, since both $s^{\mathrm{I},0}$ and t^0 are affine and $s^{\mathrm{I},0}(z^0) = t^0(z^0)$ we have

$$s^{\mathrm{I},0}(z^h) - t^0(z^h) = (s^{\mathrm{I},0}(z^{\mathrm{I}}) - t^0(z^{\mathrm{I}}))\frac{z^h - z^0}{z^{\mathrm{I}} - z^0}$$
$$= (f(z^{\mathrm{I}}) - t^0(z^{\mathrm{I}}))\frac{h}{z^{\mathrm{I}} - z^0} = \frac{f(z^{\mathrm{I}}) - t^0(z^{\mathrm{I}})}{z^{\mathrm{I}} - z^0}h$$

and therefore

$$\lim_{h \to 0} \frac{s^{\mathrm{I},0}(z^h) - t^0(z^h)}{h} = \frac{t^0(z^{\mathrm{I}}) - f(z^{\mathrm{I}})}{z^0 - z^{\mathrm{I}}} > 0.$$

Therefore for sufficiently small h we have $[s^{\mathrm{I},0}(z^h) - t^0(z^h)]/h > [f(z^h) - t^0(z^h)]/h$ and thus $s^{\mathrm{I},0}(z^h) > f(z^h)$. Take such a point and denote it z^{II}. Consider the secant $s^{\mathrm{I,II}}$ through $(z^{\mathrm{I}}, f(z^{\mathrm{I}}))$ and $(z^{\mathrm{II}}, f(z^{\mathrm{II}}))$, given by

$$s^{\mathrm{I,II}}(z) = f(z^{\mathrm{I}}) + \frac{f(z^{\mathrm{II}}) - f(z^{\mathrm{I}})}{z^{\mathrm{II}} - z^{\mathrm{I}}}(z - z^{\mathrm{I}}).$$

Since both $s^{\mathrm{I},0}$ and $s^{\mathrm{I,II}}$ are affine, $s^{\mathrm{I},0}(z^{\mathrm{I}}) = s^{\mathrm{I,II}}(z^{\mathrm{I}})$ and $s^{\mathrm{I},0}(z^{\mathrm{II}}) > f(z^{\mathrm{II}}) = s^{\mathrm{I,II}}(z^{\mathrm{II}})$ we have $s^{\mathrm{I},0}(z) > s^{\mathrm{I,II}}(z)$ for all $z \in (z^{\mathrm{I}}, z^{\mathrm{II}})$ and in particular for $z = z^0$, i.e.,

$$f(z^{\mathrm{I}}) + \frac{f(z^{\mathrm{II}}) - f(z^{\mathrm{I}})}{z^{\mathrm{II}} - z^{\mathrm{I}}}(z^0 - z^{\mathrm{I}}) = s^{\mathrm{I,II}}(z^0) < s^{\mathrm{I},0}(z^0) = f(z^0).$$

Note also that since $z^0 \in \mathrm{int}(Z)$ for sufficiently small h we also have $z^h \in Z$. □

Direct proof of classical Gibbs tangent plane criterion

Here we present a direct proof of the classical Gibbs tangent plane as a stability criterion, Theorem 5. Baker et al.[1] prove various intermediate results of thermodynamic importance and use thermodynamic concepts. Our proof is more direct and purely based on calculus, Lemma 12 and Theorem 4.

Theorem 4 (Gibbs Tangent Plane). *Consider a system containing* n *species at a given temperature* T *and pressure* P *and an overall mole number* $n_i^0 > 0$ *for the species* $i =$

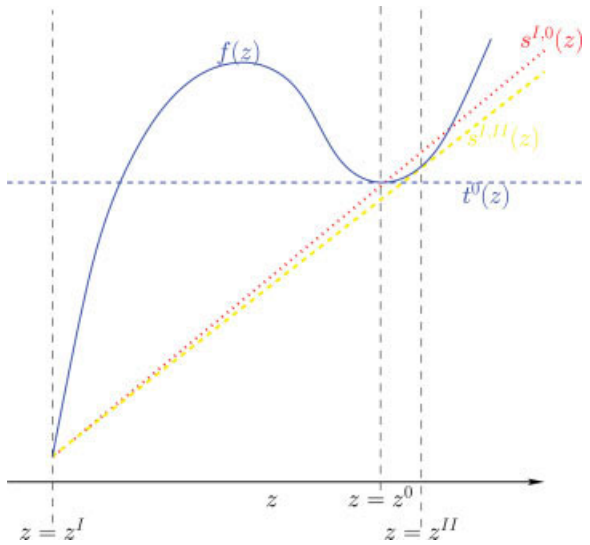

Figure A2. Illustration of Lemma 12.

[Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

$1,\ldots,n$. *Denote* $n_t^0 = \sum_{i=1}^n n_i^0$. *Consider a state described by a collection of phases with index set* J *and composition* $\mathbf{x}^j \in \text{int}(X)$ *where*

$$X = \left\{ \mathbf{x} \in [0,1]^{n-1} : \sum_{i=1}^{n-1} x_i \leq 1 \right\}.$$

and with nonzero total mole number in each phase $n^j{}_t > 0$ *such that*

$$\sum_{j\in J} n_t^j x_i^j = n_i^0, \quad \forall i = 1,\ldots,n-1.$$

Consider finally the associated tangent at $\mathbf{x}^{j^*}$ *for some* $j^* \in J$

$$T(\mathbf{x}) = G(\mathbf{x}^{j^*}) + \sum_{i=1}^{n-1} G_{x_i}(\mathbf{x}^{j^*})(x_i - x_i^{j^*}).$$

The state is stable if and only if for all $\mathbf{x} \in X$

$$T(\mathbf{x}) \leq G(\mathbf{x}) \tag{7}$$

$$T(\mathbf{x}) = G(\mathbf{x}^j) + \sum_{i=1}^{n-1} G_{x_i}(\mathbf{x}^j)(x_i - x_i^j), \quad \forall j \in J. \tag{8}$$

Proof. We first show the necessity and then the sufficiency.

1. Necessity of (7)
This part of the proof proceeds by contraposition.
Suppose first that for some $\bar{\mathbf{x}} \in X$ we have

$$T(\bar{\mathbf{x}}) > G(\bar{\mathbf{x}}).$$

Consider now the implications of Lemma 12 along the line through $\mathbf{x}^{j^*}$ and $\bar{\mathbf{x}}$. We can find a point $\hat{\mathbf{x}}$ and $\kappa \in (0,1)$ such that

$$\mathbf{x}^{j^*} = \kappa\hat{\mathbf{x}} + (1-\kappa)\bar{\mathbf{x}} \quad \text{and} \quad G(\mathbf{x}^{j^*}) > \kappa G(\hat{\mathbf{x}}) + (1-\kappa)G(\bar{\mathbf{x}}).$$

Since $\mathbf{x}^{j^*} \in \text{int}(X)$ and the point $\hat{\mathbf{x}}$ can be found arbitrarily close to $\mathbf{x}^{j^*}$ we also have $\hat{\mathbf{x}} \in X$. We can therefore reduce the Gibbs free energy by replacing phase j^* with two phases of composition $\bar{\mathbf{x}}$ and $\hat{\mathbf{x}}$. According to Theorem 4 the considered state is not stable.
2. Necessity of (8)
We will show that (8) follows from the equality of chemical potentials.
Recall that equality of chemical potentials in all phases

$$\mu_i^j = \mu_i^{j^*}, \quad \forall i = 1,\ldots,n, \quad j \in J$$

is a necessary condition for stability. From the equality of chemical potentials ($\mu_i^j = \mu_i^{j^*}$ and $\mu_n^j = \mu_n^{j^*}$) and relation (6) it follows that the slopes are equal

$$G_{x_i}(\mathbf{x}^j) = G_{x_i}(\mathbf{x}^{j^*}).$$

From (5) and $\mu_n^j = \mu_n^{j^*}$ we also obtain

$$G(\mathbf{x}^j) - \sum_{k=1}^{n-1} x_k G_{x_k}(\mathbf{x}^j) = G(\mathbf{x}^{j^*}) - \sum_{k=1}^{n-1} x_k G_{x_k}(\mathbf{x}^{j^*}).$$

Combining the last two equalities we obtain for any $j \in J$ and all $\mathbf{x}$

$$G(\mathbf{x}^j) + \sum_{i=1}^{n-1} G_{x_i}(\mathbf{x}^j)(x_i - x_i^j) = G(\mathbf{x}^{j^*}) + \sum_{i=1}^{n-1} G_{x_i}(\mathbf{x}^{j^*})(x_i - x_i^{j^*}) = T(\mathbf{x}).$$

3. Sufficient stability criterion
The (extensive) Gibbs free energy of the considered state is given by $\underline{G}^J = \sum_{j\in J} n_t^j G(\mathbf{x}^j)$. As a direct consequence of (8) we have $T(\mathbf{x}^j) = G(\mathbf{x}^j)$ for all $j \in J$ and therefore $\underline{G}^J = \sum_{j\in J} n^0{}_t T(\mathbf{x}^j) = n^0{}_t T(\mathbf{x}^0)$. Suppose now that the Gibbs tangent plane lies below the Gibbs surface. Consider any other collection of phases with index set K, composition $\mathbf{x}^k \in X$ and total mole numbers n_t^k such that

$$\sum_{k\in K} n_t^k x_i^k = n_{i,0}, \quad \forall i = 1,\ldots,n.$$

Since $\underline{G}^K = \sum_{k\in K} n_t^k G(\mathbf{x}^k) \geq \sum_{k\in K} n_t^k T(\mathbf{x}^k) = n^0{}_t T(\mathbf{x}^0)$, we obtain $\underline{G}^K \geq \underline{G}^J$ or by Theorem 4 the state indexed by $j \in J$ is stable. □

Assuming the existence of second partial derivatives of G with respect to x_i, a direct consequence of Theorem 5 is that for a stable state the Gibbs surface is locally convex, i.e., its Hessian is positive semidefinite at $\mathbf{x}^j$ for all j, see also Ref. 26. This is a necessary but not sufficient stability criterion.

Proof of duality stability criterion

Prior to a proof of the equivalence of dual solutions with the Gibbs tangent plane criterion we establish that the extrema are indeed attained as asserted in the definitions of the dual function and dual problem. We will refer to the program defining the dual function

$$\min_{\mathbf{x}\in X} \left(G(\mathbf{x}) + \sum_{i=1}^{n-1} \lambda_i (x_i^0 - x_i) \right)$$

as the *inner program*.

Proposition 13 (Existence of Minimum of the Inner Program). *The infimum of the inner program is attained for any* $\boldsymbol{\lambda} \in \mathrm{R}^{n-1}$.

Proof. The feasible set is compact and since G is by assumption continuous on the feasible set, for fixed $\boldsymbol{\lambda}$ the objective function $G(\mathbf{x}) + \sum_{i=1}^{n-1} \lambda_i (x_i^0 - x_i)$ is also continuous. Therefore, by Theorem 8 the infimum is attained. □

Proposition 14 (Dual Function is Coercive). *The negative of the dual function is coercive, i.e.,*

$$q(\boldsymbol{\lambda}) \to -\infty \quad \textit{for } \|\boldsymbol{\lambda}\| \to +\infty.$$

Proof. Note first that by the definition of the norm $\|\boldsymbol{\lambda}\| \to +\infty$ if and only if there exists $i^* \in \{1,\ldots,n-1\}$ such that $\lambda_{i^*} \to \pm\infty$. Choose $\bar{\mathbf{x}}$, such that

$$\bar{x}_{i^*} = x^0_{i^*} + \varepsilon \quad \text{and} \quad \bar{\mathbf{x}}_i = x^0_i, \quad \forall i \in \{\{1,\ldots,n-1\} : i \neq i^*\}.$$

Since $\mathbf{x}^0 \in \text{int}(X)$, for sufficiently small $|\varepsilon| > 0$ we have $\bar{\mathbf{x}} \in X$. Also

$$L(\bar{\mathbf{x}}, \boldsymbol{\lambda}) = G(\bar{\mathbf{x}}) + \sum_{i=1}^{n-1} \lambda_i (x^0_i - \bar{x}_i) G(\bar{\mathbf{x}}) - \varepsilon \lambda_{i^*}.$$

For $\lambda_{i^*} \to +\infty$ pick $\varepsilon > 0$ and for $\lambda_{i^*} \to -\infty$ pick $\varepsilon < 0$. In either case, since G is by assumption bounded on X, we obtain $L(\mathbf{x}, \boldsymbol{\lambda}) \to -\infty$. Since $q(\boldsymbol{\lambda}) = \min_{\mathbf{x}\in X} L(\mathbf{x},\boldsymbol{\lambda})$ also $q(\boldsymbol{\lambda}) \to -\infty$. □

Proposition 15 (Dual Function is Uniformly Continuous). *The dual function $q(\boldsymbol{\lambda})$ is uniformly continuous on* R^{n-1}.

Continuity follows from well-known results in parametric optimization[27] but here we present a proof that takes into account the special structure.

Proof. Without loss of generality we use the infinity norm in the definition of continuity. Take any $\bar{\boldsymbol{\lambda}} \in \mathrm{R}^{n-1}$ and any $\varepsilon > 0$. Take any $\hat{\boldsymbol{\lambda}} \in \mathrm{R}^{n-1}$ such that $\|\hat{\boldsymbol{\lambda}} - \bar{\boldsymbol{\lambda}}\| \leq \varepsilon/(n-1)$. By the definition of the infinity norm we obtain $|\hat{\lambda}_i - \bar{\lambda}_i| \leq \varepsilon/(n-1)$, $\forall i = 1,\ldots,n-1$. Take now any optimal solution points of the inner programs for $\bar{\boldsymbol{\lambda}}$ and $\hat{\boldsymbol{\lambda}}$ and denote them $\bar{\mathbf{x}}$ and $\hat{\mathbf{x}}$ respectively. By definition

$$q(\bar{\boldsymbol{\lambda}}) \leq L(\hat{\mathbf{x}}, \bar{\boldsymbol{\lambda}}) = G(\hat{\mathbf{x}}) + \sum_{i=1}^{n-1} \bar{\lambda}_i (x^0_i - \hat{x}_i)$$

and thus

$$\begin{aligned} q(\bar{\boldsymbol{\lambda}}) &\leq G(\hat{\mathbf{x}}) + \sum_{i=1}^{n-1} \bar{\lambda}_i (x^0_i - \hat{x}_i) \\ &\quad + \sum_{i=1}^{n-1} \hat{\lambda}_i (x^0_i - \hat{x}_i) - \sum_{i=1}^{n-1} \hat{\lambda}_i (x^0_i - \hat{x}_i) \\ &= G(\hat{\mathbf{x}}) + \sum_{i=1}^{n-1} \hat{\lambda}_i (x^0_i - \hat{x}_i) + \sum_{i=1}^{n-1} (\bar{\lambda}_i - \hat{\lambda}_i)(x^0_i - \hat{x}_i) \end{aligned}$$

and by the definition of $q(\hat{\boldsymbol{\lambda}})$

$$q(\bar{\boldsymbol{\lambda}}) \leq q(\hat{\boldsymbol{\lambda}}) + \sum_{i=1}^{n-1} (\bar{\lambda}_i - \hat{\lambda}_i)(x^0_i - \hat{x}_i).$$

Therefore, together with $|\hat{\lambda}_i - \bar{\lambda}_i| \leq \varepsilon/(n-1)$ and $\hat{x}_i$, $x^0_i \in [0,1]$ we obtain

$$q(\bar{\boldsymbol{\lambda}}) \leq q(\hat{\boldsymbol{\lambda}}) + \varepsilon.$$

Similarly,

$$q(\hat{\boldsymbol{\lambda}}) \leq q(\bar{\boldsymbol{\lambda}}) + \varepsilon.$$

Combining the last two inequalities we establish continuity and since the distance between $\hat{\boldsymbol{\lambda}}$ and $\bar{\boldsymbol{\lambda}}$ was chosen independently of $\bar{\boldsymbol{\lambda}}$ also uniform continuity. □

Proposition 16 (Existence of Maximum in Dual Problem). *The supremum of*

$$G^d = \sup_{\boldsymbol{\lambda}\in\mathrm{R}^{n-1}} q(\boldsymbol{\lambda}).$$

is attained.

Proof. By Proposition 14 the negative of the dual function is coercive and by Proposition 15 it is continuous. Therefore, by Theorem 8 the supremum is attained. □

We now prove the main result of our article: Theorem 6.

Theorem 6 (Dual Extremum Principle). *Consider a system of n species at a given temperature T, pressure P, an overall mole number $n^0_i > 0$ for the species $i = 1,\ldots,n$ and corresponding overall composition $\mathbf{x}^0 \in \text{int}(X)$. Take any solution $\boldsymbol{\lambda}^* \in \mathrm{R}^{n-1}$ of the dual problem (9) and any $\mathbf{x}^* \in X$, such that $L(\mathbf{x}^*, \boldsymbol{\lambda}^*) = q(\boldsymbol{\lambda}^*) = G^d$. The hyperplane*

$$T(\mathbf{x}) = G(\mathbf{x}^*) + \sum_{i=1}^{n-1} \lambda^*_i (x_i - x^*_i).$$

is a supporting hyperplane of G on *X*, i.e.

$$T(\mathbf{x}) \leq G(\mathbf{x}), \qquad \forall \mathbf{x} \in X.$$

Moreover

$$\mathbf{x}^0 \in \text{conv}(X^*),$$

where X is the set of common points between the Gibbs surface and the hyperplane, i.e.,*

$$X^* = \{\mathbf{x} \in X : T(\mathbf{x}) = G(\mathbf{x})\}.$$

Finally, a state described by a collection of phases with index set J and compositions $\mathbf{x}^j \in X$ and with nonzero total mole number in each phase $n^j_t > 0$ such that

$$\mathbf{x}^0 \in \text{conv}(X^J), \quad \textit{where } X^J = \{\mathbf{x}^j \in X : j \in J\},$$

is stable if and only if $\mathbf{x}^j \in X^*$ for all $j \in J$, or equivalently all $L(\mathbf{x}^j,\boldsymbol{\lambda}^*) = G^d$.

Proof. Note first in general, neither $\boldsymbol{\lambda}^*$ nor X^* are unique. We first show that the solution of the dual gives a supporting hyperplane, and then that the overall composition lies in the convex hull of the points common to the hyperplane and the Gibbs surface. Finally we show the equivalence of the dual solution to stability.

1. $T(\mathbf{x}) \leq G(\mathbf{x})$, $\forall \mathbf{x} \in X$.

By the definition of the Lagrangian function

$$L(\mathbf{x}, \boldsymbol{\lambda}^*) = G(\mathbf{x}) + \sum_{i=1}^{n-1} \lambda^*_i (x^0_i - x_i), \quad \forall \mathbf{x} \in X$$

and therefore since $\mathbf{x}^* \in \arg\min_{\mathbf{x}\in X} L(\mathbf{x}, \boldsymbol{\lambda}^*)$

$$G(\mathbf{x}^*) + \sum_{i=1}^{n-1} \lambda^*_i (x^0_i - x^*_i) \leq G(\mathbf{x}) + \sum_{i=1}^{n-1} \lambda^*_i (x^0_i - x_i), \quad \forall \mathbf{x} \in X,$$

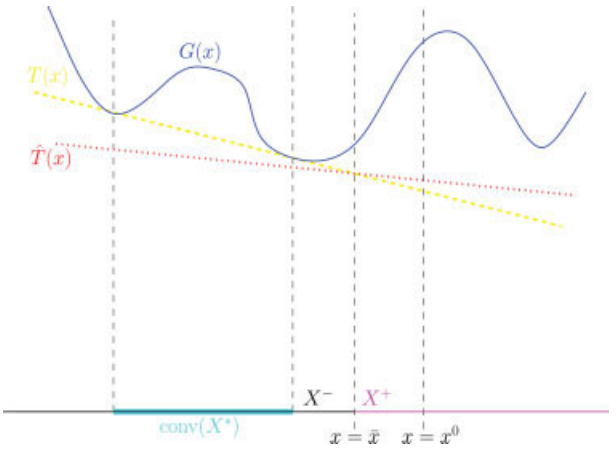

Figure A3. Illustration of the proof of $\mathbf{x}^0 \in \text{conv}(X^*)$ in Theorem 6.

[Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

and thus

$$G(\mathbf{x}^*) + \sum_{i=1}^{n-1} \lambda_i^*(x_i - x_i^*)(\mathbf{x}) \quad \forall \mathbf{x} \in X,$$

or

$$T(\mathbf{x}) \le G(\mathbf{x}) \quad \forall \mathbf{x} \in X.$$

Note also that $T(\mathbf{x}^0) = G^d$ by Proposition 3.

2. $\mathbf{x}^0 \in \text{conv}(X^*)$.

We will first argue that conv(X^*) is nonempty, convex, and compact, so that we can invoke Corollary 11. By its definition conv(X^*) is convex. By its definition X^* is compact, i.e., bounded and closed, and by Lemma 9 so is conv(X^*). The set of common points X^* cannot be empty, for otherwise since $T(\mathbf{x}) \le G(\mathbf{x})$ we would have $T(\mathbf{x}) < G(\mathbf{x})$ for all $\mathbf{x} \in X$. In that case the solution of the dual would not give the highest intercept among the supporting hyperplanes, violating Proposition 3. Since $X^* \subset \text{conv}(X^*)$, also conv($X^*$) is nonempty.

The rest of this part of the proof proceeds by contraposition, and is illustrated graphically in Figure A3.

Assume that

$$\mathbf{x}^0 \notin \text{conv}(X^*)$$

By Corollary 11 there exist $\mathbf{c}, \bar{\mathbf{x}} \in \mathrm{R}^{n-1}$ such that

$$\begin{aligned} \mathbf{c}^{\mathrm{T}}\bar{\mathbf{x}} &> \mathbf{c}^{\mathrm{T}}\mathbf{x}^0 \\ \mathbf{c}^{\mathrm{T}}\bar{\mathbf{x}} &< \mathbf{c}^{\mathrm{T}}\mathbf{x} \quad \forall \mathbf{x} \in \text{conv}(X^*). \end{aligned}$$

Define

$$\begin{aligned} X^+ &= \{\mathbf{x} \in X : \mathbf{c}^{\mathrm{T}}\bar{\mathbf{x}} \ge \mathbf{c}^{\mathrm{T}}\mathbf{x}\} \\ X^- &= \{\mathbf{x} \in X : \mathbf{c}^{\mathrm{T}}\bar{\mathbf{x}} \le \mathbf{c}^{\mathrm{T}}\mathbf{x}\}. \end{aligned}$$

Clearly $\mathbf{x}^0 \in X^+$ and $X = X^+ \cup X^-$. Also, conv(X^*) $\subset X^-$ and therefore $X^* \subset X^-$. Since $X^+ \cap X^- = \emptyset$, also $X^+ \cap X^* = \emptyset$ and therefore by definition of X^* together with $T(\mathbf{x}) \le G(\mathbf{x})$ we have

$$T(\mathbf{x}) < G(\mathbf{x}), \quad \forall \mathbf{x} \in X^+. \tag{A1}$$

Note also that X^+ is compact since it is a closed subset of X.

Since X^+ is compact by (A1) we obtain as a consequence of Theorem 8

$$\delta \equiv \min_{\mathbf{x} \in X^+} (G(\mathbf{x}) - T(\mathbf{x})) > 0.$$

Similarly, there exists Δ, such that

$$0 \le \mathbf{c}^{\mathrm{T}}\bar{\mathbf{x}} - \mathbf{c}^{\mathrm{T}}\mathbf{x} < \Delta, \quad \forall \mathbf{x} \in X^+.$$

Take $\varepsilon = \frac{\delta}{\Delta} > 0$ and the perturbed hyperplane

$$\hat{T}(\varepsilon, \mathbf{x}) = T(\mathbf{x}) + \varepsilon(\mathbf{c}^{\mathrm{T}}\bar{\mathbf{x}} - \mathbf{c}^{\mathrm{T}}\mathbf{x}) = T(\mathbf{x}) + \frac{\delta}{\Delta}(\mathbf{c}^{\mathrm{T}}\bar{\mathbf{x}} - \mathbf{c}^{\mathrm{T}}\mathbf{x}).$$

By the definition of Δ

$$\hat{T}(\varepsilon, \mathbf{x}) \le T(\mathbf{x}) + \delta \quad \forall \mathbf{x} \in X^+$$

and therefore by the definition of δ there exists $\varepsilon > 0$ such that

$$\hat{T}(\varepsilon, \mathbf{x}) \le G(\mathbf{x}) \quad \forall \mathbf{x} \in X^+$$

Moreover by definition of $\hat{T}$ and X^-

$$\hat{T}(\varepsilon, \mathbf{x}) < T(\mathbf{x}) \qquad \forall \mathbf{x} \in X^-$$

and since $T(\mathbf{x}) \le G(\mathbf{x})$ for all $\mathbf{x} \in X$

$$\hat{T}(\varepsilon, \mathbf{x}) < G(\mathbf{x}) \qquad \forall \mathbf{x} \in X^-.$$

So the perturbed hyperplane $\hat{T}$ is a supporting hyperplane for G on X. Finally $\hat{T}(\varepsilon, \mathbf{x}^0) > T(\mathbf{x}^0)$, violating the requirement of maximal intercept of Proposition 3. Therefore, $\mathbf{x}^*$, $\boldsymbol{\lambda}^*$ are not optimal solutions of the dual.

3. Stability.

By Proposition 16 a solution $(\mathbf{x}^*, \boldsymbol{\lambda}^*)$ of the dual exists. We will first show that $\bar{\mathbf{x}} \in X^*$ is equivalent to $\bar{\mathbf{x}} \in X$ being a solution of the inner program for $\boldsymbol{\lambda}^*$. Note that

$$L(\bar{\mathbf{x}}, \boldsymbol{\lambda}^*) = L(\mathbf{x}^*, \boldsymbol{\lambda}^*) = G^d,$$

is equivalent to

$$G(\bar{\mathbf{x}}) + \sum_{i=1}^{n-1} \lambda_i^*(x_i^0 - \bar{x}_i) = G(\mathbf{x}^*) + \sum_{i=1}^{n-1} \lambda_i^*(x_i^0 - x_i^*),$$

or

$$G(\bar{\mathbf{x}}) = G(\mathbf{x}^*) + \sum_{i=1}^{n-1} \lambda_i^*(\bar{x}_i - x_i^*) = T(\bar{\mathbf{x}}).$$

Consider now the sufficiency. Take any state of phases with index set J, compositions $\mathbf{x}^j \in X$ such that $G(\mathbf{x}^j) = T(\mathbf{x}^j)$, and total mole numbers n_i^j such that

$$\sum_{j\in J} n_t^j x_i^j = n_{i,0} \qquad \forall i = 1,\ldots,n. \tag{A2}$$

Such states exist since $\mathbf{x}^0 \in \text{conv}(X^*)$. The overall (extensive) Gibbs free energy of the state indexed by J is given by

$$\underline{G}^J = \sum_{j\in J} n_t^j G(\mathbf{x}^j) = \sum_{j\in J} n_t^j T(\mathbf{x}^j) = n_t^0 T(\mathbf{x}^0) = n_t^0 G^d,$$

where the third equality is a direct consequence of the linearity of T and (A2). Consider any other state comprising phases with index set K, compositions $\mathbf{x}^k \in X$ and total mole numbers n_t^k such that

$$\sum_{k\in K} n_t^k x_i^k = n_{i,0} \qquad \forall i = 1,\ldots,n. \tag{A3}$$

Since

$$\underline{G}^K = \sum_{k\in K} n_t^k G(\mathbf{x}^k) \geq \sum_{k\in K} n_t^k T(\mathbf{x}^k)$$

by the linearity of T and (A3) we obtain

$$\underline{G}^K \geq n_t^0 T(\mathbf{x}^0) = n_t^0 G^d = \underline{G}^J.$$

Since $\underline{G}^K \geq \underline{G}^J$, by Theorem 4 the state indexed by J is stable.

Consider finally the necessity. Suppose for the state indexed by K for some $k^* \in K$ we have $G(\mathbf{x}^{k^*}) > T(\mathbf{x}^{k^*})$. Similarly to above we obtain

$$\underline{G}^K = \sum_{k\in K} n_t^k G(\mathbf{x}^k) > \sum_{k\in K} n_t^k T(\mathbf{x}^k) = n_t^0 T(\mathbf{x}^0) = n_t^0 G^d = \underline{G}^J,$$

or the state indexed by K is not stable. □

Convergence proof

The convergence of Algorithm 1 is a direct consequence of the results in Ref. 13. For the sake of completeness we give a convergence proof here as well.

Theorem 17 (Finite Termination). *Algorithm 1 terminates finitely, i.e., for any $\varepsilon > 0$ after a finite number of iterations it furnishes a pair $(\mathbf{x}^*, \boldsymbol{\lambda}^*)$ that satisfies $q^* - \varepsilon \leq L(\mathbf{x}^*, \boldsymbol{\lambda}^*) \leq q^*$.*

Proof. The second part of the inequality is given by the validity of the lower bounding problem. Take any $\varepsilon > 0$ and consider now an infinite sequence of points $\{\boldsymbol{\lambda}^j\}$ furnished by the upper bounding problem. Since by the initialization we have restricted $\boldsymbol{\lambda}$ to a compact set, a convergent subsequence $\{\boldsymbol{\lambda}^{j_k}\}$ exists. Therefore by definition of convergence and the continuity of the dual function q we can find finite K_1 such that

$$|q(\boldsymbol{\lambda}^{j_k}) - q(\boldsymbol{\lambda}^{j_{K_1}})| < \varepsilon/3, \quad \forall k > K_1 \tag{A4}$$

and therefore

$$\text{LBD}^{j_k} > q(\boldsymbol{\lambda}^{j_{K_1}}) - \varepsilon/3, \quad \forall k > K_1, \tag{A5}$$

where LBD^j denotes the lower bound obtained at iteration j. Similarly there exists finite K_2 such that

$$\sum_{i=1}^{n-1} \left(\lambda_i^{j_k} - \lambda_i^{j_{K_2}}\right) x_i^{j_{K_2}} < \varepsilon/3, \quad \forall k > K_2.$$

Since for all $k > K_2$ we have $\mathbf{x}^{j_{K_2}} \in X^d$ we also obtain:

$$\text{UBD}^{j_K} < G(\mathbf{x}^{j_{K_2}}) + \sum_{i=1}^{n-1} \lambda_i^{j_{K_2}} \left(x_i^0 - x_i^{j_k}\right), \quad \forall k > K_2.$$

where UBD^j denotes the upper bound obtained at iteration j. Combining the last two inequalities:

$$\begin{aligned}\text{UBD}^{j_K} &< G(\mathbf{x}^{j_{K_2}}) + \sum_{i=1}^{n-1} \lambda_i^{j_{K_2}} \left(x_i^0 - x_i^{j_{k_2}}\right) + \varepsilon/3 \\ &= q(\boldsymbol{\lambda}^{j_{K_2}}) + \varepsilon/3, \quad \forall k > K_2.\end{aligned}$$

which together with (A4) and (A5) gives

$$\text{UBD}^{j_k} < \text{LBD}^{j_k} + \varepsilon \qquad \forall k > \max\{K_1, K_2\}. \qquad \square$$

Manuscript received Jan. 21, 2007, and revision received May 5, 2007.