Consistent thermodynamic derivative estimates for tabular equations of state

Gary A. Dilts
Continuum Dynamics Group, Los Alamos National Laboratory, Mail Stop D413, Los Alamos, New Mexico 87544, USA
(Received 22 March 2005; revised manuscript received 27 December 2005; published 13 June 2006)

A valid fluid equation of state (EOS) must satisfy the thermodynamic conditions of consistency (derivation from a free energy) and stability (positive sound speed squared). Numerical simulations of compressible fluid flow for realistic materials require a tabular EOS, but typical software interfaces to such tables based on polynomial or rational interpolants may enforce the stability conditions, but do not enforce the consistency condition and its derivatives. The consistency condition is important for the computation of various dimensionless parameters of an EOS that may involve derivatives of up to second order which are important for the development of more sensitive artificial viscosities and Riemann solvers that accurately model shock structure in regions near phase transitions. We describe a table interface based on the tuned regression method, which is derived from a constrained local least-squares regression technique. It is applied to several SESAME EOS showing how the consistency and stability conditions can be satisfied to round-off while computing first and second derivatives with demonstrated second-order convergence. An improvement of 14 orders of magnitude over conventional derivatives is demonstrated, although the method is apparently two orders of magnitude slower, due to the fact that every evaluation requires solving an 11-dimensional nonlinear system. Application is made to the computation of the fundamental derivative.

DOI: 10.1103/PhysRevE.73.066704 PACS number(s): 47.11.−j, 64.10.+h, 02.70.−c

I. INTRODUCTION

The two most common techniques for modeling shock waves in the numerical simulation of compressible flows are artificial viscosity [1] and the Riemann solver [2]. Over the years, these have seen many enhancements and variations. For comprehensive summaries, see [3–5]. Unfortunately, the Riemann solver is only well developed for the ideal gas equation of state (EOS). Some attempts have been made for more complicated analytic EOS, such as the Mie-Gruneisen EOS [6,7], but real materials in general have such a complicated EOS that it can only adequately be expressed in a table, for which there is no Riemann solver yet published. The SESAME library [8,9], which is widely used at Los Alamos National Laboratory and has been distributed throughout the world, contains tabular EOS for many elements and will be the source of our examples. Our treatment in this paper is specific to the choice of variables used in SESAME but it may be straightforwardly modified for others.

The mathematical description of the behavior of shock waves in real fluids with an arbitrary equation of state was described in detail in [10]. Four dimensionless quantities are important:

$$\gamma = -\frac{V}{P}\left.\frac{\partial P}{\partial V}\right|_S, \quad \Gamma = -\frac{V}{T}\left.\frac{\partial T}{\partial V}\right|_S,$$

$$g = \frac{PV}{T^2}\left.\frac{\partial T}{\partial S}\right|_V, \quad \mathcal{G} = \frac{1}{2}\frac{V^2}{\gamma P}\left.\frac{\partial^2 P}{\partial V^2}\right|_S. \tag{1}$$

The symbols P, V, T, and S represent pressure, specific volume, temperature, and entropy, respectively. The quantity γ is the adiabatic exponent, Γ is the Grüneisen coefficient, g is the dimensionless specific heat, and $\mathcal{G}$ is the fundamental derivative. The quantities γ and $\mathcal{G}$ represent the slope and curvature of isentropes in the P-V plane, respectively. The quantity $\mathcal{G}$ is most important for the determination of shock-wave structure. When $\mathcal{G}>0$, shocks occur in compression; when $\mathcal{G}<0$, shocks occur in rarefaction. In a numerical simulation this information must be incorporated into the switch used to turn on artificial viscosity or in the solution constructed by a Riemann solver. First- and second-order approximate Riemann solvers for real EOS would make extensive use of $\mathcal{G}$. Clearly, in order to construct these solvers we must first know how to compute physically realistic values of $\mathcal{G}$ from tables. The derivative γ is also used in numerical solution methods for shock-driven flows to determine time steps that satisfy stability conditions.

Assume for the moment we have internal energy E expressed as a function of specific volume V and entropy S. The thermodynamic definitions of pressure P and temperature T are

$$P = -\left.\frac{\partial E}{\partial V}\right|_S, \quad T = \left.\frac{\partial E}{\partial S}\right|_V \tag{2}$$

and imply that

$$\left.\frac{\partial P}{\partial S}\right|_V = -\left.\frac{\partial T}{\partial V}\right|_S. \tag{3}$$

This is the thermodynamic consistency condition, and it amounts to a differential equation that a valid equation of state must satisfy. It is equivalent to the statement that the equation of state is derivable from a thermodynamic potential, in this case the energy. In the SESAME tables, pressure and energy are expressed as functions of temperature and density. With temperature and density independent, the appropriate thermodynamic potential is the Helmholtz free energy, $F=E-TS$, which gives

$$P = -\left.\frac{\partial F}{\partial V}\right|_T, \quad S = -\left.\frac{\partial F}{\partial T}\right|_V, \tag{4}$$

from which we conclude

$$\frac{\partial P}{\partial T} = \frac{\partial S}{\partial V}. \tag{5}$$

Using the second law of thermodynamics, condition (5) takes the form

$$P = T\frac{\partial P}{\partial T} + \rho^2 \frac{\partial E}{\partial \rho}. \tag{6}$$

Thermodynamic stability requires that the Hessian of E be jointly convex in V and S, which leads to the conditions

$$\left.\frac{\partial^2 E}{\partial S^2}\right|_V \geq 0, \quad \left.\frac{\partial^2 E}{\partial V^2}\right|_S \geq 0,$$

$$\left.\frac{\partial^2 E}{\partial S^2}\right|_V \left.\frac{\partial^2 E}{\partial V^2}\right|_S \geq \left(\frac{\partial^2 E}{\partial S\,\partial V}\right)^2. \tag{7}$$

With temperature and density independent these are satisfied if

$$\frac{\partial E}{\partial T} \geq 0, \quad \frac{\partial P}{\partial \rho} \geq 0. \tag{8}$$

Equations (6) and (8) are thus constraints on any derivatives one might construct from the table data for P and E in order for them to be physically realistic. Note that if the pressure and energy are smooth, then the derivatives of Eq. (6) with respect to temperature and density are valid second-order constraints:

$$\rho^2 \frac{\partial^2 E}{\partial T\,\partial \rho} + T\frac{\partial^2 P}{\partial T^2} = 0,$$

$$\rho^2 \frac{\partial^2 E}{\partial \rho^2} + \frac{2}{\rho}\left(P - T\frac{\partial P}{\partial T}\right) + T\frac{\partial^2 P}{\partial T\,\partial \rho} - \frac{\partial P}{\partial \rho} = 0. \tag{9}$$

The consistency condition can be critical in the evaluation of the quantities in Eq. (1). For example, the computation of $\mathcal{G}$ involves second derivatives, so Eqs. (6) and (9) both need to be satisfied to make $\mathcal{G}$ physically realistic. With temperature and density independent the expression for $\mathcal{G}$ independent of the consistency conditions is

$$[-T^2E_{TT}P_T^3 + TE_TP_T^2(-\rho^2E_{T\rho} + P_T + 2TP_{TT}) + 3T\rho^2E_T^2P_TP_{T\rho} + \rho^3E_T^3(2P_\rho + \rho P_{\rho\rho})]/2E_T^2(T\rho P_T^2 + \rho^3E_TP_\rho), \tag{10}$$

where the notation $A_B = \partial A/\partial B$ is used. When the consistency condition and its derivatives with respect to temperature and density are solved for P_T, $P_{T\rho}$, and P_{TT} and applied to (10) we obtain

$$[\rho^3E_T^3(2P_\rho + \rho P_{\rho\rho}) + T^2E_{TT}S_\rho^3 - TE_TS_\rho^2(3\rho^2E_{T\rho} + S_\rho) + 3\rho E_T^2S_\rho(2P + \rho^3E_{\rho\rho} - \rho P_\rho + 2TS_\rho)]/2E_T^2(\rho^3E_TP_\rho + T\rho S_\rho^2), \tag{11}$$

where $TS_\rho = \partial S/\partial\rho|_T = \rho^2E_\rho - P$. These are radically different expressions if the consistency constraints (6) and (9) are not satisfied.

That satisfaction of these constraints is not automatic for traditional derivative evaluation schemes is illustrated by Table I, which shows the average absolute value of the *logarithmic scale* of the normalized consistency error, $\mathrm{ls}(\varepsilon) \equiv \mathrm{sgn}\,\varepsilon \ln|1+\varepsilon|$, where

$$\varepsilon = \left(-P + T\frac{\partial P}{\partial T} + \rho^2\frac{\partial E}{\partial \rho}\right) \Bigg/ \left(|P| + T\left|\frac{\partial P}{\partial T}\right| + \rho^2\left|\frac{\partial E}{\partial \rho}\right|\right), \tag{12}$$

using several common methods [11,12] for computing derivatives of SESAME table 2984 for molybdenum. Since pressure and energy vary by 6 and 12 orders of magnitude, respectively, in this example, the normalizing denominator is necessary for a fair assessment of the error. The table grid was 37×65, and the evaluation grid was 75×135. The standard deviation for all methods was approximately 1.0×10^{-15}. Thus, at most points, traditional derivatives match the consistency condition to a little less than three decimal places. The minimum value of $\partial E/\partial T$ was -0.198 Mbar cm^3/K and of $\partial P/\partial\rho$ was -33.2 Mbar cm^3/g using birational derivatives. These are well below the minimum allowed values of zero according to (8). The other methods showed similar results. Some software interfaces have options to enforce these positivity constraints [11], but it is not done in a way which *simultaneously* guarantees satisfaction of the consistency constraints. In the rest of this paper, we elucidate a technique to do precisely this.

TABLE I. Average absolute value of the logarithmic scale of the normalized consistency error (12) for various derivative methods in common use. Bilinear, biquadratic, and birational methods are described in [11]. The "bihermitian" method is the bicubic hermitian method described in [12]. The actual software package of [11] was used for the first three methods, and the author's implementation was used for the last.

Method	Error
Bilinear	0.00665
Biquadratic	0.00240
Birational	0.00175
Bi-Hermitian	0.00107

II. NUMERICAL METHODS

The *tuned regression estimator* (TRE) method [13] allows us to estimate derivatives of tabular EOS data while simultaneously guaranteeing (6) and (8). We shall summarize briefly the basic ideas of that paper and refer the reader thereto for more background, generality, detail, and examples. Here, we shall just remark that it grew out of the application of the statistical method of local regression estimators [14] to the numerical solution of differential equations.

Let us suppose we have data points in two dimensions $\{y_i\}_{i=1}^N \subset \mathbb{R}^2$ with associated m-dimensional data values $\{u_i\}_{i=1}^N \subset \mathbb{R}^m$ which we presume to sample a continuous m-valued function of two variables $u:\mathbb{R}^2 \to \mathbb{R}^m$. Suppose we want to estimate n derivatives of u at an arbitrary point x

TABLE II. Sets of monomials used for several traditional interpolation schemes [11,12]. The biquadratic method here includes three extra terms of cubic and quartic order than that described in [11], which might properly be termed the "quadratic" method.

Method	Monomials
Bilinear	$[1,x^0,x^1,x^0x^1]$
Biquadratic	$[1,x^0,(x^0)^2,x^1,x^1x^0,x^1(x^0)^2,(x^1)^2,(x^1)^2x^0,(x^1)^2(x^0)^2]$
Bicubic	$[1,x^0,(x^0)^2,(x^0)^3,x^1,x^1x^0,x^1(x^0)^2,x^1(x^0)^3,(x^1)^2,(x^1)^2x^0,$ $(x^1)^2(x^0)^2,(x^1)^2(x^0)^3,(x^1)^3,(x^1)^3x^0,(x^1)^3(x^0)^2,(x^1)^3(x^0)^3]$

$=(x^1,x^2)\in\mathbb{R}^2$. First, we describe traditional polynomial interpolation methods, in a formalism that will prepare us for local and tuned regression estimation.

Suppose the data points make up a Cartesian grid. Let a column vector of n monomials $p:\mathbb{R}^2\to\mathbb{R}^n$ be chosen from Table II. Of course, n will be restricted to 4, 9, or 16. The derivative corresponding to a monomial $(x^i)^a(x^j)^b$ is $\partial^{a+b}/(\partial x^i)^a(\partial x^j)^b$. Let $\mathcal{Z}$ be a subset of the data points of size n consisting of the most-centered subgrid that encloses x of 2×2, 3×3, and 4×4 points for the bilinear, biquadratic, and bicubic methods, respectively. Denote the elements of $\mathcal{Z}$ by $\{z_1,\ldots,z_n\}$, and let k be a map between indices of points in $\mathcal{Z}$ and data points such that $y_{k(i)}=z_i$ for $i=1,\ldots,n$. We wish to approximate $u(x)$ by $\hat{u}(x)=\zeta p(x)$, where ζ is a $m\times n$ matrix such that the data are exactly interpolated. Let $Q=[p(z_1),\ldots,p(z_n)]$, an $n\times n$ matrix, and $v=[u_{k(1)},\ldots,u_{k(n)}]$, an $m\times n$ matrix. We require $\zeta Q=v$, and thus $\zeta=vQ^{-1}$. For any integer m, and any m-valued function $f(x)$, let $J_f(x)=[f,\partial f/\partial x_0,\partial f/\partial x_1,\ldots]$ be an $m\times n$ matrix whose columns are the n derivatives corresponding to the monomials in p, which we call the *jet matrix*. Then $J_{\hat{u}}(x)=\zeta J_p(x)$. If $u_i=\Lambda p(y_i)$, where Λ is a constant $m\times n$ matrix, then $v=\Lambda Q$ and $\zeta=\Lambda$, and polynomial interpolants are said to *reproduce* the *basis* $p(x)$. It is well known that they converge with order n for smooth data, but produce oscillations near discontinuities. The great advantage of polynomial interpolation is speed, as the Q matrix depends only on the y_i and need be computed only once for all x.

The tuned regression method is a mesh-free method and, as such, notions of nearness are determined by the value of a real-valued *weight function* $w(x,y_i)$, instead of a grid. The weight function is generally smooth, centrally peaked about y_i, and has compact support. When $w(x,y_i)$ is large, x is close to y_i. When $w(x,y_i)$ is small, x is far away from y_i. Although the grids used in the SESAME tables are nonuniformly spaced Cartesian, mesh-free techniques may be applied to them. In this paper, we use the weight function

$$w(x,y_j)=\frac{N_1}{h_j^1}\frac{N_2}{h_j^2}B_4\left(\frac{y_j^1-x^1}{h_j^1}\right)B_4\left(\frac{y_j^2-x^2}{h_j^2}\right),\qquad(13)$$

where B_4 is the cubic B-spline, defined by

$$B_4(z)=\begin{cases}1-\frac{3}{2}z^2+\frac{3}{4}|z|^3, & |z|\le1,\\ \frac{1}{4}(2-|z|)^3, & 1<|z|\le2,\end{cases}$$

and N_1 and N_2 are constants such that $\int w(x,y)dx=1$. We use a vector smoothing length $h_j=[h_j^1,h_j^2]^T$.

Now let $J_p(x)$ represent the $n\times n$ jet matrix of p. It has been verified for a large set of monomial bases that the *shifted basis* has the form

$$p(x,y_i)\equiv J_p^{-1}(x)p(y_i)=Dp(y_i-x),\qquad(14)$$

where D is a constant diagonal matrix. Suppose that $\beta(x)$ is a $m\times n$ matrix whose columns are derivative estimates of u, the same derivatives that are used in $J_p(x)$. Then $\beta(x)p(x,y_i)$ is the Taylor series expansion from x to y_i. Now suppose that we want $\beta(x)$ to satisfy a set of differential constraints at x—say, $\mathcal{D}(\beta)=0$. Through the implicit function theorem, this implies that a subset of β can be eliminated or, equivalently, we can change variables to a smaller number of variables γ such that $\beta=\mathcal{E}(\gamma)$ and $\mathcal{D}(\mathcal{E}(\gamma))=0$. Our Taylor series then takes the form $\mathcal{E}(\gamma(x))p(x,y_i)$. The inverse mapping is given by $\gamma=\mathcal{F}(\beta)$ with $\mathcal{E}(\mathcal{F}(\beta))=\beta$ when $\mathcal{D}(\beta)=0$ and $\mathcal{F}(\mathcal{E}(\gamma))=\gamma$. We measure the average error of the Taylor series expansion from x to all nearby points y_j with

$$\mathcal{R}(x)=\sum_j\|u_j-\mathcal{E}(\gamma(x))p(x,y_j)\|^2w(x,y_j).\qquad(15)$$

If we optimize $\mathcal{R}(x)$ with respect to $\gamma(x)$ by solving $\partial\mathcal{R}/\partial\gamma=0$, we will obtain optimal estimates of $u(x)$ and all of its derivatives through $\beta(x)$. The constraints will be satisfied to round-off by construction: $\mathcal{D}(\beta(x))\equiv0$. This constitutes the general method of tuned regression.

The case where $\mathcal{D}$ is empty, or $\mathcal{E}(\gamma)=\gamma=\beta$, is known as the *local regression estimator* (LRE) and has an explicit solution

$$\beta(x)=\sum_i u_i\psi_i(x)^T,\quad \psi_i(x)=P^{-1}(x)p(x,y_i)w(x,y_i),$$

$$P(x)=\sum_i p(x,y_i)p(x,y_i)^Tw(x,y_i).\qquad(16)$$

It is well studied in the statistics literature [14], and it is easy to show it has the form $\beta(x)=\zeta(x)J_p(x)$, similar to polynomial interpolants, and it has the reproducing property, just like polynomial interpolants and the moving-least-squares (MLS) estimators used in the engineering literature [15]. In fact, the zeroth-derivative estimate of the LRE is identical to that of MLS [13]. The convergence rates for $x\in\mathbb{R}$ for the νth derivative are $n-\nu+1$ for $n-\nu$ odd and $n-\nu+2$ for $n-\nu$ even [14]. The moment matrix $P(x)$ in Eqs. (16) becomes singular when the data points in the neighborhood of x become coplanar or there are fewer than n of them, so the

smoothing length must be made large enough to prevent these two situations. If it is too large, however, the procedure becomes expensive, as more neighbors are included in the sums. The proper selection of smoothing length for the LRE is a fine art discussed in [14]. It is not known exactly how much of that discussion applies to the TRE, but in practice, it is seen that at least n neighbors are also required. It is also wise to monitor condition numbers in the course of solution.

The present application makes use of the following specializations: $x=[T,\rho]^T$, $y_i=[T_i,\rho_i]^T$, $u=[E,P]^T$, and

$$p=[1,T,\rho,T^2/2,T\rho,\rho^2/2]^T,$$

$$\beta \approx \begin{bmatrix} E & \partial E/\partial T & \partial E/\partial \rho & \partial^2 E/\partial T^2 & \partial^2 E/\partial T\,\partial \rho & \partial^2 E/\partial \rho^2 \\ P & \partial P/\partial T & \partial P/\partial \rho & \partial^2 P/\partial T^2 & \partial^2 P/\partial T\,\partial \rho & \partial^2 P/\partial \rho^2 \end{bmatrix},$$

$$\mathcal{D}(\beta)=\beta_{1,0}-T\beta_{1,1}-\rho^2\beta_{0,2}=0,$$

$$\mathcal{F}(\beta)=[\beta_{0,0},\ldots,\beta_{0,5},\beta_{1,1},\ldots,\beta_{1,5}]^T=\gamma,$$

$$\mathcal{E}(\gamma)=\begin{bmatrix} \gamma_0 & \gamma_1 & \cdots & \gamma_5 \\ T\gamma_6+\rho^2\gamma_2 & \gamma_6 & \cdots & \gamma_{10} \end{bmatrix}=\beta. \tag{17}$$

We have eliminated $\beta_{1,0}$, which represents the pressure, through $\mathcal{D}=0$, to define γ. The evaluation and optimization of Eq. (15) is aided by the observation that it can be rewritten as

$$\mathcal{R}=\mathrm{Tr}\ W+\mathrm{Tr}(P\mathcal{E}^T\mathcal{E}-2U^T\mathcal{E}), \tag{18}$$

where

$$P=\sum_i p_i p_i^T w_i, \quad U=\sum_i u_i p_i^T w_i, \quad W=\sum_i u_i u_i^T w_i, \tag{19}$$

and $p_i=p(x,y_i)$ and $w_i=w(x,y_i)$.

This prescription addresses the consistency condition (6), but the stability conditions (8) require further attention. We define three possible differential constraints to use:

$$\beta_{1,0}-T\beta_{1,1}-\rho^2\beta_{0,2}=0, \tag{20a}$$

$$\beta_{0,1}=0, \tag{20b}$$

$$\beta_{1,2}=0, \tag{20c}$$

which represent the consistency and stability conditions, and adopt a multipass algorithm to enforce all constraints simultaneously:

Algorithm 1

(i) Try (20a) everywhere.

(ii) Where $\beta_{0,1}<0$ in the result of (i) apply the combination of (20a) and (20b).

(iii) Where $\beta_{1,2}<0$ in the result of (i) apply the combination of (20a) and (20c).

(iv) Where both $\beta_{0,1}<0$ or $\beta_{1,2}<0$ in (i), (ii), or (iii), apply the combination of (20a)–(20c).

For the SESAME tables examined to date the number of locations where (ii)–(iv) are required is very small. The technique results in the values of $\beta_{0,1}$ and $\beta_{1,2}$ being clamped to zero in regions where passes (i)–(iii) cause them to be negative. In the software, one needs to code four possibilities for $\mathcal{E}$ corresponding to the combinations in passes (i)–(iv). For the four different passes we solve constrained systems with $\{\beta_{1,0}\}$, $\{\beta_{1,0},\beta_{0,1}\}$, $\{\beta_{1,0},\beta_{1,2}\}$, and $\{\beta_{1,0},\beta_{0,1},\beta_{1,2}\}$ eliminated, respectively, which results in solving 11×11, 10×10, and 9×9 systems, respectively. This illustrates a key feature of tuned regression: by eliminating some derivatives, we solve a smaller system with enhanced accuracy. The penalty is that the smaller system is more complicated.

Both local regression and tuned regression in regions where only the consistency constraint is enforced produce approximations with smoothness equal to that of the weight function.

A Mathematica [16] program has been written to symbolically optimize Eq. (18) for arbitrary P, U, W, and $\mathcal{E}$. It then generates code which is spliced into a C++ library called LORELI (local regression library) which is used to operate on the SESAME data. The Appendix exhibits the expressions for $\mathcal{R}$ generated by this Mathematica program in terms of the matrices in Eqs. (19) when the various combinations of constraints in Eqs. (20) are active.

III. EXAMPLE: ANALYTIC

Let $J_u=[u,\partial_T u,\partial_\rho u,\partial_T^2 u,\partial_T\partial_\rho u,\partial_\rho^2 u]$, and suppose that $J_u=\Lambda J_p$ and $\mathcal{D}(J_u)=0$. That is, u is a quadratic function that exactly satisfies the consistency and stability conditions. Such an example is given by

$$u=\begin{bmatrix} E \\ P \end{bmatrix}=\begin{bmatrix} -1+T+\rho+T^2 \\ -T+T\rho+\rho^2 \end{bmatrix}, \tag{21}$$

with

$$\Lambda=\begin{bmatrix} -1 & 1 & 1 & 2 & 0 & 0 \\ 0 & -1 & 0 & 0 & 1 & 2 \end{bmatrix}. \tag{22}$$

Now suppose that $u_i=u(y_i)=\Lambda p(y_i)$. Then,

$$\mathcal{R}(x)=\sum_i \|\Lambda p(y_i)-\mathcal{E}(\gamma(x))J_p^{-1}(x)p(y_i)\|^2 w(x,y_i) \tag{23}$$

is minimized if $\gamma=\mathcal{F}(J_u)$, because $\mathcal{E}(\gamma(x))=J_u(x)=\Lambda J_p(x)$ and $\mathcal{R}$ is identically zero. In other words, tuned regression possesses the reproducing property just like polynomial interpolation and local regression: a polynomial solution of the differential constraints evaluated at discrete points $\{y_i\}$ will be exactly reproduced at an arbitrary $\{T,\rho\}$, to round off, and can be used by a software implementation as a verification tool. The LORELI library mentioned above has been so checked on this example and does indeed reproduce to round off. On the basis of these results we surmise that the tuned regression method for SESAME data is at least third-order accurate. There is a formal proof, but its presentation is out of scope here; however, numerical examples below will confirm it.

IV. EXAMPLE: OXYGEN

We choose SESAME table 5011 for oxygen at low temperatures as our first example because the 23×51 grid is fairly

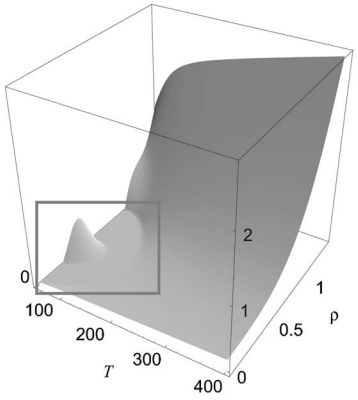

FIG. 1. TRE result for $\partial P/\partial \rho$ using SESAME EOS table 5011 for oxygen with consistency and stability constraints active. Note the flat annulus at low temperature caused by activation of the stability constraint (box). Units are K and g/cm^3.

uniform. Most SESAME tables have grids that are exponential in character to handle the many orders of magnitude variation of temperature and density required. This leads to ill-conditioned matrices in the TRE solution process and requires special treatment as described below. Table 5011, however, does not present this problem. On a 45×103 grid TRE agrees well with the input data and produces no discernable difference to the eye. Figure 1 shows the pressure derivative with respect to density, and a prominent feature is the flat annulus at low temperature. This is a region where the stability constraints were active and the algorithm did what it was supposed to. The zeroth-derivative estimates were not seen to echo this feature to the eye, which illustrates how the derivatives are estimated independently in the TRE. The estimate of the derivative is not the derivative of the estimate, as it is in finite-element or spectral methods. The two do converge, however, as the data become dense and the smoothing length goes to zero. The value of the unnormalized consistency error was everywhere less than 1.0×10^{-13}, which is close to round off, as promised. Figure 2 shows the relative error between the TRE result and the input table values when the input and output grids are identical. There is good agreement except at low temperatures where the constraints become active.

V. LOGARITHMIC FORM

As mentioned above, the exponential grids present in many SESAME tables present numerical difficulties, so a method must be devised to treat the ill-conditioned matrices $\partial^2 \mathcal{R}/\partial \gamma^2$ that appear in the Newton solver for the equations $\partial \mathcal{R}/\partial \gamma = 0$. These occur because the wide range of powers that appear in the moment matrix $P(x)$ in Eq. (18) get drastically out of balance when applied to very large numbers. For example, assume 8 decades of range in table coordinates and 45 points, which gives a ratio of about 1.5 in the size of successive intervals. Adjusting h so that there are 49 neighbor points it is easy to verify the condition number of $P \approx 10^{18}$ when $x \approx 10^4$ and $\approx 10^{25}$ when $x \approx 10^6$. One way to restore good conditioning is to use a preconditioner in the solver, which is under investigation. Another way is to logarithmically scale the variables. To scale the independent variables we use the following transformation:

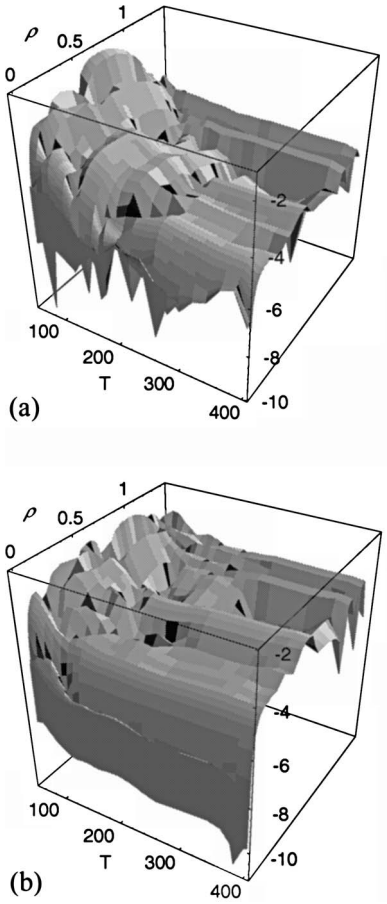

FIG. 2. Base-10 logarithm of absolute value of the error of the tuned regression estimate for oxygen table 5011 when the input grid equals the output grid. (a) Energy. (b) Pressure.

$$\tau = \ln T, \quad r = \ln \rho, \quad \varepsilon = E\rho. \tag{24}$$

In terms of these variables, the consistency and stability conditions become

$$P + \varepsilon = \frac{\partial P}{\partial \tau} + \frac{\partial \varepsilon}{\partial r}, \tag{25a}$$

$$\frac{\partial \varepsilon}{\partial \tau} \geq 0, \tag{25b}$$

$$\frac{\partial P}{\partial r} \geq 0. \tag{25c}$$

These we refer to as *semilog* constraints. It is the simplest form of the consistency condition and is linear, just like the original consistency condition, and thus requires only one 11×11 linear solve. This coordinate change was found to work for a few tables, but a further step was required to get satisfactory behavior, because even though the new grid is not exponential, the data for pressure and energy have be-

come exponential, and the problem of ill-conditioned matrices still appears. We now transform the dependent variables by means of

$$\varepsilon_s = \min \varepsilon - \varepsilon_0, \quad P_s = \min P - P_0,$$

$$\eta = \ln(\varepsilon - \varepsilon_s), \quad \zeta = \ln(P - P_s). \tag{26}$$

The new energy and pressure minima ε_0 and P_0 are arbitrary but must be positive. In the rest of this paper we set them equal to 1 so that the quantities ζ and η have a minimum value of zero and are always positive otherwise. The larger we make ε_0 and P_0, the flatter the transformed data surfaces become in the large. The consistency and stability conditions for ζ and η are

$$\exp(\zeta)\left(\frac{\partial \zeta}{\partial \tau} - 1\right) + \exp(\eta)\left(\frac{\partial \eta}{\partial r} - 1\right) = P_s + \varepsilon_s, \tag{27a}$$

$$\frac{\partial \zeta}{\partial r} \geq 0, \tag{27b}$$

$$\frac{\partial \eta}{\partial \tau} \geq 0. \tag{27c}$$

Notice now that the consistency constraint is nonlinear, whereas previously it was linear. We refer to these as *log-log* constraints. We use the LRE solution (16) as the initial condition for a Newton solver or steepest-descent solver to optimize $\mathcal{R}$. In practice, we typically see convergence in three to five Newton iterations with this initial condition. The same four-pass strategy of Algorithm 1 for enforcing the stability constraints applies.

The LORELI library thus contains 12 separate encodings of the residual function and its derivatives: for each type of constraint (flat, semilog, and log-log), there are four versions corresponding to the combination of consistency and stability constraints listed in the four passes following Eqs. (20). The Appendix exhibits these residual functions.

The expressions for the dimensionless derivatives when logarithmic transformations are employed are given by

$$\gamma = \frac{e^{\zeta}(e^{\zeta-\eta}\zeta_\tau^2 + \zeta_r \eta_\tau)}{(P_s + e^{\zeta})\eta_\tau},$$

$$\Gamma = \frac{e^{\zeta-\eta}\zeta_\tau}{\eta_\tau},$$

$$g = \frac{e^{-\eta}(P_s + e^{\zeta})}{\eta_\tau},$$

$$\begin{aligned}\mathcal{G} = e^{-\eta}[&e^{2(r+\zeta)}(2\eta_\tau\zeta_\tau - \eta_{\tau\tau})\zeta_\tau^3 + 3e^{r+\zeta+\eta}\zeta_\tau\zeta_{\tau r}\eta_\tau^2 \\ &+ e^{2\eta}(\zeta_r + \zeta_r^2 + \zeta_{rr})\eta_\tau^3 - e^{\zeta}\zeta_\tau^2\eta_\tau(-2e^{2r+\zeta}\zeta_{\tau\tau} \\ &+ (e^{2r+\zeta} - 2e^{\eta} + 2e^{r+\eta} + 2e^{r}\varepsilon_s + e^{2r}P_s - 3e^{r+\eta}\zeta_r)\eta_\tau \\ &+ e^{r+\eta}\eta_{\tau r})]/[2\eta_\tau^2(e^{\zeta}\zeta_\tau^2 + e^{\eta}\zeta_r\eta_\tau)].\end{aligned} \tag{28}$$

The reciprocal factors of η_τ and η_τ^2 become significant if the specific heat becomes small as discussed below. The expression for $\mathcal{G}$ makes use of the consistency constraint (27a). One could further incorporate its derivatives with respect to τ and r:

$$e^{-\tau}\{e^{\zeta}(\zeta_\tau - \zeta_\tau^2 - \zeta_{\tau\tau}) - e^{\eta}[(\eta_r - 1)\eta_\tau + \eta_{\tau r}]\} = 0,$$

$$e^{-r}\{-e^{\zeta}[\zeta_r(\zeta_\tau - 1) + \zeta_{\tau r}] - e^{\eta}(-\eta_r + \eta_r^2 + \eta_{rr})\} = 0,$$

but the derivation of the log-log TRE method would have to be modified to add these two equations as constraints to those of Eqs. (27). As it stands, the expression for $\mathcal{G}$ in Eqs. (28) is consistent with the log-log TRE method implied by Eqs. (27) above.

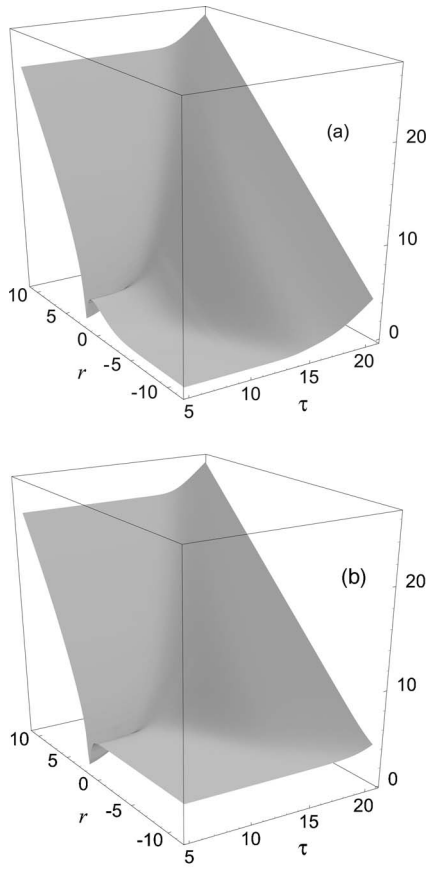

FIG. 3. Results of the log-log TRE on table 2984. 75×135 grid. (a) Logarithm of shifted energy. (b) Logarithm of shifted pressure.

VI. EXAMPLE: MOLYBDENUM

The log-log TRE method was applied to SESAME table 2984 for molybdenum and the results are shown in Fig. 3. The input grid size was 37×65, the output grid was 75×135, and once again there was no discernable difference to the eye between the two. The normalized log-log consistency error, given by

$$\begin{aligned}&(e^{\zeta} + e^{\eta} + \varepsilon_s + P_s - e^{\zeta}\zeta_\tau - e^{\eta}\eta_r)/(e^{\zeta} + e^{\eta} + |\varepsilon_s| + |P_s| \\ &\quad + e^{\zeta}|\zeta_\tau| + e^{\eta}|\eta_r|),\end{aligned} \tag{29}$$

was smaller in magnitude than 2×10^{-16} at all points as re-

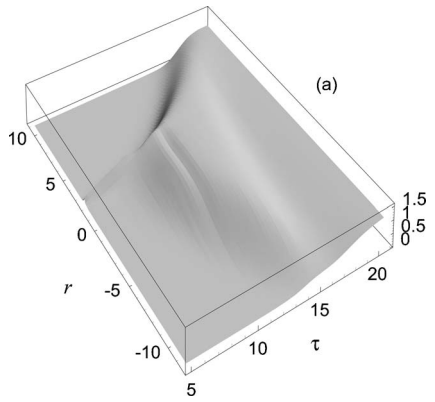

FIG. 4. $\partial\eta/\partial\tau$ from the log-log TRE on table 2984.

quired and is 14 orders of magnitude smaller than the values of Table I obtained by traditional derivatives, a substantial improvement. The condition number of the final iteration of the Newton solver at each evaluation point was everywhere less than 10^5, and the number of Newton iterations required to converge to a tolerance of 1.0×10^{-13} was at all points between one and four iterations, a very reasonable number for such a nonlinear problem. The condition numbers are rather high, but it appears possible to reduce them considerably by a simple scaling procedure which may be reported in subsequent publications.

The dimensionless quantities of Eqs. (28) were computed using the TRE derivatives. At high temperatures, the values of γ and $\mathcal{G}$ approach the theoretical values for monatomic ideal gases of 5/3 and 4/3, respectively, which gives some confidence to the calculations. On the other hand, at lower temperatures, there seem to be rather large divergent regions which correlate well with the flat regions at lower temperature in the plot of $\partial\eta/\partial\tau$ in Fig. 4 which have low values ($\leq1.0\times10^{-3}$). This is significant because the expressions for γ, Γ, and g in Eqs. (28) have $\partial\eta/\partial\tau$ in the denominator and $\mathcal{G}$ has $(\partial\eta/\partial\tau)^2$. To test that this is the cause of the divergent regions, the dimensionless derivatives are multiplied by the appropriate power of $\partial\eta/\partial\tau$ and plotted in Fig. 5. The divergent behavior has been mostly eliminated, and in the plot of γ one can see the outline of several phase boundaries. At low T and high ρ we expect to find the solid phase, and at low T and low ρ we expect to find the mixed phase. In these two phases, the theory leading to the definition of the dimensionless derivatives is incomplete because it does not include the effects of deviatoric strains or stresses and thus nonsensical results may be inescapable. Also, there are known jump conditions on γ and Γ that may come into play across phase transitions, and these have not been enforced. One of the main points of all these calculations is the determination of the sign of $\mathcal{G}$, which is negative mostly in the mixed-phase region, so the ability to provide reliable guidance to numerical methods for shock waves in these regions is clouded. In light of these observations, it seems an appropriate approach is to include the explicit phase boundaries in the EOS evaluator and make reasonably correct estimates of the dimensionless derivatives when they are crossed. It is also possible that some of the divergent behavior seen with the TRE is caused by numerical difficulties in early iterations of the Newton solver, and this should be investigated.

Figure 6 shows the γ calculated by the birational method

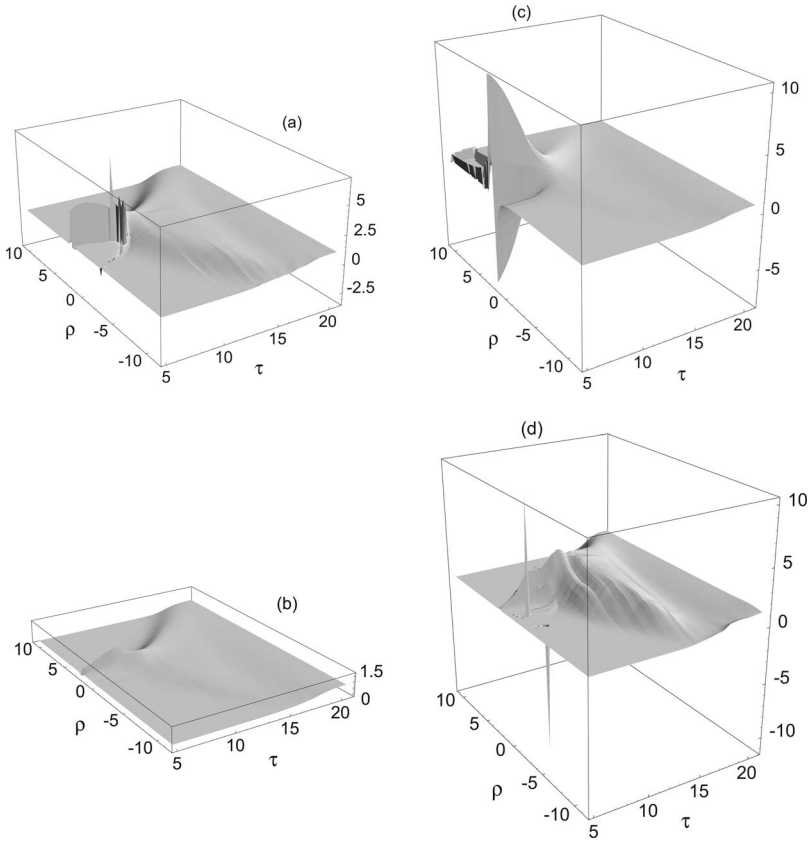

FIG. 5. Dimensionless derivatives by the log-log TRE on table 2984 multiplied by $\partial\eta/\partial\tau$ or its square. (a) Logarithmic scale of $\partial\eta/\partial\tau$ γ. (b) Logarithmic scale of $\partial\eta/\partial\tau$ Γ. (c) Logarithmic scale of $\partial\eta/\partial\tau$ g. (d) Logarithmic scale of $(\partial\eta/\partial\tau)^2$ $\mathcal{G}$.

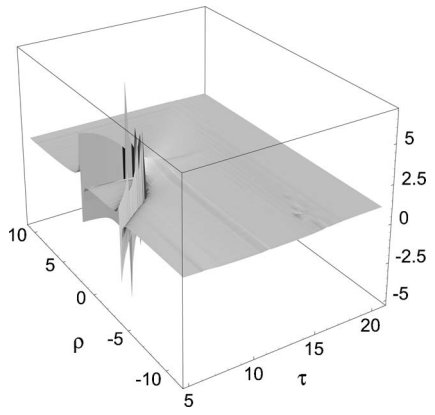

FIG. 6. Plot of γ using EOSPAC derivatives for table 2984.

of the EOSPAC library [11], which does not contain any divergent regions like the TRE result. Figure 7 shows a comparison of γ by birational and TRE at three different temperatures, roughly 2 eV, 1 keV, and 100 keV. In general, for positive $\ln \rho$, the two agree fairly well, except at the high-density boundary of the table. At the highest temperatures, both results approach the theoretical value for an ideal gas.

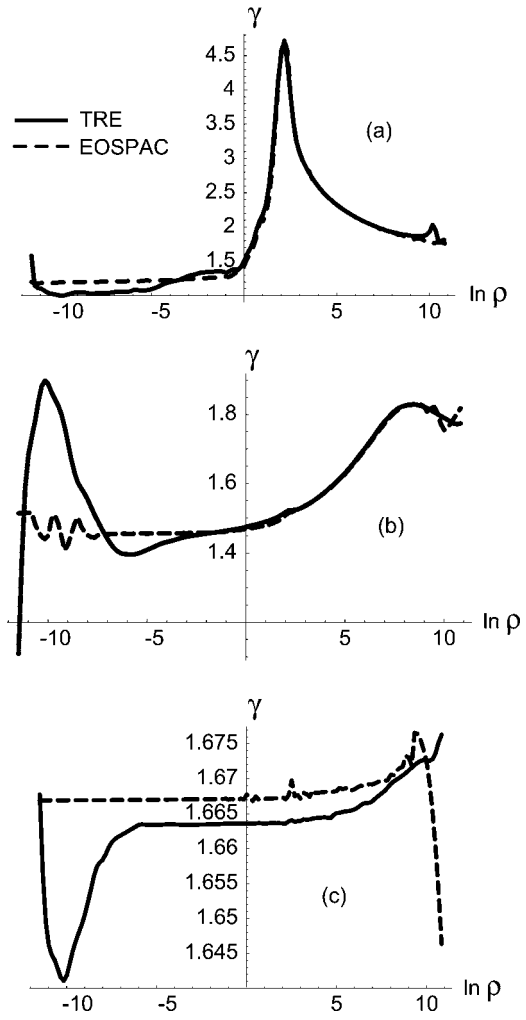

FIG. 7. Comparison of EOSPAC and TRE values of γ at $T=2$ eV (a), 1 keV (b), and 100 keV (c).

For negative $\ln \rho$, the two results are always in disagreement. Perhaps this is because the constraints are more active in that region. This too should be further investigated.

To compare the computational cost of each method, a grid of $750 \times 1350 = 1\,012\,500$ points was constructed. The EOSPAC birational evaluation on a 1.7-GHz Pentium IV system took 5.7 sec. The time for LRE was 172 sec, which involved neighbor finding using a general two-dimensional binning algorithm and solving and performing one 6×6 linear solve. The log-TRE method doing a single 11×11 linear solve took 249 sec, and the log-log-TRE method doing multiple 11×11 solves took 634 sec, which are 40 and 100 times slower than EOSPAC, respectively. This is disappointing, but not unexpected, because in addition to the linear solves involved, the algebra for the TRE is much more complicated than for EOSPAC or even the LRE. In practice, this computational cost can be avoided by evaluating all derivatives of an EOS on a fine grid and storing them for later evaluation by normal means of interpolation, such as the LRE. Presumably the consistency condition would not be violated too much. This too needs further investigation. It must be observed that a linear TRE formulation is possible which would guarantee the consistency condition and involve only a 5×5 solve. Presumably, this would be competitive with the LRE and EOSPAC from a performance viewpoint, but one would not be able to use it to compute the fundamental derivative, which requires a quadratic TRE at a minimum.

The LORELI implementation of the log-log TRE was applied to tables for copper (3333), aluminum (3719), and tin (2160) with similar results, except for some anomalous divergences in one corner which seem to be due to a poor choice of smoothing length. When the smoothing length is too small, there are too few neighbors, and the LRE or TRE methods develop ill-conditioned matrices. It becomes an issue near a table boundary because there are fewer neighbors than in the interior. If the smoothing length is too large, it may not be possible to satisfy the constraints with finite values. For the LRE there is a fairly well-developed methodology for choosing the smoothing length, but more research is needed to do the same for the TRE.

The molybdenum table 2984 does not appear to contain Maxwell constructions for the removal of van der Waals loops which lead to discontinuous derivatives. When the log-log TRE method is applied to tables that seem to have Maxwell constructions, such as gold (SESAME 2700), there are severe convergence problems in the vicinity thereof. This may be because second-order interpolation methods generally sustain oscillations in the vicinity of discontinuities and, perhaps because of the exponentials in the TRE method, these oscillations cause serious ill conditioning in the solvers, both Newton and steepest descent. Some tables seem to have apparently arbitrary abrupt transitions at the edges which also cause a similar problem. These tables may require more physical adjustments at the edges before the TRE method is robust on them. Clearly the issue of Maxwell constructions requires more research. Perhaps they can be detected, by a linear LRE estimate, for example, which always seems to be monotone (although a proof is unknown to the author), and then a separate technique applied. It is also pos-

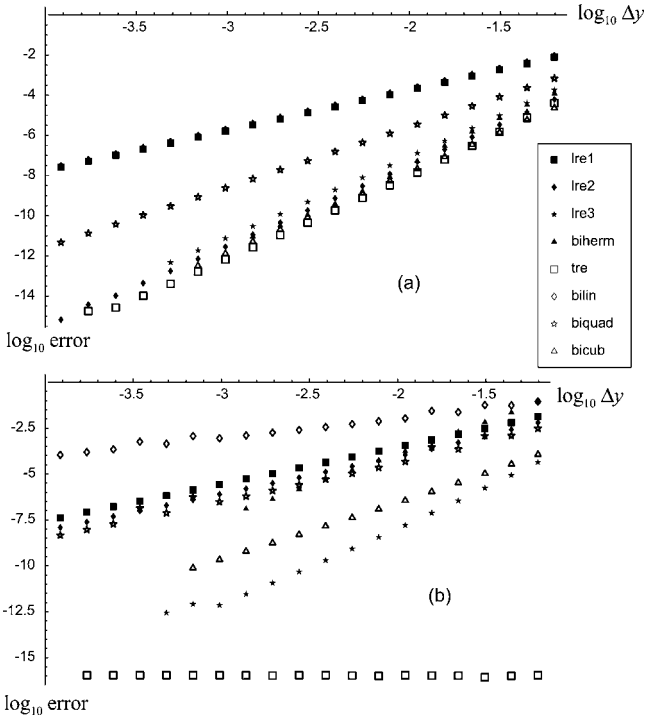

FIG. 8. Error of various methods on analytic EOS of (25) as a function of mesh spacing in the input table for ζ and consistency. (a) Error in pressure. (b) Error in consistency measured by Eq. (12).

sible that a linear TRE technique would handle such discontinuities, in the same way that the linear LRE seems to be robust with respect to discontinuities. These speculations are left for future investigations.

VII. CONVERGENCE AND ACCURACY

To test convergence of the TRE method, we use the following biquartic EOS:

$$E = (\rho^3 + \rho^2 + \rho + 1)(T^4 + T^3 + T^2 + 1) + T,$$

$$P = -\tfrac{1}{6}\rho^2(3\rho^2 + 2\rho + 1)(2T^4 + 3T^3 + 6T^2 - 6), \quad (30)$$

which satisfies the consistency and stability constraints. It is not reproducible by either quadratic or cubic polynomial, LRE, or TRE methods. This EOS was sampled on a grid of 17×17 points centered at $(0.5,0.5)$ with spacing $\Delta y = 2^{-4-k/2}$, $k=0,\ldots,18$, in each dimension to generate a table of energy and pressure that was input to the various estimation procedures operating on 5×5 grid centered at $(0.5,0.5)$ with spacing $\Delta y/4$. The error in the estimates was measured, and two sample results for pressure are plotted in Fig. 8. The curves have been truncated on the left where convergence ceased for each method. In particular, the linear methods were the most robust (working at smaller mesh spacings), followed by the quadratic and cubic methods. The bicubic hermitian method was the least robust. Table III shows the convergence rates for the various methods coded by the author. The local regression methods converge in keeping with the theoretical rates given in Sec. II above. The tuned regression zeroth derivatives converge at the same rates as the quadratic local regression estimator, which is encouraging. For zeroth derivatives the convergence is fourth order, which is remarkable since only a quadratic polynomial is used in the modeling. Table IV shows the $\ln \Delta y = 0$ intercept of the convergence curves, which gives an indication of the relative accuracy of the various methods. The various regression methods trade advantages in different derivatives with their competitors of like polynomial order. The bi-Hermitian method, as coded by the author, seems to have a markedly higher intercept than the other cubic methods, implying that a finer table is required to get the same accuracy as a bicubic or cubic LRE method could get. The tuned regression estimator does the best job on consistency, of course,as evidenced by the intercept value. The other methods all converge in consistency error, as they must if they converge at all, but unless you build consistency into the algorithm, you cannot guarantee it.

VIII. MESH-FREE ILLUSTRATION

Finally, in Fig. 9(a), we show the molybdenum table 2984 sampled at 21 583 random points uniformly distributed

TABLE III. Convergence rate for various method on biquartic EOS (30). Linear local regression cannot compute second derivatives accounting for vacancies in column 2. Tuned regression gets consistency to round off, so no convergence figure appears in its column. "Bi-Herm" denotes bicubic Hermitian.

	LRE1	LRE2	LRE3	TRE	Bilinear	Biquad	Bicubic	Bi-Herm
E	2	4	4	3.9	2	3	4	4.3
P	2	4	4	4.1	2	3	4	4.4
$\partial E/\partial T$	2	2	3.9	2	1	2	3	3.3
$\partial P/\partial T$	2	2	4	2	1	2	3	3.4
$\partial E/\partial \rho$	2	2	3.9	2	1	2	3.9	3.4
$\partial P/\partial \rho$	2	2	4	2	1	2	3	3.6
$\partial^2 E/\partial T\partial \rho$		2	2	2	1	2	2.9	2.2
$\partial^2 P/\partial T\partial \rho$		2	2	2	1	2	2.9	2.4
Consistency	2	2	4		1	2.1	3.1	3.4

TABLE IV. Extrapolated $\ln \Delta y=0$ intercept of convergence curves similar to those of Fig. 8. For methods of like order of convergence, these figures indicate relative accuracy. "Bi-Herm" denotes bicubic Hermitian.

	LRE1	LRE2	LRE3	TRE	Bilinear	Biquad	Bicubic	Bi-Herm
E	0.4	0.4	0.9	0.4	0.4	0.4	0	1.2
P	0.2	0.6	1	0.4	0.3	0.4	0.1	1.4
$\partial E/\partial T$	0.9	1	1	1	0.9	0.8	0.6	2.8
$\partial P/\partial T$	0.6	0.8	1.3	0.6	0.3	0.1	−0.2	3.3
$\partial E/\partial \rho$	0.6	0.8	0.8	0.9	0.6	0.3	0.1	3
$\partial P/\partial \rho$	0.8	0.7	1.5	0.7	0.9	0.9	0.7	3.2
$\partial^2 E/\partial T\partial \rho$	0.9	1.1	1.4	1.1	1.3	1.1	0.7	2.3
$\partial^2 P/\partial T\partial \rho$	0.7	1.3	1.6	1.3	1.4	1.2	1	2.9
Consistency	0.5	0.2	0.1	−15.5	0.1	−0.1	−0.3	3

across the full range of temperature and density using the log-log TRE of Sec. V with $\varepsilon_S=P_S=-10^3$, which flattens the graphs of η and ζ compared to Fig. 3. Each point is shaded with the value of ζ obtained from the TRE. These data are then used as an input "table" for resampling at a finer uniform random distribution of 64 749 points, again using the log-log TRE. The results are plotted in Fig. 9(b) and exhibit the expected behavior. The consistency error was zero to 15 decimal places as with the other TRE examples. The significance of the mesh-free nature of the TRE for EOS data is that traditional rectangular-grid tables may be supplemented with extra data points near Maxwell constructions or phase change boundaries to achieve greater resolution near these discontinuities with no loss of consistency or accuracy.

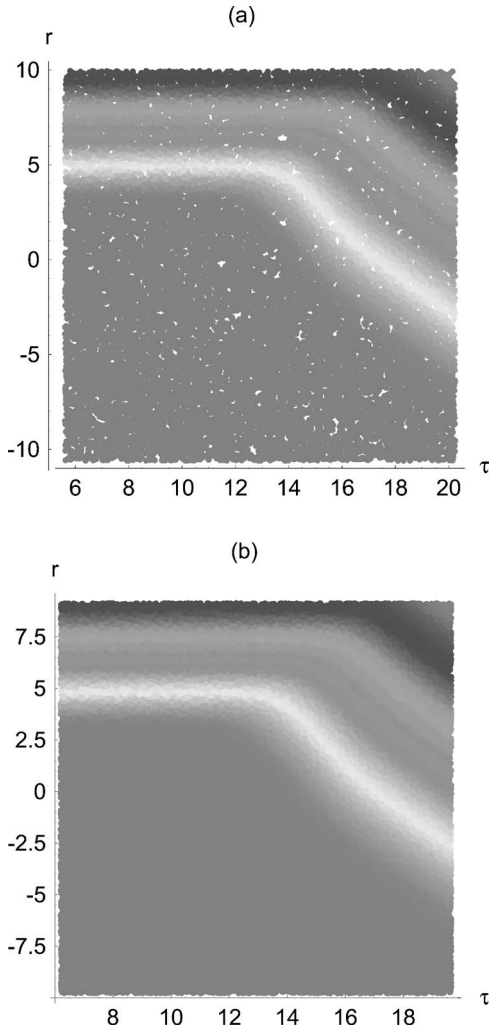

FIG. 9. (a) SESAME table 2984 sampled at 21583 uniformly distributed random points using the log-log TRE. (b) Sampling of the data from (a) at 64 749 different uniformly distributed random points using the log-log TRE. In these examples $\varepsilon_S=P_S=-10^3$. Each point is colored with the value of the interpolant.

IX. TABLE INVERSION AND ENTROPY

A common use of thermodynamic tables is in the solution of the equations of motion involving pressure as a function of energy and density. Tables such as SESAME which express pressure as functions of temperature and density must be "inverted," which is possible analytically because of the stability constraint $C_V \geq 0$. In this section we examine how such a process can proceed with tuned regression. Suppose we have an input grid of $\{T_i,\rho_i\}$ upon which are defined energies and pressures $\{E_i,P_i\}$. The set of points $\{E_i,\rho_i\}$ with data $\{T_i,P_i\}$ can be used to interpolate $T(E,\rho)$ and $P(E,\rho)$. While the input grid $\{T_i,\rho_i\}$ has a Cartesian configuration, the points $\{E_i,\rho_i\}$ have an irregular character as shown in Fig. 10. The mesh-free nature of the LRE and TRE makes them perfectly suited for interpolation on such point sets without nonlinearly interpolating the table data to a new Cartesian grid, as is typically done. A consistency condition is still required, as before. From the second law,

$$\left.\frac{\partial S}{\partial E}\right|_\rho = \frac{1}{T(E,\rho)}, \qquad \left.\frac{\partial S}{\partial \rho}\right|_E = -\frac{P(E,\rho)}{\rho^2 T(E,\rho)}. \tag{31}$$

By equating cross derivatives, this leads to the consistency condition

$$\rho^2 \left.\frac{\partial T}{\partial \rho}\right|_E + P \left.\frac{\partial T}{\partial E}\right|_\rho - T \left.\frac{\partial P}{\partial E}\right|_\rho = 0, \tag{32}$$

which may be enforced in the TRE in the same way that Eq. (6) was enforced. Transformations to other sets of variables

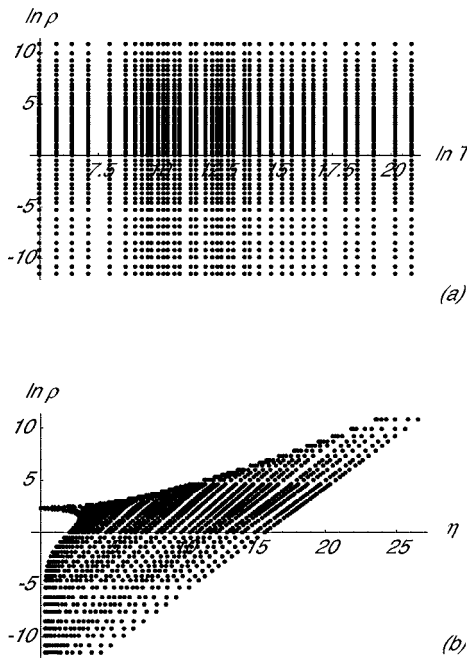

FIG. 10. (a) Temperature-density grid. (b) Energy-density grid. Logarithmic transformations applied. Data from SESAME table 2984 for molybdenum.

and thermodynamic consistency conditions can be handled similarly.

If we were to implement such a regression scheme, it is then possible to compute an entropy based on the original table data. Many tables in SESAME and other libraries provide a Helmholtz free energy from which one could compute an entropy based on Eqs. (4) using any scheme. For those situations where the free energy is not available, however, we first form a regression residual analogous to Eq. (15),

$$\mathcal{R}(S) = \sum_j [S_j - \mathcal{E}(S)p(x,y_j)]^2 w(x,y_j),$$

$$\mathcal{E}(S) = \left[S(E,\rho), \frac{1}{T(E,\rho)}, \frac{P(E,\rho)}{\rho^2 T(E,\rho)} \right],$$

$$x = [E,\rho]^T, \quad y_j = [E_j,\rho_j]^T, \quad p = [1, E_j - E, \rho_j - \rho]^T. \tag{33}$$

The temperature and pressure are computed with the TRE based on Eq. (32). A problem arises, however: we do not have any of the data S_j to put into the optimization of Eqs. (33). To resolve the issue we optimize

$$\sum_i \mathcal{R}(S_i) = \sum_{ij} [S_j - \mathcal{E}(S_i)p(x_i,y_j)]^2 w(x_i,y_j) \tag{34}$$

to obtain a global implicit set of nonlinear equations to solve for the entropies. Such a procedure is beyond the scope of this paper and must be the subject of future research. Solutions would satisfy the thermodynamic relations (31) and (32) to round off.

X. SOME CONSIDERATIONS

It might be questioned whether it is thermodynamically permissable to impose $\beta_{0,1}=0$ and $\beta_{1,2}=0$ independently or simultaneously in algorithm 1. From [10] we have $\beta_{0,1}=C_V$ and $\beta_{1,2}=1/(\rho K_T)$, where C_V is the specific heat at constant volume and K_T is the isothermal compressibility. Let C_P and K_S denote the specific heat at constant pressure and isentropic compressibility, respectively. From [10] we have the identity

$$K_S/K_T = C_V/C_P \tag{35}$$

and the alternate stability constraints

$$C_V^{-1} \geq C_P^{-1} \geq 0, \quad K_S^{-1} \geq K_T^{-1} \geq 0. \tag{36}$$

With T and V independent, we have

$$K_S = \frac{C_V}{\frac{C_V}{K_T} + V \left.\frac{\partial P}{\partial T}\right|_V \left(\left.\frac{\partial E}{\partial V}\right|_T + P \right)} \tag{37}$$

and

$$C_P = C_V + V K_T \left.\frac{\partial E}{\partial V}\right|_T \left.\frac{\partial P}{\partial T}\right|_V. \tag{38}$$

Thus, if $\beta_{0,1}=0$, then $C_V=0$ and $K_S=0$ by Eq. (37) and (35) is satisfied for $\beta_{1,2}\neq 0$. If $\beta_{1,2}=0$, then $K_T\to\infty$ and $C_P\to\infty$ by Eq. (38) and (35) is satisfied if $\beta_{0,1}\neq 0$. If both $\beta_{0,1}=0$ and $\beta_{1,2}=0$, then $K_S=0$ and $C_P\to\infty$ and Eq. (35) is satisfied. So it is thermodynamically legitimate to impose the stability constraints either independently or simultaneously.

Although the numerical method described in Sec. II may enforce the consistency and stability constraints, it is not always correct to do so. The consistency condition (6) is valid only for a fluid. Pressure and volume are defined for a solid but do not provide a complete thermodynamic representation. A solid has multiple strains that determine the thermodynamic state instead of a single volume for which other consistency constraints must be satisfied [17], and thus a different tuned regression estimator has to be used. Thus enforcement of (6) in a solid region of the EOS is unrealistic. In a rigorous implementation of a tuned regression table interpolation scheme one would first want to make sure that the consistency condition (6) is only enforced in fluid regions, which requires knowledge of the phase diagram for the material modeled.

The derivative $\partial P/\partial\rho|_T$ is zero usually only in a mixed region corresponding to a phase change. It is therefore unrealistic to enforce $\partial P/\partial\rho|_T=0$ outside of a mixed region. In a solid this derivative can legitimately be negative (as with ice), and so in a solid region it may be incorrect to force it to be zero. Again, knowledge of the phase diagram and a different treatment for solids are required to make a rigorous implementation.

The specific heat C_V is zero at a temperature of absolute zero, as can be seen from the expression $C_V=T\partial S/\partial T|_V$. For nonzero temperatures it is not realistic to enforce $C_V=0$ as in algorithm 1.

Nonetheless, it may be fitting to enforce the stability constraints (6) and (8) at points where it is physically unrealistic as described above. For example, it has been the practice for many decades by developers and users of simulation codes modeling shock-wave propagation to assume that hydrostatic stress dominates the shear components of the stress deviator tensor in a solid under the influence of high explosives or hypervelocity impact, so that even though an equation of state may contain a solid region, it is used as if it were representing only a fluid [10]. The objective is then to obtain derivatives of quantities such as pressure and energy under this equivalent fluid approximation even though the equation of state may well represent solid material. This is routinely done, for example, when computing a sound speed for use in satisfying a Courant stability condition. Therefore in practice derivatives are taken across solid-liquid as well as liquid-gas phase boundaries, where an EOS table may have explicit or apparent discontinuities of various orders. Explicit discontinuities of slope (kinks) occur at phase boundaries according the relations [10]

$$\frac{c^2-c_m^2}{c_m^2}=\frac{\gamma-\gamma_m}{\gamma_m}=(\gamma g-\Gamma^2)\left[\frac{T}{V}\left.\frac{dS}{dP}\right|_{sat}\right]\geq 0,$$

$$\frac{\xi-\gamma}{\Gamma}=\frac{\xi-\gamma_m}{\Gamma_m},\quad \xi=\frac{V}{P}\left.\frac{dP}{dV}\right|_{sat}, \tag{39}$$

where the notation $dX/dY|_{sat}$ means the slope of the saturation boundary and the subscript m denotes evaluation in the mixed region. A rigorous implementation would use another tuned regression estimator to enforce these relations at phase boundaries, which represent discontinuities of the first derivatives of the EOS. Maxwell constructions in the table designed to eliminate negative $\partial P/\partial\rho|_T$ also introduce kinks. An apparent discontinuity occurs when the continuous function that the table samples makes a transition that occurs in less space than the width of adjacent data points in the table. If the table had been created with a finer mesh, the apparent discontinuity would seem like a smooth transition. Some tables may also have erroneous explicit discontinuities at table boundaries (cliffs), caused by arbitrarily setting quantities to zero at $T=0$, for example.

Regardless of whether the table discontinuities are explicit, apparent, or erroneous, polynomial approximations to EOS functions across them can sustain unphysical oscillations whose derivatives grow exponentially with the order of derivative and the order of the polynomial and violate the stability constraints for purely numerical reasons. In such situations it is imperative to control these oscillations to restore thermodynamic fidelity without sacrificing accuracy, as when evaluating the fundamental derivative $\mathcal{G}$, for example. The tuned regression method does this in a manner reminiscent of that of monotone slope limiters used in high-resolution methods for fluid flow [4,5]. In the enforcement of the stability constraints, it limits the amplitude of the oscillations so that the result is thermodynamically acceptable: no negative bulk modulus or specific heat. If the table is at fault, the best approach is to correct it or add more resolution, but few table users have the knowledge or ability to generate a new table, so as a practical matter it is imperative to seek numerical methods for table interpolation that get as much information as possible from the available table data without violating thermodynamic constraints. The tuned regression method fills this role by finding the optimal local polynomial fit to nearby data that satisfies the constraints.

The tuned regression method results in derivatives that match the thermodynamic constraints, but it will be less accurate than a method that has no such constraints if applied to an EOS which does not satisfy the constraints because it is trying to force specific behavior of derivatives on a function for which it is inappropriate. A tuned regression estimator is only good for data that are sampled from a function whose behavior at least roughly matches the differential constraints it was designed to satisfy.

Why compute and adjust separate estimates of derivatives instead of "simply" taking derivatives of an estimate? For data sets sampling a smooth function this is natural and effective. For data sets sampling functions discontinuous in either the zeroth or first derivatives the fact that polynomial interpolants frequently produce large oscillations near discontinuities forces us to reconsider.

The problem of EOS table interpolation is to construct functions and derivatives from a finite number of data points, for which there is no unique solution. Consider an approximate function $\tilde{P}(T,\rho)$ interpolating pressure from table data on a regular grid. One can add the function

$$\delta(T,\rho)=\varepsilon\sin\left(\frac{2\pi k}{\Delta T}T\right)\sin\left(\frac{2\pi k}{\Delta\rho}\rho\right)$$

to P for k an arbitrary integer and ε arbitrary but small enough and still have an approximation to the pressure every bit as accurate as $\tilde{P}$ and which matches the data exactly on the grid points, but the (n_1,n_2)th derivative of $\tilde{P}+\delta$ will have oscillations of size $\varepsilon(2\pi k/\Delta T)^{n_1}(2\pi k/\Delta\rho)^{n_2}$ which can be arbitrarily large. Polynomial interpolants develop oscillations near discontinuities that also grow exponentially with the order of the polynomial and the order of derivative. Clearly, it is not always best to estimate derivatives by explicit differentiation, for the behavior of the derivatives is not easily controlled.

If one were to try to define a continuous function to estimate pressure and energy from the table data and apply the consistency and stability constraints to its derivatives, it would likely be much more expensive to solve than the tuned regression technique outlined above, involving a continuous solution to a possibly nonlinear system of partial differential equations (the consistency condition plus a regularization condition, which is needed because solutions of the consistency condition are not unique) and boundary conditions (the stability and boundedness conditions) at each evaluation and table point. The tuned regression technique avoids this expense by approximating the derivative at any point as the coefficient of the linear or quadratic term in a polynomial that optimally fits nearby data, much the same as the notion of a Frechet derivative in real analysis, which is identical to the familiar Leibnitz difference quotient definition for real-valued functions. The set of continuous coefficients defines

continuous approximations to a finite number of derivatives. We are trading a finite number of conditions on an infinite number of derivatives of a single function for a finite number of conditions on a finite number of functions. The former creates a difficult nonlocal problem, while the latter creates a simple local problem.

One could obtain the consistency condition by interpolating a tabular free energy where provided and taking explicit derivatives. All the consistency conditions would be automatically satisfied, but the stability constraints would now involve second derivatives and the computation of $\mathcal{G}$ would require third derivatives of the free energy. These would oscillate more strongly near discontinuities than first and second derivatives of tabular values of pressure and energy, which nominally are derived from explicit derivatives of the analytic functions used in modeling the EOS. Ensuring the stability conditions on a continuous interpolant would become more difficult. The errors would also be larger because of the generic loss of accuracy with derivative order in polynomial-based interpolants, including tuned and local regression. Furthermore, getting good third derivatives is more expensive and less robust in general because the condition number of the matrix inversions involved generally goes up rapidly with polynomial interpolant order.

XI. CONCLUSION

We have shown that traditional numerical derivatives of equation-of-state tables do not simultaneously satisfy the thermodynamic consistency and stability conditions and that a tractable method to compute them can be developed from the tuned regression estimator. The LORELI implementation has demonstrated the reproducing property for quadratic analytic EOS which satisfy the consistency and stability conditions. Trials on a few SESAME tables have shown that the consistency and stability constraints can be simultaneously enforced to round off without sacrificing accuracy and that theoretical values for γ and $\mathcal{G}$ are approached at high temperature. Versions of the theory and software were developed for flat, semilogarithmic, and log-log coordinates, the last being necessary to handle tables with exponential grids. The convergence rates follow those of the statistical local regression estimator, giving fourth order for zeroth derivatives and second order for first and second derivatives. The TRE method is apparently much more expensive than traditional derivatives, however. The construction of tables of derivatives could mitigate this problem. The mesh-free character of the TRE method was convincingly demonstrated and provides a basis for augmenting conventional tables with extra data points near physical discontinuities. There are outstanding issues regarding phase boundaries, Maxwell constructions, and table-edge drop-offs which require further research before the technique can be made into a fully robust tool.

APPENDIX: RESIDUAL EXPRESSIONS

The residual for the consistency constraint given in (20a) is

$$\begin{aligned}
&\frac{W_{11}}{2}+\frac{W_{22}}{2}+\gamma_0\left(-U_{11}+\frac{P_{11}\gamma_0}{2}+P_{12}\gamma_1+P_{13}\gamma_2+P_{14}\gamma_3+P_{15}\gamma_4+P_{16}\gamma_5\right)+\gamma_1\left(-U_{12}+\frac{P_{22}\gamma_1}{2}+P_{23}\gamma_2+P_{24}\gamma_3\right.\\
&\left.+P_{25}\gamma_4+P_{26}\gamma_5\right)+\gamma_3\left(-U_{14}+\frac{P_{44}\gamma_3}{2}+P_{45}\gamma_4+P_{46}\gamma_5\right)+\gamma_4\left(-U_{15}+\frac{P_{55}\gamma_4}{2}+P_{56}\gamma_5\right)+\gamma_5\left(\frac{P_{66}\gamma_5}{2}-U_{16}\right)\\
&+\gamma_2\left[-U_{21}\rho^2+P_{13}\gamma_7\rho^2+P_{14}\gamma_8\rho^2+P_{15}\gamma_9\rho^2+P_{16}\gamma_{10}\rho^2-U_{13}+\left(\frac{P_{11}\rho^4}{2}+\frac{P_{33}}{2}\right)\gamma_2+P_{34}\gamma_3\right.\quad+\quad P_{35}\gamma_4+P_{36}\gamma_5\\
&\left.+(TP_{11}\rho^2+P_{12}\rho^2)\gamma_6\right]+\gamma_6\left[-TU_{21}-U_{22}+\left(\frac{P_{11}T^2}{2}+P_{12}T+\frac{P_{22}}{2}\right)\gamma_6\quad+\quad(TP_{13}+P_{23})\gamma_7+(TP_{14}+P_{24})\gamma_8+(TP_{15}+P_{25})\gamma_9\right.\\
&\left.+(TP_{16}+P_{26})\gamma_{10}\right]+\gamma_7\left(-U_{23}+\frac{P_{33}\gamma_7}{2}+P_{34}\gamma_8+P_{35}\gamma_9+P_{36}\gamma_{10}\right)+\gamma_8\left(-U_{24}+\frac{P_{44}\gamma_8}{2}+P_{45}\gamma_9+P_{46}\gamma_{10}\right)\\
&+\gamma_9\left(-U_{25}+\frac{P_{55}\gamma_9}{2}+P_{36}\gamma_{10}\right)+\gamma_{10}\left(\frac{P_{66}\gamma_{10}}{2}-U_{26}\right),
\end{aligned}\tag{A1}$$

with

$$\mathcal{E}=\begin{pmatrix}\gamma_0 & \gamma_1 & \gamma_2 & \gamma_3 & \gamma_4 & \gamma_5\\ \gamma_2\rho^2+T\gamma_6 & \gamma_6 & \gamma_7 & \gamma_8 & \gamma_9 & \gamma_{10}\end{pmatrix}.\tag{A2}$$

The residual for the constraints (20a) and (20b) active is

$$\frac{W_{11}}{2}+\frac{W_{22}}{2}+\gamma_0\left(-U_{11}+\frac{P_{11}\gamma_0}{2}+P_{13}\gamma_1+P_{14}\gamma_2+P_{15}\gamma_3+P_{16}\gamma_4\right)+\gamma_2\left(-U_{14}+\frac{P_{44}\gamma_2}{2}+P_{45}\gamma_3+P_{46}\gamma_4\right)$$
$$+\gamma_3\left(-U_{15}+\frac{P_{55}\gamma_3}{2}+P_{56}\gamma_4\right)+\gamma_4\left(\frac{P_{66}\gamma_4}{2}-U_{16}\right)+\gamma_1\Big[-U_{21}\rho^2+P_{13}\gamma_6\rho^2+P_{14}\gamma_7\rho^2+P_{15}\gamma_8\rho^2+P_{16}\gamma_9\rho^2-U_{13}$$
$$+\left(\frac{P_{11}\rho^4}{2}+\frac{P_{33}}{2}\right)\gamma_1+P_{34}\gamma_2\quad+P_{35}\gamma_3+P_{36}\gamma_4+(TP_{11}\rho^2+P_{12}\rho^2)\gamma_5+\gamma_5\Big[-TU_{21}-U_{22}+\left(\frac{P_{11}T^2}{2}+P_{12}T+\frac{P_{22}}{2}\right)\gamma_5$$
$$+(TP_{13}+P_{23})\gamma_6+(TP_{14}+P_{24})\gamma_7+(TP_{15}+P_{25})\gamma_8+(TP_{16}+P_{26})\gamma_9\Big]+\gamma_6\left(-U_{23}+\frac{P_{33}\gamma_6}{2}+P_{34}\gamma_7+P_{35}\gamma_8+P_{36}\gamma_9\right)$$
$$+\gamma_7\left(-U_{24}+\frac{P_{44}\gamma_7}{2}+P_{45}\gamma_8+P_{46}\gamma_9\right)+\gamma_8\left(-U_{25}+\frac{P_{55}\gamma_8}{2}+P_{56}\gamma_9\right)+\gamma_9\left(\frac{P_{66}\gamma_9}{2}-U_{26}\right),\tag{A3}$$

with

$$\mathcal{E}=\begin{pmatrix}\gamma_0 & 0 & \gamma_1 & \gamma_2 & \gamma_3 & \gamma_4\\ \gamma_1\rho^2+T\gamma_5 & \gamma_5 & \gamma_6 & \gamma_7 & \gamma_8 & \gamma_9\end{pmatrix}.\tag{A4}$$

The residual for the constraints (20a) and (20c) active is

$$\frac{W_{11}}{2}+\frac{W_{22}}{2}+\gamma_0\left(-U_{11}+\frac{P_{11}\gamma_0}{2}+P_{12}\gamma_1+P_{13}\gamma_2+P_{14}\gamma_3+P_{15}\gamma_4+P_{16}\gamma_5\right)+\gamma_1\left(-U_{12}+\frac{P_{22}\gamma_1}{2}+P_{23}\gamma_2+P_{24}\gamma_3\right.$$
$$\left.+P_{25}\gamma_4+P_{26}\gamma_5\right)+\gamma_3\left(-U_{14}+\frac{P_{44}\gamma_3}{2}+P_{45}\gamma_4+P_{46}\gamma_5\right)+\gamma_4\left(-U_{15}+\frac{P_{55}\gamma_4}{2}+P_{56}\gamma_5\right)+\gamma_5\left(\frac{P_{66}\gamma_5}{2}-U_{16}\right)$$
$$+\gamma_2\left[-U_{21}\rho^2+P_{14}\gamma_7\rho^2+P_{15}\gamma_8\rho^2+P_{16}\gamma_9\rho^2-U_{13}+\left(\frac{P_{11}\rho^4}{2}+\frac{P_{33}}{2}\right)\gamma_2+P_{34}\gamma_3+P_{35}\gamma_4+P_{36}\gamma_5+(TP_{11}\rho^2+P_{12}\rho^2)\gamma_6\right]$$
$$+\gamma_6\left[-TU_{21}-U_{22}+\left(\frac{P_{11}T^2}{2}+P_{12}T+\frac{P_{22}}{2}\right)\gamma_6+(TP_{14}+P_{24})\gamma_7+(TP_{15}+P_{25})\gamma_8+(TP_{16}+P_{26})\gamma_9\right]$$
$$+\gamma_7\left(-U_{24}+\frac{P_{44}\gamma_7}{2}+P_{45}\gamma_8+P_{46}\gamma_9\right)+\gamma_8\left(-U_{25}+\frac{P_{55}\gamma_8}{2}+P_{56}\gamma_9\right)+\gamma_9\left(\frac{P_{66}\gamma_9}{2}-U_{26}\right),\tag{A5}$$

with

$$\mathcal{E}=\begin{pmatrix}\gamma_0 & \gamma_1 & \gamma_2 & \gamma_3 & \gamma_4 & \gamma_5\\ \gamma_2\rho^2+T\gamma_6 & \gamma_6 & 0 & \gamma_7 & \gamma_8 & \gamma_9\end{pmatrix}.\tag{A6}$$

The residual for the constraints (20a)–(20c) active is

$$\frac{W_{11}}{2}+\frac{W_{22}}{2}+\gamma_0\left(-U_{11}+\frac{P_{11}\gamma_0}{2}+P_{13}\gamma_1+P_{14}\gamma_2+P_{15}\gamma_3+P_{16}\gamma_4\right)+\gamma_2\left(-U_{14}+\frac{P_{44}\gamma_2}{2}+P_{45}\gamma_3+P_{46}\gamma_4\right)+\gamma_3\left(-U_{15}+\frac{P_{55}\gamma_3}{2}\right.$$
$$\left.+P_{56}\gamma_4\right)+\gamma_4\left(\frac{P_{66}\gamma_4}{2}-U_{16}\right)+\gamma_1\Big[-U_{21}\rho^2+P_{14}\gamma_6\rho^2+P_{15}\gamma_7\rho^2+P_{16}\gamma_8\rho^2-U_{13}+\left(\frac{P_{11}\rho^4}{2}+\frac{P_{33}}{2}\right)\gamma_1\quad+\quad P_{34}\gamma_2+P_{35}\gamma_3$$
$$+P_{36}\gamma_4+(TP_{11}\rho^2+P_{12}\rho^2)\gamma_5\Big]+\gamma_5\Big[-TU_{21}-U_{22}+\left(\frac{P_{11}T^2}{2}+P_{12}T+\frac{P_{22}}{2}\right)\gamma_5+(TP_{14}+P_{24})\gamma_6+(TP_{15}+P_{25})\gamma_7+(TP_{16}$$
$$+P_{26})\gamma_8\Big]+\gamma_6\left(-U_{24}+\frac{P_{44}\gamma_6}{2}+P_{45}\gamma_7+P_{46}\gamma_8\right)+\gamma_7\left(-U_{25}+\frac{P_{55}\gamma_7}{2}+P_{56}\gamma_8\right)+\gamma_8\left(\frac{P_{66}\gamma_8}{2}-U_{26}\right),\tag{A7}$$

with

$$\mathcal{E}=\begin{pmatrix}\gamma_0 & 0 & \gamma_1 & \gamma_2 & \gamma_3 & \gamma_4\\ \gamma_1\rho^2+T\gamma_5 & \gamma_5 & 0 & \gamma_6 & \gamma_7 & \gamma_8\end{pmatrix}.\tag{A8}$$

[1] J. Von Neumann and R. D. Richtmeyer, J. Appl. Phys. **21**, 232 (1950).
[2] S. K. Godunov, Mat. Sb. **47**, 271 (1959).
[3] C. Hirsch, *Numerical Computation of Internal and External Flows* (John Wiley & Sons, New York, 1990), Vol. 2.
[4] E. F. Toro, *Riemann Solvers and Numerical Methods for Fluid Dynamics* (Springer-Verlag, New York, 1997).
[5] R. J. Leveque, *Finite Volume Methods for Hyperbolic Problems* (Cambridge University Press, Cambridge, England, 2002).
[6] G. H. Miller and E. G. Puckett, J. Comput. Phys. **128**, 134 (1996).
[7] K. M. Shyue, J. Comput. Phys. **171**, 678 (2001).
[8] K. S. Holian (unpublished).
[9] S. P. Lyon and J. D. Johnson (unpublished).
[10] R. Menikoff and B. J. Plohr, Rev. Mod. Phys. **61**, 75 (1989).
[11] D. A. Pimentel (unpublished).
[12] F. N. Fritsch (unpublished).
[13] G. A. Dilts, A. Haque, and J. Wallin, in *Meshfree Methods for Partial Differential Equations*, edited by M. Griebel and M. A. Schweitzer, Lecture Notes in Computational Science and Engineering, Vol. 26 (Springer-Verlag, Berlin, 2002).
[14] J. Fan and I. Gijbels, *Local Polynomial Modeling and Its Applications* (Chapman and Hall, London, 1996).
[15] T. Belytschko, Y. Y. Lu, and L. Gu, Int. J. Numer. Methods Eng. **37**, 229 (1994).
[16] Wolfram Research, Inc., Mathematica, version 5.0, Champaign, IL, 2003.
[17] D. C. Wallace, *Thermodynamics of Crystals* (Dover, Mineola, NY, 1998).