LA-UR-17-20806
Approved for public release; distribution is unlimited.

Title: Introduction to Shock Waves and Shock Wave Research

Author(s): Anderson, William Wyatt

Intended for: Internal Short Course

Issued: 2017-02-02

INTRODUCTION TO SHOCK WAVES AND SHOCK WAVE RESEARCH

William W. Anderson

I. INTRODUCTION

M-9 and a number of other organizations at LANL and elsewhere study materials in dynamic processes. Often, this is described as "shock wave research," but in reality is broader than is implied by that term. Most of our work is focused on dynamic compression and associated phenomena, but you will find a wide variety of things we do that, while related, are not simple compression of materials, but involve a much richer variety of phenomena. This tutorial will introduce some of the underlying physics involved in this work, some of the more common types of phenomena we study, and common techniques. However, the list will not be exhaustive by any means.

Fundamental physical constants.

Most of the material discussed in these notes does not require use of the fundamental physical constants. However, some discussions do have a few of these constants in them and, although the values of the constants are not generally important to the discussions in the notes, they are important for the actual application of the material. Hence, Table I-1 gives the current accepted values of those constants that appear in these notes. A full listing of the fundamental physical constants and their accepted values can be found at http://physics.nist.gov/cuu/Constants/index.html.

Notation.

These notes use notation that is fairly standard in shock wave research in the US. However, some variations exist, even in the US, depending on the field from which a research is coming. There can also be differences in notation that depend on the nationality of the researcher. Such differences in notation can a problem in many areas of research that cut across disciplines. The notation in this handout comes from the geosciences, which is traditionally the biggest non-defense user of shock wave techniques. There will likely be instances where the notation used here differs from what you may have seen in your own fields of study. The important point to remember is that you should be sure you understand the meaning of symbols used any time you are reading published literature or lecture notes such as these. Some of the more common confusing notations are given in Table I-2. The specific notation used in these notes is given in Table I-3.

Table I-1 Fundamental Physical Constants

Quantity	Symbol	Value	Units
Speed of Light in Vacuum	c	299492458	$m\ s^{-1}$
Planck Constant	h	6.62606957×10^{-34}	$J\ s$
Boltzmann Constant	k	1.30806488×10^{-23}	$J\ K^{-1}$
Molar Gas Constant	R	8.3144621	$J\ mol^{-1}\ K^{-1}$

Table I-2-Confusing Notation.		
Quantity	**Symbols used here**	**Other common symbols**
Pressure	P	p
Internal energy	E	U, ε
Bulk modulus	K	B, κ
Compressibility	κ	K, k, β
Shock wave speed	U_s	D, U
Sound or elastic wave speed	c	C
Specific heat	C	c
Young's modulus	Y	q
Shear modulus	G	μ
Helmholtz free energy	F	A
Gibbs free energy	G	g
Thermodynamic Grüneisen parameter	γ	Γ, Γ_{th}, γ_{th}
Lattice Grüneisen parameter	γ_L	γ, Γ
Volume thermal expansion coefficient	α	β, α_V
Linear thermal expansion coefficient	α_L	α
Molar mass	μ	M

Organization.

These notes are organized into three major subject areas—Materials and Material Behavior, Waves and Shock Waves, and Shock Wave Research Techniques. Depending on your background, you will likely have already seen some of this material before, but much of it will be new.

Table I-3-Symbols and Units.		
Quantity	**Symbol**	**Commonly Used Units**
Density	ρ	kg/m^3, g/cm^3
Pressure	P	Pa, GPa, kbar
Stress	σ	Pa, GPa, kbar
Isentropic and isothermal bulk modulus	K_S, K_T	Pa, GPa, kbar
Pressure derivative of K_S and K_T	K_S', K_T'	
Isobaric and isochoric specific heat	C_P, C_V	J/kgK
Volume thermal expansion coefficient	α	K^{-1}
Linear thermal expansion coefficient	α_L	K^{-1}
Thermodynamic Grüneisen parameter	γ	
Specific internal energy	E	J/kg
Specific enthalpy	H	J/kg
Specific Gibbs free energy	G	J/kg
Specific Helmholtz free energy	F	J/kg
Shock wave speed	U_s, U	km/s, m/s

Particle speed	u_p, u	km/s, m/s
Eulerian longitudinal sound speed	c_l	km/s, m/s
Lagrangian longitudinal sound speed	${}^{L}c_l$	km/s, m/s
Eulerian bulk sound speed	c_b	km/s, m/s
Lagrangian bulk sound speed	${}^{L}c_b$	km/s, m/s
Eulerian shear sound speed	c_s	km/s, m/s
Hugoniot intercept	C_0	km/s, m/s
Hugoniot slope	s	
Hugoniot quadratic coefficient	q	s/km, s/m
Strain	ε	
Specific entropy	S	J/kgK
Temperature	T	K
Molar mass	μ	kg/mol
Specific volume	V	m^3/kg
Mass flux	j	kg/m^2s
Atomic or molar fraction	x	
Mass fraction	m	
Mass	m	kg
Heat	Q	J
Work	W	J
First and second Lamé constants	λ, μ	Pa, GPa, kbar
Frequency	ν	Hz, s^{-1}
Wavelength	λ	m
Spectral emissivity	ε	
Planck Function	I, I'	J $m^{-2}sr^{-1}$, W $m^{-3}sr^{-1}$
Electrical potential	V	V
Current	$\mathbf{I}$	A
Magnetic Field	$\mathbf{B}$	Gauss
Acoustic impedance	z	kg $m^{-2}s^{-1}$
Shear modulus	G	Pa, GPa, kbar
Number	N	
Positions	x, y, z	m
Displacements	u, v, w	m
Length	L	m
Elastic constant	c	Pa, GPa, kbar
Young's modulus	Y	Pa, GPa, kbar
Poisson's ration	ν	
Force	F, $\mathbf{F}$	N
Acceleration	$\mathbf{a}$	m/s^2
Time	t	s
Parameters for Birch-Murnaghan Eqn.	ξ_1, ξ_2	
Interatomic Potential	φ	J
Compression	η	
Rotation (in strain tensor)	ϖ	radians

Function (in wave description)	F	
Variables (in wave description)	$\xi,\ \eta$	
Displacement	χ	m
Velocity (in discussions of electromagnetics)	$\mathbf{V}$	m/s
Isothermal and isentropic Anderson-Grüneisen parameter	$\delta_T,\ \delta_S$	
$(\partial \ln\gamma/\partial \ln V)_T$	q	
Polytropic exponent	γ	

II. MATERIALS AND MATERIAL BEHAVIOR

Everything we deal with on a daily basis—the air we breathe, the chair we sit on, our own body, is made up of matter, with a variety of forms nebulously called materials. Everything about these materials and how they behave are controlled by their properties, which, for our purposes, can be lumped into the headings of thermodynamic, mechanical, and transport properties. Most of shock wave research is involved with probing material properties. Shock wave studies are also affected by those properties.

Thermodynamics.

Thermodynamics deals with the relationship between work and heat and the consequences for variables such as temperature, pressure, internal energy, and volume, known as thermodynamic potentials. These variables may be either intensive, meaning the value doesn't depend on how much material is present, i.e., can't be expressed in terms of "so much per unit of mass", or extensive, which does depend on the amount of material present. The fundamentally intensive quantities are pressure (P) and temperature (T). Volume, energy, and other variables are extensive, but are usually recast into intensive properties by dividing them by the mass of material. Intensive quantities that are derived in this way are known as *specific quantities*. Thus, the intensive form of volume is the *specific volume*. In practice, with the notable exception of the volume, the word "specific" is almost never used because it is very rare to actually use extensive properties in calculations.

It is important to remember that thermodynamics is concerned with bulk behavior of materials. When dealing with individual molecules or atoms, the concepts of pressure, temperature, and so on break down and are not strictly applicable. Thermodynamics in physics is often called statistical mechanics or statistical thermodynamics, from the concept that a statistical average of the behavior of a large enough ensemble of atoms is what shows up in the macroscopic world as thermodynamics.

Now, consider a block of your favorite substance (Figure II-1). That block has some volume (an extensive property), related to the amount of the substance and the density (an intensive property). What, other than adding more of the substance, might we do to change the volume? The most obvious is that we can apply pressure—i.e., squeeze the block. Pressure is the natural driving force for changing the volume. However, applying pressure and compressing the block also affects other variables. The action of pressure to compress the volume is a form of work—the force due to pressure acts on the boundaries of the block to move them, and work is force applied through a distance. The units of work are energy—you are adding energy to the block when you compress it.

One can also transfer energy into (or out of) the block. In classical thermodynamics discussions, the transfer is always discussed as if it were by transfer of heat through conduction, but this can also encompass deposition of electromagnetic radiation, or change in energy through flow of electrical current with attendant ohmic dissipation. The net result is the same—the deposited energy (eventually) ends up as thermal energy—manifested as the microscopic motion of the atoms.

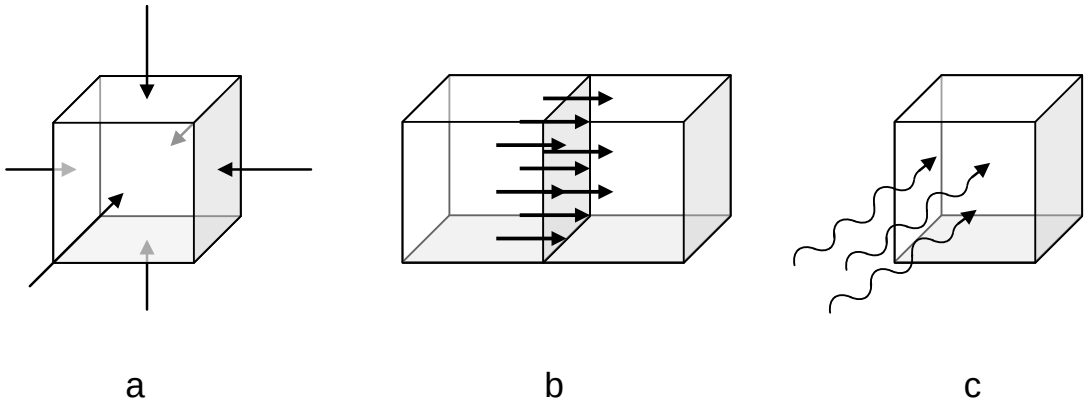

Figure II-1. Energy may be added to an object by (a) performing work in compression, (b) allowing diffusion of heat into the object, and (c) deposition of thermal radiation.

The First and Second Laws of Thermodynamics.

We don't have time to go into a lengthy discussion of the laws of thermodynamics, but a brief introduction is in order. From the foregoing, we see that the energy of the block (the "system") can be changed by performing work and/or by transferring heat. The energy in this case is usually called the *internal energy*, *E*. The *First Law of Thermodynamics* states that the change in the internal energy of a system is equal to the work done on the system plus the heat added to the system, or

$$dE = dW + dQ\,, \tag{II.1}$$

where *W* is work and *Q* is heat.

Now, consider a heat engine, which takes heat from one reservoir to another reservoir and uses the heat to do work. The basic concept of the *Second Law of Thermodynamics* is that a heat engine can do no work if the two reservoirs are the same temperature and that heat will not spontaneously flow from one reservoir to a second that is at the same or higher temperature. Effectively, the second law states that in a *reversible* process, the work done is exactly equal to the difference between the heat removed from the hotter reservoir and the heat added to the cooler reservoir.

Now, let's delve a little deeper into the second law. In the case of a reversible heat engine, the French engineer Sadi Carnot came to the conclusion that the efficiency of such an engine was independent of the design of the engine. Hence, conclusions drawn from the behavior of one example of an "ideal" engine apply to all ideal engines. Now, let's consider a perfect gas. Such a gas obeys the relation

$$P = \frac{NkT}{V}\,, \tag{II.2}$$

where *N* is the number of molecules, *k* is a constant of proportionality (Boltzmann's constant), and *V* here is the extensive volume. The fundamental property of a perfect gas is that its molecules are infinitesimal and therefore never interact with one another. This

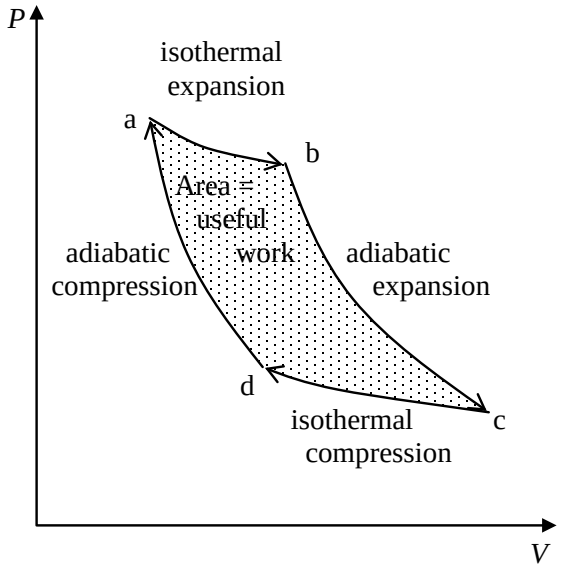

Figure II-2. The Carnot cycle.

has the interesting consequence that the internal energy explicitly depends on the temperature and nothing else. To see what this means for an ideal heat engine, consider the cycle of expansion and compression shown in figure II-2. Starting at point a, the gas expands, doing work on its surroundings

$$W = \int_{V_a}^{V_b} P dV \,, \tag{II.3}$$

but the expansion takes place at a constant temperature T_1, so that amount of heat Q_1 removed from a reservoir is $Q_1 = W$ and, from the above expressions,

$$Q_1 = \int_a^b NkT_1 \frac{dV}{V} = NkT_1 \ln \frac{V_b}{V_a} \,. \tag{II.4}$$

Now, if we cause the gas to expand further to V_c, but without any addition or removal of heat, the ideal gas must come to a different temperature given by

$$T_1 V_b^{\gamma-1} = T_2 V_c^{\gamma-1}, \tag{II.5}$$

where the quantity γ as used here is called the polytropic exponent. Some work has been done, but with no heat transfer, so the energy and temperature have now changed. Now, we isothermally compress the gas, with the heat given up being

$$Q_2 = NkT_2 \ln \frac{V_d}{V_c}. \tag{II.6}$$

Finally, let's compress adiabatically (no heat transfer) back to V_a:

$$T_1 V_a^{\gamma-1} = T_2 V_d^{\gamma-1}. \tag{II.7}$$

This complete cycle is known as the Carnot cycle, and shows that the heat changes and temperatures involved in completing the cycle are related by

$$\frac{Q_1}{T_1} = \frac{Q_2}{T_2}. \tag{II.8}$$

Note that, in any case where the system is not ideal—i.e., not reversible, the foregoing expression will be an inequality and the amount of work done will be less than the net heat gained by the system.

A vital concept that was alluded to above, but we now state, is that the thermodynamic state of a material, i.e., the particular pressure-temperature condition (along with the phase), is ideally independent of the path taken to arrive at that state. Hence, when doing a calculation of a change from one set of conditions to another, you can use whatever paths make the problem easiest to calculate. This is why the efficiency of an ideal heat engine is independent of the details of the engine.

Thermodynamic Temperature and Entropy.

The discussion above gives us the opportunity to develop a thermodynamic definition of temperature, where temperature is proportional to heat added:

$$Q = ST. \tag{II.9}$$

Note that we have introduced a new quantity: the entropy, *S*. This quantity, as expected, is a constant only in a reversible system. It turns out that the change in *S* is always positive in any irreversible system. Since the universe is not reversible, this is why the second law of thermodynamics is often cast into the expression that the total entropy of the universe is increasing. More rigorously, the total entropy of any closed system (i.e., one that has no communication with the remainder of the universe) cannot decrease.

Now, let's cast the above into differential form:

$$dS = \frac{dQ}{T}. \tag{II.10}$$

If we now combine this with the differential form for work, *-PdV*, then the first law of thermodynamics becomes

$$dE = TdS - PdV. \tag{II.11}$$

Note this means that

$$\left(\frac{\partial E}{\partial S}\right)_V = T\,, \tag{II.12}$$

$$\left(\frac{\partial E}{\partial V}\right)_S = -P\,. \tag{II.13}$$

Thus E is a function of S and V, so we say that S and V are the natural variables of E. Using the internal energy as a starting point, a set of thermodynamic potentials, all with units of energy but having different natural variables, can be defined. These are given in Table II-1.

Only the internal energy has a simple physical meaning, but all of these quantities have vital practical importance. In particular, the two free energies will be seen more as we go through this course. The best way to illustrate F and G is to mention some of their unique properties and how they are used. The most important consideration in F is that, in any constant temperature (isothermal) process, the change in F is the reversible work. More practically, when doing theoretical calculations of thermodynamic processes, the simplest independent variables to work in are V and T, which are the natural variables of F. Typically, when a physicist mentions "free energy" without reference to which one, it is the Helmholtz free energy being discussed.

The Gibbs free energy has P and T as its natural variables. Since these are the most easily controlled quantities in typical experiments, particularly in benchtop chemistry (where P is usually constant and T is easily controlled), G is particularly important in chemistry. Hence, when a chemist mentions a "free energy" without specifying which one, it is almost always the Gibbs free energy that is being discussed. A particularly important feature of G is that, in any system at equilibrium, i.e., in a state where the net thermodynamic driving force is zero, the value of G is minimized for the given combination of P and T.

Table II-1. Thermodynamic Energies.

Name and Symbol	Relationship to E	Differential Form	Derivatives
Internal Energy E	$E = E$	$dE = TdS - PdV$	$\left(\frac{\partial E}{\partial S}\right)_V = T \quad \left(\frac{\partial E}{\partial V}\right)_S = -P$
Enthalpy H	$H \equiv E + PV$	$dH = TdS + VdP$	$\left(\frac{\partial H}{\partial S}\right)_P = T \quad \left(\frac{\partial H}{\partial P}\right)_S = V$
Helmholtz Free Energy F	$F \equiv E - TS$	$dF = -SdT - PdV$	$\left(\frac{\partial F}{\partial T}\right)_V = -S \quad \left(\frac{\partial F}{\partial V}\right)_T = -P$
Gibbs Free Energy G	$G \equiv H - TS$ $= E + PV - TS$	$dG = -SdT + VdP$	$\left(\frac{\partial G}{\partial T}\right)_P = -S \quad \left(\frac{\partial G}{\partial P}\right)_T = V$

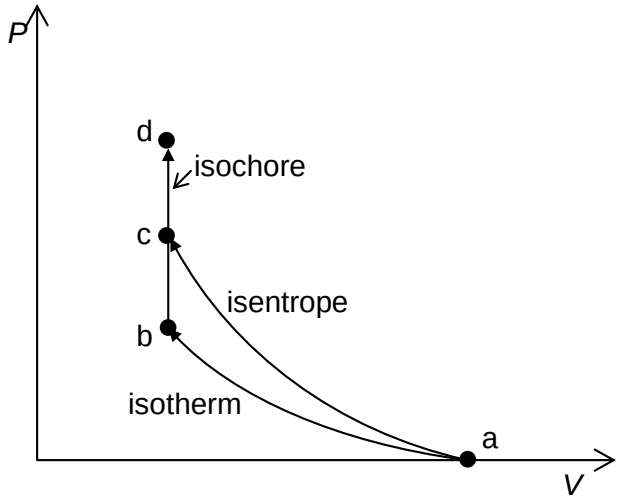

Figure II-3. Example of taking different paths to calculate the same change.

In terms of thermodynamic variables, the specific state achieved is independent of the path taken, as noted before during the discussion of the reversible heat engine. Hence, getting to a given *P*-*T* condition could be done by any number of processes (figure II-3), including isobaric, then isothermal; or isothermal, then isobaric; or isentropic, then isobaric, or isentropic, then isochoric (constant density or volume), etc. Most important is that, to keep the calculations tractable, we will always choose to hold one natural variable constant, while changing the other, then holding the other constant, while changing the first. That way, the calculations involve single derivatives for any given segment of the path. All the different properties of a material that relate the thermodynamic variables are defined (by necessity) as a given property when holding some variable constant.

The Equation of State.

The relationship between given thermodynamic quantities for a specific material can be described by a mathematical expression (or set of expressions) called the equation of state, or EOS. There are many EOS expressions in use and all are approximations, although often quite good ones. Most are empirical. The EOS has material-specific numerical parameters that quantify the thermodynamic derivatives. The most important of these usually involve descriptions of the bulk modulus (either isothermal, K_T, or isentropic, K_S), the specific heat (either isobaric, C_P, or isochoric, C_V), and the volume thermal expansion coefficient, α. These quantities are derivatives of the thermodynamic potentials and their definitions are given in Table II-2.

Along with the density, the quantities in Table II-2 are combined in a derived parameter that is used almost constantly in practice, the thermodynamic Grüneisen parameter, γ:

Table II-2. Common Thermodynamic Derivatives.			
Quantity	Symbol	Constant	Defining Formula
Isothermal Bulk Modulus	K_T	Temperature	$K_T \equiv -V\left(\frac{\partial P}{\partial V}\right)_T$
Isentropic Bulk Modulus	K_S	Entropy	$K_S \equiv -V\left(\frac{\partial P}{\partial V}\right)_S$
Isochoric Specific Heat	C_V	Specific Volume	$C_V \equiv \left(\frac{\partial E}{\partial T}\right)_V$
Isobaric Specific Heat	C_P	Pressure	$C_P \equiv \left(\frac{\partial H}{\partial T}\right)_P$
Volume Thermal Expansion Coefficient	α	Pressure	$\alpha \equiv \frac{1}{V}\left(\frac{\partial V}{\partial T}\right)_P$

$$\gamma \equiv V\left(\frac{\partial P}{\partial E}\right)_V = \frac{\alpha K_T}{\rho C_V} = \frac{\alpha K_S}{\rho C_P}. \tag{II.14}$$

The advantage of this particular quantity is that it allows calculations to be done without explicitly including temperature, which is usually not known in shock wave experiments. Most EOS formulations in use are partial equations of state in that they do not contain all of the expressions required to fully constrain all of the thermodynamic potentials. Usually, the most important unconstrained quantity is the temperature. The most common EOS forms used in shock wave work consist of a reference compression curve and an expression to describe thermal deviations from that curve.

The Murnaghan equation.

The simplest reference curve formulation that can reasonably describe the compression of a material over an extended pressure range is the Murnaghan equation, which has a bulk modulus that changes linearly with pressure:

$$K = K_0 + K'P\,. \tag{II.15}$$

Given the definition of *K*, this leads to the following expressions:

$$K = K_0\left(\frac{\rho}{\rho_0}\right)^{K'}, \tag{II.16}$$

$$P = P_0 + \frac{K_0}{K'}\left[\left(\frac{\rho}{\rho_0}\right)^{K'} - 1\right]. \tag{II.17}$$

The Birch-Murnaghan Equation.

The main shortcoming of the Murnaghan equation is that *K* is a more complicated function than can be expressed by a linear expression in *P*. An expression that attempts to give a more realistic description of the compression behavior is the Birch-Murnaghan equation:

$$P = \frac{3}{2} K_0 \left(x^7 - x^5 \right) \left[1 + \xi_1 - \xi_1 x^2 + \xi_2 \left(x^2 - 1 \right)^2 \right], \tag{II.18}$$

where

$$x = \left(\frac{\rho}{\rho_0} \right)^{1/3}, \tag{II.19}$$

$$\xi_1 = \frac{3}{4} \left(4 - K' \right), \tag{II.20}$$

and

$$\xi_2 = \frac{3}{8} K_0 K'' + \frac{3}{8} K' \left(K' - 7 \right) + \frac{143}{24}. \tag{II.21}$$

Where K′ and K″ are the first and second pressure derivatives of the bulk modulus. For an isothermal reference curve at 0 K, (II.18) corresponds to an interatomic potential of the form

$$\varphi(r) = -A \left(\frac{a}{r} \right)^2 + B \left(\frac{a}{r} \right)^4 + C \left(\frac{a}{r} \right)^6 + \cdots. \tag{II.22}$$

For the special case of ξ_1 and higher order coefficients equal to zero, this corresponds to coefficients in eq. (II.22) of all terms but the first two being zero. A nonzero value of ξ_1 corresponds to a nonzero value of *C* in (II.22) and the expression can be carried to higher order, though very seldom past the terms specifically given in (II.18). The Birch-Murnaghan EOS is an example of a finite strain model. As such, it is nominally formulated specifically as an isotherm. However, except that the relevant interatomic potential no longer is described by (II.22), there is no prohibition against using (II.18) to describe an isentrope. This expression is probably the most commonly used reference curve model in the earth sciences. Equation (II.18) is commonly referred to as the fourth-order Birch-Murnaghan EOS (BMEOS). In this version, it is explicitly assumed that the higher derivatives of *K* have values that make coefficients of higher order terms than those in (II.18) equal to zero. If ξ_2 is equal to zero, then (II.18) is known as the third-order BMEOS. For certain specific applications, the BMEOS is particularly useful.

Because of this fact, I will go ahead and give you the *PdV* integral solution corresponding to (II.18):

$$-\int_{V_0}^{V_1} PdV = \frac{9}{2}V_0K_0\left[(\xi_1+1)\left(\frac{x^4}{4}-\frac{x^2}{2}+\frac{1}{4}\right)-\xi_1\left(\frac{x^6}{6}-\frac{x^4}{4}+\frac{1}{12}\right)+\xi_2\left(\frac{x^8}{8}-\frac{x^6}{2}+\frac{3x^4}{4}-\frac{x^2}{2}+\frac{1}{8}\right)\right] \tag{II.23}$$

Choice of Reference Curve.

The decision of whether to use an isotherm or an isentrope as the reference curve depends on whether you are working with the internal energy *E* or the Helmholtz free energy *F*. For theoretical EOS work, the most common approach is to use *F*, which then causes one to use an isothermal reference curve. However, we will find that internal energy is particularly easy to deal with when considering shock wave experiments, in which case one would use the isentrope. It should be noted that K_T is not equal to K_S and, in fact, the two are related by the expression

$$\frac{K_S}{K_T}=\frac{C_P}{C_V}=1+\alpha\gamma T\,. \tag{II.24}$$

The Mie-Grüneisen equation.

To calculate states off the reference curve, the transformation is usually done by an isochoric (i.e., constant density or constant volume) internal energy change (if you are working in *F*, the transformation to *E* is made using the definition in Table II-1, then the energy changed, then the transformation back to *F*). Recalling eq. (II.14), we can get

$$\left(\frac{\partial P}{\partial E}\right)_V=\rho\gamma\,. \tag{II.25}$$

This is the basis of a partial EOS known as the Mie-Grüneisen EOS, which in its most general form states

$$E_2-E_1=V\int_{P_1}^{P_2}\frac{dP}{\gamma}\ ,\quad \text{(constant } V\text{)}. \tag{II.26}$$

In the vast majority of studies, γ is assumed to explicitly depend only on ρ, so that (II.26) reduces to the simple expression:

$$E_2-E_1=\frac{V_r}{\gamma(V_r)}\left(P_2-P_1\right). \tag{II.27}$$

The most common assumption is that γ can be described by

$$\gamma = \gamma_0 \left(\frac{V}{V_0}\right)^n \quad . \tag{II.28}$$

It should be noted that there are other parameters that are also called Grüneisen parameters. Most of these are related in some way to the thermodynamic parameter. In particular, there are the so-called mode Grüneisen parameters, which describe the frequency shift of vibrational modes in a lattice, the lattice Grüneisen parameter, which describes the phonon contribution to the thermodynamic value, and the electronic Grüneisen parameter, which is related to the electronic energy in a metal. Of these, the lattice value is most closely associated with the thermodynamic value and, in fact, the two should be the same for insulators in the harmonic approximation. The lattice value, in turn, is related to the appropriate average of the mode parameters. Mode Grüneisen parameters in the harmonic approximation exhibit the behavior in (II.28), and the lattice parameter does so well above the Debye temperature, which is why (II.28) is often a reasonable behavior to assume. That said, the thermodynamic parameter includes all the warts, such as anharmonicity, that isn't captured by the harmonic approximation. Thus, for almost everything, the reality is more complicated. The reason that the expression in (II.28) is so widely used is the simple fact that, except where it can be calculated using (II.14), the thermodynamic Grüneisen parameter is poorly constrained by most types of data. We will see later that, in the case of shock wave compression, γ enters the computation only inside an integral, eq. (II.26), and so can only be constrained as an average over P at a given density. Another important point that is often overlooked is that, when constraining a complete EOS, the behavior of γ cannot be defined independently of the specific heat.

The shock wave EOS.

For many high-pressure applications, a particularly simple EOS is the shock wave EOS, which is simply the algebraic expression of the relationship between the speed, U_s, of a shock wave and the material speed behind the shock wave, u_p. We will discuss shock wave and this relationship, known as the Hugoniot curve, in detail later. For now, we will simply state the way this is used as an EOS. For many materials, the U_s-u_p relationship is linear:

$$U_s = C_0 + s u_p \quad . \tag{II.29}$$

It turns out that, taking the zero-pressure density, ρ_0, the pressure, P_H, compressed state density, ρ_H, and specific internal energy (relative to the initial state), E_H, can be easily calculated as

$$P_H = \rho_0 U_s u_p \ , \tag{II.30}$$

$$\rho_H = \rho_0 \frac{U_s}{U_s - u_p} \quad , \tag{II.31}$$

$$E_H = \tfrac{1}{2} u_p^2 \quad . \tag{II.32}$$

Typically, the Mie-Grüneisen EOS is used to calculate states away from the Hugoniot curve described in these equations.

A complete EOS: The (augmented) Vinet EOS.

There is one complete EOS that is gaining currency in the high pressure research community. It is known variously as the universal, Rose, or Vinet equation of state. Its functional form is purely empirical, but it does quite well at describing the behaviors of many materials, over a much wider range of pressure than with the Murnaghan or Birch-Murnaghan equations. Originally, the Vinet EOS was a partial EOS, but its originators later incorporated an explicit temperature dependence which, with a specific heat, makes it a complete EOS. As it now exists, this EOS is expressed by

$$P(T,V) \approx \frac{3K_0(T_0)}{X^2}\left(1-X\right)\exp\left[\eta_0(T_0)\left(1-X\right)\right] + \alpha_0(T_0)K_0(T_0)\left(T-T_0\right), \tag{II.33}$$

$$X \equiv \left(\frac{V}{V_0(T_0)}\right)^{1/3}, \tag{II.34}$$

$$\eta_0(T_0) = \frac{3}{2}\left[\left(\frac{\partial K}{\partial P}\right)_0 (T_0) - 1\right], \tag{II.35}$$

where the subscript 0 refers to a reference condition. Note that (II.33) is shown as an approximation. The actual definition of the reference isotherm (the first term) is somewhat more subtle, but is well approximated in practice by the expression shown.

Phase Changes.

When we discussed the Gibbs free energy, *G*, we noted that this quantity is minimized by a system at equilibrium. One of the consequences of this fact is that the stable phase of a material will always be the one with the lowest value of *G*. This being the case, an interesting problem arises. Although we can calculate changes in *E*, *G*, *F*, and *H*, we haven't said anything about their absolute values. In fact, in practice, the absolute value of *E* for some material phase at some set of conditions is arbitrary. However, once the value has been set at some reference condition for the stable phase at that condition, this constrains the values of all phases at all conditions. Usually, the reference value of *E* is simply set to zero at the reference condition for the most commonly considered phase (usually whatever phase happens to be stable at 1 bar and 298 K). Then, energies of transition from this reference condition and phase are given to the reference conditions of each independent phase.

The coexistence line, also called the phase boundary, between any two phases is determined by the locus of states at which the values of G are equal for the two phases being considered. Note that this boundary is a line. The number of degrees of freedom in a thermodynamic system is the number of components minus the number of phases, plus 2 ($f = c - p + 2$, called the Phase Rule). Thus, for a single component system (practically speaking, a system with a single pure substance and no chemical reactions), the boundary between two phases is a line in any chosen space. If three phases are involved, there are no degrees of freedom and a unique point, called the triple point, is constrained. Although equilibrium dictates that two phases can coexist only on the phase boundary, real life is a little more complicated. There is also a Gibbs free energy of mixing between two phases (or substances). For simple physical mixing of two phases of the same substance, this quantity is given by

$$\Delta G_{mix} = \frac{RT}{\mu}\left(x_1 \ln x_1 + x_2 \ln x_2\right), \tag{II.36}$$

where the quantities x are the mole fractions of the two phases. This expression assumes mixing in the atomic limit and thus is really a lower bound (the value is always negative). It is always minimized at equal proportions. The origin of this quantity is in the fact that the entropy of a mixed system, being a measure of disorder, is higher than that of a pure system. The entropy of mixing is simply the negative of the Gibbs energy of mixing divided by the temperature. It should be noted that eq. (II.36) can be extended to an arbitrary number of phases or substances simply by adding more terms in the parentheses.

Another complication in reality is that, sometimes, although a particular phase is at a higher Gibbs energy than some other phase, the transformation is frustrated so there is no path that can be followed in transforming from one to the other with the available energy. In that case, the higher-energy phase is said to be metastable. The most common example in everyday life is silica glass. Silica glass has a higher Gibbs energy than the crystalline phase with the same composition. However, in order to drive a transition in a short period of time, one would need to provide substantial energy by heating the glass.

Mechanical properties.

Now we will turn to a set of material behaviors and properties that are not thermodynamic. Thermodynamic properties, by virtue of the path independence noted earlier, are, in and of themselves, only state-dependent and are not sensitive to the path taken to achieve that state. Now we will consider properties, some of which are intimately tied to the thermodynamic properties, that can be affected by the history they have seen.

Stress and strain.

In order to discuss mechanical properties, we must consider the concepts of stress and strain. In the discussion of thermodynamics, we dealt extensively with pressure. If you place a cube in a pressure vessel, surrounded by a fluid, and applied pressure, the force per unit area seen by the cube is everywhere the same and everywhere directed inward normal to the surface of the cube. What if you placed the cube in a vise so that two

opposite faces were completely covered by the jaws of the vise? Now there is an additional force applied over the area of those two surfaces, so that the total force per unit area applied is not equal on all surfaces. This gets at the idea of stress.

Stress is the generalization of the force over area concept to the situation where both the magnitude and the vector direction of the force per unit area may depend on the orientation of the surface. Consider one surface of the cube and apply an arbitrary force **F** to that surface. Force is a vector and has three orthogonal components, one of which we can choose to be normal to the surface and two parallel to the surface, but orthogonal to one another. Thus, there are both normal and tangential forces, which means there are normal and tangential stresses.

Now consider our cube as a whole, with applied forces (and stresses) on each face (figure II-4). Any given stress component has a magnitude equal to the corresponding force component divided by the area over which the force is applied. Conversely, the force applied to a given surface area is equal to the stress multiplied by the area. The total stress, containing all of these components, is a tensor:

$$\boldsymbol{\sigma} = \begin{bmatrix} \sigma_{xx} & \sigma_{xy} & \sigma_{xz} \\ \sigma_{yx} & \sigma_{yy} & \sigma_{yz} \\ \sigma_{zx} & \sigma_{zy} & \sigma_{zz} \end{bmatrix}. \tag{II.37}$$

This tensor is a symmetric matrix, so that $\sigma_{xy} = \sigma_{yx}$. Now consider a randomly chosen point in the cube, with some position (x,y,z) and displace it to a new position $(x+u,y+v,z+w)$, so that the displacement has components u, v, and w, corresponding to the x, y, and z directions. Now, suppose an adjacent point starts at a position $(x+\delta x,y+\delta y,z+\delta z)$ and is displaced by $u+\delta u$, $v+\delta v$, $w+\delta w$ components. If the displacements are infinitesimal, then the components can be written as

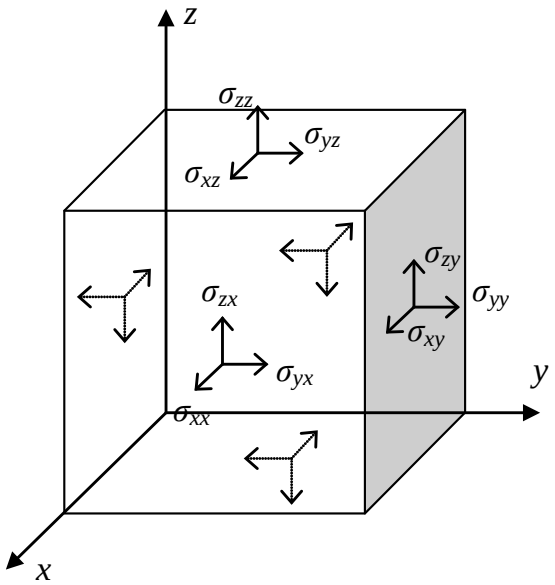

Figure II-4. Infinitesimal element with stress components acting on surfaces.

$$\left.\begin{aligned}\delta u &= \frac{\partial u}{\partial x}\delta x + \frac{\partial u}{\partial y}\delta y + \frac{\partial u}{\partial z}\delta z \\ \delta v &= \frac{\partial v}{\partial x}\delta x + \frac{\partial v}{\partial y}\delta y + \frac{\partial v}{\partial z}\delta z \\ \delta w &= \frac{\partial w}{\partial x}\delta x + \frac{\partial w}{\partial y}\delta y + \frac{\partial w}{\partial z}\delta z\end{aligned}\right\}. \tag{II.38}$$

The above expression contains nine derivatives of displacement with respect to position. These derivatives define the strain, which is the distortion of the material, and the solid body rotation of the material:

$$\left.\begin{aligned} \varepsilon_{xx} &= \frac{\partial u}{\partial x}, & \varepsilon_{yy} &= \frac{\partial v}{\partial y}, & \varepsilon_{zz} &= \frac{\partial w}{\partial z} \\ \varepsilon_{yz} &= \frac{1}{2}\left(\frac{\partial w}{\partial y} + \frac{\partial v}{\partial z}\right) = \varepsilon_{zy}, & \varepsilon_{zx} &= \frac{1}{2}\left(\frac{\partial u}{\partial z} + \frac{\partial w}{\partial x}\right) = \varepsilon_{xz}, & \varepsilon_{xy} &= \frac{1}{2}\left(\frac{\partial v}{\partial x} + \frac{\partial u}{\partial y}\right) = \varepsilon_{yx} \\ 2\bar{\omega}_x &= \frac{\partial w}{\partial y} - \frac{\partial v}{\partial z}, & 2\bar{\omega}_y &= \frac{\partial u}{\partial z} - \frac{\partial w}{\partial x}, & 2\bar{\omega}_z &= \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}\end{aligned}\right\}. \tag{II.39}$$

The top row in (II.39) defines the longitudinal, or normal, strains, which are the dilatational distortions (expansion or contraction) parallel to the three axes we have defined. The second row defines shear strains, which distort the shape of our infinitesimal cube without changing its volume. The final row defines the rotations of the cube around the three axial directions. Note that the linear displacement of an object is not described by (II.39), because the quantities involved are derivatives of displacement, which are identically zero for solid body displacement. Rather, these are the rotation and distortion that are superimposed on the displacement.

Obviously, the quantities in (II.39) are differential quantities, but a real object may undergo macroscopic deformation, so let's consider how macroscopic strains are measured. First, let's focus on the longitudinal strain, considering the component ε_{xx}. The simple algebraic version of (II.39), which is known as engineering strain, is

$$\varepsilon_{xx} = \frac{\Delta L_x}{L_{x0}}, \tag{II.40}$$

where L_x is the length of the object in the x direction. However, the more appropriate approach is to integrate the distortion, resulting in the logarithmic strain, often called true strain:

$$\varepsilon_{xx} = \int \frac{dL_x}{L_x} = \Delta \ln L_x = \ln \frac{L_{x0} + \Delta L_x}{L_{x0}} . \qquad \text{(II.41)}$$

Note that we have, in (II.40) and (II.41), made the assumption that the object has not been rotated. When rotation has occurred, the appropriate approach is to, for the purpose of calculating the strain, rotate the coordinate system with the object.

The strain due to shear stress can also be defined either as an engineering strain or a true strain. If we consider our block of material, the engineering strain is the amount of offset of the surface to which the stress is applied, relative to the opposite surface, divided by the separation distance:

$$\varepsilon_{yx} = \frac{\Delta y}{L_x} + \frac{\Delta x}{L_y} \qquad \text{(II-42)}$$

This expression is the engineering shear strain and is better written as angular values (in radians), as noted in figure II-5. Note a vitally important feature here—there is a factor of 2 difference between the expressions in (II.39) and (II.42). The engineering strain combines the terms above and below the diagonal of the tensor, while the true strain does not. That is the only difference and is vital because most tabulations of material properties involving shear strain are based on the engineering strain. You should be careful when performing quantitative calculations to assure that you understand which type of strain you should be using with a given set of published parameters. It should also be noted that many reference works will treat longitudinal strain using true strain, but will use the engineering version of shear strain. Because the only difference is a factor of 2, this is not a serious inconsistency, although it is mathematically not strictly correct with the treatment of strain as a tensor.

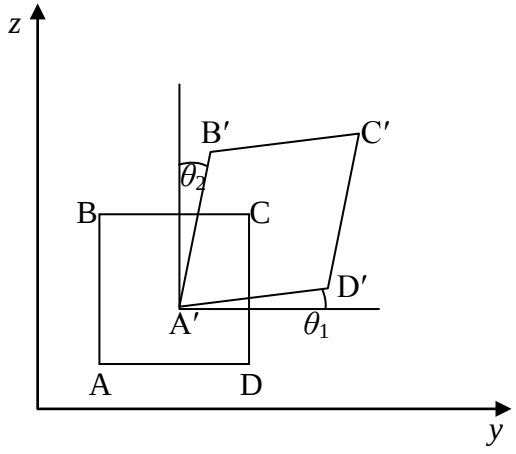

Figure II-5. Shear strain given as angular quantities.

Hooke's Law.

Hooke's Law states that, for sufficiently small deformations, stress and strain are proportional. This law can be generalized to state that the nine components of the stress tensor can be written as linear combinations of the nine components of the strain tensor. Now we can see why the engineering version of the shear strain is usually used—the tensors are symmetric, so that combining the upper and lower symmetric terms reduces the number of terms required to write Hooke's law. We will follow this convention and now present the equations describing the generalized form of Hooke's law:

$$\left.\begin{aligned}
\sigma_{xx} &= c_{11}\varepsilon_{xx} + c_{12}\varepsilon_{yy} + c_{13}\varepsilon_{zz} + c_{14}\varepsilon_{yz} + c_{15}\varepsilon_{zx} + c_{16}\varepsilon_{xy} \\
\sigma_{yy} &= c_{21}\varepsilon_{xx} + c_{22}\varepsilon_{yy} + c_{23}\varepsilon_{zz} + c_{24}\varepsilon_{yz} + c_{25}\varepsilon_{zx} + c_{26}\varepsilon_{xy} \\
\sigma_{zz} &= c_{31}\varepsilon_{xx} + c_{32}\varepsilon_{yy} + c_{33}\varepsilon_{zz} + c_{34}\varepsilon_{yz} + c_{35}\varepsilon_{zx} + c_{36}\varepsilon_{xy} \\
\sigma_{yz} &= c_{41}\varepsilon_{xx} + c_{42}\varepsilon_{yy} + c_{43}\varepsilon_{zz} + c_{44}\varepsilon_{yz} + c_{45}\varepsilon_{zx} + c_{46}\varepsilon_{xy} \\
\sigma_{zx} &= c_{51}\varepsilon_{xx} + c_{52}\varepsilon_{yy} + c_{53}\varepsilon_{zz} + c_{54}\varepsilon_{yz} + c_{55}\varepsilon_{zx} + c_{56}\varepsilon_{xy} \\
\sigma_{xy} &= c_{61}\varepsilon_{xx} + c_{62}\varepsilon_{yy} + c_{63}\varepsilon_{zz} + c_{64}\varepsilon_{yz} + c_{65}\varepsilon_{zx} + c_{66}\varepsilon_{xy}
\end{aligned}\right\}. \qquad \text{(II.43)}$$

The constants in (II.43) are the elastic constants of the material. In many materials, inherent or effective symmetries reduce the number of independent terms in (II.43). The simplest case is that of an isotropic material. Since untextured polycrystalline materials are effectively isotropic, even when the underlying crystalline structures are much less symmetric, we will limit our discussion to this particular case. Complete description of an isotropic material's elastic behavior requires only two parameters, known as the Lamé constants λ and μ:

$$\left.\begin{aligned}
c_{12} = c_{13} = c_{21} = c_{23} = c_{31} = c_{32} &= \lambda \\
c_{44} = c_{55} = c_{66} &= \mu \\
c_{11} = c_{22} = c_{33} &= \lambda + 2\mu
\end{aligned}\right\}, \qquad \text{(II.44)}$$

with the remaining elastic constants equal to zero. Inserting these into (II.43) gives

$$\left.\begin{aligned}
&\sigma_{xx} = \lambda\Delta + 2\mu\varepsilon_{xx} \quad \sigma_{yy} = \lambda\Delta + 2\mu\varepsilon_{yy} \quad \sigma_{zz} = \lambda\Delta + 2\mu\varepsilon_{zz} \\
&\sigma_{yz} = \mu\varepsilon_{yz} \qquad \sigma_{zx} = \mu\varepsilon_{zx} \qquad \sigma_{xy} = \mu\varepsilon_{xy}
\end{aligned}\right\}, \qquad \text{(II.45)}$$

where

$$\Delta = \varepsilon_{xx} + \varepsilon_{yy} + \varepsilon_{zz}. \qquad \text{(II.46)}$$

Although only two constants are required, we usually use four constants that are more readily related to experimentally measured stress-strain relationships. These are Young's

modulus, Y, the isentropic bulk modulus, K_S, the shear modulus, G, and the Poisson ratio, ν.

Young's modulus is the ratio of longitudinal stress, say, σ_{xx}, to the longitudinal strain, ε_{xx}, when the two orthogonal stresses are set to zero. Thus, the first line of (II.45) becomes

$$\left.\begin{aligned}\sigma_{xx} &= (\lambda+2\mu)\varepsilon_{xx} + \lambda(\varepsilon_{yy}+\varepsilon_{zz}) \\ 0 &= (\lambda+2\mu)\varepsilon_{yy} + \lambda(\varepsilon_{xx}+\varepsilon_{zz}) \\ 0 &= (\lambda+2\mu)\varepsilon_{zz} + \lambda(\varepsilon_{xx}+\varepsilon_{yy})\end{aligned}\right\}. \tag{II.47}$$

Solving for the strains, we get

$$\left.\begin{aligned}\varepsilon_{xx} &= \frac{\lambda+\mu}{\mu(3\lambda+2\mu)}\sigma_{xx} \\ \varepsilon_{yy} = \varepsilon_{zz} &= -\frac{\lambda}{\mu(3\lambda+2\mu)}\sigma_{xx}\end{aligned}\right\}. \tag{II.48}$$

Y thus becomes

$$Y = \frac{\sigma_{xx}}{\varepsilon_{xx}} = \frac{\mu(3\lambda+2\mu)}{\lambda+\mu}. \tag{II.49}$$

The Poisson ratio is the ratio of the longitudinal and lateral strains in the foregoing:

$$\nu = -\frac{\varepsilon_{yy}}{\varepsilon_{xx}} = \frac{\lambda}{2(\lambda+\mu)}. \tag{II.50}$$

The bulk modulus is the ratio of the volume stress, or pressure, to the volume strain in the case where the longitudinal stresses are all equal and thus equal to the pressure:

$$K = \frac{P}{\Delta} = \lambda + \frac{2\mu}{3}. \tag{II.51}$$

Finally, the shear modulus is identical to the second Lamé constant:

$$G = \mu. \tag{II.52}$$

Strength.

One of the consequences of Hooke's Law is that when a stress is applied and then removed, the strain due to that stress is also reversed, so that the final strain state is

identical to the initial strain state. This is elastic deformation. However, when a sufficiently large stress is applied, a point is reached where some portion of the deformation is not reversible. This phenomenon is called plastic yielding and the permanent deformation is the plastic strain. Figure II-6 shows a stress-strain curve that corresponds to this behavior. The stress magnitude at the onset of plastic yielding is the strength of the material. As might be expected, the strength depends on the details of the stress field, so that there is no single strength. However, certain common stress situations are encountered in practice, or easily set up in tests, so that the strengths of materials in these special situations are commonly measured and used. In particular, there is tensile strength, which is the case where σ_{xx} is tensile and σ_{yy} and σ_{zz} are both zero. A similar quantity can be defined for compression.

The underlying cause of elasticity is that, up to a certain level of deformation, particularly in solids with crystalline symmetry, the atomic configuration in the no-strain state is the lowest energy configuration and any distortion of the crystal lattice leads to restoring forces that will return the lattice to its original configuration. However, a sufficiently large distortion results in the creation or activation of defects in the lattice such that a new local minimum in the energy surface is encountered, so that the restoring force goes to zero before the lattice has returned to its original state. This point is where plastic yielding begins. As might be expected, in the crystalline lattice, the ability of defects to form and propagate depends critically on the lattice symmetry, so that low-symmetry structures may have different elastic behaviors and different strengths for identical stresses or strains applied in different directions.

The discussion in the previous paragraph is put in the context of the crystalline lattice, which assumes a single continuous crystal. Polycrystalline materials or noncrystalline materials can be thought of as crystals with particularly large numbers of preexisting defects and so having much lower, but finite, strengths than the pure ideal crystals of the same composition. It should be remembered that inertial effects and finite time required

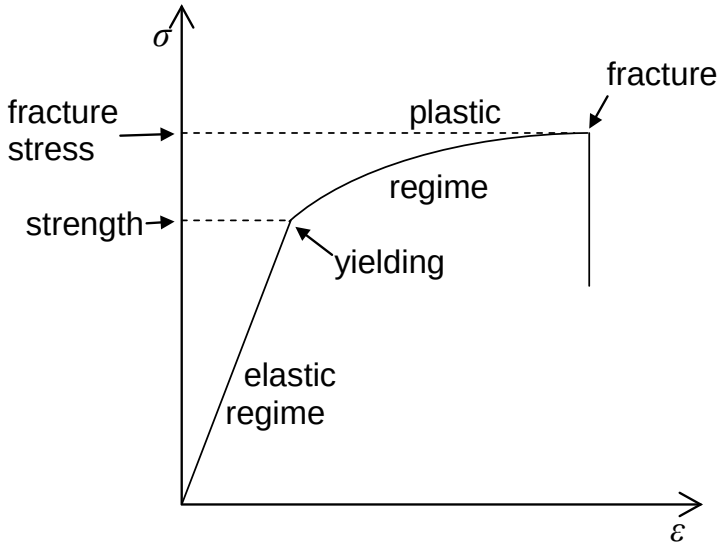

Figure II-6. Stress-strain curve of a material showing important features.

for information to travel across a lattice means that the strength and even the effective elastic constants of a material can depend on the strain rate. Also, because strain results in rearrangements of lattices, strength and elastic constants can also depend on the magnitude of the strain. Hence, it is common to encounter discussions of rate-dependence or to hear one speak of work-hardening or work-softening of a material. The hardening or softening, used in this context, is speaking of changes in the yield strength of the material.

Fracture toughness.

Eventually, at sufficiently high strains, particularly in shear or in tension, the number of defects at a localized plane in the lattice becomes so great that the lattice is no longer continuous across the plane and the two sides of the plane no longer exert any siginificant attractive forces on one another, so that the material separates into two independent pieces. This behavior is known as fracture and the stress at which it occurs is known as the fracture toughness. Often, the term strength is used in place of toughness, but strictly speaking, the term strength applies to plastic yielding. However, the usage of strength to apply to fracture is so common as to be considered a legitimate use of the term by most researchers.

As in the discussion of yield strength, we have been considering a single crystal, but again polycrystalline materials or noncrystalline materials can be thought of as crystals with particularly large amounts of preexisting defects, often with many of these defects already aligned on localized planes. A note concerning noncrystalline materials is in order here. Many noncrystalline materials are noncrystalline because they exhibit strongly covalent bonding character that allows them to form chains or networks that extend well beyond the size scale typical of a crystalline unit cell. In such materials, the yield strength may be negligible and the restoring forces that do occur are the result of straining the bonds in these polymeric constructs away from their lowest energy configurations. The ability of polymeric chains and networks to deform without significant bond breakage is why many polymers are able to absorb large amounts of plastic strain before fracturing.

III. WAVES AND SHOCK WAVES

Planar Waves and the Wave Equation.

Imagine a block of some substance, which has an (as yet unspecified) equation of state and, initially, has a uniform density throughout. Suppose that, on one side of the block, we apply an infinitesimal time-varying deviation in the pressure P:

$$P(t) = P_0 + P'(t). \tag{III.1}$$

At the surface where the pressure is applied, the density is related to the pressure by

$$\rho(t) = \rho_0 + \rho'(t) = \rho_0 + P'(t)\left(\frac{\partial \rho}{\partial P}\right)_s. \tag{III.2}$$

Alternatively, we can write the pressure as

$$P(t) = P_0 + \rho'(t)\left(\frac{\partial P}{\partial \rho}\right)_S \quad . \tag{III.3}$$

Writing the derivative in terms of commonly used EOS parameters and taking advantage of the infinitesimal nature of the disturbance,

$$P(t) = P_0 + \rho'(t)\frac{K_{S0}}{\rho_0}. \tag{III.4}$$

Taking the surface as an infinitesimally thick layer, then the pressure is now applied to the immediately adjacent infinitesimally thick layer, and so on. The atoms at one position must be accelerated toward their neighbors, resulting in a change in density as the atoms approach their neighbors. Because this requires finite time, there will be an increasing time lag associated with the pressure as one moves farther away from the surface where the pressure was originally applied. Thus, we have a varying pressure and density field that depends both on position and time—this is a wave.

Since the force is normal to the surface where it is applied, the atoms are accelerated normal to the surface. However, the disturbance also is propagating that same direction, so that the wave propagation direction and the induced atomic motions are parallel. This is known as a compressional, or longitudinal, wave. We are most familiar with longitudinal waves as sound waves. If the applied stress had been shear, then the atomic motions would be perpendicular to the propagation direction and the wave would be a transverse, or shear, wave. All materials are able to propagate longitudinal waves, while only materials that can support shear stress are able to propagate a shear wave. Now let's consider the mathematical description of the wave we are discussing.

Consider the density, which depends on both the time and the position, $\rho(x,t)$, where propagation is in the x direction. Since the pressure variation is applied uniformly to a planar surface, there is no lateral variation in the pressure, so the wave is planar. Now, at time = 0, consider two closely spaced points in the undisturbed material, x and $x + dx$.

Since there is no net lateral force, the material does not spread laterally and we can think of a section of the wave having a constant area *A*. The mass between *x* and *x* + *dx* traversed by that section of the wave is

$$m = \rho_0 A dx \quad . \tag{III.5}$$

Starting at time zero, the wave passes, imparting motion to the material and causing time- and position-dependent displacements χ at time *t* to the new positions

$$\begin{matrix} x + \chi\left(x,t\right) \\ \\ x + dx + \chi(x + dx, t) \end{matrix} \quad , \tag{III.6}$$

where the difference in displacements is given by

$$\chi(x + dx, t) - \chi(x, t) = \left(\frac{\partial \chi}{\partial x}\right)_t dx \quad . \tag{III.7}$$

The mass remains constant, so that the new density ρ_1 must satisfy

$$\rho_1 A[x + dx + \chi(x + dx, t) - x - \chi(x, t)] = \rho_0 A dx \quad . \tag{III.8}$$

This leads to the densities being related by

$$\rho_1 = \rho_0 \left(1 - \left(\frac{\partial \chi}{\partial x}\right)_t\right). \tag{III.9}$$

Putting this back into equation (III.2), we find at time *t*,

$$\rho' = -\rho_0 \left(\frac{\partial \chi}{\partial x}\right)_t . \tag{III.10}$$

Newton's second law gives the force on the mass as

$$F = \rho_0 A dx \left(\frac{\partial^2 \chi}{\partial t^2}\right)_x . \tag{III.11}$$

The force on the layer at *x* is *P*(*x*,*t*)*A*, while the force at *x* + *dx* is -*P*(*x* + *dx*,*t*)*A*. The net force, then, is

$$F = P(x,t)A - P(x + dx, t)A = -A dx \left(\frac{\partial P'}{\partial x}\right)_t . \tag{III.12}$$

This means that

$$\rho_0 A dx \left(\frac{\partial^2 \chi}{\partial t^2} \right)_x = -A dx \left(\frac{\partial P'}{\partial x} \right)_t , \tag{III.13}$$

which is equivalent to

$$\rho_0 \left(\frac{\partial^2 \chi}{\partial t^2} \right)_x = -\left(\frac{K_{S0}}{\rho_0} \right) \left(\frac{\partial \rho'}{\partial x} \right)_t . \tag{III.14}$$

Taking the derivative of ρ' from (III.10) and eliminating ρ_0 on both sides, we get

$$\left(\frac{\partial^2 \chi}{\partial t^2} \right)_x = \left(\frac{K_{S0}}{\rho_0} \right) \left(\frac{\partial^2 \chi}{\partial x^2} \right)_t . \tag{III.15}$$

Now, let's define (dropping the subscripts 0 for simplicity),

$$c^2 \equiv \frac{K_s}{\rho} . \tag{III.16}$$

Then we finally rearrange to get

$$\left(\frac{\partial^2 \chi}{\partial x^2} \right)_t = \frac{1}{c^2} \left(\frac{\partial^2 \chi}{\partial t^2} \right)_x . \tag{III.17}$$

This is the wave equation for sound, but the form is more general. With different parameters defining *c*, this equation describes any type of wave, including electromagnetic radiation and water waves.

Now, let's find the solution to this equation. Since the derivatives on both sides are of the same order, we can write new variables that are a linear combination of the variables *x* and *t*. We might also consider that the proportionality constant *c* should be used, so we introduce the new variables $\xi = x - ct$ and $\eta = x + ct$. The same applies to a second arbitrary function $F_2(\eta)$, so that the most general form for χ is

$$\chi = F_1(x - ct) + F_2(x + ct) . \tag{III.18}$$

Wave speed and the speed of sound.

Note that the function F_1 describes a disturbance propagating with speed *c* in the positive *x* direction, while F_2 desribes propagation in the negative *x* direction, again at speed *c*, and that the functions F_1 and F_2 are arbitrary and independent. Given that the speed of the wave is *c*, we now know the speed of sound in terms of material properties:

$$c = \sqrt{\frac{K_S}{\rho}}\,. \tag{III.19}$$

It turns out that the foregoing is much more general than the derivation above might seem to indicate. If the force due to pressure is replaced by an elastic restoring force, then this equation holds equally well for arbitrary (small) deformation in materials that show strength. In such cases, remember that there are two different types of stress and strain—longitudinal and shear. When the shear strength of a material is not zero, the longitudinal sound speed is modified by the fact that longitudinal deformation has both dilatational and shear strain components. When a longitudinal strain is applied to a laterally unconfined material, the longitudinal strain is given by

$$\varepsilon_x = \frac{\sigma_x}{Y}\,, \tag{III.20}$$

where Y is Young's modulus. The lateral strains are related to this by Poisson's ratio:

$$\varepsilon_y = \varepsilon_z = -\nu\varepsilon_x = -\nu\frac{\sigma_x}{Y} \tag{III.21}$$

Note how this makes the relations between the stress and strain come out:

$$\left.\begin{aligned}
\varepsilon_x &= \frac{1}{Y}\left[\sigma_x - \nu\left(\sigma_y + \sigma_z\right)\right] \\
\varepsilon_y &= \frac{1}{Y}\left[\sigma_y - \nu\left(\sigma_z + \sigma_x\right)\right] \\
\varepsilon_z &= \frac{1}{Y}\left[\sigma_z - \nu\left(\sigma_x + \sigma_y\right)\right]
\end{aligned}\right\} \tag{III.22}$$

However, if the material is confined so that it cannot expand laterally, such as when the material has an effectively infinite lateral extent, the stress field must adjust itself to make the lateral strains exactly zero, which is accomplished if

$$\sigma_y = \sigma_z = \frac{\nu\sigma_x}{1-\nu}, \tag{III.23}$$

and the longitudinal strain becomes

$$\varepsilon_x = \frac{\sigma_x}{Y}\left[1 - \frac{2\nu^2}{\left(1-\nu\right)}\right]. \tag{III.24}$$

Let us define a modulus *m* that has the property

$$m = \frac{\sigma_x}{\varepsilon_x} = \frac{Y(1-\nu)}{(1-2\nu)(1+\nu)}. \tag{III.25}$$

It turns out that this quantity can be related to the bulk and shear moduli by their relations to Young's modulus:

$$K_S = \frac{Y}{3(1-2\nu)}, \tag{III.26}$$

$$G = \frac{Y}{2(1+\nu)}, \tag{III.27}$$

$$m = K_S + \frac{4}{3}G. \tag{III.28}$$

Thus, the longitudinal elastic wave in a medium with strength will travel at a speed

$$c_L = \sqrt{\frac{K_S + \frac{4}{3}G}{\rho}}. \tag{III.29}$$

As you might suspect, the speed of a shear wave in a material with strength is

$$c_S = \sqrt{\frac{G}{\rho}}. \tag{III.30}$$

Shock Waves.

Origin of shock waves.

As noted above, the wave equation allows arbitrary wave profiles, so the profile can be a discontinuous jump, i.e., a shock wave. We know that shock waves exist, but why do they exist? The answer lies in the details of the equation of state.

With a few special exceptions over a restricted pressure range, materials are characterized by an isentropic bulk modulus with a positive pressure derivative K_S', so that K_S increases with pressure. If you apply a small increase in pressure, say in the leading edge of a wave, both K_S and ρ will increase:

$$K_S = K_{S0} + K_S'\delta P = K_{S0}\left(1 + K_S'\frac{\delta P}{K_{S0}}\right), \tag{III.31}$$

$$\rho = \rho_0 + \rho_0 \frac{\delta P}{K_{S0}} = \rho_0 \left(1 + \frac{\delta P}{K_{S0}} \right). \qquad \text{(III.32)}$$

These lead to

$$c = c_0 \left(1 + K_S' \frac{\delta P}{K_{S0}} \right)^{1/2} \left(1 + \frac{\delta P}{K_{S0}} \right)^{-1/2} \quad , \qquad \text{(III.33)}$$

where c_0 is the initial sound speed. The consequence of this equation is that the next small part of the wave has a higher speed and can overtake the leading edge of the wave. Now, consider the entire wave, which can be thought of as little wavelets with each part moving faster than all the parts that went before (Figure III-1). The entire wave steepens. Because the trailing portions of the wave cannot pass the leading portions (they would slow down upon reaching the uncompressed material ahead of the wave), the wave steepens into a discontinuity and the resulting wave is a shock wave. In fact, because the wave also compresses the material so that the distance between two points in the material is decreased, shock waves can even form in materials where K_S' is slightly negative. Strictly speaking, as long as the expression

$$\left(1 + K_S' \frac{\delta P}{K_{S0}} \right) \left(1 + \frac{\delta P}{K_{S0}} \right) > 1 \qquad \text{(III.34)}$$

holds, a wave will steepen into a shock wave if allowed to propagate far enough.

It should be noted that a few specific phases of a few materials have relatively large negative values of K_S', at least over some pressure range. Among these are fused quartz and the γ phase of cerium. These phases, over the appropriate pressure range, cannot form shock waves in compression.

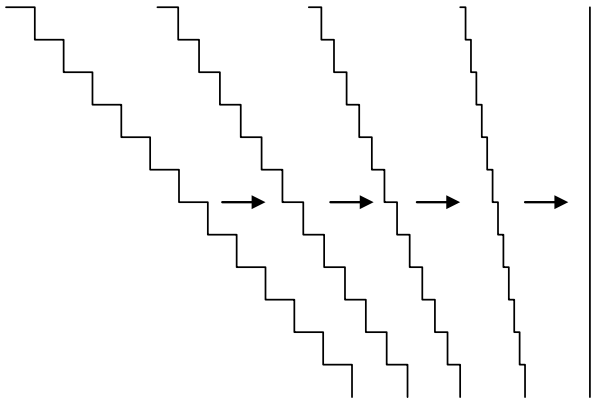

Figure III-1. Steepening of a wave into a shock wave.

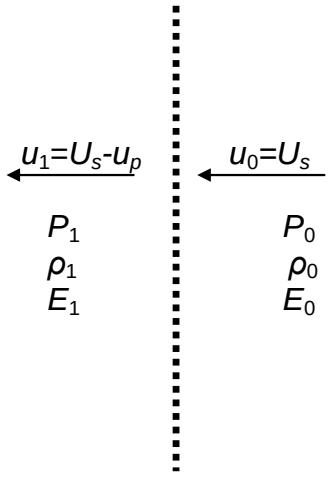

Figure III-2. Thermodynamic potentials and velocities are changed by passage through the shock wave.

The Conservation Equations.

A shock wave changes the state of the material through which it passes (Figure III-2) and the state of material emerging from the shock wave is related to the state of the material entering the shock wave by the conservation of mass, momentum, and energy. The equations describing the conservation relations are known as the Rankine-Hugoniot equations. The Rankine-Hugoniot equations are most easily derived if they are placed in a reference frame in which the shock wave itself is at rest.

Conservation of Mass.

Conservation of mass requires that the mass flux $j = \rho u$ be constant, where u is the speed of the material perpendicular to the wave, so that the relationship between the flux out of the wave and that into the wave is

$$\rho_1 u_1 = j = \rho_0 u_0 . \tag{III.35}$$

where $u_1 = U_s - u_p$, with u_p being the velocity change imparted by the wave to the material and U_s being the speed of the shock wave relative to the unshocked material.

Conservation of Momentum (Newton's second law).

The momentum density of the material is equal to ρu, so that the momentum flux is ρu^2. As a consequence of (III.35), any nontrivial solution requires that the momentum flux must change. This, then, requires an acceleration and an accelerating force in accordance with Newton's second law. The force is provided by a pressure difference, $P_1 - P_0$. If the material spends an arbitrarily short time Δt in the shock wave, then the amount of material in area A of the shock wave is $Aj\Delta t$. The velocity difference $u_1 - u_0$ must be equal to

$$u_1 - u_0 = \frac{F}{Aj\Delta t}\Delta t = \frac{P_0 - P_1}{j} . \tag{III.36}$$

We can rearrange this and substitute for j to give

$$P_1 + \rho_1 u_1^2 = P_0 + \rho_0 u_0^2 . \tag{III.37}$$

Conservation of energy.

An ideal shock wave is a discontinuity, so the material remains in the shock wave only an infinitesimal amount of time. Hence the process is adiabatic because there is no time for heat transfer, so the energy flux into the shock wave must equal that coming out. The total energy flux is equal to the mass flux times the sum of enthalpy and kinetic energy, so

$$\rho_1 u_1 \left(H_1 + \frac{1}{2} u_1^2 \right) = \rho_0 u_0 \left(H_0 + \frac{1}{2} u_0^2 \right). \tag{III.38}$$

Since ρu is constant, we can make use of $H = E + PV = E + P/\rho$ to get

$$E_1 + \frac{P_1}{\rho_1} + \frac{1}{2} u_1^2 = E_0 + \frac{P_0}{\rho_0} + \frac{1}{2} u_0^2 . \tag{III.39}$$

Since the specific volume $V = 1/\rho$,

$$u = jV . \tag{III.40}$$

Then, we can rewrite (III.37) as

$$P_1 + j^2 V_1 = P_0 + j^2 V_0 \tag{III.41}$$

and solve for j^2:

$$j^2 = \frac{P_1 - P_0}{V_0 - V_1} . \tag{III.42}$$

Using this with (III.40), rearranging (III.39) and simplifying, we get

$$E_1 - E_0 = \frac{1}{2} \left(P_0 + P_1 \right) \left(V_0 - V_1 \right) . \tag{III.43}$$

Thus, we have an expression of the internal energy change.

Laboratory Reference Frame.

The conservation equations are typically more useful in practice if we recast them into a reference frame in which the material is initially at rest, since this is the situation

that apples to measurements in the usual laboratory experiment. Reminding ourselves that $U_s = -u_0$ and setting

$$u_p = u_0 - u_1 , \tag{III.44}$$

we rearrange (III.35):

$$\frac{\rho_1}{\rho_0} = \frac{U_s}{U_s - u_p} . \tag{III.45}$$

From (III.37), we get

$$P_1 - P_0 = \rho_0 U_s u_p . \tag{III.46}$$

Now, note that (III.43) is unaffected by the change in reference frame. However, a convenient consequence of this change is to make use of (III.41) and (III.44) to get

$$E_1 - E_0 = P_0 V_0 \frac{u_p}{U_s} + \frac{1}{2} u_p^2 . \tag{III.47}$$

In the typical experimental case, where $P_0 \approx 0$, this is simply

$$E_1 - E_0 = \frac{1}{2} u_p^2 . \tag{III.49}$$

Effective Compressive Modulus

An important note is that, although the shock wave obeys the wave equation, the wave speed is not that of a sound wave, because the sound speed was derived under the assumption of infinitesimal pressure and density change. Because the changes are finite (and can be large), the effective modulus that would be used in the wave speed formula is different and is, in fact, dependent on the amplitude of the shock wave. Going back to the wave equation and the calculation of the wave speed, we see that the effective modulus is going to be

$$K = \rho_0 U_s^2 , \tag{III.50}$$

where U_s is the speed of the shock wave. An interesting problem is that, ideally, the jump from the undisturbed state to the final state in a shock wave is instantaneous, so (III.50) is the only value of the modulus for the entire compression process. Now, let's recast this in terms of V by using

$$\left.\frac{d\rho}{dV}\right|_{V_0} = -\rho_0^2 : \qquad \text{(III.51)}$$

$$\frac{\partial P}{\partial V} = -\rho_0^2 U_s^2 \qquad . \qquad \text{(III.52)}$$

If we look at the Rankine-Hugoniot equations (III.45) and (III.46), we find that the entire compression, from start to finish, has

$$\frac{\Delta P}{\Delta V} = -\frac{\rho_0 U_s u_p}{u_p \big/ \rho_0 U_s} = -\rho_0^2 U_s^2 , \qquad \text{(III.53)}$$

which is exactly the same expression as in (III.52), even though we are now dealing with a finite quantity. Thus the actual path taken by the material during the shock compression process is a straight line in *P*-*V* space. This straight line is called the Rayleigh line.

Shock Wave Stability.

Shock waves do not always form and, if formed may be unstable and break up. One condition for stability is that the flow behind the shock wave be subsonic. To see why this is, suppose we consider the shock wave in a coordinate system where the shock wave is stationary. Then we see material approaching from the right at velocity U_s and leaving to the left at velocity U_s-u_p, as if Figure III-2. If U_s-$u_p > c$, then any disturbance at the shock front will propagate away from the shock front and the shock wave will break up. Thus, a necessary condition for shock wave stability is

$$U_s \le u_p + c \qquad . \qquad \text{(III.54)}$$

The shock Hugoniot.

Because the shock wave properties of a material are ultimately controlled by the equation of state (at low pressures also by the mechanical strength of the material), the shock wave response is a property of the material. It is useful to identify a curve that is defined as the locus of end states of the shock process, which is known as the Hugoniot curve. The Hugoniot curve can be cast into any number of spaces, but the most common are shock wave speed versus particle speed, both measured in the rest frame of the unshocked material (U_s-u_p); pressure versus specific volume (*P*-*V*); and pressure versus particle speed, again with the speed measured in the rest frame of the unshocked material (*P*-u_p). A common mistake is to think of the Hugoniot curve as a thermodynamic path, but it isn't. It is simply the curve defined by the end states of the shock process for progressively stronger shocks, all starting from the same initial condition. If the initial

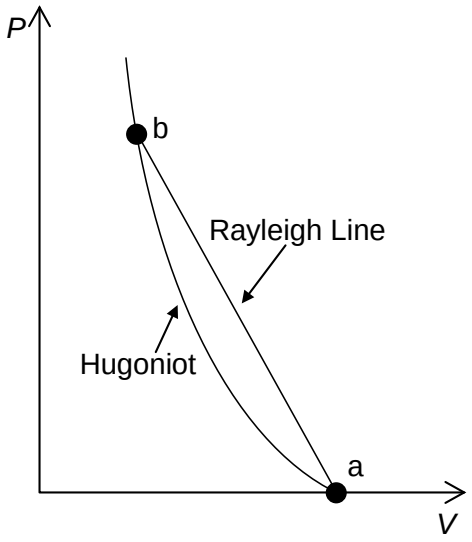

Figure III-3. Simple concave upward Hugoniot curve in P-V space, with a Rayleigh line connecting the starting and ending states (a and b) for a shock wave.

condition is (effectively) zero pressure and 298 K in the full-density stable phase at that condition, then the Hugoniot is known as the principal Hugoniot.

For a simple material without any complicated behavior, the *P*-*V* Hugoniot typically looks like the curve in figure III-3, i.e., a simple concave-upward line. We know that the actual path taken by the material is the Rayleigh line. The negative slope of the Rayleigh line, from equation (III.53) is the square of the shock impedance, $\rho_0 U_s$.

If the material has finite strength, the yield strength of the material is never reached at low stresses and the wave is an elastic wave. It may run faster than the longitudinal sound speed if the stress dependence of the elastic constants gives the modulus *m* in equation (III.28) a positive derivative with stress, resulting in an elastic shock. Typically, in such cases, the speed is close to the sound speed because the stress is low. If the final stress state is above the yield strength, the material will yield, with greater volume strain. The Hugoniot at that point will change slope as seen at point b in figure III-4. The longitudinal stress at that point is known as the Hugoniot elastic limit, or HEL. However, the change in slope means that the Rayleigh line required to get to the region of the Hugoniot just above point b has a smaller slope and, hence, a lower wave speed. This violates the basic condition for a shock wave, namely that the higher pressure portions of the wave travel faster than the lower pressure portions. Such a wave cannot exist and the only allowed path is actually two waves, with the first taking the material to the yield point and the second taking the material to the final state, as shown by the path a-b-c in figure III-3. Now, what about path a-d in figure III-4, where the Rayleigh line falls above the lower portion of the Hugoniot? In such a case, the wave speed required by the slope of the Rayleigh line is higher than that for any lower portion of the Hugoniot and there is again only a single wave taking the material all the way to the final state. We say in such a case that the elastic wave is overdriven. The stress of the final shock wave required to overdrive the elastic wave is usually substantially higher than the value of the HEL. The wave profiles associated with the paths shown in figure III-3 are given in figure III-5.

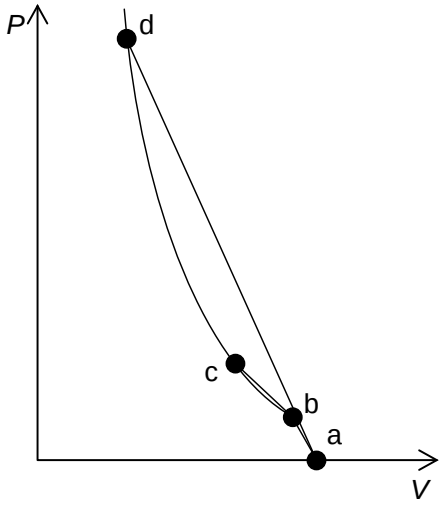

Figure III-4. Hugoniot curve of a material with finite strength, showing compression paths involving two waves (a-b-c) and one wave (a-d).

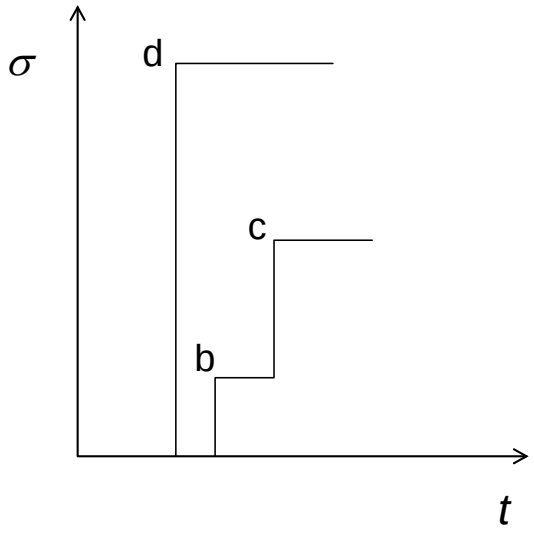

Figure III-5. Stress histories of compression paths shown in Figure III-4. The path taking the material to states b and c require two separate waves, with two jumps in stress separated in time, while the path to state d requires only a single wave.

A similar behavior can be seen when the material undergoes a pressure-induced phase transition to a lower volume phase, although the details are different. In fact, it is possible to have multiple upward convex regions of a Hugoniot, so that to reach some parts of the Hugoniot may require three or even more individual waves. A particularly good example is the mineral calcite ($CaCO_3$), which can have as many as four individual waves required to reach portions of the Hugoniot.

An empirical observation is that the Hugoniot curves of many materials are linear, or piecewise linear, within the resolution of the data, in the U_s-u_p plane:

$$U_s = C_0 + su_p \quad . \tag{III.55}$$

We noted this behavior earlier when we discussed the shock wave equation of state. This provides a convenient means of describing the Hugoniot curve of a material with only two parameters plus the initial density.

The Hugoniot elastic limit.

As noted before, at sufficiently low pressures, materials that show finite strength will exhibit an elastic wave that precedes the shock wave. This elastic precursor has a stress known as the Hugoniot elastic limit. The longitudinal stress of this wave is, very simply, given by

$$\sigma_x = \rho_{00} C_{el} u_p \quad . \tag{III.64}$$

where C_{el} is the elastic wave speed, which might not be quite the same as the sound speed you would measure in the material with an ultrasonic apparatus.

IV. CALCULATIONS INVOLVING SHOCK WAVES

Reference Frames.

For many people, reference frames in shock wave work are a source of confusion, frustration, and error. A number of published papers, particularly those dealing with complicated shock geometries, are invalidated by the fact that calculations were done in the wrong reference frames. There are several different reference frames that you might encounter. Each has its own applications and some are used more than others.

All of the different reference frames are important for the appropriate applications. The danger that most researchers face is in using data referenced to one frame in equations that were developed in a different frame, thus producing erroneous results. Even more confusing is that, in many experiments, calculations are best done using equations that require the use of two different reference frames. In such cases, the equations are chosen for the most simple implementation and, often, to minimize difficulties in propagation of errors.

Eulerian reference frames.

The reference frames we are accustomed to, where some point (the origin) is fixed relative to some reference, with constant lengthscales, are known as Eulerian reference frames. When we use a ruler to measure an object, the dimension we measure is an Eulerian dimension. If we measure the position of that object relative to some fixed point in the room, then we have determined the position in an Eulerian reference frame. If the object is moving, and we use the same ruler and a stopwatch to determine the velocity of the object, the velocity we obtain is an Eulerian velocity. The most important characteristic of Eulerian reference frames is that the ruler we use doesn't change. Hence, if we measure the length of a relaxed spring and then stretch the spring, we use the same ruler, to measure it again. The fact the object (in this case, the spring) has changed position or dimension does not affect the reference frame.

<u>Laboratory Frame.</u>

The laboratory reference frame is an Eulerian inertial (sometimes called Newtonian) frame in that it is tied to a set spatial scale and the reference frame has no fictitious forces due to acceleration. The laboratory frame, or lab frame, is the easiest to comprehend, simply because it is the frame we typically use in the course of our everyday lives. The important aspect of the laboratory frame is that, typically, it is the rest frame for the instrumentation used in experiments, so that, before an experiment begins (usually), all objects in the entire system making up the experiment are at rest. Hence, experimental data are typically referenced to the laboratory frame. Note that the origin of the frame, i.e., the position designated as "zero," is completely arbitrary and may be chosen for convenience. However, when position is important, it is vital that the same origin be used throughout the experiment.

<u>Shock Wave Reference Frame.</u>

Another Eulerian reference frame is the frame in which the shock wave is at rest and, usually, is at the origin. An important assumption that is made is that the shock wave is steady, so that the reference frame is not accelerating and the reference frame is therefore

inertial. This reference frame is also Eulerian, so that the lengthscale is constant. As a result, transformation from the laboratory frame to the shock wave frame is accomplished by a Galilean transformation. The most important application of the shock wave reference frame is in the derivation of the shock wave jump conditions.

Shocked material reference frame.

The last Eulerian reference frame we consider is one in which the material emerging from a shock wave is at rest. The Eulerian nature of the reference frame means that the lengthscale is constant. Also, the assumption is made that the shock wave is steady, so that the reference frame is also an inertial frame. This reference frame is important in situations where more than one wave traverses the material during an experiment, so that the ending state of one wave provides the initial condition for the next wave.

Lagrangian reference frames.

In our discussion above, we used the example of a spring. Now suppose that the spring itself is the ruler. As the spring is stretched or compressed, the ruler changes dimension by the same amount such that the measured length of the spring never changes and the position of a given coil in the spring is constant. This is the fundamental characteristic of a Lagrangian reference frame, in which the material is always at the same position. In such a reference frame, every atom is at a fixed position. The material always has a velocity of zero in the Lagrangian reference frame. This is inherently a noninertial reference frame. In the typical experiment, the material frame initially coincides with the laboratory frame, possibly except for the position of the origin, but the two begin to diverge at the onset of the experiment. The Lagrangian reference frame is particularly useful in experimental data analysis because experimental measurements are usually made at specific positions, such as interfaces, that change position and velocity in other reference frames but stay constant in the Lagrangian reference frame. Since the material velocity is always zero, only waves have a velocity in the Lagrangian reference frame.

A note on wave speeds.

One thing to be careful of when dealing with waves is that characteristic wave speeds that are a property of the material, such as the speed of sound, only have meaning when measured relative to the rest frame of the material. However, such speeds may be expressed as either Eulerian or Lagrangian values.

Position-time diagrams.

A common tool in shock wave work is the position-time diagram, more often called the *x-t* diagram. Actually, there are two separate types of such diagrams and different researchers use different terminologies for these diagrams. The most common diagram is the Lagrangian position-time diagram, which is what is most commonly meant by *x-t* diagram. This diagram, an example of which is in figure IV-1(a), shows the positions of interfaces and waves. Since this is a Lagrangian diagram, material interfaces are at a constant position and only the waves propagate. Some researchers color-code this diagram with different colors to distinguish interfaces, fractures, compressional and

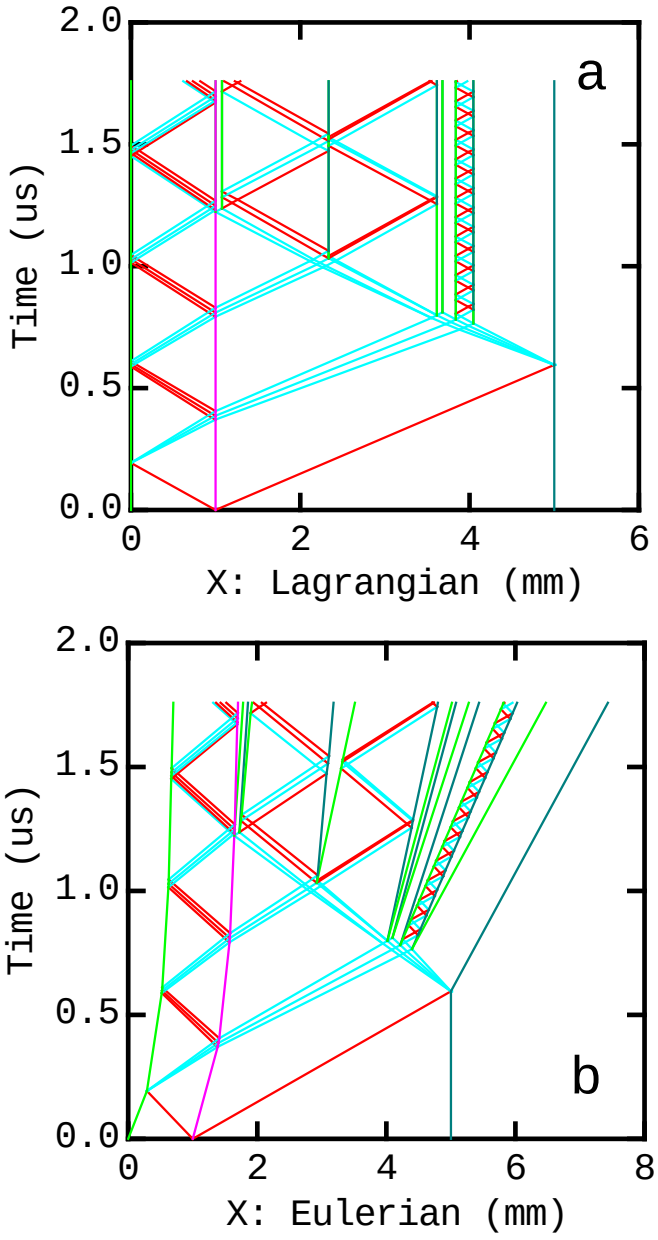

Figure IV-1. Position-time diagrams of impact of a 1 mm stainless steel 316 plate onto a 3 mm aluminum 6061 plate at 1.5 km/s. Red lines indicate compressional waves, while light blue lines indicate dilatational waves. Green lines indicate left-facing free surfaces, dark blue lines indicate right-facing free surfaces, and pink or purple lines indicate material interfaces. (a) Lagrangian diagram. (b) Eulerian diagram.

dilatational waves. This type of diagram is particularly useful in visualizing wave interactions. Because measured wave speeds in experiments are usually presented as the Lagrangian values, this is often the most relevant diagram for discussing such measurements. The static nature of interfaces also makes this a particularly easy diagram to draw and makes it conceptually simple.

The position-time diagram can also be drawn with Eulerian coordinates, as in Figure IV-1(b). This is sometimes called a *y*-*t* diagram, but is just as often called an *x*-*t* diagram (there are also some researchers who call the Lagrangian diagram the *y*-*t* diagram). The

Eulerian diagram is most important when one needs to visualize material interface motion.

Calculating Hugoniot points from an EOS.

In some applications, one may have an equation of state (either partial or complete) in hand and want to calculate the Hugoniot of the material. The calculation is straightforward and is based on the requirement that the internal energy of the Hugoniot state be that achieved by adiabatic compression along the Raleigh line.

Consider an equation of state for some phase ϕ_H consisting of a reference condition, (P_0,T_0), a reference compression curve, a Mie-Grüneisen thermal component, and an energy of transition, E_{tr}, required to get from the initial condition of the experiment, $(P_{00},T_{00},\phi_{00})$, to the reference condition and phase of interest. Remember that the change in internal energy given by the Rankine-Hugoniot conservation equations can be written as

$$E_H - E_{00} = \frac{1}{2}\left(P_H + P_{00}\right)\left(V_{00} - V_H\right) \tag{IV.1}$$

where the subscripts 00 and *H* refer to the initial condition and the Hugoniot state, respectively. Now, we require that the energy change using the EOS be equal to that given by (IV.1):

$$E_{tr} + E_c + V_H \int_{P_c}^{P_H} \frac{dP}{\gamma} = \frac{1}{2}\left(P_H + P_{00}\right)\left(V_{00} - V_H\right) \tag{IV.2}$$

where

$$E_c = \int_{V_0}^{V_H} \left(\alpha K_T T - P\right)_{T=T_0} dV \quad , \quad E_c = -\int_{V_0}^{V_H} P dV \tag{IV.3}$$

for a reference isotherm or isentrope, respectively. In (IV.2), the first term on the left hand side is the energy taken up by getting to the reference state of the phase of interest, the second term is the energy of compression along the reference compression curve, and the third term is the energy due to isochoric heating at the Hugoniot volume to get from the reference compression curve to the Hugoniot. P_c is the pressure on the reference compression curve at volume V_H.

Equation (IV.3) is not, in general, analytic, but some simple approximations commonly used in many EOS fits make it both analytic and simple to arrange to obtain

the value of P_H for a given Hugoniot volume V_H. First, if we take the common assumption that the Grüneisen parameter explicitly depends only on the volume (or density), then (IV.3) becomes

$$E_{tr} + E_c + \frac{V_H}{\gamma_H}\left(P_H - P_c\right) = \frac{1}{2}\left(P_H + P_{00}\right)\left(V_{00} - V_H\right), \tag{IV.4}$$

where γ_H is the value of γ at volume V_H. This expression is then rearranged to give

$$P_H = \frac{E_{tr} + E_c - \frac{V_{00} - V_H}{2} P_{00} - \frac{V_H}{\gamma_H} P_c}{\frac{V_{00} - V_H}{2} - \frac{V_H}{\gamma_H}}. \tag{IV.5}$$

Equilibrium Hugoniot temperature.

Normally, temperature is not known in shock wave experiments, but it can be calculated based on certain assumptions. The most important assumption is that the material is in thermal equilibrium and must be in a single phase. The validity of this assumption is open to debate, but it can be useful. Starting at the reference temperature, T_0, if we are using an isotherm as the reference compression curve, then we simply require equality between the two integrals:

$$\int_{T_0}^{T_H} C_V dT = V_H \int_{P_c}^{P_H} \frac{dP}{\gamma}, \tag{IV.6}$$

with the integrals being evaluated on the V_H isochore. If the reference compression curve is an isentrope, then we must do the calculation in two steps, the first being the temperature rise on the isentrope:

$$T_S = T_0 \exp\left(-\int_{V_0}^{V_H} \frac{\gamma}{V} dV \right). \tag{IV.7}$$

The second step is to perform the calculation in (IV.6), replacing T_0 with T_S:

$$\int_{T_S}^{T_H} C_V dT = V_H \int_{P_c}^{P_H} \frac{dP}{\gamma}. \tag{IV.8}$$

As noted earlier, this is the equilibrium temperature in a single phase. This is not what you will get if the Hugoniot state is a mixed phase or if the shock process does not deposit energy uniformly. The latter is often the case when the sample is initially porous

or when the sample undergoes nonuniform deformation, such as localized shearing, during compression.

Impedance matching.

One of the most important concepts in practical shock wave physics is impedance matching. This concept defines how a mechanical wave, such as a shock wave, is transmitted and/or reflected at an interface between two materials. It is a bit more than that, however, in that it is really a description of how momentum is transported across an interface, whether that momentum arrives in the form of a wave or as momentum carried by an impactor (with the interface being the interface of the impactor with the target). The basic concept here is that of mechanical impedance *z*, which is effectively defined by the initial density multiplied by the wave speed:

$$z = \rho c \quad , \tag{IV.9}$$

In the case of an acoustic wave, *c* is the longitudinal sound speed. In the case of a shock wave, *c* is the shock wave speed, U_s.

Impedance matching is based on the requirement that both normal stress and normal material velocity component be continuous across an interface. These continuity requirements are easy to understand by considering the consequences of violating them. In the first case, if the material velocity normal to the interface is not continuous, then the materials will either separate (target faster than the projectile) or will interpenetrate (projectile faster than the target). Obviously, the latter is unphysical, while the former cannot occur unless the normal stress is exactly zero. A discontinuity in the stress implies a force that is accelerating the interface. However, there is no source for such a force, so such a discontinuity cannot exist. Now that we have established the continuity requirements, relating two of the shocked state variables between the two materials, the requirement that the shocked states also satisfy the Rankine-Hugoniot conditions with the specific Hugoniot parameters of the two materials makes the system fully constrained.

Impacts.

It is easiest to begin by considering what happens when an impactor impacts a target at some normal velocity u_d. The conditions attained in the impact are governed by the properties of the materials—specifically the shock Hugoniots—and the continuity requirements. To satisfy the continuity equation for material velocity, we must use a common reference frame for the both the impactor and the target. We choose to use the laboratory frame as our rest frame, in which case, we need to make a transformation on the material velocity of the impactor, so that u_p becomes u_d-$u_{p,t}$, where $u_{p,t}$ is the material velocity of the target. Thus, the velocity decrease in the impactor is exactly equal to the difference between the impact velocity and the particle velocity in the target. That said, and substituting the new expressions of the particle speed into the momentum equation (III.46), we require

$$\rho_{0,i} U_{s,i} \left(u_d - u_{p,t} \right) = \rho_{0,t} U_{s.t} u_{p,t} \quad . \tag{IV.10}$$

If the material Hugoniot curves can be expressed in the usual way as

$$U_s = C + su_p \quad , \tag{IV.11}$$

then the solution for the particle velocity in the target can be written as

$$u_{p,t} = \frac{-b - \sqrt{b^2 - 4ac}}{2a} \quad , \tag{IV.12}$$

where

$$a = s_i \rho_{0,i} - s_t \rho_{0,t} \quad , \tag{IV.13}$$

$$b = -C_i \rho_{0,i} - 2s_i \rho_{0,i} u_d - C_t \rho_{o,t} \quad , \tag{IV.14}$$

$$c = \rho_{0,i} u_d (C_i + s_i u_d) \quad . \tag{IV.15}$$

When the target Hugoniot is unknown, then the impact velocity is combined with a second measurement, usually of the shock velocity U_s in the target, to get

$$u_{p,t} = u_d + a' - \left[a'^2 + \frac{\rho_{0,t}}{\rho_{0,i}} \frac{U_s u_d}{s_i} \right]^{1/2} \quad , \tag{IV.16}$$

where

$$a' = \frac{1}{2s_i} \left(C_i + U_s \frac{\rho_{0,t}}{\rho_{0,i}} \right) . \tag{IV.17}$$

A special case arises if the impactor and the target are the same material, which requires that the value of $u_{p,t}$ must be $\frac{1}{2} u_d$. This second case is the situation of the traditional experiment to determine the shock Hugoniot of a material. It should be noted that the forms of the equations can result in large uncertainties for density in experiments with large compressions, even though the uncertainties in the impact and shock velocities might be considerably better than 1%.

The foregoing can often be made clearer by a graphical representation (Figure IV-2). If we imagine the target P-u_p Hugoniot beginning at rest in the laboratory reference frame, then the rest frame of the impactor is moving at the impact velocity, with the sense of direction reflected, so that the origin of the impactor Hugoniot is at a velocity equal to the impact velocity, u_d, with the Hugoniot extending in the negative velocity direction of the lab frame. In this construction, the conditions achieved by the impact are determined

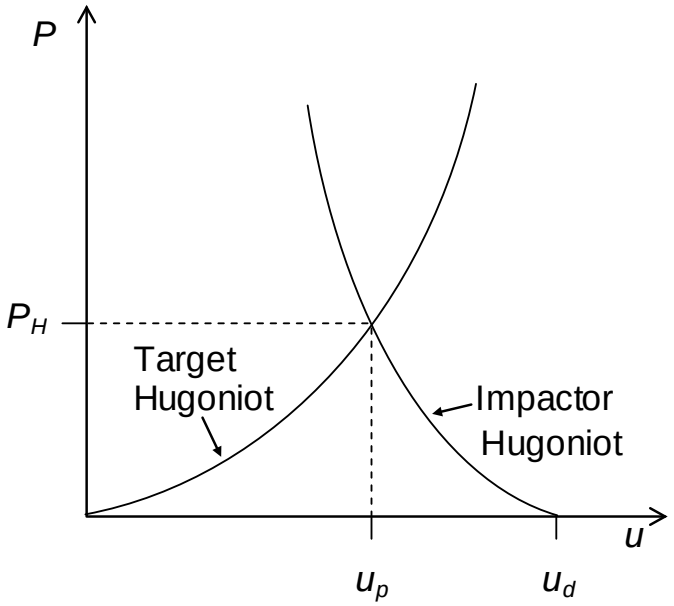

Figure IV-2. The state achieved in an impact is governed by the Hugoniot curves of the impactor and the target. The initial state of the impactor is at rest in a reference frame moving at the impact velocity relative to the laboratory frame, with the positive direction being opposite to that defined by the laboratory frame. The rest frame of the target is the laboratory frame.

by the intersection of the two Hugoniots. This graphical construction is a useful tool for understanding the evolution of a material state in complicated experiments. Release adiabats and reshock Hugoniots can also be represented on such a diagram, so that the evolution of materials as they interact with other materials at interfaces can be mapped out. Typically, this is done to obtain a good conceptual grasp of the interactions that are taking place so that the calculations to obtain accurate descriptions of the states achieved can be set up properly. To tie the graphical construction back to the concept of the material impedance, we note that the slope of the Hugoniot cast in P-u_p space is equal to the impedance. We say that a material has a higher impedance if the Hugoniot has a greater slope in P-u_p space.

Transmitted waves.

Although the most common (and most easily calculated) impedance matching problem is that of impact, many situations require the application of impedance matching to a wave that is transmitted to an interface inside a material. In that case, the continuity equations still apply, but the donor material through which the arriving wave has travelled is no longer at the principal Hugoniot starting state, so that the post-interaction state will fall either on a secondary Hugoniot anchored to the initial shock state or on a release adiabat anchored to the initial shock state. The impedance mismatch determines whether an impinging shock wave is reflected into the donor material as a release wave or as a reshock. If the acceptor material has a higher impedance than the donor, then the reflected wave will be a reshock and the transmitted wave will have a higher stress than the impinging wave. Similarly, a lower impedance in the acceptor material will result in

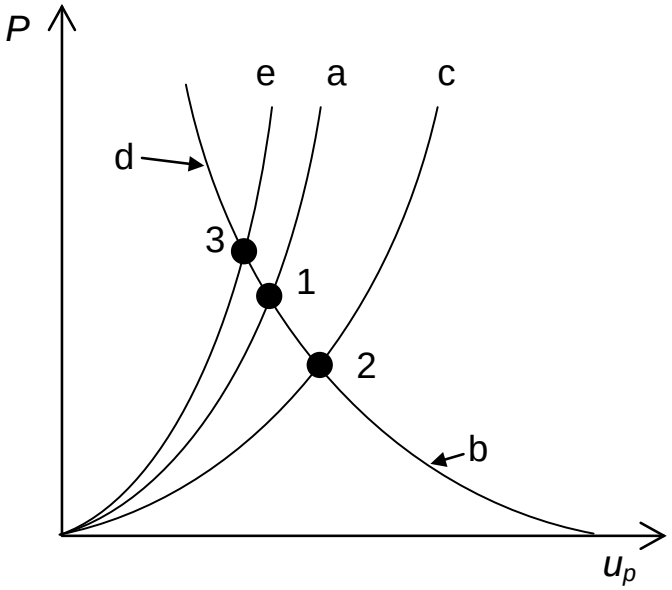

Figure IV-3. Impedance matching with a transmitted wave. The incident wave brings the donor material to state 1 on its principal Hugoniot (a). If the acceptor material has a lower impedance, then the donor will follow a release adiabat (b) to match state 2 on the acceptor Hugoniot (c). If the acceptor has a higher impedance, then the donor will be reshocked to state 3 at the intersection of its reshock Hugoniot with the principal Hugoniot of the acceptor (e).

the wave being reflected as a partial release and the transmitted wave will carry a lower stress than the impinging wave.

The graphical representation of this problem is shown in Figure IV-3. To do the calculation with complete fidelity requires a more complete EOS of the donor, because the final state for the reflected wave is not on the principal Hugoniot.

Wave interactions.

Interaction of two waves propagating through a single medium is also an impedance matching problem. Now, instead of two separate materials, the matching occurs with the same material that has been placed into two independent P-u_p states by two separate waves. Although the two waves are independent of one another, once they pass through each other, the stress and laboratory-frame (or any common reference frame) particle velocity is common between the two end states. The basic situation is shown in Figure IV-4. Because both waves have already placed the material at states on the principal Hugoniot, the passage of one wave through material already shocked by the other wave places the material at a state on a Hugoniot anchored at the first shock state on the principal Hugoniot, so one needs more than just the principal Hugoniot of the material to perform this calculation. Additionally, the material may not be at the same thermodynamic condition as that produced by the passage of the second wave through material already shocked by the first. The reason is that although the pressure (or stress) is the same, it is the laboratory frame in which the particle velocities must match. There is no requirement that the particle velocity be such that it gives the same density (and thus any other thermodynamic variable) in the two states. Hence, there will usually still effectively be an interface at which the actual *in situ* properties of the material change.

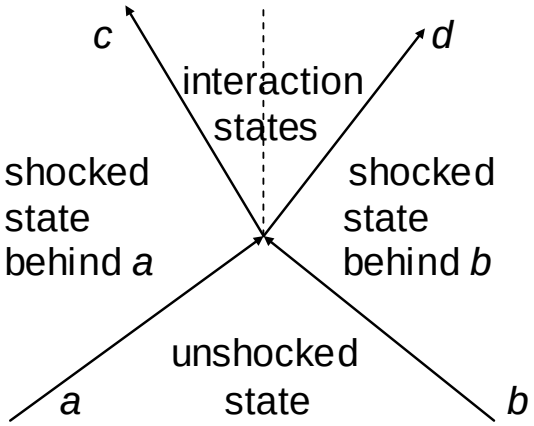

Figure IV-4. Trajectories of two shock waves interacting in a Lagrangian *x*-*t* diagram. both waves (the incoming waves, *a* and *b*) propagate into unshocked material. Interaction produces two outgoing waves, *c* and *d*. The interaction states, produced by the propagation of the outgoing waves through material already shocked by the incoming waves, must have the same pressure (or longitudinal stress) and particle velocity in the laboratory frame, but can have different thermodynamic states. The two thermodynamic states interface at the vertical dashed line.

The only exception to this is when the two waves that interact have the same amplitude prior to interacting.

V. EXPERIMENT DESIGN AND TECHNIQUES

A wide variety of experimental techniques are used in shock wave studies. New techniques are continually being developed and existing techniques are continually being refined and improved, so no static document such as these lecture notes can hope to be either comprehensive or current. Hence, the approach taken here is to present a survey of historical and current techniques, but focus on those few that have long been the standards for work in this field. In particular, shortcomings with historical methods that researchers using older data should be aware of are noted.

Generating Shock Waves.

Ultimately, shock waves are generated experimentally by applying a sudden pressure increase to some part of a target, so that the disturbance propagates away as a wave that either starts out as a shock wave or becomes one as it propagates. The important feature of the process is its dynamic nature, as opposed to the essentially static nature of compression techniques using diamond anvil cells, piston-cylinder presses, and multianvil presses, for instance.

Traditionally, the pressure increase has been provided either by the detonation of a high explosive charge or the impact of a projectile. However, any means of depositing a substantial amount of energy in a short time, including laser irradiation, irradiation by ionizing or neutron radiation, or ramp loading by application of the pressure associated with a changing magnetic field can also be used. The primary feature of shock waves generated for fundamental material studies is that they are planar in the region where the material behavior is being probed, so that the one-dimensional spatial derivatives can be used during analysis.

Explosives.

The original approach to shock wave generation was to use high explosives to drive the experiment. In some ways, this is a little circular, since the detonation front in a high explosive is itself a kind of shock wave. In the earliest experimental designs, a planar detonation wave was generated by a plane wave lens and either propagated directly into the target or though a booster charge into the target (Figure V-1). The explosive was in direct contact with the target and the detonation wave propagated into the target as a shock wave. This technique was used extensively until the mid-late 1960's and is still used for some purposes, especially for understanding material response in situations where the material would be driven directly by an explosive charge anyway. The primary shortcoming of the in-contact explosive drive for many applications is that the detonation products expand from the back of the charge, allowing the pressure to drop so that the shock wave in the target decays as it propagates. This complicates data analysis. The advantage of the technique is that explosives are relatively easy to use, charges can be made quite large, and not much highly expensive apparatus is involved in the experiments.

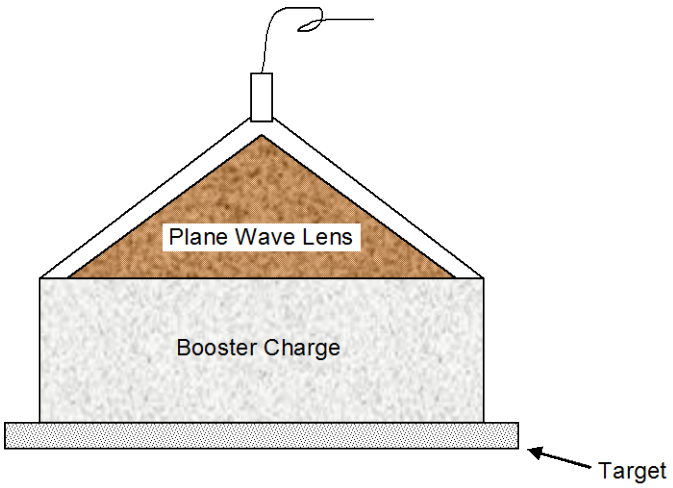

Figure V-1. In-contact explosive drive system.

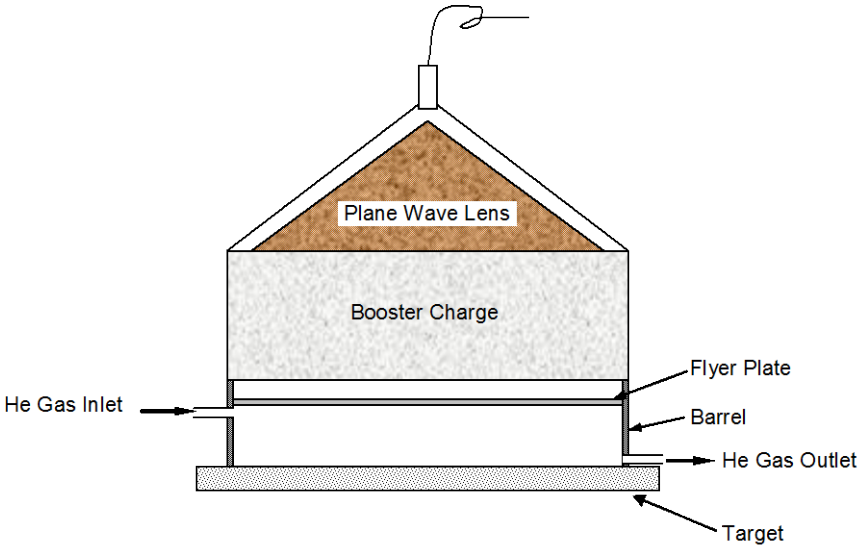

Figure V-2. Explosively driven flyer system.

The problem of shock wave decay can be solved if the explosive charge is used to drive a flat plate that impacts the target, generating a shock wave (Figure V-2). Because the impact generates the shock wave and the same amplitude shock propagates into the plate as well as the target, the shock wave does not decay but propagates with a constant amplitude. The impacting plate, called the flyer plate or the driver plate, can be heated by the acceleration process, thus changing its properties in a poorly known way, but heating is minimized by allowing the detonation products of the driving charge to expand across a gap between the driving charge and the flyer plate to "gently" accelerate the plate without subjecting it to strong shocks. This is the preferred method for driving

experiments when equation of state data are the goal of the experiments and the samples are too large for use with other driving techniques. Depending on details, the impacting plate can be driven up to velocities ranging from ~ 2 km/s to ~ 8 km/s. In addition to possible heating of the flyer plate, the primary uncertainty in experiments driven is this way is the accuracy with which the impact speed of the plate is known. Until very recently, accuracies better than 2% were very difficult to achieve and the accuracy was often significantly worse. An interesting modification of this technique eliminates the flyer plate, so that the expanding detonation products impinge directly on the target, subjecting it to a ramp loading rather than a shock wave.

Guns.

Starting in the mid- to late 1960's guns became the means of choice for driving planar shock waves. Guns are still considered the gold standard for driving shock waves for most types of measurements. A variety of gun designs exist, but all make use of a nominally cylindrical sabot that carries a planar flyer plate in the nose of the projectile. There are two basic types of guns. Single-stage guns (Figure V-3) are what you normally think of when you think of a gun. A projectile is accelerated down a barrel by a working fluid, which can be either compressed gas or the combustion products of a propellant charge. Depending on details, such guns produce projectile speeds ranging from a few 10's of m/s up to ~2.7 km/s. Two-stage guns (Figure V-4) are significantly more complicated and are designed for extremely high performance. The basic idea of the two stage gun is to compress the working fluid, typically H_2 or, less frequently, He (hence the use of the name light gas gun by some researchers) to a much higher pressure than normal and to allow it to expand rapidly, driving the projectile. To accomplish this, a piston is driven, usually by a propellant charge, but sometimes by a compressed gas, down a large diameter pump tube containing the working fluid, compressing it. The pump tube is separated from the barrel, called the launch tube, by a tapered section known as the central breech. The purpose of the taper is to stop the piston by allowing it to trade kinetic energy for plastic deformation as it extrudes itself into the taper. The downstream end of the taper is separated from the launch tube by one or more metal discs designed to rupture at a predetermined pressure. The projectile sits immediately downstream of the burst disc, also known as the rupture diaphragm. Upon rupture of the disc, the working fluid expands, accelerating the projectile down the launch tube. Because of the high density and temperature of the working fluid under these conditions, the sounds speed in the working fluid is very high, allowing higher limiting velocities. Additionally, a shock wave is set up in the fluid that reverberates between the projectile and the piston, providing additional acceleration over what would be provided by the gas pressure alone. In single stage guns, the primary concern is that the projectile tilt be acceptable. Typically, the tilt ranges from near zero to 2 mrad. In two stage guns, the acceleration process places sufficient stress on the projectile that the flyer plate may also undergo deformation. This deformation is usually assumed to be axisymmetric and parabolic or spherical, so it is common to talk about impactor bow in addition to projectile tilt.

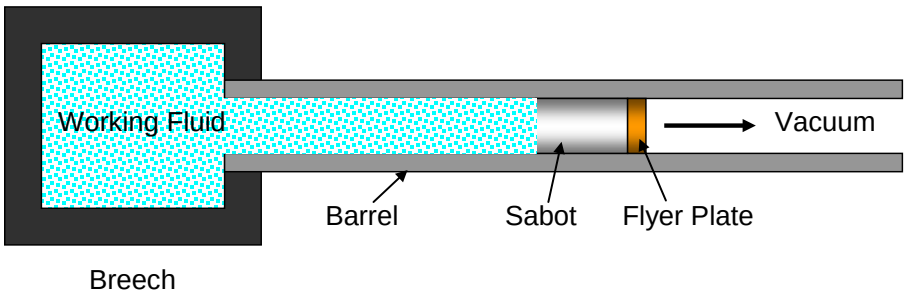

Figure V-3. Single-stage gun for accelerating projeciles.

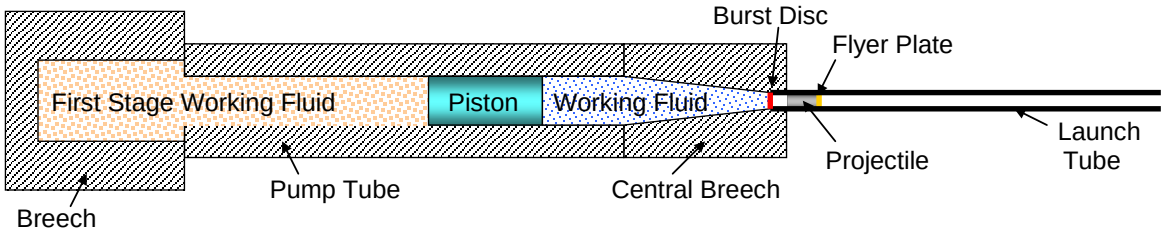

Figure V-4. Two-stage gun for accelerating projectiles to very high velocities.

Radiation-Driven Methods.

A relatively recent development has been the advent of laser systems that can deposit sufficient power in a spatially uniform manner over an extended area that such lasers can be used to drive shock wave experiments. A variety of methods for accomplishing this exist, but they tend to fall into one of three types: direct drive, hohlraum drive, and flyer launch. The most common approach with lower power lasers is to deposit the laser radiation in an ablator that is in contact with the target (Figure V-5). The ablator may be either confined by a window or unconfined. Confinement is limited by the window's ability to remain transparent to the laser radiation. The high pressure generated by the essentially instantaneous deposition of the laser energy into the ablator is transmitted into the target. Depending on the history of the deposition, the resulting wave may be either a ramp rave or a shock wave. The target, as used in this discussion, may be the sample under study or it can also be a flyer plate that is accelerated against a sample to produce a shock wave by impact onto the sample.

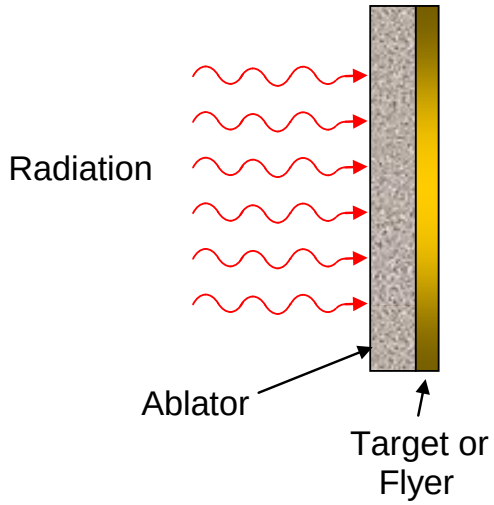

Figure V-5. Direct drive laser-driven experiment.

Holhraum drive is the preferred drive for very high energy multi-beam systems such as the National Ignition Facility. In this approach, laser energy is deposited on the inner walls of a very small hohlraum, heating the walls to sufficient temperatures to generate soft x-rays. The x-rays, in turn, are absorbed by an ablator material to produce high pressures that are, most often, transmitted directly into a sample through a shield (required to keep the x-rays from reaching the sample and heating it prior to wave arrival).

Although we have concentrated on laser radiation here, any radiation that can be absorbed by a material can be used to generate high pressures. This was actually used during the days of underground testing, when the radiation flux from a nuclear explosion was used to generate extremely high shock pressures.

Magnetic Drive.

Another recent development is the ability to use large current sources to magnetically drive shock wave experiments. Magnetic drive makes use of the magnetic pressure caused by a magnetic field produced by and interacting with a rapidly rising current. A typical drive geometry is shown in Figure V-6. In this configuration, a current passes through the cathode, and is short-circuited to the anode. The magnetic field **B** produced by the current loop interacts with the current **I** to produce a force per unit length of current path on the conductors through the relation

$$\mathbf{F} = \mathbf{I} \times \mathbf{B} \qquad \text{(V.1)}$$

This force accelerates the conductors. This can be used in two different ways. In the first way, because the current starts at zero and rises over a finite time, the force on the inner surface of the loop rises, producing a ramp compression history. This can be used to shocklessly compress the conductors and other materials placed in contact with the conductors. In the second way, the conductors are accelerated, becoming impactors that

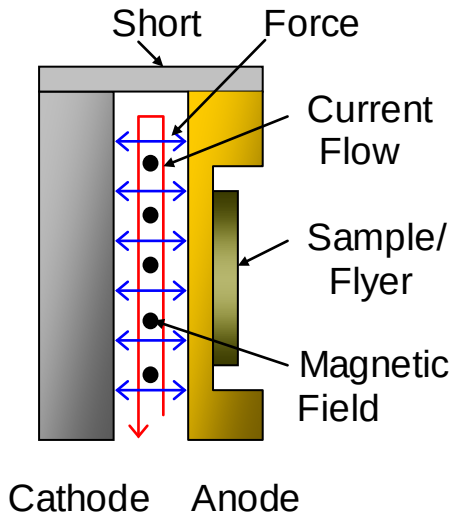

Figure V-6. Configuration of a magnetically-driven experiment.

can be used to generate very high shock pressures. The Z-Machine at Sandia can use this technique to accelerate plates to velocities in excess of 30 km/s.

Diagnostics for experiments.

Once we have the ability to drive shock waves in an experiment, we must then be able to make useful measurements on the experiment, which means fielding some suite of diagnostic instrumentation. There is now a wide variety of different measurement techniques. Often, the choice of what particular technique to use comes down the cost and required resolution. In a few cases, only one particular type of measurement can give the required information. Here, we will discuss some of the various techniques that are available, although only a relatively restricted set of these is widely used on a regular basis. This listing is not intended to be exhaustive, both because there are many types of measurements that might be made and because new techniques are constantly being developed. The ones discussed in depth are those that are considered mature and tend to be used by researchers at more than one institution.

Mechanical response measurements.

The most mature measurements are those that measure the mechanical behavior of the material understudy, usually by detecting arrivals of shock waves or measuring changes in velocity or position with time. These techniques include optical and electrical techniques and, because of their mature and fundamental nature, some subset of them is found in every shock wave laboratory in the world.

Pins.

A variety of pins can be used to detect shock wave arrival at a surface and also to determine impact times, and, by detecting passage of a projectile at different positions,

measure projectile speed. Two particular types, the electrical shorting pin and the piezoelectric pin, are very widely used (Figure V-7). Shorting pins exist in two different configurations. The simplest is a bare wire conductor that is held at a potential relative to ground. Contact with a grounded conductor, which could be the projectile (through simultaneous contact with a grounded pin) or a conducting target, or electrical breakdown of an insulator between the pin and a grounded conductor, which can be caused by passage of a shock wave, allows current to flow, which is detected by appropriate recording electronics. Details of the circuits involved depend on the purpose of the measurement and the time resolution needed. The more complicated type of pin is the coaxial self-shorting pin. In this case, a central conductor is held at a potential relative to a surrounding grounded shield. An end cap over the pin is impacted, driving it against the central conductor and shorting the pin, allowing the current to flow. In the most sensitive pins of this type, the cap is actually already in contact with the conductor, but an anodized layer insulates it from the conductor. Passage of a shock wave through the cap causes either electrical or mechanical breakdown of the anodized layer, allowing the pin to short. A common feature of shorting pins is the requirement for a power supply and RLC circuit with a tailored time constant.

A different type of electrical pin that is widely used is a piezoelectric pin. In this type of pin, a piezoelectric crystal or ceramic is held in contact with the central and outer conductors of a coaxial pin, with electrical contact made in such a way that the passage of a stress pulse through the piezoelectric material causes a momentary potential and current flow. Probably the most common piezoelectric material in this application is lead zirconium titanate (PZT). Such pins are often used to determine impact time of a projectile and as a source of a trigger signal for recording diagnostics.

A less commonly used pin is the optical pin, which generates an optical signal when shocked. There are two types of such pins that have been used. The first consists of an optical fiber with a glass microballoon at the end. Collapse of the microballoon causes a brief flash of light that is detected by fast optical detectors. The other type of pin consists

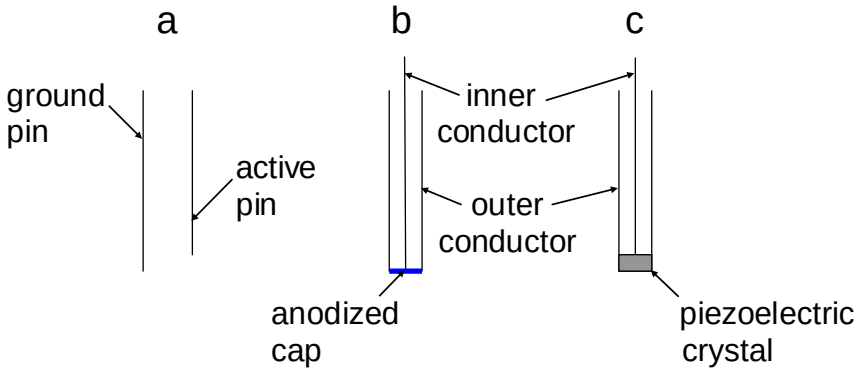

Figure V-7. Electrical pins commonly used to detect impact or shock arrival. (a) Bare shorting pin (shown with grounding pin); (b) Coaxial shorting pin; (c) Piezoelectric pin.
of an optical fiber with a polished and mirrored end. Movement of the mirrored surface is detected by one of the Doppler techniques described below.

Optical Shock Breakout.

Optical methods have been, for many years, the preferred alternative to pins for high-resolution determination of shock arrival times and, also, for obtaining free-surface velocities. There are two different approaches.

Flash gaps.

When a shock wave arrives at an interface, it is transmitted and reflected in accordance with the impedance matching principle discussed earlier. If the acceptor material is a gas, then it may be very strongly heated by the shock wave to a temperature sufficient to emit thermal radiation as visible light. This is the basic principle of the flash gap (Figure V-8), in which a thin gap (typically 50-100 μm), filled with a gas (usually argon or krypton) is placed against the surface of the donor material. The gap is usually formed by a groove machined into an acrylic block that serves as a window. Arrival of the shock wave results in a brief flash of light that can be recorded. Arrays of such gaps could be placed against a target having surfaces at different levels and on different materials and be viewed through slits by a streak camera (to be discussed below), allowing shock arrival timing to be obtained that is spatially resolved in one direction.

Mirror extinction.

A second approach to optically detecting arrival of a shock wave is to place a rear-surface mirror against the surface of a donor material (Figure V-9) and observing extinction of light reflected from that surface when the shock wave arrives. This approach has been used for many years in situations where space constraints preclude the use of flash gaps. One can also obtain position-time histories of free surfaces by inclining the mirror so that

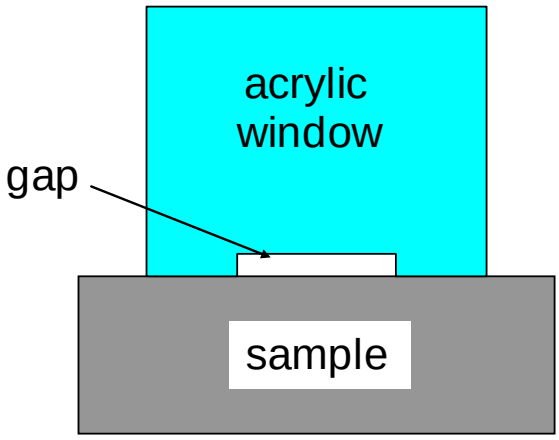

Figure V-8. Flash gap used to detect shock arrival at the surface of a sample.

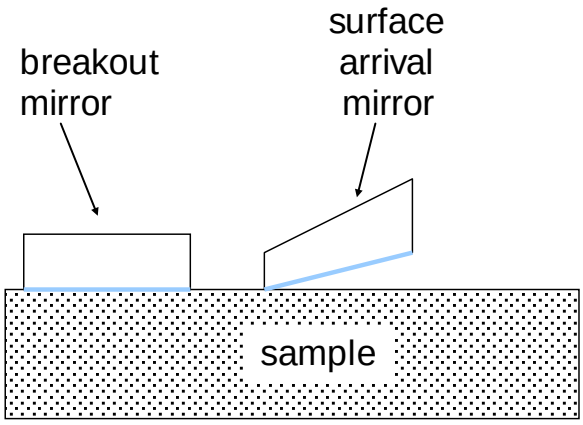

Figure V-9. Rear-surface mirrors used to detect shock arrival at the surface (breakout mirror) or arrival of a moving sample surface (surface arrival mirror)).

it extinguished progressively as the surface arrives at the mirror. The signals from mirrors are typically recorded using streak cameras.

Doppler velocimetry.

Doppler velocimetry has become one of the most common approaches to obtaining time-resolved velocity histories of interfaces when such interfaces can be accessed optically and can reflect light. Such interfaces can be either free surfaces or interfaces with transparent windows. The most common windows are 100-cut LiF, PMMA, z-cut sapphire, and z-cut quartz. There are two basic approaches to Doppler velocimetry.

VISAR.

VISAR (Velocity Interferometer System for Any Reflector) has been in use since the 1970's and has become the workhorse technique in many laboratories. In this technique, monochromatic light is reflected from the surface and enters a modified Michelson interferometer in which one leg of the interferometer has a known delay relative to the other. The delay is usually achieved by insertion of an etalon of known properties. Motion of the surface parallel to the optical path of the light being reflected from it induces a Doppler shift. In the VISAR, the light coming from the surface is split and sent through the two legs of the interferometer and then recombined. If there has been a change in the line of sight velocity, this results in a shift of the frequency of the delayed light relative to the undelayed light, so that recombination of the two causes modulation of the light intensity with a beat frequency, so that the velocity change can be determined. The advantage of this approach is that only the changes in velocity are determined, which are small enough that the beat frequency can be readily detected with relatively inexpensive recorders. Additionally, the sensitivity of the system can be adjusted by changing the amount of delay in the delay leg of the interferometer. The major shortcoming of VISAR is that it cannot properly deal with situations when more than one velocity is in the field of view.

PDV.

A relatively new technique that has been made possible by optical telecommunications technology is a technique that is known variously as photonic Doppler velocimetry (PDV), Heterodyne velocimetry (Het-v), or, simply, Doppler velocimetry. This is true Doppler velocimetry and simply mixes the shifted light reflected from the surface being probed with unshifted light from the source laser. The mixed signal, containing a modulation due to the heterodyne beat frequency, is recorded. The great advantages of this technique are that it is self-calibrating (it depends only on the frequency of the laser) and it can admit multiple velocities in the field of view. The biggest drawback is that it requires fast recorders that are substantially more expensive than those needed for VISAR. However, the cost is offset somewhat by much of the rest of the instrumentation being less expensive, since many of the components, including the laser, are standard items used in the telecommunications industry.

Gauges.

The major drawback of both pins and Doppler techniques is that they can only probe events at mechanically or optically accessible interfaces or surfaces. The most common approach to probing the interior of a sample is through the use of embedded gauges. Gauges consist of a transducer with lead wires or ribbons, typically contained in a flat package. In a typical gauged experiment, the sample is cut into two pieces and the two pieces glued back together with the gauge package between them (Figure V-10). Multiple gauge packages may be placed at different levels in the same target to provide information at different points, which is especially useful if the wave is not steady. There are several types of gauges in use.

Quartz gauges.

Quartz is a piezoelectric material and quartz crystals may be thinned sufficiently for use as a gauge. Passage of a stress wave a results in a potential across the gauge due to the piezoelectric effect. At sufficiently low stresses where the quartz remains elastic, the stress history may be followed as the potential changes in response to the changing stress amplitude. Such gauges must be carefully calibrated in order to obtain quantitatively accurate results.

Manganin gauges.

Manganin is the most commonly used gauge material. This alloy is peizoresistive, so that changes in the stress change the resistance of a gauge in a nominally linear fashion. Other materials, such as carbon, are also used, but manganin is by far the most common type of piezoresistive gauge.

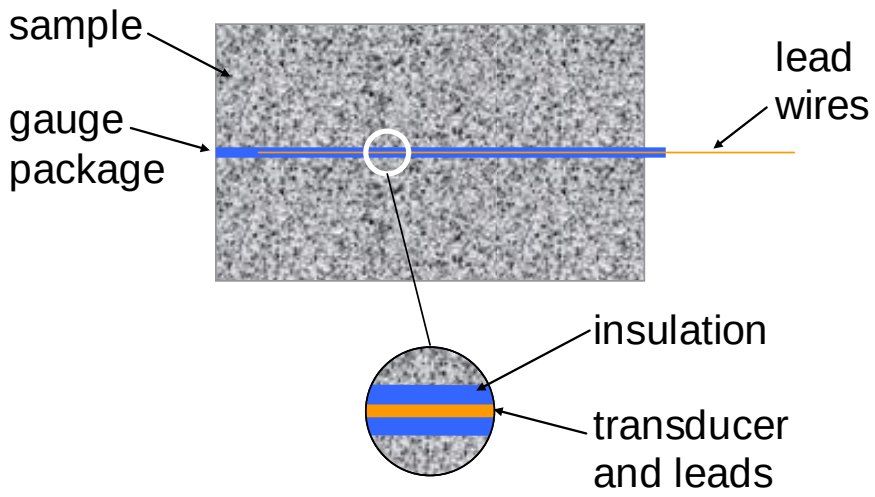

Figure V-10. Embedded gauges are placed into a sample that has been cut open and then glued back together with the gauge package between the two parts.

Induction gauges.

A special type of gauge is the inductance gauge. This gauge makes use of the fact that a conductor moving perpendicular to the direction of a magnetic field will develop a potential across it that is given by

$$V = \mathbf{l} \bullet (\mathbf{v} \times \mathbf{B}) \qquad \text{(V.2)}$$

where **l** is the vector describing the orientation and length of the conductor, **v** is the velocity vector of the conductor and **B** is the magnetic field vector. If the motion, field, and conductor are all mutually perpendicular, then the potential is equal to the product of the scalar length, speed, and field strength. These gauges require that there be an applied field of known strength and that the material in which the gauge is embedded be sufficiently insulating to not interfere with the field. Induction gauge packages have several advantages. It is common for a gauge package configuration similar to that shown in Figure V-11 to be placed inclined in the target. As the shock wave progresses through the sample, it accelerates the different elements of the shock tracker, allowing the speed of the wave to be followed as a function of time. Simultaneously, the velocity history at different levels in the target are measured by the single-element gauges, known as stirrup gauges, in the middle of the package. This is useful in situations in which the wave is not steady and its speed is therefore changing.

Radiography.

Another type of probe that can be used to investigate the interior of the target is radiography. The most common form is x-radiography. A variety of x-ray sources exist that can be used, from small flash-x-ray machines to large facilities like DARHT. The common feature is that they produce a very short, intense flash of x-rays. The x-rays are used to measure the path-length absorption between the x-ray source and an imaging

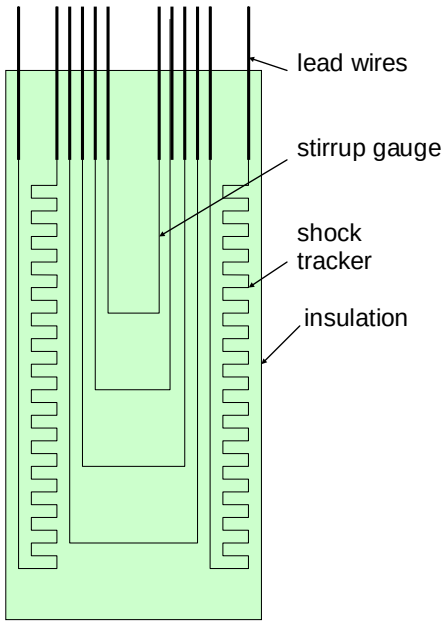

Figure V-11. Common type of induction gauge package, with stirrup gauges to measure velocity history and shock trackers to measure shock wave position as a function of time. In use, the magnetic field is oriented so that only the portions of the gauge element that are horizontal in this diagram produce an induced current.

detector. The resulting images allow the path-integrated densities to be obtained and can be used to image density changes within the target. Multiple flashes can be used to watch progression of an experiment. Multiple simultaneous flashes from different orientations can also be used to produce crude three-dimensional tomographic reconstructions. The main drawback with x-rays is that they cannot be focused, which limits resolution. An alternative to x-rays is protons. Because protons are charge particles, they can be focused, allowing a plane within the target to be imaged. Also, because the interactions of protons with the atoms in a material depend on the charge density in the atoms, protons are sensitive to composition. The Proton Radiography facility at LANSCE is capable of making up to 32 images in quick succession in a single experiment.

Non-Mechanical Measurements.

Pyrometry.

If one thinks back to our discussions of shock waves and equations of state, one thing that becomes apparent is that temperature does not appear in the Rankine-Hugoniot equations. This means that normal probes of mechanical response in shock waves will not provide information on the temperature. However, temperature is a vital factor in studies of phase transformations and material strength, so that one must either use a different set of probes that can measure temperature or resort to theoretical models to estimate temperature. There are several possible ways to measure temperature, but most are either so specialized or so slow that we will not consider them here. The one

technique we will address is optical and infrared pyrometry. A surface at temperature T will radiate energy as a function of frequency in accordance with the Planck blackbody function and spectral emissivity:

$$I(\nu,T) = \varepsilon(\nu)\frac{2h\nu^3}{c^2}\frac{1}{e^{h\nu/kT}-1} \tag{V.3}$$

where I is the power emitted per unit area per unit frequency per unit solid angle, ε is the spectral emissivity, h is Planck's constant, c is the speed of light in vacuum, ν is frequency, and T is temperature. It is more common to cast this equation into one dependent on wavelength, λ:

$$I'(\lambda,T) = \varepsilon(\lambda)\frac{2hc^2}{\lambda^5}\frac{1}{e^{hc/\lambda kT}-1} \tag{V.4}$$

where I' is the power emitted per unit area per unit wavelength per unit solid angle. Figure V-12 shows the appearance of this function at three different temperatures for a blackbody, which has an emissivity of 1.0. The resulting radiation can be measured as a function of wavelength, usually employing a multichannel pyrometer which has detectors, such as photomultiplier tuber or photodiodes, that view the incident radiation through narrow band-pass filters. Fitting the Planck function using some assumption about the emissivity allows the temperature of the emitting surface to be estimated. If, now, the surface is placed against a transparent window, then it can remain at high pressure after passage of a shock wave, allowing the temperature state at high pressure to be obtained. The major shortcoming of this technique is that, for opaque samples such as metals, this is a surface (interface) technique. The temperature is not as easily related to the bulk temperature of the interior of the sample as is the particle velocity. A thermal diffusion profile is set up between the sample and the window that depends on the thermal properties of the sample and window, neither of which is well-known. However, to the extent that the properties of the components are known, the bulk temperature in the sample can be obtained from the interface temperature. The second problem is that the state that one obtains from such a calculation is the reshocked or released state due to impedance matching against the window. This is related to the initial shock state, but, again, requires calculations with built-in assumptions.

The one exception to these difficulties is if the state of the sample after impedance matching against the window is in a mixed phase, lying on a phase boundary. In such a case, the thermal diffusion into the window, instead of setting up a temperature gradient in the sample, changes the relative abundances of the two phases, supplying the heat though latent heat of transformation. In such a case, the temperature observed is the phase transition temperature of the material at the pressure achieved by the impedance matching process. This is the only situation in which the interpretation of the measured temperature is relatively straightforward.

Another simplification exists if the sample being observed is transparent. In such cases, the emission from the shocked material behind the shock wave can be observed as the shock wave is propagating through the material. The primary difficulty here is that

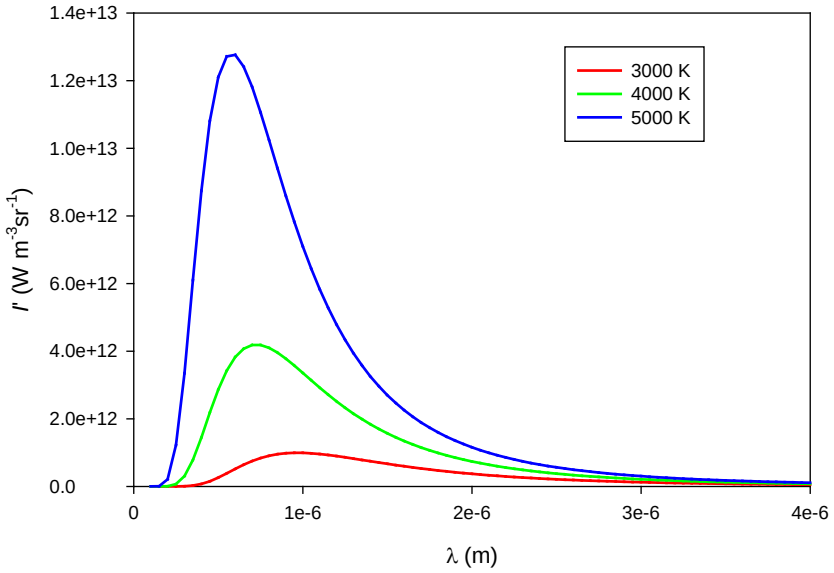

Figure V-12. Blackbody Planck function vs. wavelength for temperatures of 3000 K, 4000K, and 5000 K.

the emission is a volume emission, so that the effects of emissivity and absorption in the sample become more complicated. However, the fact that there are no wave-interface interactions means that the observed state is the principal Hugoniot state. One caution is in order. Particularly at shock stresses below 50 GPa in materials with covalent bonding character, localized deformation can give rise to nonuniform temperature distributions behind the shock wave. Typically, analysis of pyrometer records makes the assumption that the observed state is at a single uniform temperature.

X-ray Diffraction.

One of the difficulties in shock wave work is that most materials can undergo phase changes when shocked. If the new phase is a solid, unless there is auxiliary data from static high pressure experiments, we usually have no information on the crystal structure of the new phase. Under some circumstances, x-ray diffraction can be fielded on shock wave experiments. Because most x-rays used for diffraction studies have relatively low energies, the region probed is within a few microns of the sample surface. As a result, an anvil that can keep the sample at high pressure for sufficient time to obtain the data must be used. Such anvils must be transparent to the x-rays and, ideally, should be amorphous so as to avoid spurious peaks. This technique is still at a very low level of maturity and is an active area of research and development.

Spectroscopy.

If the sample can be accessed optically, particularly if the sample is transparent, then it can be probed spectroscopically. The easiest spectra to obtain are emission spectra in the optical or infrared, but absorption and stimulated emission spectra can be obtained. As with x-ray diffraction, the use of spectroscopy in dynamic experiments is not a mature technique and is an active area of research and development.

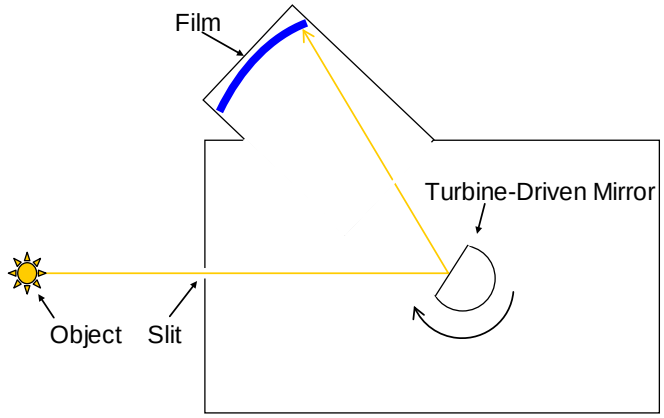

Figure V-13. Schematic of a rotating mirror streak camera. This type has been used extensively at LANL.

Recorders.

Regardless of the type of measurement, the data must be recorded in some way. This may be on image plates, film, or with digitizers or oscilloscopes. We will not discuss these recorders. There is one type of recorder, however, that has a long history of use in shock wave experimentation. This is the streak camera. Today, there are two basic types of streak camera. The first and most important is the rotating mirror streak camera (Figure V-13), in which a mirror is turned by a turbine or electric motor, allowing the image of one or more slits to be smeared out in time onto a strip of film. Thus, the image that results has one spatial dimension representing position along the slit and the other representing time. A variety of designs exist for such cameras. Although such cameras are being replaced with electronic systems for some applications, they are still widely used. The second type of streak camera is the electronic image converter streak camera. In this case, the image is electronically smeared out in time over a detector. Such systems can record faster events than is typically possible with rotating mirror cameras and, often, the detectors are more sensitive than film. Often, image converter streak cameras suffer from nonuniform writing rates that require a periodic time signal be recorded simultaneously with the light coming from the experiment.

Designing a successful experiment.

Here is some basic guidance for designing a successful one-dimensional shock wave experiment. The most important goals in experiment design are to assure the experiment stays one-dimensional and, usually, to assure the shock wave is steady. Also important is to strive for maximum accuracy, but all the accuracy in the world won't help if the basic experiment is invalid because of a bad design.

Edge effects.

Probably the most common mistake made in experimental shock wave physics is design of an experiment in which one-dimensionality is destroyed before the important part of the experiment is complete. This occurs because waves emanating from lateral interfaces in the target/projectile package and traveling laterally through the target reach

position(s) where measurements are being made while the phenomenon of interest is still occurring.

To avoid edge effects polluting the experiment, it is important to assure that the total experiment (target plus impactor) has sufficient lateral extent to allow all the longitudinal waves of interest to be observed prior to the arrival of waves propagating from the edge of assembly. There are two issues to be aware of. First, in a given material, waves will travel in toward the center from the edge. In the usual situation, the "edge wave" originates at the edge of the sample on the impact surface (Figure V-14). At this point, the impedance mismatch between the sample and whatever the rest of the target assembly is made of results in a lateral pressure difference on impact and this initiates a disturbance that propagates inward. Simultaneously, the shock wave in the sample is propagating in the longitudinal direction. Keep in mind that the sound speed behind the shock wave is usually significantly greater than the shock wave speed relative to the shocked material, so the edge wave travels faster than the shock wave. However, since it must travel along a slant path, it takes more time to reach the center of the target at the shock front. A common rule of thumb is that the boundary of the one-dimensional shock wave moves inward with a 45º angle from the edge at the impact surface.

If you have reasonable estimates of the Eulerian sound speeds in the shocked target and impactor, then you can, with a high degree of accuracy, predict the time of arrival of edge effects. If the sound speed in the target is faster, then the path of least time is simply a straight line from the edge of the target or impactor (whichever has the smaller diameter) at the impact surface to the observation point. This path must be computed in the Eulerian reference frame, so that the target is foreshortened due to the one-dimensional compression by the shock wave. If the impactor has the higher sound speed in the shocked state, then the path of least time initially follows the interface between the impactor and target and then takes off at an angle determined by the ratio of the sound speeds (Figure V-15). Specifically, the path of departure from the interface makes an angle with the normal to the interface whose sine is the ratio of the target sound speed to the impactor sound speed:

$$\theta = \sin^{-1} \frac{c_{l,t}}{c_{l,i}} \quad . \tag{V.5}$$

The first problem that arises is that one is often interested in more than just the initial shock wave. One might, for instance, be interested in the arrival of a reflected wave from the rear of the impactor. In such a case, the conservative approach is to calculate the total time from impact until arrival of the last feature of interest at the point of observation, and then assure that the time for a longitudinal sound wave in the shocked target to reach the center of the target at the impact surface is longer. This builds in a margin to account for the approximations that usually must be made in the calculations.

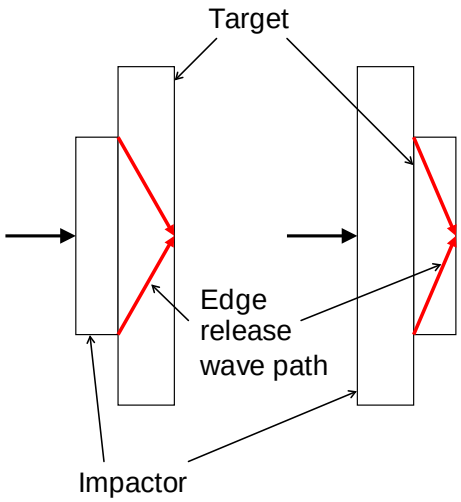

Figure V-14. Edge releases originate at the lateral interface nearest the centerline of the experiment and follow the path of least time to the observation point. The situation shown here is for a target with higher sound speed in the shocked state than the impactor.

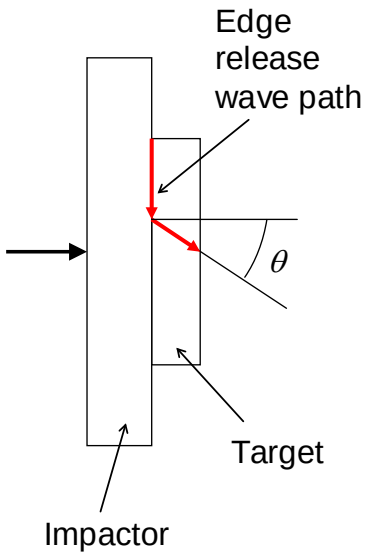

Figure V-15. If the impactor has a higher sound speed, the edge release path is initially along the impactor-target interface.

Overtake.

A second issue that has occasionally caused problems with Hugoniot measurements is that, since the Lagrangian sound speed in the shocked material is substantially greater than the speed of the shock wave itself, a release wave from the rear of the impactor can, given sufficient running distance, overtake the shock wave in the target as in Figure V-16. Thus, one must exercise care to assure that the impactor is thick enough to avoid this happening. Usually, choosing an impactor with a thickness such that the shock wave

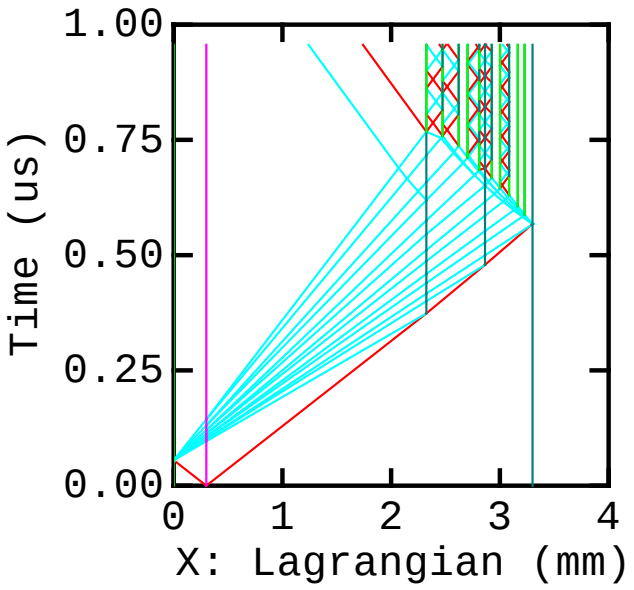

Figure V-16. If the impactor is too thin or the target too thick, release waves from the rear of the impactor may overtake the shock wave in the target before it reaches the surface of the target.

reaches the back of the impactor at least half way through the expected transit time of the shock wave through the target is sufficiently conservative. This might not be the case if the impactor is very compressible. Occasionally, overtake inside the sample is desired. In such cases, careful calculations should be performed to assure that the correct target and impactor thicknesses are used.

Multiple waves.

In some cases, the existence of multiple waves, due to strength or phase transitions, can cause problems. This is particularly true in experiments when a simple wave arrival diagnostic, such as pins or flash gaps, is used. Flash gaps are not sensitive to low-amplitude waves and may therefore miss early arrivals. Both pins and, to some extent, flash gaps tend to be desensitized by the arrival of a strong wave, so that later wave arrivals may not be detected. In both cases, the problem is that a single wave is detected, but there are really other waves. The conditions achieved in a single shock wave are different from those achieved by multiple waves, so that the data analysis in such experiments gives the wrong answer.

Some Tools for Estimation.

Estimation of the Hugoniot of a Material.

Often, one needs to estimate the shock Hugoniot curve of a material or mixture that has not been characterized, either to use in a model or as a guide for designing experiments. There are several approaches that one can take, each with its own pros and cons. The simplest possible case, if the sound speed has been measured, is to assume that the Hugoniot intercept is the bulk sound speed and then assume an appropriate value of s, depending on the type of material. Highly incompressible materials and metals often

have $s \approx 1.25$. Many of the less refractory metals have $s \approx 1.5$. Liquids and polymers typically have $s > 1.5$. Very porous materials often have $s \approx 1$. A slightly better approach is, if you think the material is similar to some other material that has a known Hugoniot, just to use the Hugoniot parameters of that other material with the density of the material you are interested in.

Hugoniots of Mixtures.

Probably the most common, and surprisingly accurate, methods of estimating the Hugoniot of a mixture is simple volume addition of Hugoniots. Select a couple of pressures in the range of interest and determine the density of each constituent on its own Hugoniot at that pressure. Then obtain a mass-weighted average of the densities using

$$\frac{1}{\rho} = \sum \frac{m_i}{\rho_i} \tag{V.6}$$

to obtain the mixture Hugoniot point. Then take the two points and obtain U_s and u_p from

$$u_p = \left[P\left(\frac{1}{\rho_0} - \frac{1}{\rho_H} \right) \right]^{1/2} \tag{V.7}$$

and

$$U_s = \frac{P}{\rho_0 u_p}. \tag{V.8}$$

From these, C_0 and s can be estimated. If you have some information, such as a single Hugoniot point, it is best to adjust the value of s to bring your estimate in line with the data. The advantages to this approach are that it is easy and surprisingly accurate.

The underlying reason for the relatively good agreement this method gives with data is that, as long as the pressures aren't too high, the thermal component of the pressure is relatively modest and the fact that thermal equilibrium isn't achieved can be safely ignored.

Equation of state of a mixture.

If a material is a mixture of two or more materials with different equations of state, then the EOS of the mixture can be estimated using a conservation of volume, mass, and energy approach. This approach works best if the materials are mixed at a fine enough scale that thermal equilibrium is achieved at shorter timescales than the processes of interest. The mixing is also assumed to be ideal. Hence, if we are dealing with a solution, rather than a simple mechanical mixture, then nonideal solutions will be poorly modeled.

The basic idea comes from the principle that a collection of objects made up of different materials take up the same total volume and have the same total mass whether they are together or separated. If we take the same amounts of the materials, but now

divide them finely and mix them, taking care to squeeze out any porosity in the mixture, the total volume and mass will still be the same. Then, any quantity that can be cast as a specific quantity, or the derivative of the specific quantity, can be found for the mixture by a mass-weighted average of that quantity. Hence

$$\frac{1}{\rho} = \sum \frac{m_i}{\rho_i} \tag{V.9}$$

$$\frac{\alpha}{\rho} = \sum \frac{m_i \alpha_i}{\rho_i} \tag{V.10}$$

$$\frac{1}{K_{T0}\rho} = \sum \frac{m_i}{K_{T0,i}\rho_i} \tag{V.11}$$

$$\frac{K_T' + 1}{K_{T0}^2 \rho} = \sum \frac{m_i \left(K_{T,i}' + 1 \right)}{K_{T0,i}^2 \rho_i} \tag{V.12}$$

$$C_V = \sum m_i C_{V,i} \tag{V.13}$$

With these averaged quantities in hand, other parameters, such as γ, can be obtained. One important note should be considered here. For this approach to work, the different equation of state parameters being used in the calculation should have been fit to the same type of EOS. Also, the mixture EOS parameters should be used in the same EOS formulation used to obtain the individual material EOS parameters. Finally, if you are using the third order Birch-Murnaghan formulation, the mixture EOS should be taken to the next order. Keep in mind that the value of K'' in the third order BMEOS is not zero, but instead is the value required to give a zero value of the parameter ξ_2.

Estimation of material paths in experiments.

One common need is to estimate the states that will be achieved by different components of an experiment. To see how we might do this, let us return to the graphical impedance matching technique. For the moment, let us restrict our discussion to materials that are not highly compressible. In such cases, in the P-u_p plane, the release adiabat is quite often reasonably well approximated by the Hugoniot, reflected about the point from which the release is initiated. This is because, as noted above, the thermal component of the Hugoniot pressure isn't very large at moderate pressures, so that the Hugoniot actually is a reasonable approximation of the isentrope, as long as large volume changes are not involved. It is often common to assume that reshock paths are also well-

represented by the same curve. Now, imagine that you want to estimate what pressure will be achieved in a material that is sitting on a baseplate, so that the baseplate, rather than the sample is impacted directly by the impactor. First, the impact sets up a shock wave in the baseplate. However, when the shock wave reaches the interface between the baseplate and the sample, it will be reflected and transmitted in accordance with the impedance matching between the two materials. Now, since we are using the reflected Hugoniot of the baseplate to approximate the release or reshock of the baseplate, this is fully equivalent to the unshocked baseplate impacting the sample at a velocity that is exactly twice the particle velocity imparted to the baseplate by the initial shock wave. So we just reflect the Hugoniot of baseplate and then allow it to interact with the sample. As you might guess, this technique can be extended to an arbitrary number of layers, treating the interaction at each interface in turn as an impact by the donor material at twice the particle velocity in the donor. While this approach is tedious if done by hand, it is easily performed if you use a program to automate the impedance matching calculations. This technique generally works well with simple materials, but will fail rather spectacularly if porous materials or phase changes are involved.

Estimation of times in experiments.

The estimation of times in an experiment is best done if you are able to calculate the sound speeds in the compressed state. This requires knowledge of the Grüneisen parameters and Poisson ratios of the materials involved. Barring such knowledge or the ability to make reasonable estimates, a reasonable approach is to assume the Eulerian bulk sound speed behind the shock wave is equal to the shock wave speed through the unshocked material and that the Poisson ratio is $^1/_3$. Then, it is a simple matter to estimate the time required for a wave to transit a material based on the compressed thickness of the material and the estimated wave speed.

Some Notes on Uncertainties.

Incompletely documented uncertainties in references.

Shock wave researchers often have the bad habit of publishing data without listing the uncertainties. Less often, uncertainties are stated, but on closer examination are found to be assumptions based on past experience, rather than actual evaluated uncertainties. When you run into such a situation, it is often difficult, if not impossible, to determine what the real uncertainties are. Deciding how to proceed can be difficult, but my own approach is, when possible, to fit the data to some simple form such as a polynomial, and use the standard deviation from the fit as an estimate of the experimental uncertainties. Then, if I can estimate systematic uncertainties (say, from the uncertainties on the standards used, etc.), I can add those uncertainties.

Another problem, although less often encountered, is that the Hugoniots or other parameterized fits for standards are listed without covariances to the fit parameters, so that the uncertainties you would calculate are not the full story. In such cases, if possible, I use the more conservative of two possible approaches. The first approach is to use the standard deviation of the fit. Otherwise, I use the stated parametric uncertainties when I know that the governing equations would likely lead to smaller uncertainties if the covariances were actually used in the error estimation.

Calibration.

Beware of uncalibrated instruments. While instruments may be perfectly good, you may have some requirements for quality assurance that can only be proved during an audit if the instruments you use are in calibration. Along those same lines, you can be surprised by systems you thought were well-characterized. I have run into situations where careful measurements have been made to characterize a system, only to find out that the measurements were made with instruments that had unknown offsets in their response.

Outliers.

When you conduct a series of measurements, you are likely confronted with outliers. While many would have you believe that all the data have to be included or you are committing fraud, this is not true. You should *present* all of the data, but it is perfectly reasonable to reject outliers from data fits, if you have a good reason to do so. This is where keeping good notes comes in handy. If there is an observed and documented anomaly that is the probable cause of an apparently erroneous datum, that datum can be rejected out of hand with a note as to why. More often, though, you may have an outlier and not have a "smoking gun" to blame for it. What to do in such cases is more problematic. There are well-documented statistical tests that can be used to determine when an outlier should be omitted from a fit, although the exact criterion may be dependent on the details of the data type. The most important thing is to thoroughly document everything you do in your notes and in your reports and papers. The readers should be able to understand what you did and why. They can then evaluate the total product, both the data and the fit to the data, fairly.