

The Computation of Constrained and Unconstrained
Equilibrium Compositions of Ideal Gas Mixtures
using Gibbs Function Continuation

by

Stephen B. Pope

FDA 03-02

December, 2003

The Computation of Constrained and Unconstrained Equilibrium Compositions of Ideal Gas Mixtures using Gibbs Function Continuation

by

Stephen B. Pope

Abstract

We consider the computation of the chemical-equilibrium state of an ideal gas mixture of given elemental composition, with and without linear constraints on the species present. While computer programs to solve these problems have been available for over thirty years, it has been found that they are not always successful in determining the chemical equilibrium state when constraints are present. In this paper the constrained equilibrium problem is examined in detail. It is shown that the specified constraints can be replaced by a (generally simpler) set of reduced constraints. For fixed temperature and pressure, the equilibrium composition is that which minimizes the Gibbs function subject to these constraints. The solution to this minimization problem is known in terms of Lagrange multipliers (or constraint potentials), which are determined as the solution to a set of non-linear equations. It is shown that Newton's method can fail to obtain the solution because the iteration matrix can be singular. The new method of *Gibbs function continuation* is introduced which is guaranteed to obtain the solution. Based on this, a practical algorithm is presented which has been implemented and tested in a Fortran code. It is claimed that this algorithm is assured to determine the (constrained or unconstrained) equilibrium state for all well-posed problems.

Contents

Abstract	2
Nomenclature	5
1 Introduction	8
1.1 Problem Considered	8
1.2 Motivation and Previous Work	8
1.3 Objective	9
1.4 Outline of the Paper	9
1.5 Attribution of Ideas	10
2 Formulation of Constrained Equilibrium Problems	10
2.1 Species and Elements	10
2.2 Basic Constraints	11
2.3 Modified and Reduced Constraints	12
2.4 The Feasible Region	15
2.5 Thermodynamic Properties	17
2.6 Constrained Equilibrium	18
3 Existence and Uniqueness	19
4 Constraint Potentials	20
4.1 Basic Constraints	20
4.2 Reduced Constraints	22
4.3 Normalization Conditions	24
5 Solution for the Constraint Potentials	25
5.1 Equations to be Solved	25
5.2 Infinitesimals	26
5.3 Newton's Method	27
5.4 Gibbs Function Continuation	27
6 Solution Algorithm	30
6.1 Overview	30
6.2 Well-Posedness and Perturbation	31
6.3 Initial Conditions	33
6.4 Time Stepping	34

6.5	Newton's Method	36
7	Fixed Enthalpy and Pressure	38
8	Conclusions	39
A	The Max-Min Composition	40
B	Demonstration that D is Strictly Positive	41
C	Composition of Air and Methane	43

Nomenclature

Roman

A	modified constraint transformation matrix ($n_b \times n_c$).
a	atom vector: species k consists of a_k atoms
\mathbf{a}^u	atom vector for undetermined species
B	basic constraint matrix ($n_s \times n_c$)
$\tilde{\mathbf{B}}$	modified constraint matrix ($n_s \times n_b$)
$\hat{\mathbf{B}}$	reduced constraint matrix ($n_{su} \times n_{rc}$)
\mathbf{B}^g	general linear constraint matrix ($n_s \times n_g$)
\mathcal{B}	n_b -dimensional constraint subspace
\mathcal{B}^\perp	orthogonal complement of \mathcal{B}
$C_{p,k}$	molar constant-pressure specific heat of species k : ergs/(mole K)
\bar{C}_{pc}	constrained molar specific heat of the mixture: ergs/(mole K)
\bar{C}_p	molar constant-pressure specific heat of the mixture: ergs/(mole K)
c	basic constraint vector ($n_c \times 1$)
$\tilde{\mathbf{c}}$	modified constraint vector ($n_c \times 1$)
$\hat{\mathbf{c}}$	reduced constraint vector ($n_{rc} \times 1$)
D	denominator defined by Eq.(108)
E_{kj}	number of atoms of element j in a molecule of species k
E	element matrix ($n_s \times n_e$) with components E_{kj}
\mathcal{E}	n_e -dimensional element subspace
\mathbf{e}_k	basis vector for species k
\mathcal{F}	feasible region
$\bar{\mathcal{F}}$	interior feasible region
G	Gibbs function of the mixture: ergs
\tilde{G}	normalized Gibbs function $G/(\mathcal{R}T)$
$\tilde{\mathbf{g}}$	normalized molar Gibbs function of species ($n_s \times 1$)
$\tilde{\mathbf{g}}^d$	normalized molar Gibbs function of determined species ($n_{sd} \times 1$)
$\tilde{\mathbf{g}}^u$	normalized molar Gibbs function of undetermined species ($n_{su} \times 1$)
$\bar{\mathbf{g}}(s)$	normalized molar Gibbs function of undetermined species in pseudo-time ($n_{su} \times 1$)

\mathbf{H}	matrix ($n_{su} \times n_{rc}$) defined by Eq.(87)
H	enthalpy of the mixture: ergs
\bar{H}	molar enthalpy of the mixture: ergs/mole
$h_k(T)$	molar enthalpy of species k : ergs/mole
$\mathbf{I}_{(m \times n)}$	the ($m \times n$) identity matrix
\mathbf{N}	moles of species ($n_s \times 1$): mole
\mathbf{N}^{CE}	constrained equilibrium composition ($n_s \times 1$): mole
\mathbf{N}^d	moles of determined species ($n_{sd} \times 1$): mole
\mathbf{N}^e	moles of atoms of element ($n_e \times 1$): mole
\mathbf{N}^g	min-g composition: mole
\mathbf{N}^{mm}	max-min composition: mole
\mathbf{N}^u	moles of undetermined species ($n_{su} \times 1$): mole
\mathbf{N}^{up}	upper bound on species moles: mole
\bar{N}	moles of all species: mole
\bar{N}^d	moles of all determined species: mole
\bar{N}^u	moles of all undetermined species: mole
N^a	moles of all atoms: mole
N^{au}	moles of atoms in undetermined species: mole
N_{\min}	minimum moles of species in the max-min composition: mole
n_b	number of linearly independent basic constraints; rank(\mathbf{B})
n_c	number of basic constraints
n_{cs}	number of constrained species
n_e	number of elements
n_{ed}	number of determined elements
n_{eu}	number of undetermined elements
n_g	number of general linear constraints
n_{gi}	number of independent general linear constraints
n_{rc}	number reduced constraints
n_s	number of species
n_{sd}	number of determined species
n_{su}	number of undetermined species
p	pressure: dynes/cm ²
p_{atm}	pressure of one standard atmosphere: dynes/cm ²
\mathbf{Q}	normalization condition vector ($n_{su} \times 1$)
q	normalization condition residual
\mathcal{R}	universal gas constant: ergs/(mole K)

\mathbf{r}	residual vector ($n_{rc} \times 1$)
S	entropy: ergs/K
$S_k(T, p)$	molar specific entropy of species k : ergs/(mole K)
$S_k^o(T)$	standard state molar specific entropy of species k : ergs/(mole K)
\mathcal{S}	n_s -dimensional species space
s	psuedo-time
T	temperature: K
$\mathbf{v}(s)$	constraint vector defined by Eq.(80) ($n_{rc} \times 1$)
\mathbf{X}	mole fractions of species ($n_s \times 1$)
\mathbf{X}^u	mole fractions of undetermined species ($n_{su} \times 1$)
$\bar{\mathbf{X}}(s)$	mole fractions of undetermined species in pseudo-time ($n_{su} \times 1$)
$\mathbf{Y}(s)$	diagonal matrix formed from $\mathbf{y}(s)$ ($n_{su} \times n_{su}$)
$\mathbf{y}(s)$	square root of mole fractions of undetermined species in pseudo-time ($n_{su} \times 1$)

Greek

Δs	pseudo-time step
ε_e	lower bound on normalized element moles
ε_s	lower bound on normalized species moles
ε_{tol}	error tolerance
$\boldsymbol{\lambda}$	basic constraint potential (or Lagrange) vector ($n_c \times 1$)
$\hat{\boldsymbol{\lambda}}$	reduced constraint potential (or Lagrange) vector ($n_{rc} \times 1$)
$\bar{\boldsymbol{\lambda}}(s)$	reduced constraint potential in pseudo-time ($n_{rc} \times 1$)

1 Introduction

1.1 Problem Considered

The problem considered is the numerical determination of constrained equilibrium compositions of ideal gas mixtures. Three types of equality constraints are considered:

1. Constraints on the moles of some specified species
2. Constraints on the moles of atoms of each element in the system
3. General linear equality constraints on the moles of species.

In addition to the specified constraints, the constrained equilibrium composition depends on two thermodynamic state variables. The primary problem considered is that in which the pressure and temperature are specified. In this case, by definition, the constrained equilibrium mixture is that which satisfies the constraints and minimizes the Gibbs function. The second problem considered is that in which the pressure and enthalpy are specified. In this case, the constrained equilibrium mixture is that which maximizes the entropy subject to the constraints.

1.2 Motivation and Previous Work

If only the second type of constraint (on the elements) is imposed, then the problem becomes the standard one of determining the (unconstrained) chemical equilibrium composition. Several approaches have been developed to solve this problem, notably those of Gordon and McBride (1971) and Reynolds (1986). The latter is implemented in the code STANJAN.

The computation of constrained equilibrium compositions is required in the rate-controlled constrained equilibrium (RCCE) approach of Keck and Gillespie (1971), which is reviewed by Keck (1990), and which has been further developed by Hamiroune et al. (1998) and Tang and Pope (2002). STANJAN allows for constraints of the second type (on species), but not of the third type (general). Both STANJAN and the NASA equilibrium code (Gordon and McBride 1971) have been extended to general constraints and tested by Bishnu et al. (1997). In the algorithm used in STANJAN, the computational work increases with the number of constrained species.

However, as shown in Section 2.3, it is possible to devise an algorithm in which the work *decreases* as the number of constrained species increases.

The numerical task of determining chemical equilibrium compositions is notoriously ill-conditioned; and the difficulties can be greater when constraints are present. Bishnu et al. (1997) report that both the NASA code and STANJAN fail to determine the equilibrium composition in some cases. Similarly, in previous work on RCCE (Tang and Pope 2002) we encountered well-posed problems for which STANJAN is unable to determine the equilibrium composition.

1.3 Objective

Based on the above motivation, the objective of the work described here is to develop an algorithm for determining constrained equilibrium compositions of ideal gas mixtures which:

1. Allows for all three types of constraints
2. Treats constrained species efficiently
3. Is guaranteed to solve all well-posed problems.

1.4 Outline of the Paper

In Section 2 we develop the mathematical formulation of the constrained-equilibrium problem. The problem is posed in terms of minimizing the Gibbs function subject to the given constraints. It is shown that the constraints can be simplified to a *reduced constraint equation*. In Section 3 it is shown that a unique constrained equilibrium mixture exists for every well-posed problem, and a method for determining well-posedness is presented. The minimization of the Gibbs function subject to constraints is considered in Section 4 using the method of Lagrange multipliers, which are referred to as *constraint potentials*. This reduces the problem to solving a set of non-linear equations for the constraint potentials. The numerical solution of these equations is considered in Section 5. It is shown that Newton's method can fail, because the iteration matrix can be singular. The method of *Gibbs function continuation* is introduced which is guaranteed to obtain the constrained equilibrium solution. An algorithm implementing Gibbs function continuation is described in Section 6. This algorithm has been implemented in a Fortran code (Pope

2003) and tested comprehensively. The extension of this algorithm to determine the constrained equilibrium mixture at constant pressure and enthalpy is described in Section 7. Conclusions are drawn in the final section.

1.5 Attribution of Ideas

The formulation of the unconstrained equilibrium problem in terms of Lagrange multipliers (or *element potentials*) is due to Reynolds (1986) (and to earlier works cited therein). The extension of this formulation to the constrained equilibrium problem is due to Keck (1990) (see also Hamiroune et al. 1998). For completeness, this formulation is re-presented here (in Sections 2–4).

The principal original contributions of the present work are the reduction of the constraints (Sections 2.3 and 4.2), and the introduction of the method of Gibbs function continuation (Sections 5.4 and 6).

2 Formulation of Constrained Equilibrium Problems

2.1 Species and Elements

The ideal gas mixture consists of n_s species. The number of moles of species k is denoted by N_k , and the total moles of species is

$$\bar{N} \equiv \sum_{k=1}^{n_s} N_k. \quad (1)$$

The mole fractions of the species are then given by

$$X_k \equiv N_k/\bar{N}. \quad (2)$$

Obviously, from this definition it follows that the species mole fractions sum to unity:

$$\sum_{k=1}^{n_s} X_k = 1. \quad (3)$$

The species space \mathcal{S} is defined to be the real n_s -dimensional Euclidean space with orthonormal basis vectors \mathbf{e}_k , $k=1,2,\dots,n_s$. Thus the species moles are represented by the vector

$$\mathbf{N} \equiv \mathbf{e}_k N_k, \quad (4)$$

where the summation convention applies.

There are n_e different elements in the system. The number of atoms of element j in one molecule of species k is denoted by E_{kj} . The element matrix \mathbf{E} , whose general component is E_{kj} , is an $n_s \times n_e$ matrix with non-negative integer components. Every row of \mathbf{E} has at least one non-zero entry. The j -th column of \mathbf{E} is a vector in \mathcal{S} corresponding to the j -th element; and the element subspace is defined by

$$\mathcal{E} \equiv \text{span}(\mathbf{E}). \quad (5)$$

The number of moles of atoms of element j in the mixture is

$$N_j^e = E_{kj} N_k, \quad (6)$$

or equivalently

$$\mathbf{N}^e = \mathbf{E}^T \mathbf{N}. \quad (7)$$

The number of atoms in a molecule of species k is denoted by

$$a_k \equiv \sum_{j=1}^{n_e} E_{kj}. \quad (8)$$

The total number of moles of atoms in the system is

$$\begin{aligned} N^a &\equiv \sum_{j=1}^{n_e} N_j^e \\ &= \mathbf{a}^T \mathbf{N}. \end{aligned} \quad (9)$$

2.2 Basic Constraints

The moles of species \mathbf{N} are subject to the three types of constraints mentioned in Section 1.1. These are here expressed in terms of linear equations.

First, the moles of n_{cs} specified species are given, where $0 \leq n_{cs} \leq n_s$. The corresponding constraints can be written

$$(\mathbf{S}^c)^T \mathbf{N} = \mathbf{N}^c, \quad (10)$$

where \mathbf{S}^c is an $n_s \times n_{cs}$ matrix, every column of which is zero, except for a 1 in the row corresponding to a constrained species; and the corresponding row of \mathbf{N}^c is the specified number of moles of that species.

Second, there are n_e constraints on the moles of elements, which can be written

$$\mathbf{E}^T \mathbf{N} = \mathbf{N}^e, \quad (11)$$

where \mathbf{N}^e is specified.

Third, there are n_g general linear constraints ($n_g \geq 0$) which are written

$$(\mathbf{B}^g)^T \mathbf{N} = \mathbf{c}^g. \quad (12)$$

The total number of basic constraints is $n_c \equiv n_{cs} + n_e + n_g$. These are combined into the *basic constraint equation*

$$\mathbf{B}^T \mathbf{N} = \mathbf{c}, \quad (13)$$

where the $n_s \times n_c$ *basic constraint matrix* \mathbf{B} is

$$\mathbf{B} = [\mathbf{S}^c \ \mathbf{E} \ \mathbf{B}^g] \quad (14)$$

and the *basic constraint vector* \mathbf{c} (of length n_c) is

$$\mathbf{c} = [(\mathbf{N}^c)^T \ (\mathbf{N}^e)^T \ (\mathbf{c}^g)^T]^T. \quad (15)$$

The rank of \mathbf{B} , i.e., the number of linearly independent columns, is denoted by n_b . If \mathbf{B} has full column rank, then $n_b = n_c$; otherwise $1 \leq n_b < n_c$.

If \mathbf{B} is rank deficient ($n_b < n_c$) then the basic constraint vector must satisfy a solvability condition (that \mathbf{c} is in the range of \mathbf{B}^T) in order for Eq.(13) to have a solution. This condition is guaranteed to be satisfied if \mathbf{c} is evaluated as $\mathbf{B}^T \mathbf{N}^*$, where \mathbf{N}^* is any n_s -vector. Henceforth it is assumed that this solvability condition is satisfied.

2.3 Modified and Reduced Constraints

The constraints are fundamentally determined by the n_b -dimensional subspace \mathcal{B} spanned by the columns of \mathbf{B} , not by \mathbf{B} itself. Let \mathbf{A} be an arbitrary full-rank $n_b \times n_c$ matrix. Then pre-multiplying both sides of Eq.(13) by \mathbf{A} we obtain

$$\tilde{\mathbf{B}}^T \mathbf{N} = \tilde{\mathbf{c}}, \quad (16)$$

where

$$\tilde{\mathbf{B}} = \mathbf{B} \mathbf{A}^T, \quad (17)$$

and

$$\tilde{\mathbf{c}} = \mathbf{A}\mathbf{c}. \quad (18)$$

Equations (13) and (16) apply exactly the same constraint to \mathbf{N} . Note that the columns of $\tilde{\mathbf{B}}$ are linear combinations of the columns of \mathbf{B} , and so $\text{span}(\tilde{\mathbf{B}}) = \text{span}(\mathbf{B}) = \mathcal{B}$.

We now construct an $n_s \times n_b$ *modified constraint matrix* $\tilde{\mathbf{B}}$ (whose columns span the same n_b -dimensional subspace as \mathbf{B}) which is more convenient than \mathbf{B} . This construction involves: identifying “determined” and “undetermined” species and elements; removing redundant columns of \mathbf{B} (if it is rank deficient); and modifying the rows of the resultant matrix, but always ensuring that its columns span \mathcal{B} .

The k -th species is either “determined” or “undetermined” according to whether or not the value of N_k is determined by Eq.(13). Obviously constrained species are determined, but it is possible that other species (e.g., inerts) can also be determined.

The general solution to the under-determined system Eq.(13) (or, equivalently, Eq.(16)) can be written

$$\mathbf{N} = \mathbf{N}^{\mathcal{B}} + \mathbf{N}^{\mathcal{B}^\perp}, \quad (19)$$

where $\mathbf{N}^{\mathcal{B}}$ is in the subspace \mathcal{B} and is uniquely determined by Eq.(13); and $\mathbf{N}^{\mathcal{B}^\perp}$ is in the orthogonal complement of \mathcal{B} , denoted by \mathcal{B}^\perp , and which is the null space of \mathbf{B}^T . The species k is uniquely determined if, and only if, the null space of \mathbf{B}^T is orthogonal to the k -direction, for then $\mathbf{N}^{\mathcal{B}^\perp}$ is zero.

Whether or not this orthogonality condition is satisfied can be determined from the SVD of \mathbf{B} , which is written

$$\mathbf{B} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T = [\tilde{\mathbf{U}} \hat{\mathbf{U}}] \begin{bmatrix} \tilde{\mathbf{\Sigma}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} [\mathbf{V}]^T. \quad (20)$$

Here $\tilde{\mathbf{\Sigma}}$ is the $n_b \times n_b$ diagonal matrix of singular values. The columns of the $n_s \times (n_s - n_b)$ matrix $\hat{\mathbf{U}}$ span the null space of \mathbf{B}^T . Hence if the k -th row of $\hat{\mathbf{U}}$ is zero, then the null space of \mathbf{B}^T is orthogonal to the k -direction, and so species k is uniquely determined by Eq.(13). We denote by n_{sd} ($n_{sd} \geq n_{cs}$) the number of determined species, and by n_{su} the number of undetermined species ($n_s = n_{sd} + n_{su}$).

We now start to form the modified constraint matrix $\tilde{\mathbf{B}}$ (whose columns span the same subspace \mathcal{B} as do the columns of \mathbf{B}). Let the determined

species be put first in the ordering of species, and let the $n_s \times n_{sd}$ matrix \mathbf{D} be

$$\mathbf{D} = \begin{bmatrix} \mathbf{I}_{(n_{sd} \times n_{sd})} \\ \mathbf{0} \end{bmatrix}. \quad (21)$$

Then the columns of $[\mathbf{D} \ \mathbf{E} \ \mathbf{B}^g]$ span \mathcal{B} . Let $\hat{\mathbf{E}}$ and $\hat{\mathbf{B}}^g$ denote the matrices \mathbf{E} and \mathbf{B}^g with their first n_{sd} rows set to zero. Clearly the columns of $[\mathbf{D} \ \hat{\mathbf{E}} \ \hat{\mathbf{B}}^g]$ also span \mathcal{B} .

Under several circumstances the matrix $[\mathbf{D} \ \hat{\mathbf{E}} \ \hat{\mathbf{B}}^g]$ may be rank deficient, in which case columns of the matrix can be removed. A column of the matrix $\hat{\mathbf{E}}$ can be removed if (a) it is zero, or (b) it is linearly dependent upon the preceding columns (which can be determined by using the QR algorithm with pivoting). The element corresponding to a removed column is said to be “determined,” and n_{ed} and n_{eu} denote the number of determined and undetermined elements ($n_{ed} + n_{eu} = n_e$). It is assumed that the determined elements are first in the ordering.

The element matrix \mathbf{E} can be partitioned

$$\mathbf{E} = \begin{bmatrix} \mathbf{E}^{dd} & \mathbf{E}^{du} \\ \mathbf{0} & \mathbf{E}^{uu} \end{bmatrix}, \quad (22)$$

where \mathbf{E}^{uu} is an $n_{su} \times n_{eu}$ matrix. (The lower left block is zero, since undetermined species cannot contribute to determined elements.) The matrix $\hat{\mathbf{E}}$ with redundant columns removed is

$$\hat{\mathbf{E}} = \begin{bmatrix} \mathbf{0} \\ \mathbf{E}^{uu} \end{bmatrix}. \quad (23)$$

The matrix $\hat{\mathbf{B}}^g$ has the form

$$\hat{\mathbf{B}}^g = \begin{bmatrix} \mathbf{0} \\ \bar{\mathbf{B}}^g \end{bmatrix}, \quad (24)$$

where $\bar{\mathbf{B}}^g$ is $n_{su} \times n_g$. If $[\mathbf{E}^{uu} \ \bar{\mathbf{B}}^g]$ does not have full column rank, then linearly dependent columns of $\bar{\mathbf{B}}^g$ can be removed to yield the $n_{su} \times n_{gi}$ matrix $\tilde{\mathbf{B}}^g$, where n_{gi} is the number of independent general constraints.

In summary:

1. The species are ordered so that the n_{sd} determined species are first, followed by the n_{su} undetermined species.

2. The elements are ordered so that the n_{ed} determined elements are first, followed by the n_{eu} undetermined elements.
3. The basic $(n_s \times n_c)$ constraint matrix $\mathbf{B} = [\mathbf{S}^c \ \mathbf{E} \ \mathbf{B}^g]$ spans the n_b -dimensional constraint subspace \mathcal{B} .
4. The modified $(n_s \times n_b)$ constraint matrix $\tilde{\mathbf{B}}$, which also spans \mathcal{B} , is defined by

$$\tilde{\mathbf{B}} = \begin{bmatrix} \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{E}^{uu} & \tilde{\mathbf{B}}^g \end{bmatrix} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \hat{\mathbf{B}} \end{bmatrix}, \quad (25)$$

where: \mathbf{I} is the $n_{sd} \times n_{sd}$ identity; \mathbf{E}^{uu} is the $(n_{su} \times n_{eu})$ undetermined element matrix; $\tilde{\mathbf{B}}^g$ is the $(n_{su} \times n_{gi})$ modified general constraint matrix; and $\hat{\mathbf{B}} \equiv [\mathbf{E}^{uu} \ \tilde{\mathbf{B}}^g]$ is the $(n_{su} \times n_{rc})$ *reduced constraint matrix*, where $n_{rc} = n_{eu} + n_{gi} = n_b - n_{sd}$.

5. With $\tilde{\mathbf{B}}$ given by Eq.(25), the *modified constraint equation* Eq.(16) is

$$\tilde{\mathbf{B}}^T \mathbf{N} = \begin{bmatrix} \mathbf{N}^d \\ \hat{\mathbf{B}}^T \mathbf{N}^u \end{bmatrix} = \tilde{\mathbf{c}} = \begin{bmatrix} \mathbf{N}^d \\ \hat{\mathbf{c}} \end{bmatrix}, \quad (26)$$

where: \mathbf{N}^d and \mathbf{N}^u are the moles of determined and undetermined species (\mathbf{N}^d being specified); and $\hat{\mathbf{c}}$ is the $(n_{rc} \times 1)$ *reduced constraint vector*.

6. Given \mathbf{B} and $\tilde{\mathbf{B}}$, the implied *modified constraint transformation matrix* \mathbf{A} can be deduced from Eq.(17). (If \mathbf{B} is rank deficient, then \mathbf{A} is not unique; but $\tilde{\mathbf{c}} = \mathbf{A}\mathbf{c}$ is unique, provided that \mathbf{c} satisfies the solvability condition.)
7. The first n_{sd} rows of the modified constraint equation, Eq.(26), explicitly set the values of the determined species. The remaining n_{su} rows form the *reduced constraint equation*

$$\hat{\mathbf{B}}^T \mathbf{N}^u = \hat{\mathbf{c}}. \quad (27)$$

2.4 The Feasible Region

In this Section we review the lower and upper bounds on the species, and identify the feasible region of the species space corresponding to realizable compositions which satisfy the constraints.

The species moles are of course non-negative, $N_k \geq 0$. This is the lower bound. Given the element moles \mathbf{N}^e , there is also an upper bound on each species, namely,

$$N_k^{up} \equiv 1 / \max_j (E_{kj} / N_{(j)}^e), \quad (28)$$

which corresponds to all atoms of some element j being in the form of species k . (Subscripts in brackets are excluded from the summation convention.) From these observations we obtain:

Result 1 *Given the constraint equation Eq.(13) with \mathbf{c} finite, the moles of each species is bounded by $0 \leq N_k \leq N_k^{up}$.*

Given a constraint matrix \mathbf{B} and vector \mathbf{c} , the *feasible region* \mathcal{F} of the species space \mathcal{S} is defined by

$$\mathcal{F} \equiv \{\mathbf{N} : N_k \geq 0, \mathbf{B}^T \mathbf{N} = \mathbf{c}\}. \quad (29)$$

The *interior feasible region* $\bar{\mathcal{F}}$ is defined by

$$\bar{\mathcal{F}} \equiv \{\mathbf{N} : N_k > 0, \mathbf{B}^T \mathbf{N} = \mathbf{c}\}, \quad (30)$$

and the difference $\partial\mathcal{F} = \mathcal{F} - \bar{\mathcal{F}}$ is the boundary of the feasible region, on which at least one species is zero. A *proper feasible region* is defined to be a feasible region with a non-empty interior.

Given \mathbf{B} , whether or not \mathcal{F} is a proper feasible region depends upon \mathbf{c} ; and if \mathcal{F} is a proper feasible region, then \mathbf{c} is said to be a *proper constraint vector*.

It is obvious that:

Result 2 *If N_k is strictly positive for all k , then $\mathbf{c} \equiv \mathbf{B}^T \mathbf{N}$ is a proper constraint vector.*

The converse is established through the solution to the *max-min problem* (see Appendix A). The max-min bound N_{\min} and the max-min moles \mathbf{N}^{mm} are defined as the solution to the problem: given \mathbf{B} and \mathbf{c} , find the composition \mathbf{N}^{mm} which maximizes N_{\min} subject to $N_k^{mm} \geq N_{\min}$ (for all k) and $\mathbf{B}^T \mathbf{N}^{mm} = \mathbf{c}$. If N_{\min} is strictly positive, then so also is N_k^{mm} . Hence, we obtain:

Result 3 *Given \mathbf{B} and \mathbf{c} , if N_{\min} is strictly positive, then \mathbf{c} is a proper constraint vector.*

Henceforth, unless otherwise stated, we consider only proper constraint vectors, so that a proper feasible region \mathcal{F} exists.

We now consider the geometry of the feasible region \mathcal{F} . The points \mathbf{N} satisfying $\mathbf{B}^T \mathbf{N} = \mathbf{c}$ define the *feasible affine space*, the general point in which is

$$\mathbf{N} = \mathbf{N}^{\mathcal{B}} + \mathbf{N}^{\mathcal{B}^\perp}, \quad (31)$$

where $\mathbf{N}^{\mathcal{B}}$ is the component in the constraint subspace (which is determined by the constraint equation), and $\mathbf{N}^{\mathcal{B}^\perp}$ is an arbitrary vector in \mathcal{B}^\perp . The feasible region \mathcal{F} is a convex polytope lying in this affine space.

We use $\Delta \mathbf{N}$ to denote a feasible increment vector, that is, the difference between two feasible compositions. From Eq.(31) we obtain:

Result 4 *Every feasible increment vector $\Delta \mathbf{N}$ is an element of \mathcal{B}^\perp . Hence $\mathbf{B}^T \Delta \mathbf{N} = 0$, and $\Delta \mathbf{N}$ is orthogonal to every vector in \mathcal{B} .*

A result that will be used below is:

Result 5 *A non-zero feasible increment vector $\Delta \mathbf{N}$ has both positive and negative components.*

This stems from the observation that the atom vector \mathbf{a} is in \mathcal{B} and that each of its components is strictly positive. Hence, if $\Delta \mathbf{N}$ is non-zero, the equation $\mathbf{a}^T \Delta \mathbf{N} = 0$ can be satisfied only if $\Delta \mathbf{N}$ has both positive and negative components.

2.5 Thermodynamic Properties

We now define the relevant thermodynamic properties of the ideal gas mixture at temperature T and pressure p .

The enthalpy H is given by

$$H = \sum_{k=1}^{n_s} N_k h_k(T) = N_k h_k = \mathbf{N}^T \mathbf{h}, \quad (32)$$

where $h_k(T)$ is the molar specific enthalpy, and the summation convention applies.

The entropy S is given by

$$S = N_k S_k(T, p), \quad (33)$$

where the molar specific entropy is

$$S_k(T, p) = S_k^o(T) - \mathcal{R} \ln X_k - \mathcal{R} \ln(p/p_{\text{atm}}), \quad (34)$$

where $S_k^o(T)$ is the standard state molar entropy, \mathcal{R} is the universal gas constant, and p_{atm} is the standard state pressure.

The Gibbs function G is defined as

$$G = H - TS, \quad (35)$$

and its normalized value as

$$\tilde{G} \equiv \frac{G}{\mathcal{R}T}. \quad (36)$$

From the above equations we obtain

$$\tilde{G} = N_k[\tilde{g}_k(T, p) + \ln X_k], \quad (37)$$

where the normalized molar specific Gibbs function is

$$\tilde{g}_k = \frac{h_k}{\mathcal{R}T} - \frac{S_k^o}{\mathcal{R}} + \ln\left(\frac{p}{p_{\text{atm}}}\right). \quad (38)$$

2.6 Constrained Equilibrium

At given fixed temperature T and pressure p , and given the basic constraint matrix \mathbf{B} and the proper constraint vector \mathbf{c} , the *constrained equilibrium composition* \mathbf{N}^{CE} is the species moles vector \mathbf{N} which satisfies the basic constraint equation, Eq.(13), and minimizes the Gibbs function G , Eq.(35).

Note that at fixed T and p the normalized species Gibbs functions \tilde{g}_k are fixed, and that minimizing G is equivalent to minimizing \tilde{G} , Eq.(37), which can be written

$$\begin{aligned} \tilde{G} &= N_k(\tilde{g}_k + \ln X_k) \\ &= \sum_{k=1}^{n_s} N_k(\tilde{g}_k + \ln N_k - \ln \sum_{i=1}^{n_s} N_i). \end{aligned} \quad (39)$$

Hence \mathbf{N}^{CE} is the value of \mathbf{N} which minimizes \tilde{G} subject to the constraint $\mathbf{B}^T \mathbf{N} = \mathbf{c}$.

3 Existence and Uniqueness

In this section we establish the important result:

Result 6 *Given a constraint matrix \mathbf{B} and a proper constraint vector \mathbf{c} , then there exists a unique constrained equilibrium composition \mathbf{N}^{CE} which is in the interior of the feasible region.*

The existence of \mathbf{N}^{CE} follows simply from the facts that the feasible region is non-empty, and the constrained equilibrium state is defined as the point (or points) in \mathcal{F} at which G (or, equivalently, \tilde{G}) is minimum.

The gradient of \tilde{G} , obtained by differentiating Eq.(39), is

$$\frac{\partial \tilde{G}}{\partial N_k} = \tilde{g}_k + \ln X_k. \quad (40)$$

On the boundary of the feasible region, $\partial \mathcal{F}$, $\partial \tilde{G}/\partial N_k$ is infinitely negative for each species k which is zero. Let \mathbf{N}^∂ be a point on the boundary $\partial \mathcal{F}$, and let \mathbf{N}^I be any point in the interior $\bar{\mathcal{F}}$ (so that $N_k^I > 0$). Then the gradient in the direction $\mathbf{N}^I - \mathbf{N}^\partial$ is

$$\frac{\partial \tilde{G}}{\partial s} = \frac{(N_k^I - N_k^\partial)}{|\mathbf{N}^I - \mathbf{N}^\partial|} \frac{\partial \tilde{G}}{\partial N_k}. \quad (41)$$

Now for any component k for which N_k^∂ is zero, we have $N_k^I - N_k^\partial > 0$. Hence $\partial \tilde{G}/\partial s$ is negatively infinite, from which it follows that there is no minimum of \tilde{G} on the boundary $\partial \mathcal{F}$.

The uniqueness of \mathbf{N}^{CE} stems from the fact that \tilde{G} is convex in the interior feasible region, which is established by showing that the second derivative of \tilde{G} in every feasible direction is strictly positive.

The second derivative of \tilde{G} is

$$J_{ij} \equiv \frac{\partial^2 \tilde{G}}{\partial N_i \partial N_j} = \frac{\partial \ln X_i}{\partial N_j} = \frac{\delta_{ij}}{N_{(j)}} - \frac{\delta_{i(i)} \delta_{j(j)}}{\bar{N}}. \quad (42)$$

This can be re-written

$$\mathbf{J} = \bar{\mathbf{Y}}^{-1}(\mathbf{I} - \mathbf{u}\mathbf{u}^T)\bar{\mathbf{Y}}^{-1}, \quad (43)$$

where

$$\bar{y}_i \equiv \sqrt{N_i}, \quad (44)$$

$$\mathbf{u} \equiv \bar{\mathbf{y}}/|\bar{\mathbf{y}}|, \quad (45)$$

and

$$\bar{\mathbf{Y}} \equiv \text{diag}(\bar{\mathbf{y}}). \quad (46)$$

With \mathbf{v} being a unit vector, the second derivative of \tilde{G} in the direction of \mathbf{v} is

$$\begin{aligned} \frac{\partial^2 \tilde{\mathbf{G}}}{\partial s^2} &= \mathbf{v}^T \mathbf{J} \mathbf{v} = \mathbf{v}^T \bar{\mathbf{Y}}^{-2} \mathbf{v} - \mathbf{v}^T \bar{\mathbf{Y}}^{-1} \mathbf{u} \mathbf{u}^T \bar{\mathbf{Y}}^{-1} \mathbf{v} \\ &= |\bar{\mathbf{Y}}^{-1} \mathbf{v}|^2 - |\mathbf{u}^T \bar{\mathbf{Y}}^{-1} \mathbf{v}|^2. \end{aligned} \quad (47)$$

It follows from the Cauchy-Schwartz inequality that $\partial^2 \tilde{G}/\partial s^2$ is non-negative. It is zero only if \mathbf{v} is parallel to $\bar{\mathbf{Y}}\mathbf{u}$, the latter being a vector with all positive components.

The above result pertains to any unit vector \mathbf{v} in the species space \mathcal{S} . To show that \tilde{G} is convex in the interior feasible region $\bar{\mathcal{F}}$, we now restrict \mathbf{v} to be a unit vector in the direction of a feasible increment vector $\Delta \mathbf{N}$, i.e., \mathbf{v} is in \mathcal{B}^\perp . In this case, \mathbf{v} has both positive and negative components (Result 5). Hence, \mathbf{v} (with both positive and negative components) cannot be parallel to $\bar{\mathbf{Y}}\mathbf{u}$ (which has all positive components). Hence, $\partial^2 \tilde{G}/\partial s^2$ is strictly positive, showing that \tilde{G} is convex in $\bar{\mathcal{F}}$, and the result is established.

4 Constraint Potentials

4.1 Basic Constraints

At the minimum of \tilde{G} , Eq.(39), subject to the basic constraints $\mathbf{B}^T \mathbf{N} = \mathbf{c}$, the quantity

$$\bar{G} \equiv \tilde{G} - \boldsymbol{\lambda}^T (\mathbf{B}^T \mathbf{N} - \mathbf{c}), \quad (48)$$

is stationary with respect to infinitesimal changes $d\mathbf{N}$ and $d\boldsymbol{\lambda}$, where $\boldsymbol{\lambda}$ is the $n_c \times 1$ vector of Lagrange multipliers, or *basic constraint potentials*. Thus, at the minimum

$$\begin{aligned} 0 &= d\bar{G} \\ &= (\tilde{g}_k + \ln X_k) dN_k - \lambda_j B_{kj} dN_k - d\lambda_j (B_{kj} N_k - c_j) \\ &= (\ln X_k + \tilde{g}_k - B_{kj} \lambda_j) dN_k, \end{aligned} \quad (49)$$

the last step following because the multiplier of $d\lambda_j$ is zero. The multiplier of dN_k must also be zero, leading to

$$X_k = \exp(-\tilde{g}_k + B_{kj}\lambda_j), \quad (50)$$

or, equivalently,

$$\mathbf{X} = \exp(-\tilde{\mathbf{g}} + \mathbf{B}\boldsymbol{\lambda}). \quad (51)$$

Thus we have the important result

Result 7 *In the constrained equilibrium state, the n_s species mole fractions \mathbf{X} are determined in terms of n_c basic constraint potentials $\boldsymbol{\lambda}$ by Eq.(51).*

A different expression of the same result, stemming from the equation

$$\hat{g}_k \equiv \ln X_k + \tilde{g}_k = B_{kj}\lambda_j, \quad (52)$$

is

Result 8 *In the constrained equilibrium state, the n_s -vector $\hat{\mathbf{g}}$ is in the (n_b -dimensional) range of the basic constraint matrix \mathbf{B} , i.e., $\hat{\mathbf{g}}$ is in the constraint subspace \mathcal{B} .*

Observing that the above equations are in terms of the mole fractions (independent of \bar{N}), we rewrite the constraint equation as

$$\mathbf{B}^T \mathbf{X} = \mathbf{c}/\bar{N}. \quad (53)$$

Consider $\boldsymbol{\lambda}$ now being an arbitrary n_c -vector, and define

$$\bar{X}_k = \exp(-\tilde{g}_k + B_{kj}\lambda_j). \quad (54)$$

Then we have

Result 9 *Given \mathbf{B} and $\boldsymbol{\lambda}$, $\bar{\mathbf{X}}$ corresponds to a constrained equilibrium mixture if, and only if, $\bar{\mathbf{X}}$ satisfies the normalization condition*

$$\sum_{k=1}^{n_s} \bar{X}_k = 1. \quad (55)$$

If this condition is satisfied, then the basic constraint vector is

$$\mathbf{c} = \bar{N}\mathbf{B}^T\bar{\mathbf{X}}, \quad (56)$$

where \bar{N} is indeterminate, and $\boldsymbol{\lambda}$ is said to be a *proper Lagrange vector*.

With \mathbf{c} being a proper constraint vector, the constraint direction is defined by the unit vector

$$\bar{\mathbf{c}} \equiv \mathbf{c}/|\mathbf{c}|. \quad (57)$$

It then follows from the above

Result 10 *Given \mathbf{B} , the mole fractions in the constrained equilibrium state are uniquely determined by the constraint direction $\bar{\mathbf{c}}$, independent of $|\mathbf{c}|$.*

Result 11 *If the constraint equation is replaced by*

$$\mathbf{B}^T\mathbf{N} = \alpha\mathbf{c}, \quad (58)$$

where α is a positive parameter, then (a) in the constrained equilibrium state \mathbf{X} and $\boldsymbol{\lambda}$ are independent of α (b) the total number of moles \bar{N} is linearly proportional to the α (c) there is a particular value of α , denoted by $\bar{\alpha}$, such that $\bar{N} = 1$, and (d) the above constraint equation can alternatively be written

$$\mathbf{B}^T\mathbf{X} = \bar{\alpha}\mathbf{c}. \quad (59)$$

4.2 Reduced Constraints

The development above shows that the n_s mole fractions in the constrained equilibrium state can be determined from the n_c basic constraint potentials, $\boldsymbol{\lambda}$. Here, by considering the reduced constraints, we show that a knowledge of n_{rc} *reduced constraint potentials* is sufficient.

We recall that \mathbf{N}^d and \mathbf{N}^u denote the moles of the n_{sd} determined species, and of the n_{su} undetermined species, respectively. The total moles of determined and undetermined species are denoted by

$$\bar{N}^d \equiv \sum_{k=1}^{n_{sd}} N_k^d \quad \text{and} \quad \bar{N}^u \equiv \sum_{k=1}^{n_{su}} N_k^u, \quad (60)$$

and so the total moles of species is $\bar{N} = \bar{N}^d + \bar{N}^u$.

The normalized Gibbs function, Eq.(39), can be decomposed as

$$\begin{aligned}
\tilde{G} &= \sum_{k=1}^{n_s} N_k (\tilde{g}_k + \ln N_k - \ln \bar{N}) \\
&= \sum_{k=1}^{n_{sd}} N_k^d (\tilde{g}_k^d + \ln N_k^d) + \sum_{k=1}^{n_{su}} N_k^u (\tilde{g}_k^u + \ln N_k^u) \\
&\quad - (\bar{N}^d + \bar{N}^u) \ln(\bar{N}^d + \bar{N}^u),
\end{aligned} \tag{61}$$

where $\tilde{\mathbf{g}}^d$ and $\tilde{\mathbf{g}}^u$ are the normalized Gibbs functions of the determined and undetermined species. Note that \mathbf{N}^d and \bar{N}^d are known and fixed, whereas \mathbf{N}^u is subject to the n_{rc} reduced constraints $\hat{\mathbf{B}}^T \mathbf{N}^u = \hat{\mathbf{c}}$, Eq.(27). Hence, in the constrained equilibrium state, the quantity

$$\begin{aligned}
\hat{G} &\equiv \sum_{k=1}^{n_{su}} N_k^u (\tilde{g}_k^u + \ln N_k^u) - (\bar{N}^d + \bar{N}^u) \ln(\bar{N}^d + \bar{N}^u) \\
&\quad - \sum_{j=1}^{n_{rc}} \hat{\lambda}_j \left(\sum_{k=1}^{n_{su}} \hat{B}_{kj} N_k^u - \hat{c}_j \right),
\end{aligned} \tag{62}$$

is stationary with respect to infinitesimal changes $d\mathbf{N}^u$ and $d\hat{\boldsymbol{\lambda}}$, where $\hat{\boldsymbol{\lambda}}$ are n_{rc} Lagrange multipliers, or the *reduced constraint potentials*. Thus

$$\begin{aligned}
d\hat{G} &= 0 \\
&= (\tilde{g}_k^u + \ln X_k^u - \hat{B}_{kj} \hat{\lambda}_j) dN_k^u \\
&\quad - d\hat{\lambda}_j (\hat{B}_{kj} N_k^u - \hat{c}_j),
\end{aligned} \tag{63}$$

where $\mathbf{X}^u \equiv \mathbf{N}^u / \bar{N}$ is the mole fraction of undetermined species, and the summation convention applies.

Similar to the development from Eq.(49), the multiplier of $d\hat{\lambda}_j$ in Eq.(63) is zero, and the requirement that the multiplier of dN_k^u be zero leads the result

$$X_k^u = \exp(-\tilde{g}_k^u + \hat{B}_{kj} \hat{\lambda}_j), \tag{64}$$

or, equivalently,

$$\mathbf{X}^u = \exp(-\tilde{\mathbf{g}}^u + \hat{\mathbf{B}} \hat{\boldsymbol{\lambda}}). \tag{65}$$

4.3 Normalization Conditions

In terms of the basic constraints, the problem of determining the constrained equilibrium composition has been reduced to solving the $n_c + 1$ non-linear equations

$$\bar{N}\mathbf{B}^T\mathbf{X} = \mathbf{c}, \quad (66)$$

and

$$\sum_{k=1}^{n_s} X_k = 1, \quad (67)$$

where

$$\mathbf{X} = \exp(-\tilde{\mathbf{g}} + \mathbf{B}\boldsymbol{\lambda}), \quad (68)$$

for the $n_c + 1$ unknowns $\boldsymbol{\lambda}$ and \bar{N} . Equation (67) is the normalization condition.

In terms of the reduced constraints, there are $n_{rc} + 1$ equations

$$\bar{N}\hat{\mathbf{B}}^T\mathbf{X}^u = \hat{\mathbf{c}}, \quad (69)$$

and

$$\sum_{k=1}^{n_{su}} X_k^u + \frac{\bar{N}^d}{\bar{N}} = 1, \quad (70)$$

where

$$\mathbf{X}^u = \exp(-\tilde{\mathbf{g}}^u + \hat{\mathbf{B}}\hat{\boldsymbol{\lambda}}), \quad (71)$$

for the $n_{rc} + 1$ unknowns $\hat{\boldsymbol{\lambda}}$ and \bar{N} . The normalization condition Eq.(70) is more complicated than Eq.(67) in that it involves the unknown \bar{N} . The purpose of this section is to obtain an explicit normalization condition for \mathbf{X}^u .

Let $\hat{\mathbf{w}}$ be a vector of length n_{rc} , and we define the n_{su} -vector \mathbf{w} by

$$\mathbf{w} \equiv \hat{\mathbf{B}}\hat{\mathbf{w}}. \quad (72)$$

Now consider the scalar $\mathbf{w}^T\mathbf{X}^u$ which, using Eq.(72) and Eq.(69), can be written

$$\mathbf{w}^T\mathbf{X}^u = \hat{\mathbf{w}}^T\hat{\mathbf{B}}^T\mathbf{X}^u = \frac{\hat{\mathbf{w}}^T\hat{\mathbf{c}}}{\bar{N}}. \quad (73)$$

Providing that $\hat{\mathbf{w}}^T\hat{\mathbf{c}}$ is non-zero, we then obtain an expression for \bar{N} in terms of \mathbf{X}^u :

$$\frac{1}{\bar{N}} = \frac{\mathbf{w}^T\mathbf{X}^u}{\hat{\mathbf{w}}^T\hat{\mathbf{c}}}. \quad (74)$$

This can be used to eliminate \bar{N} from the normalization condition Eq.(70), which can then be written simply as

$$\mathbf{Q}^T \mathbf{X}^u = 1, \quad (75)$$

where the n_{su} -vector \mathbf{Q} is defined by

$$\mathbf{Q} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} + \frac{\bar{N}^d}{\hat{\mathbf{w}}^T \hat{\mathbf{c}}} \begin{bmatrix} \mathbf{w} \end{bmatrix}. \quad (76)$$

A reasonable choice of $\hat{\mathbf{w}}$ corresponds to atoms:

$$\hat{\mathbf{w}} = \begin{bmatrix} \mathbf{I}_{(n_{eu} \times 1)} \\ \mathbf{0}_{(n_{gi} \times 1)} \end{bmatrix}, \quad (77)$$

so that \mathbf{w} is equal to \mathbf{a}^u , where a_k^u ($k = 1, 2, \dots, n_{su}$) is the number of atoms in a molecule of the undetermined species k . Then the moles of atoms in the undetermined species, N^{au} , is given by

$$N^{au} = \mathbf{w}^T \mathbf{N}^u = \bar{N} \mathbf{w}^T \mathbf{X}^u = \hat{\mathbf{w}}^T \hat{\mathbf{c}}. \quad (78)$$

The requirement that $\hat{\mathbf{w}}^T \hat{\mathbf{c}}$ be non-zero is clearly met, since N^{au} is strictly positive.

5 Solution for the Constraint Potentials

5.1 Equations to be Solved

We now express in final form the non-linear equations to be solved to determine the constrained equilibrium composition at fixed temperature and pressure. The given quantities are $\hat{\mathbf{B}}$, $\hat{\mathbf{c}}$, $\tilde{\mathbf{g}}^u$ and \mathbf{Q} ; the unknowns are the reduced constraint potential vector $\hat{\lambda}$ and the total moles \bar{N} ; and the equations to be satisfied are Eq.(69) and Eq.(75). These equations are now re-expressed in terms of variables $\bar{\lambda}$ and $\bar{\mathbf{g}}$, and residuals \mathbf{r} and q , such that the solution corresponds to $\hat{\lambda} = \bar{\lambda}$ provided that $\bar{\mathbf{g}} = \tilde{\mathbf{g}}^u$, $\mathbf{r} = 0$ and $q = 0$. To this end we define

$$\bar{\mathbf{X}} \equiv \exp(-\bar{\mathbf{g}} + \hat{\mathbf{B}}\bar{\lambda}), \quad (79)$$

$$\mathbf{v} \equiv \hat{\mathbf{B}}^T \bar{\mathbf{X}}, \quad (80)$$

$$\mathbf{r} \equiv \frac{\mathbf{v}}{|\mathbf{v}|} - \frac{\hat{\mathbf{c}}}{|\hat{\mathbf{c}}|}, \quad (81)$$

and

$$q \equiv 1 - \mathbf{Q}^T \bar{\mathbf{X}}. \quad (82)$$

It is evident that, for $\bar{\boldsymbol{\lambda}} = \hat{\boldsymbol{\lambda}}$ and $\bar{\mathbf{g}} = \tilde{\mathbf{g}}^u$, then $\bar{\mathbf{X}} = \mathbf{X}^u$ (Eq.(71)), and that $\mathbf{r} = 0$ and $q = 0$ are equivalent to the satisfaction of Eq.(69) and Eq.(75).

5.2 Infinitesimals

The essence of the numerical problems involved in these equations, and the key to overcoming them, are revealed by considering infinitesimal changes in the variables. Specifically, we consider infinitesimal changes $d\bar{\boldsymbol{\lambda}}$ and $d\bar{\mathbf{g}}$ (in $\bar{\boldsymbol{\lambda}}$ and $\bar{\mathbf{g}}$), and examine the corresponding changes $d\bar{\mathbf{X}}$, $d\mathbf{v}$, $d\mathbf{r}$ and dq in the other quantities.

From Eq.(79) we obtain

$$d\bar{X}_i = \bar{X}_{(i)}(-d\bar{g}_i + \hat{B}_{ik}d\bar{\lambda}_k), \quad (83)$$

where bracketed suffixes are excluded from the summation convection. It happens that natural variables are

$$y_i \equiv \sqrt{\bar{X}_i}, \quad (84)$$

and the $n_{su} \times n_{su}$ diagonal matrix

$$\mathbf{Y} \equiv \text{diag}(\mathbf{y}). \quad (85)$$

In terms of these variables, Eq.(83) becomes

$$d\bar{\mathbf{X}} = -\mathbf{Y}^2 d\bar{\mathbf{g}} + \mathbf{Y}\mathbf{H}d\bar{\boldsymbol{\lambda}}, \quad (86)$$

where the $n_{su} \times n_{rc}$ matrix \mathbf{H} is defined by

$$\mathbf{H} \equiv \mathbf{Y}\hat{\mathbf{B}}, \quad (87)$$

and \mathbf{v} , Eq.(80), can alternatively be written

$$\mathbf{v} = \hat{\mathbf{B}}^T \bar{\mathbf{X}} = \mathbf{H}^T \mathbf{y}. \quad (88)$$

From Eq.(80) we then obtain

$$\begin{aligned} d\mathbf{v} &= \hat{\mathbf{B}}^T d\bar{\mathbf{X}} = 2\mathbf{H}^T d\mathbf{y} \\ &= -\hat{\mathbf{B}}^T \mathbf{Y}^2 d\bar{\mathbf{g}} + \hat{\mathbf{B}}^T \mathbf{Y} \mathbf{H} d\bar{\boldsymbol{\lambda}} \\ &= -\mathbf{H}^T \mathbf{Y} d\bar{\mathbf{g}} + \mathbf{H}^T \mathbf{H} d\bar{\boldsymbol{\lambda}}. \end{aligned} \quad (89)$$

5.3 Newton's Method

It is natural to consider Newton's method (or a variant thereof) in order to solve Eq.(81) and Eq.(82). That is, $\bar{\boldsymbol{\lambda}}$ is set to an initial guess, $\bar{\mathbf{g}}$ is set to $\tilde{\mathbf{g}}^u$ and is not changed (i.e., $d\bar{\mathbf{g}} = 0$), and then $\bar{\boldsymbol{\lambda}}$ is changed iteratively in order to reduce the residuals to zero. In this case $d\bar{\mathbf{g}}$ is zero, and Eq.(89) reduces to

$$d\mathbf{v} = \mathbf{H}^T \mathbf{H} d\bar{\boldsymbol{\lambda}}. \quad (90)$$

Clearly, the success of this approach depends crucially on the conditioning of the matrix $\mathbf{H}^T \mathbf{H}$: for if it is singular, then there is no value of $d\bar{\boldsymbol{\lambda}}$ that can effect a change in all components of $d\mathbf{v}$, and hence reduce the residual \mathbf{r} .

In infinite-precision arithmetic, the matrix $\mathbf{H} \equiv \mathbf{Y} \hat{\mathbf{B}}$ has full column rank, and hence $\mathbf{H}^T \mathbf{H}$ is symmetric positive definite. This follows from the facts that $\hat{\mathbf{B}}$ has full column rank, and that \mathbf{Y} is non-singular, since it is a diagonal matrix with strictly positive diagonal components, $y_i = \sqrt{\bar{X}_i}$. However, in chemical equilibrium, it is not unusual to have extremely small species mole fractions, e.g., $X_i = 10^{-40}$, and hence $\mathbf{H}^T \mathbf{H}$ can be numerically singular.

Suppose, for example, that the first two columns of $\hat{\mathbf{B}}$ are identical except for one row, denoted by k . Suppose further that \bar{X}_k is extremely small, e.g., 10^{-40} . Then, numerically, the first two columns of \mathbf{H} are identical. As a consequence, any value of $d\bar{\boldsymbol{\lambda}}$ has the same effect on the first two constraints: dv_1 and dv_2 cannot be changed independently.

5.4 Gibbs Function Continuation

Referring to Eq.(89), Newton's method corresponds to $d\bar{\mathbf{g}} = 0$, and seeking $d\bar{\boldsymbol{\lambda}}$ to effect a change $d\mathbf{v}$. The alternative method developed here—*Gibbs function continuation*—is based on the observation that, for arbitrary $d\bar{\mathbf{g}}$, a

value of $d\bar{\lambda}$ can always be stably determined which yields $d\mathbf{v} = 0$, even if \mathbf{H} is rank deficient.

This observation is evident when Eq.(89) is rewritten as

$$d\mathbf{v} = \mathbf{H}^T(\mathbf{H}d\bar{\lambda} - \mathbf{Y}d\bar{\mathbf{g}}). \quad (91)$$

For arbitrary $d\bar{\mathbf{g}}$, $d\bar{\lambda}$ is taken as the least-squares solution to

$$\min \|\mathbf{H}d\bar{\lambda} - \mathbf{Y}d\bar{\mathbf{g}}\|_2. \quad (92)$$

It is a general property of such least-squares solutions that the residual

$$d\gamma \equiv \mathbf{H}d\bar{\lambda} - \mathbf{Y}d\bar{\mathbf{g}}, \quad (93)$$

is in the null space of \mathbf{H}^T , i.e.,

$$\mathbf{H}^T d\gamma = \mathbf{H}^T(\mathbf{H}d\bar{\lambda} - \mathbf{Y}d\bar{\mathbf{g}}) = 0. \quad (94)$$

Thus the right-hand side of Eq.(91) is zero, yielding $d\mathbf{v} = 0$.

The general idea of Gibbs function continuation is to start with values of $\bar{\lambda}$ and $\bar{\mathbf{g}}$ ($\bar{\mathbf{g}} \neq \tilde{\mathbf{g}}^u$) such that the residuals \mathbf{r} and q are zero, and then to change $\bar{\mathbf{g}}$ towards $\tilde{\mathbf{g}}^u$ while changing $\bar{\lambda}$ in such a way that the residuals remain zero.

We introduce a pseudo-time variable s , and denote by $\bar{\lambda}(s)$, $\bar{\mathbf{g}}(s)$, $\bar{\mathbf{X}}(s)$ and $\mathbf{v}(s)$ the variables considered at time s : $s = 0$ corresponds to the initial condition; and $s = 1$ to the solution, so that $\bar{\mathbf{g}}(1) = \tilde{\mathbf{g}}^u$, $\bar{\lambda}(1) = \hat{\lambda}$ and $\bar{\mathbf{X}}(1) = \hat{\mathbf{X}}$.

The initial condition $\bar{\mathbf{X}}(0)$ is set to a feasible composition. This can always be done, for example, based on the max-min composition \mathbf{N}^{mm} . The residuals \mathbf{r} and q are therefore zero initially (and they remain so for all s). However, it is necessary to ensure that $\bar{\mathbf{X}}(0)$ is given in terms of $\bar{\lambda}(0)$ and $\bar{\mathbf{g}}(0)$ by Eq.(79). Any initial value of $\bar{\lambda}$ can be specified, for example $\bar{\lambda}(0) = 0$. Then, in order that Eq.(79) be satisfied, $\bar{\mathbf{g}}(0)$ is specified as

$$\bar{\mathbf{g}}(0) = \hat{\mathbf{B}}\bar{\lambda}(0) - \ln \bar{\mathbf{X}}(0). \quad (95)$$

Thus a consistent set of initial conditions are specified with zero residuals.

The pseudo-Gibbs function $\bar{\mathbf{g}}(s)$ is prescribed to vary linearly in time as

$$\bar{\mathbf{g}}(s) = \bar{\mathbf{g}}(0) + s[\tilde{\mathbf{g}}^u - \bar{\mathbf{g}}(0)], \quad (96)$$

so that $\bar{\mathbf{g}}(1) = \tilde{\mathbf{g}}^u$. The variation of $\bar{\lambda}(s)$ is then determined by the requirement that the residuals remain zero.

With \mathbf{r} being zero, Eq.(81) implies that the vector $\mathbf{v}(s)$ is at all times parallel with the fixed vector $\hat{\mathbf{c}}$. It follows then that $d\mathbf{v}(s)/ds$ is also parallel to $\hat{\mathbf{c}}$ and $\mathbf{v}(s)$. Thus we have

$$\frac{d\mathbf{v}(s)}{ds} = \alpha(s)\mathbf{v}(s), \quad (97)$$

where $\alpha(s)$ is a scalar that is determined below. From Eq.(89) we obtain

$$\frac{d\mathbf{v}}{ds} = \mathbf{H}^T \left(-\mathbf{Y} \frac{d\bar{\mathbf{g}}}{ds} + \mathbf{H} \frac{d\bar{\boldsymbol{\lambda}}}{ds} \right), \quad (98)$$

while the definition of \mathbf{v} , Eq.(80), can be reexpressed as

$$\mathbf{v}(s) \equiv \hat{\mathbf{B}}^T \bar{\mathbf{X}}(s) = \mathbf{H}^T(s)\mathbf{y}(s). \quad (99)$$

Hence Eq.(98) can be re-expressed as

$$\begin{aligned} 0 &= \frac{d\mathbf{v}}{ds} - \alpha\mathbf{v} \\ &= \mathbf{H}^T \left(\mathbf{H} \frac{d\bar{\boldsymbol{\lambda}}}{ds} - \mathbf{Y} \frac{d\bar{\mathbf{g}}}{ds} - \alpha\mathbf{y} \right). \end{aligned} \quad (100)$$

Thus, for arbitrary $\alpha(s)$ and $d\bar{\mathbf{g}}/ds$, Eq.(98) is satisfied if $d\bar{\boldsymbol{\lambda}}/ds$ is taken as the least squares solution to

$$\min \left\| \mathbf{H} \frac{d\bar{\boldsymbol{\lambda}}}{ds} - \mathbf{Y} \frac{d\bar{\mathbf{g}}}{ds} - \alpha\mathbf{y} \right\|_2, \quad (101)$$

for then the last line of Eq.(100) is zero. This solution can be written

$$\frac{d\bar{\boldsymbol{\lambda}}}{ds} = \dot{\boldsymbol{\lambda}}^g + \alpha\dot{\boldsymbol{\lambda}}^y, \quad (102)$$

where $\dot{\boldsymbol{\lambda}}^g$ is the least-squares solution to

$$\min \left\| \mathbf{H}\dot{\boldsymbol{\lambda}}^g - \mathbf{Y} \frac{d\bar{\mathbf{g}}}{ds} \right\|_2, \quad (103)$$

and $\dot{\boldsymbol{\lambda}}^y$ is the least-squares solution to

$$\min \|\mathbf{H}\dot{\boldsymbol{\lambda}}^y - \mathbf{y}\|_2. \quad (104)$$

Note that $d\bar{\mathbf{g}}/ds$ is known from Eq.(96), i.e.,

$$\frac{d\bar{\mathbf{g}}}{ds} = \tilde{\mathbf{g}}^u - \bar{\mathbf{g}}(0). \quad (105)$$

The value of the scalar $\alpha(s)$ is determined by the condition that the residual q Eq.(82) remain zero. Differentiating this equation (with $q = 0$), and using Eq.(86) and Eq.(102) we obtain

$$\begin{aligned} 0 &= \mathbf{Q}^T \frac{d\bar{\mathbf{X}}}{ds} \\ &= \mathbf{Q}^T \left(-\mathbf{Y}^2 \frac{d\bar{\mathbf{g}}}{ds} + \mathbf{Y}\mathbf{H} \frac{d\bar{\boldsymbol{\lambda}}}{ds} \right) \\ &= \mathbf{Q}^T \left(-\mathbf{Y}^2 \frac{d\bar{\mathbf{g}}}{ds} + \mathbf{Y}\mathbf{H}\dot{\boldsymbol{\lambda}}^g + \alpha \mathbf{Y}\mathbf{H}\dot{\boldsymbol{\lambda}}^y \right). \end{aligned} \quad (106)$$

Hence α is obtained as

$$\alpha = \left(\mathbf{Q}^T \mathbf{Y}^2 \frac{d\bar{\mathbf{g}}}{ds} - \mathbf{Q}^T \mathbf{Y}\mathbf{H}\dot{\boldsymbol{\lambda}}^g \right) / D, \quad (107)$$

where D is defined by

$$D \equiv \mathbf{Q}^T \mathbf{Y}\mathbf{H}\dot{\boldsymbol{\lambda}}^y. \quad (108)$$

It is shown in Appendix B that the denominator D is strictly positive, so that $\alpha(s)$ is well-conditioned.

In summary, the above development describes the Gibbs function continuation method and establishes that it is guaranteed to yield the correct solution for the constraint potentials $\hat{\boldsymbol{\lambda}}$ even if the Newton matrix $\mathbf{H}^T\mathbf{H}$ is singular. From realizable initial conditions at $s = 0$, Eq.(102) is integrated to $s = 1$ to yield $\hat{\boldsymbol{\lambda}} = \bar{\boldsymbol{\lambda}}(1)$. The quantities appearing in Eq.(102) are determined stably: $\dot{\boldsymbol{\lambda}}^g$ and $\dot{\boldsymbol{\lambda}}^y$ as least-squares solutions (Eq.(103) and Eq.(104)), and α from Eq.(107) in which D is strictly positive.

6 Solution Algorithm

6.1 Overview

The Gibbs function continuation method described in the previous section provides the basis for the algorithm described here. There are, however, further important steps needed to produce a robust and efficient algorithm, and

these are described in this section. For the determination of the constrained equilibrium composition at fixed temperature and pressure, these steps are

1. To ensure that the problem is well-posed (i.e., that \mathbf{c} is a proper constraint vector)
2. To determine initial conditions $\bar{\mathbf{X}}(0)$, $\bar{\boldsymbol{\lambda}}(0)$ and $\bar{\mathbf{g}}(0)$ and
3. To integrate the ODE Eq.(102) for the constraint potentials, to obtain the required solution $\hat{\boldsymbol{\lambda}} = \bar{\boldsymbol{\lambda}}(1)$.

The algorithms to perform these steps are described in Sections 6.2-6.5.

For problems in which the enthalpy is specified, the fixed-temperature algorithm is used, with an outer iteration on temperature. This is described in Section 7.

6.2 Well-Posedness and Perturbation

We claim that the algorithm described here converges to the correct solution for all well-posed constrained-equilibrium problems. The first step is to examine the specified constraint vector \mathbf{c} to determine the problem's well-posedness. There are three possibilities:

1. The constraint vector \mathbf{c} corresponds to a well-posed problem.
2. With an acceptably small perturbation to \mathbf{c} , the problem is well-posed.
3. There is no acceptably small perturbation to \mathbf{c} that yields a well-posed problem.

The third outcome corresponds to “bad user input”: there is no composition (with non-negative species) that satisfies the constraints.

There are two issues connected with the second outcome, pertaining to infinite and finite-precision arithmetic, respectively. The specified constraint vector \mathbf{c} could correspond to a point on the boundary of the feasible region, so that one or more of the species is identically zero. The problem is well posed, but the mole fraction of the absent species cannot be represented in terms of (finite) constraint potentials. There are several ways to overcome this difficulty: we take the simple approach of perturbing \mathbf{c} slightly so that the originally-absent species are strictly positive, but with chemically negligible concentrations.

The second issue pertains to round-off error. Let \mathbf{U} be an $n_s \times n_s$ orthogonal matrix (so that $\mathbf{U}\mathbf{U}^T = \mathbf{I}$), and let \mathbf{N} denote the moles of species, with the sum \bar{N} being of order unity. Suppose that N_k is zero, for some k , and in the finite-precision arithmetic we form

$$\mathbf{N}' = \mathbf{U}(\mathbf{U}^T\mathbf{N}).$$

Whereas in infinite-precision arithmetic $\mathbf{N}' = \mathbf{N}$ and $N'_k = 0$, in finite precision arithmetic N'_k is of order the unit roundoff—either positive or negative. In finite-precision arithmetic there is, therefore, an inevitable uncertainty about whether or not \mathbf{c} corresponds to a feasible constraint vector.

Given the specified constraint vector \mathbf{c} , our approach to these issues is to perturb \mathbf{c} (if necessary) to form a perturbed constraint vector \mathbf{c}^+ corresponding to a well-posed problem (i.e., \mathbf{c}^+ corresponds to an interior point of the feasible region, in which all species are strictly positive). If the perturbation is above a specified threshold, then it is deemed that the specified problem is not well posed, i.e., \mathbf{c} corresponds to “bad user input.” Otherwise, the equilibrium calculation is performed for the possibly-perturbed constraint vector.

The construction of the perturbed constraint vector \mathbf{c}^+ is now described. The moles of atoms of the elements \mathbf{N}^e is known from the specified constraint vector \mathbf{c} , see Eq.(15). We define

$$N_{\max}^e \equiv \max_j(N_j^e), \quad (109)$$

and then specify a lower limit on the moles of each element as

$$N_{\text{low}}^e \equiv \varepsilon_e N_{\max}^e, \quad (110)$$

with $\varepsilon_e = 10^{-9}$, say. Then the (possibly) perturbed moles of each element is defined by

$$N_j^{e+} \equiv \max(N_j^e, N_{\text{low}}^e). \quad (111)$$

With \mathbf{N}^{e+} thus determined, an upper bound on the moles of species k is

$$N_k^{up} \equiv 1 / \max_j(E_{kj}/N_{(j)}^{e+}). \quad (112)$$

That is, for $N_k = N_k^{up}$, species k contains all of the atoms of some element and hence a larger value of N_k is impossible.

For a determined species, k , the specified moles N_k are (possibly) perturbed to

$$N_k^+ \equiv \max(N_k, 0). \quad (113)$$

For the undetermined species, the max-min problem is solved to yield the feasible species moles N_k^{mm} . The (possibly) perturbed moles of the undetermined species k are then defined as

$$N_k^+ \equiv \max(N_k^{mm}, \varepsilon_s N_k^{up}), \quad (114)$$

with $\varepsilon_s = 10^{-9}$, say. The (possibly) perturbed constraints are defined by

$$\mathbf{c}^+ \equiv \mathbf{B}^T \mathbf{N}^+. \quad (115)$$

It is important to appreciate that the only lower bound imposed on the constrained equilibrium species moles \mathbf{N}^{CE} is zero. The quantities \mathbf{N}^+ are defined merely to obtain \mathbf{c}^+ . For an unconstrained equilibrium problem, the perturbation is to the moles of elements in the undetermined species, and is of relative magnitude $\max(\varepsilon_e, \varepsilon_s)$, e.g., 10^{-9} . As discussed in Appendix C, such small perturbations are not significant.

6.3 Initial Conditions

Initial estimates are first made of the moles \mathbf{N} , leading to the mole fractions $\bar{\mathbf{X}}(0)$ of the undetermined species.

The moles of the determined species are known (as the first n_{sd} elements of the modified constraint vector $\tilde{\mathbf{c}} = \mathbf{A}\mathbf{c}$, see Eq.(26)). Two linear-programming problems are solved to obtain feasible values of the moles of undetermined species, \mathbf{N}^u . First, the max-min composition \mathbf{N}^{mm} is determined (see Appendix A) as part of the perturbation treatment described in the preceding subsection. Second, the *min-g* composition \mathbf{N}^g is defined as the solution to

$$\text{minimize } \sum_{k=1}^{n_{su}} N_k^u \tilde{g}_k^u, \quad (116)$$

subject to

$$\hat{\mathbf{B}}^T \mathbf{N}^u = \hat{\mathbf{c}} \quad \text{and} \quad N_k^u \geq 0. \quad (117)$$

In view of the fact that Eq.(117) imposes n_{rc} equality constraints, at most n_{rc} components of \mathbf{N}^g are non-zero. This min-g composition \mathbf{N}^g is the initial

estimate used in STANJAN. The initial condition we use is a blending of \mathbf{N}^g and \mathbf{N}^{mm} , i.e.,

$$\mathbf{N}^u = \mathbf{N}^g + f_{mm}(\mathbf{N}^{mm} - \mathbf{N}^g), \quad (118)$$

with $0 < f_{mm} \leq 1$ (e.g., $f_{mm} = 0.1$). Note that many components of \mathbf{N}^g are zero, whereas all components of \mathbf{N}^{mm} are strictly positive. Hence, with $f_{mm} > 0$, all components of \mathbf{N}^u are strictly positive.

The initial condition for $\bar{\mathbf{X}}(0)$ —the mole fraction of undetermined species—is then evaluated from the above specifications of \mathbf{N}^d and \mathbf{N}^u . The initial condition for the reduced constraint potentials $\bar{\boldsymbol{\lambda}}(0)$ is taken to be the least-squares solution to

$$\min \| -\tilde{\mathbf{g}}^u + \hat{\mathbf{B}}\bar{\boldsymbol{\lambda}}(0) - \ln \bar{\mathbf{X}}(0) \|_2. \quad (119)$$

This is based on the observation (from Eq.(65)) that the constraint potential satisfies the equation

$$-\tilde{\mathbf{g}}^u + \hat{\mathbf{B}}\hat{\boldsymbol{\lambda}} - \ln \mathbf{X}^u = 0. \quad (120)$$

Finally the initial condition $\bar{\mathbf{g}}(0)$ is set as

$$\bar{\mathbf{g}}(0) = \hat{\mathbf{B}}\bar{\boldsymbol{\lambda}}(0) - \ln \bar{\mathbf{X}}(0). \quad (121)$$

The properties of these initial conditions are that $\bar{\mathbf{X}}(0)$ is consistently given by

$$\bar{\mathbf{X}}(0) = \exp[-\bar{\mathbf{g}}(0) + \hat{\mathbf{B}}\bar{\boldsymbol{\lambda}}(0)], \quad (122)$$

(on account of Eq.(121)), and that the residuals \mathbf{r} and q (Eq.(81) and Eq.(82)) are zero. This is so because both \mathbf{N}^g and \mathbf{N}^{mm} are feasible solutions, and therefore so also is \mathbf{N}^u given by Eq.(118).

6.4 Time Stepping

From the initial conditions at $s = 0$, the ODE for $\bar{\boldsymbol{\lambda}}(s)$, Eq.(102), is numerically integrated forwards in pseudo-time to $s = 1$, to yield the required solution $\hat{\boldsymbol{\lambda}} = \bar{\boldsymbol{\lambda}}(1)$. The straightforward approach is to employ an accurate ODE solver to obtain the solution, but there are reasons to seek an alternative. In the straightforward ODE integration approach, the numerical error in $\hat{\boldsymbol{\lambda}}$ is the global error incurred in integrating the ODE from $s = 0$ to $s = 1$. Hence an accurate ODE solution, involving many small steps, is needed to obtain an accurate solution. But for a given value of s , $\bar{\boldsymbol{\lambda}}(s)$ may

be determinable (for example, by using Newton’s method) from the algebraic equations Eq.(79)-Eq.(82), independent of the solution $\bar{\lambda}(s')$ at earlier times $s' < s$. However, Newton’s method is not guaranteed to converge, either because the initial guess is too far from the solution, or for the reasons given in Section 5.3.

These considerations lead to a hybrid method as follows.

1. Initially s is zero and the time step Δs is unity.
2. The explicit Euler method is used to integrate the ODE Eq.(102) numerically (in a single time step of size Δs) to yield an approximation to $\bar{\lambda}(s + \Delta s)$.
3. Newton’s method (described in the next subsection) is used to attempt to obtain an accurate solution for $\bar{\lambda}(s + \Delta s)$ (i.e., by reducing the residuals below a specified error tolerance).
4. If the Newton iteration fails to achieve an accurate solution, then the time step is rejected, Δs is divided by 4, and then step (2) is repeated.
5. If the Newton iteration succeeds, the solution $\bar{\lambda}(s + \Delta s)$ is accepted, s is reset to $s + \Delta s$, and Δs is set to $\min(2\Delta s, 1 - s)$.
6. If s is less than unity, step (2) is repeated. Otherwise, the solution $\hat{\lambda} = \bar{\lambda}(1)$ is obtained.

In the implementation of Newton’s method, the first operation performed is the evaluation of the residuals. If (initially or on subsequent iterations) the residuals are below the specified tolerance (i.e., $|\mathbf{r}| < \varepsilon_{\text{tol}}$, with $\varepsilon_{\text{tol}} = 10^{-9}$, say), then the iteration is terminated successfully. Otherwise the iterations are terminated if the residual $|\mathbf{r}|$ is greater than 90% of its value on the previous iteration.

The criterion for accepting the time step (in step (4)) requires careful consideration. For very difficult cases, the Newton iterations may repeatedly fail, so that the method reduces to ODE integration by the explicit Euler scheme. The local error incurred in the explicit Euler step can be reduced to the level of round-off error by reducing Δs repeatedly (in step (4) when the time step is rejected). But the residuals measure the global error, and the ODE integration cannot be expected to reduce this to below its level at the

start of the time step. Hence the criterion used for accepting the time step is

$$|\mathbf{r}| < |\mathbf{r}|_{\text{last}} + \max(\varepsilon_{\text{tol}}, 0.05|\mathbf{r}|_{\text{last}}), \quad (123)$$

where $|\mathbf{r}|_{\text{last}}$ is the residual on the last accepted step. In other words, the step is accepted if the local error is less than $\max(\varepsilon_{\text{tol}}, 0.05|\mathbf{r}|_{\text{last}})$. Provided that ε_{tol} is not too small compared to the unit round-off, this criterion can always be met by a sufficiently small time step. Hence the method is guaranteed to converge.

If a succession of Euler steps are performed (without a convergent Newton iteration), the global error can increase exponentially (by a factor of 1.05 on each step). But it is generally found that Newton's method becomes convergent before s reaches unity, in which case the residuals are controlled below ε_{tol} .

More accurate time-stepping strategies (e.g., 2nd and 4th order Runge-Kutta schemes) were investigated, but it was found that these generally lead to a deterioration in efficiency.

6.5 Newton's Method

For given fixed s , $\bar{\mathbf{g}}(s)$ and an initial guess for $\bar{\boldsymbol{\lambda}}(s)$, Newton's method is used to attempt to reduce the residuals \mathbf{r} and q (Eq.(81) and Eq.(82)) below the specified tolerance ε_{tol} .

From the given values of $\bar{\mathbf{g}}$ and $\bar{\boldsymbol{\lambda}}$, $\bar{\mathbf{X}}$ and \mathbf{v} are evaluated from Eq.(79) and Eq.(80). We then seek an increment in \mathbf{v} , $\delta\mathbf{v}$, so that the residual \mathbf{r} is zero, i.e.,

$$0 = \frac{\mathbf{v} + \delta\mathbf{v}}{|\mathbf{v} + \delta\mathbf{v}|} - \frac{\hat{\mathbf{c}}}{|\hat{\mathbf{c}}|}, \quad (124)$$

which we re-write as

$$\delta\mathbf{v} = (1 + \delta\alpha)\delta\mathbf{w} + \delta\alpha\mathbf{v}, \quad (125)$$

with

$$\delta\mathbf{w} \equiv |\mathbf{v}| \left(\frac{\hat{\mathbf{c}}}{|\hat{\mathbf{c}}|} - \frac{\mathbf{v}}{|\mathbf{v}|} \right) = -|\mathbf{v}|\mathbf{r}, \quad (126)$$

and

$$\delta\alpha = \frac{|\mathbf{v} + \delta\mathbf{v}|}{|\mathbf{v}|} - 1. \quad (127)$$

Note that at convergence $\delta\mathbf{w}$ is zero, and that close to convergence \mathbf{v} and $\delta\mathbf{w}$ are almost orthogonal.

An infinitesimal change $d\mathbf{v}$ in \mathbf{v} is related to an infinitesimal change $d\bar{\boldsymbol{\lambda}}$ in $\bar{\boldsymbol{\lambda}}$ by Eq.(90):

$$d\mathbf{v} = \mathbf{H}^T \mathbf{H} d\bar{\boldsymbol{\lambda}}. \quad (128)$$

Now if the vectors $\boldsymbol{\lambda}^v$ and $\delta\boldsymbol{\lambda}^w$ satisfy

$$\mathbf{H}^T \mathbf{H} \boldsymbol{\lambda}^v = \mathbf{v}, \quad (129)$$

and

$$\mathbf{H}^T \mathbf{H} \delta\boldsymbol{\lambda}^w = \delta\mathbf{v}, \quad (130)$$

then, to a linear approximation,

$$\delta\bar{\boldsymbol{\lambda}} \equiv (1 + \delta\alpha) \delta\boldsymbol{\lambda}^w + \delta\alpha \boldsymbol{\lambda}^v, \quad (131)$$

yields the change in \mathbf{v} , $\delta\mathbf{v}$, given by Eq.(125). Thus, the implementation of Newton's method involves computing changes $\delta\bar{\boldsymbol{\lambda}}$ from Eq.(131).

Since \mathbf{v} is given by $\mathbf{v} = \mathbf{H}^T \mathbf{y}$, Eq.(129) is satisfied by the least-squares solution

$$\min \|\mathbf{H} \boldsymbol{\lambda}^v - \mathbf{y}\|_2, \quad (132)$$

whether or not \mathbf{H} has full column rank. On the other hand, if \mathbf{H} is rank deficient, there may be no vector $\delta\boldsymbol{\lambda}^w$ which satisfies Eq.(130). We therefore take $\delta\boldsymbol{\lambda}^w$ to be the minimum norm solution, which is obtained through the SVD of \mathbf{H} .

The remaining unknown, $\delta\alpha$, is determined by requiring $q + \delta q = 0$ (to within a linear approximation). Now, from Eq.(82) and Eq.(86) (with $d\bar{\mathbf{g}} = 0$) we obtain

$$\begin{aligned} dq &= -\mathbf{Q}^T d\bar{\mathbf{X}} \\ &= -\mathbf{Q}^T \mathbf{Y} \mathbf{H} d\bar{\boldsymbol{\lambda}} \\ &= -\mathbf{P}^T d\bar{\boldsymbol{\lambda}}, \end{aligned} \quad (133)$$

with the definition

$$\mathbf{P} \equiv \mathbf{H}^T \mathbf{Y} \mathbf{Q}. \quad (134)$$

Thus, to a linear approximation, Eq.(131) and Eq.(133) yield

$$-\delta q = (1 + \delta\alpha) \mathbf{P}^T \delta\boldsymbol{\lambda}^w + \delta\alpha \mathbf{P}^T \boldsymbol{\lambda}^v, \quad (135)$$

and so the requirement $q + \delta q = 0$ yields

$$\delta\alpha = \frac{q - \mathbf{P}^T \delta\boldsymbol{\lambda}^w}{\mathbf{P}^T (\boldsymbol{\lambda}^v + \delta\boldsymbol{\lambda}^w)}. \quad (136)$$

At convergence $\mathbf{P}^T \delta\boldsymbol{\lambda}^w$ is zero and $\mathbf{P}^T \boldsymbol{\lambda}^v$ is unity, Hence, for robustness without sacrificing asymptotic convergence, in practice the denominator in Eq.(136) is limited to be greater than $\frac{1}{2}$.

In summary, the implementation of Newton's method consists of incrementing $\bar{\boldsymbol{\lambda}}$ by $\delta\bar{\boldsymbol{\lambda}}$ given by Eq.(131), where $\delta\alpha$ is given by Eq.(136) (possibly with a modified denominator). If the iteration converges (i.e., $\delta\bar{\boldsymbol{\lambda}}$ tends to zero), then the residual q tends to zero, and the residual \mathbf{r} tends to a limit which is zero if \mathbf{H} has (numerical) full rank, but may otherwise be non-zero.

7 Fixed Enthalpy and Pressure

The algorithm described above determines the constrained equilibrium composition at specified fixed pressure p and temperature T . For the fixed enthalpy case, an outer iteration over temperature is performed, with each successive temperature being a refined estimate of the temperature at which the enthalpy is that specified. Typically excellent convergence is achieved in approximately 4 iterations.

The outer iteration is straightforward—based on Newton's method. The only point worthy of mention is the determination of the *constrained specific heat*, \bar{C}_{pc} , defined below.

The molar specific enthalpy of the constrained equilibrium mixture is

$$\bar{H} = h_k(T)X_k, \quad (137)$$

where $h_k(T)$ is the molar specific enthalpy of species k . The molar constant-pressure specific heat of the mixture \bar{C}_p is defined as

$$\bar{C}_p \equiv \left(\frac{\partial \bar{H}}{\partial T} \right)_{p, \mathbf{X}} = \frac{dh_k}{dT} X_k = C_{p,k} X_k, \quad (138)$$

where $C_{p,k}$ is the molar constant-pressure specific heat of species k . In contrast, the *constrained molar specific heat* \bar{C}_{pc} is defined by

$$\bar{C}_{pc} \equiv \left(\frac{\partial \bar{H}}{\partial T} \right)_{p,c}, \quad (139)$$

i.e., the rate of change of \bar{H} with T at constant pressure and constant constraints, \mathbf{c} .

From Eq.(137) we obtain

$$\begin{aligned}\bar{C}_{pc} &= \frac{dh_k}{dT}X_k + h_k \left(\frac{\partial X_k}{\partial T} \right)_{p,c} \\ &= \bar{C}_p + h_k \left(\frac{\partial X_k}{\partial T} \right)_{p,c}.\end{aligned}\tag{140}$$

For the undetermined species, from Eq.(86) we obtain

$$\left(\frac{\partial \mathbf{X}^u}{\partial T} \right)_{p,c} = -\mathbf{Y}^2 \left(\frac{\partial \tilde{\mathbf{g}}^u}{\partial T} \right)_p + \mathbf{YH} \frac{\partial \hat{\boldsymbol{\lambda}}}{\partial T},\tag{141}$$

and $\partial \hat{\boldsymbol{\lambda}}/\partial T$ can be obtained from the analogue to Eq.(102).

Thus \bar{C}_{pc} can be obtained for each fixed T constrained equilibrium solution and used to obtain the value of T for the next iteration.

8 Conclusions

An efficient methodology has been developed to compute the constrained equilibrium composition of ideal gas mixtures.

The constraints considered are: on individual species; on elements; and, general linear constraints. The *basic constraint equation*, Eq.(13), stemming from these constraints is equivalent to the *reduced constraint equation*, Eq.(27), which generally involves fewer constraints and only the undetermined species. For fixed temperature and pressure, the constrained equilibrium composition is that which minimizes the Gibbs function, subject to the constraints. It is shown that, for well-posed constraints, there exists a unique constrained equilibrium composition.

The mole fractions of the undetermined species which minimize the Gibbs function are given explicitly in terms of Lagrange multipliers, or the *reduced constraint potentials*, Eq.(71). These constraint potentials are determined by the non-linear equation system consisting of the reduced constraints, Eq.(69), and the normalization condition, Eq.(75).

An analysis of this non-linear equation system (in Section 5) shows that Newton's method can fail, because the iteration matrix can be singular. The

method of *Gibbs function continuation* is introduced, which is guaranteed to determine the solution for all well-posed problems. A practical algorithm based on this method is described in Section 6, and this has been implemented in a Fortran code (Pope 2003). It is indeed found that the code succeeds in determining the constrained equilibrium composition for a comprehensive range of test problems.

Acknowledgements

I am grateful to Zhuyin Ren for valuable suggestions and for testing the constrained-equilibrium code, and to Dr. R.S. Barlow for providing the data in Table 3.

A The Max-Min Composition

Given the species moles \mathbf{N} , the minimum species moles is

$$N_{\min} \equiv \min_k(N_k). \quad (142)$$

By definition, the max-min composition $\mathbf{N} = \mathbf{N}^{mm}$ is that which maximizes N_{\min} , subject to the constraints

$$\mathbf{B}^T \mathbf{N} = \mathbf{c}. \quad (143)$$

If the max-min value N_{\min}^{mm} is negative, then no realizable compositions satisfy the constraints. The max-min composition can be determined as the solution of the following linear programming problem. Define

$$n = n_s + 1, \quad (144)$$

$$x_k = N_k, \text{ for } k = 1, n_s \quad (145)$$

$$x_n = N_{\min}, \quad (146)$$

$$\mathbf{f}_{(n \times 1)} = [0 \ 0 \ \dots \ 0 \ -1]^T, \quad (147)$$

$$\mathbf{A}_{(n_s \times n)} = \begin{bmatrix} -1 & 0 & \dots & 1 \\ 0 & -1 & & 1 \\ \vdots & & \ddots & \vdots \\ & & & -1 & 1 \end{bmatrix}, \quad (148)$$

$$\mathbf{A}_{(n_c \times n)}^e = [\mathbf{B}^T \ 0], \quad (149)$$

$$\mathbf{b}_{(n_s \times 1)} = 0. \quad (150)$$

Then determine \mathbf{x} that solves

$$\min_{\mathbf{x}} (\mathbf{f}^T \mathbf{x}), \quad (151)$$

subject to

$$\mathbf{A}\mathbf{x} \leq \mathbf{b}, \quad (152)$$

and

$$\mathbf{A}^e \mathbf{x} = \mathbf{c}. \quad (153)$$

Equation (151) calls for the maximization of N_{\min} ; Eq.(152) corresponds to $N_k \geq N_{\min}$; and Eq.(153) enforces the equality constraint equation, Eq.(143).

B Demonstration that D is Strictly Positive

The purpose of this Appendix is to show that denominator in Eq.(107), namely,

$$D \equiv \mathbf{Q}^T \mathbf{Y} \mathbf{H} \dot{\lambda}^y, \quad (154)$$

is strictly positive.

The vector \mathbf{Q} (Eq.(76)) has components

$$Q_i = \delta_{i(i)} + (\bar{N}^d / N^{au}) \mathbf{w},$$

and so we have

$$\mathbf{Y} \mathbf{Q} = \mathbf{y} + (\bar{N}^d / N^{au}) \mathbf{Y} \mathbf{w}. \quad (155)$$

Now, from Eq.(72) and Eq.(87) we obtain

$$\mathbf{Y} \mathbf{w} = \mathbf{Y} \hat{\mathbf{B}} \hat{\mathbf{w}} = \mathbf{H} \hat{\mathbf{w}},$$

and hence (from Eq.(155))

$$\mathbf{Q}^T \mathbf{Y} = \mathbf{y}^T + (\bar{N}^d / N^{au}) \hat{\mathbf{w}}^T \mathbf{H}^T. \quad (156)$$

Using this result, D is re-expressed as

$$\begin{aligned} D &= \mathbf{Q}^T \mathbf{Y} ([\mathbf{H} \dot{\boldsymbol{\lambda}}^y - \mathbf{y}] + \mathbf{y}) \\ &= \mathbf{y}^T \mathbf{H} \dot{\boldsymbol{\lambda}}^y + (\bar{N}^d / N^{au}) \hat{\mathbf{w}}^T \mathbf{H}^T (\mathbf{H} \dot{\boldsymbol{\lambda}}^y - \mathbf{y}) + (\bar{N}^d / N^{au}) \hat{\mathbf{w}}^T \mathbf{H}^T \mathbf{y} \\ &= D_1 + D_2 + D_3, \end{aligned} \quad (157)$$

where D_1 , D_2 and D_3 denote the three terms in the middle line.

We immediately observe that D_2 is zero, since the residual $\mathbf{H} \dot{\boldsymbol{\lambda}}^y - \mathbf{y}$ is orthogonal to \mathbf{H} .

Multiplying Eq.(155) by \mathbf{y}^T we obtain

$$\mathbf{y}^T \mathbf{Y} \mathbf{Q} = \bar{\mathbf{X}}^T \mathbf{Q} = 1 = \mathbf{y}^T \mathbf{y} + (\bar{N}^d / N^{au}) \mathbf{y}^T \mathbf{H} \hat{\mathbf{w}}. \quad (158)$$

The final term is just D_3 , while we have

$$\mathbf{y}^T \mathbf{y} = \sum_{k=1}^{n_{su}} X_i^u = \bar{N}^u / \bar{N} = 1 - \bar{N}^d / \bar{N}. \quad (159)$$

We thus obtain

$$D_3 = 1 - \mathbf{y}^T \mathbf{y} = \bar{N}^d / \bar{N} \geq 0. \quad (160)$$

In order to evaluate D_1 , we express $\dot{\boldsymbol{\lambda}}^y$ in terms of the SVD of \mathbf{H} :

$$\mathbf{H} = \mathbf{U} \boldsymbol{\Sigma} \mathbf{V}^T, \quad (161)$$

where \mathbf{U} is an $n_{su} \times n_{rc}$ matrix with orthonormal columns; \mathbf{V} is an $(n_{rc} \times n_{rc})$ orthogonal matrix; and $\boldsymbol{\Sigma}$ is the $n_{rc} \times n_{rc}$ diagonal matrix of singular values $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_{n_{rc}} \geq 0$. The least squares solution for $\dot{\boldsymbol{\lambda}}^y$ is then given by

$$\dot{\boldsymbol{\lambda}}^y = \mathbf{V} \tilde{\boldsymbol{\Sigma}} \mathbf{U}^T \mathbf{y}, \quad (162)$$

where $\tilde{\boldsymbol{\Sigma}}$ denotes the pseudo-inverse of $\boldsymbol{\Sigma}$. For D_1 we then obtain

$$\begin{aligned} D_1 &\equiv \mathbf{y}^T \mathbf{H} \dot{\boldsymbol{\lambda}}^y = \mathbf{y}^T \mathbf{U} \boldsymbol{\Sigma} \mathbf{V}^T \mathbf{V} \tilde{\boldsymbol{\Sigma}} \mathbf{U}^T \mathbf{y} \\ &= (\mathbf{U}^T \mathbf{y})^T (\boldsymbol{\Sigma} \tilde{\boldsymbol{\Sigma}}) (\mathbf{U}^T \mathbf{y}). \end{aligned} \quad (163)$$

In order to compare it to D_1 , we consider

$$|\mathbf{v}|^2 = \mathbf{y}^T \mathbf{H} \mathbf{H}^T \mathbf{y} = (\mathbf{U}^T \mathbf{y})^T \boldsymbol{\Sigma}^2 (\mathbf{U}^T \mathbf{y}). \quad (164)$$

Table 1: Approximate composition of air at 20°C, 1 bar, 40% relative humidity. (Adapted from McMurray & Fay 1998.)

Species	Conc. (ppm)	Mole fraction
N ₂	780,800	7.81 x 10 ⁻¹
O ₂	209,550	2.10 x 10 ⁻¹
H ₂ O	9,360	9.36 x 10 ⁻³
Ar	9,300	9.30 x 10 ⁻³
CO ₂	360	3.6 x 10 ⁻⁴
Ne	18	1.8 x 10 ⁻⁵
He	5	5 x 10 ⁻⁶
CH ₄	2	2 x 10 ⁻⁶
Kr	1	1 x 10 ⁻⁶

Now each diagonal component of $\Sigma\tilde{\Sigma}$ is unity, unless the corresponding singular value is zero. Thus each component of $\sigma_1^2\Sigma\tilde{\Sigma}$ is greater than or equal to Σ^2 . From this fact, and from the positive definite forms of Eq.(163) and Eq.(164) we obtain

$$D_1 \geq |\mathbf{v}|^2/\sigma_1^2. \quad (165)$$

Note that both $|\mathbf{v}|^2$ and σ_1^2 are well-conditioned, strictly positive quantities.

In summary: it has been shown that D is strictly positive by virtue of the facts that D_1 is strictly positive, D_2 is zero, and D_3 is non-negative.

C Composition of Air and Methane

To illustrate the fact that the perturbation possibly made to the reduced constraint vector (see Section 6.2) to ensure well-posedness is of little physical or chemical significance, we examine here the elemental compositions of air and 99% pure methane.

Table 1 shows the composition of air under typical atmospheric conditions. (The composition of dry air is taken from McMurray and Fay 1998 and humidity is added.) From this the elemental composition is evaluated and given in Table 2. As may be seen, the mole fractions of H and C atoms are greater than 10⁻⁴ those of N and O.

Table 2: Elemental composition of air.

Element	Conc. (ppm)	Mole fraction
N	773,378	7.73×10^{-1}
O	212,550	2.13×10^{-1}
H	9,275	9.28×10^{-3}
C	179	1.79×10^{-4}
inerts	4,618	4.62×10^{-3}

Table 3: Approximate composition of 99% pure methane (Matheson CP grade).

Species	Conc. (ppm)	Mole fraction
CH ₄	990,000	9.9×10^{-1}
N ₂	6,000	6.0×10^{-3}
CO ₂	2,000	2.0×10^{-3}
C ₂ H ₆	1,200	1.2×10^{-3}
C ₃ H ₈	300	3×10^{-4}
O ₂	50	5×10^{-5}
H ₂ O	5	5×10^{-6}

Similarly, Table 3 shows the composition of 99% pure methane (Matheson CP grade), which is the purest form typically used in combustion experiments. The corresponding elemental composition is shown in Table 4. As may be seen, the mole fractions of N and O atoms are greater than 10^{-4} times those of H and C.

Thus in any mixture of air and 99% pure methane, the mole fractions of C, H, N, and O are all greater than 10^{-4} ; and so a perturbation of order 10^{-9} is negligible in comparison.

Table 4: Approximate elemental composition of 99% pure methane.

Element	Conc. (ppm)	Mole fraction
H	796,948	7.97×10^{-1}
C	199,819	2.00×10^{-1}
N	2,409	2.4×10^{-3}
O	824	8.2×10^{-4}

References

- Bishnu, P. S., D. Hamiroune, M. Metghalchi, and J. C. Keck (1997). Constrained-equilibrium calculations for chemical systems subject to generalized linear constraints using the NASA and STANJAN equilibrium programs. *Combust. Theory Modelling 1*, 295–312.
- Gordon, S. and B. J. McBride (1971). Computer program for calculations of complex chemical equilibrium compositions, rocket performances, incident and reflected shocks, and Chapman-Jouguet detonations. Technical report, NASA SP-273.
- Hamiroune, D., P. Bishnu, M. Metghalchi, and J. C. Keck (1998). Rate-controlled constrained-equilibrium method using constraint potentials. *Combust. Theory Modelling 2*, 81–94.
- Keck, J. C. (1990). Rate-controlled constrained-equilibrium theory of chemical reactions in complex systems. *Prog. Energy Combust. Sci. 16*, 125–154.
- Keck, J. C. and D. Gillespie (1971). Rate-controlled partial-equilibrium method for treating reacting gas-mixtures. *Combust. Flame 17*, 237.
- McMurray, J. and R. C. Fay (1998). *Chemistry* (2nd ed.). New Jersey: Prentice Hall.
- Pope, S. B. (2003). CEQ: A Fortran library to compute equilibrium compositions using Gibbs function continuation. <http://mae.cornell.edu/~pope/CEQ>.
- Reynolds, W. C. (1986). The element potential method for chemical equilibrium analysis: implementation in the interactive program STANJAN. Technical report, Stanford University.

Tang, Q. and S. B. Pope (2002). Implementation of combustion chemistry by *in situ* adaptive tabulation of rate-controlled constrained equilibrium manifolds. *Proc. Combust. Inst* 29, 1411–1417.