Lecture T5 **DRAFT**

Stability, Convexity, and the Limits of Equilibrium

Stephen Whitcomb
whitcomb.s@northeastern.edu

The equilibrium lecture found the condition that two bodies in thermal contact satisfy when the total entropy is stationary: their temperatures are equal. This note asks what that condition established, and develops the curvature tests that decide whether a stationary state is one the system will stay in. We then allow several exchange variables at once, separate a stable phase from a merely metastable one, and examine the assumptions the standard criteria quietly require. We assume the potentials, Legendre transformations, Maxwell relations, and response functions of the two preceding units; matrix quadratic forms are introduced where they are needed.

Contents

1 What a Stationary Point Leaves Open

Let two bodies be separated by a rigid, impermeable, diathermal wall, with the composite isolated from everything else. Energy can cross the wall; volume and particles cannot. The individual volumes and particle numbers are fixed, $E_{\mathrm{tot}} = E_1 + E_2$ is fixed, and E_1 is the only free quantity. Taking the entropy of the composite to be the sum of the entropies of its parts,

$$S_{\mathrm{tot}}(E_1) = S_1(E_1, V_1, N_1) + S_2(E_{\mathrm{tot}} - E_1, V_2, N_2), \tag{1}$$

and setting the derivative to zero with $(\partial S/\partial E)_{V,N} = 1/T$ gives

$$\frac{dS_{\mathrm{tot}}}{dE_1} = \frac{1}{T_1} - \frac{1}{T_2} = 0 \qquad \Longrightarrow \qquad T_1 = T_2. \tag{2}$$

It is worth being exact about what (2) settles. A differentiable function has zero derivative at any interior extremum, so equality of temperature is a *necessary* condition for an interior maximum. Three questions remain. The same equation holds at a minimum and at an inflection, so we do not know which kind of stationary point we have. Even a genuine local maximum can be beaten by a state far away in the space of divisions, or by a state that is not a division of this kind at all. And nothing in the calculation refers to time: it identifies the state of largest entropy among those compared, not whether the system reaches it or how long it takes. The first question occupies Sections 2–5, the second Section 6, and the third Section 7.

Standing Assumptions

Unless stated otherwise, each body has a fundamental relation $S(E, V, N)$ that is twice differentiable and homogeneous of degree one; the entropy of a composite is the sum of the entropies of its parts, with no interaction or surface contribution; each body is a single homogeneous phase in internal equilibrium; and $T > 0$. Every criterion below inherits these. Section 8 examines what happens when the second and third fail.

Every criterion below is an inequality among derivatives, so it matters which variables are held fixed and how each symbol is normalised. Table 1 records those choices once. We use the physical entropy S rather than S/k_B, so that $1/T$ rather than $\beta = 1/(k_B T)$ appears as the derivative of the entropy with respect to energy, and the distinction between κ_T and κ_S is never dropped: it is the whole content of several results below.

2 The Second Variation

A curvature test is a statement about a particular set of disturbances, so we must say which disturbances the constraints allow before writing one down. For the wall above the only admissible variation is $\delta E_1 = -\delta E_2$, a one-dimensional family; a wall free to move would add $\delta V_1 = -\delta V_2$ and make it two-dimensional. The same pair of bodies can be stable against one set and unstable against a larger one. We should also be clear about what is varied: the states compared are *constrained equilibrium states*, each body separately in internal equilibrium with a definite energy,

Symbol	Definition	Dimensions
C_V	$(\partial E/\partial T)_{V,N} = T(\partial S/\partial T)_{V,N}$	energy / temperature
C_P	$T(\partial S/\partial T)_{P,N}$	energy / temperature
κ_T	$-(1/V)(\partial V/\partial P)_{T,N}$	inverse pressure
κ_S	$-(1/V)(\partial V/\partial P)_{S,N}$	inverse pressure
α	$(1/V)(\partial V/\partial T)_{P,N}$	inverse temperature
μ	$(\partial U/\partial N)_{S,V} = -T(\partial S/\partial N)_{E,V}$	energy / particle
$\mathcal{A}$	$U - T_0 S + P_0 V - \mu_0 N$, with reservoir T_0, P_0, μ_0	energy

Table 1: Symbols, the variables held fixed in each definition, and dimensions. Superscripts (1) and (2) label the two bodies.

so the entropy we differentiate is the equilibrium fundamental relation and not a functional of an arbitrary nonequilibrium state.

Differentiating (1) again, and using

$$\left(\frac{\partial^2 S}{\partial E^2}\right)_{V,N} = \left(\frac{\partial(1/T)}{\partial E}\right)_{V,N} = -\frac{1}{T^2}\left(\frac{\partial T}{\partial E}\right)_{V,N} = -\frac{1}{T^2 C_V}, \tag{3}$$

with the two temperatures equal at the stationary point, the second-order term in the Taylor expansion of S_{tot} is

$$\delta^2 S_{\text{tot}} = -\frac{1}{2T^2}\left(\frac{1}{C_V^{(1)}} + \frac{1}{C_V^{(2)}}\right)(\delta E_1)^2. \tag{4}$$

Thermal Stability of Two Bodies

For two bodies exchanging only energy at fixed individual volumes and particle numbers, and with a common temperature $T > 0$, the stationary point of S_{tot} is a strict local maximum if and only if

$$\frac{1}{C_V^{(1)}} + \frac{1}{C_V^{(2)}} > 0,$$

with both heat capacities evaluated at that temperature. The criterion assumes twice differentiable fundamental relations, additive entropies, and variations restricted to energy transfer alone.

This is not the same as requiring both heat capacities to be positive, and the difference is instructive. If one body is much larger than the other, $C_V^{(2)} \to \infty$ and the condition collapses to $C_V^{(1)} > 0$: a body in contact with a reservoir is stable exactly when its own heat capacity is positive. But if $C_V^{(1)} < 0 < C_V^{(2)}$, the sum of reciprocals is positive whenever $C_V^{(2)} < |C_V^{(1)}|$. A body with a negative heat capacity can sit at a genuine entropy maximum provided its partner is small enough, as Figure 1 shows. Such a body has a convex entropy, which the standing assumptions exclude for an additive system; Section 8 identifies the systems for which it is possible, so the figure is a model rather than a property of an ordinary fluid.

The familiar requirement $C_V > 0$ follows by choosing the partner carefully. Take one homogeneous body and divide it, conceptually, into parts containing fractions x and $1 - x$ of the material. Nothing physical has been done; we have only decided to describe it as a composite. The criterion then applies with $C_V^{(1)} = xC_V$ and $C_V^{(2)} = (1-x)C_V$, giving $1/[C_V x(1-x)] > 0$, so that stability against this redistribution for every x requires $C_V > 0$. Three ingredients went into that argument,

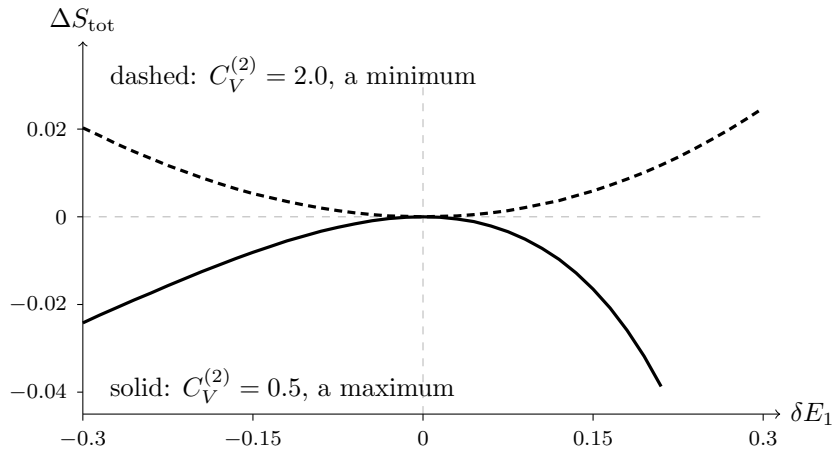

Figure 1: Total entropy near the stationary point for two bodies exchanging only energy. Body 1 has $T_1 = T_* e^{-(E_1 - E_*)/\Lambda}$, so $C_V^{(1)} = -\Lambda/T_1 = -1$ at the stationary point; body 2 has a constant $C_V^{(2)}$. Energies are in units of Λ and entropies in units of Λ/T_*, both measured from the stationary point, where $T_1 = T_2$ on either curve. The solid curve is drawn only as far as body 2 retains energy for the plotted range; it continues downward.

and Section 8 removes them one at a time: additivity, so that the entropy of the body is the sum of the entropies of the parts with no surface term; extensivity, so that a part containing a fraction x has heat capacity xC_V; and homogeneity, so that each part obeys a fundamental relation of the same form. The condition $C_V > 0$ is a consequence of stability *together with* these structural properties, not a separate law.

3 Concavity, Convexity, and a Zero Mode

Repeating the subdivision argument with arbitrary variations of energy, volume, and particle number gives a statement about the whole fundamental relation. If a homogeneous additive system is stable against every redistribution among its parts, then for any two states $X_a = (E_a, V_a, N_a)$ and X_b of the same substance and any $\lambda \in [0, 1]$,

$$S\big(\lambda X_a + (1-\lambda)X_b\big) \;\geq\; \lambda S(X_a) + (1-\lambda)S(X_b), \tag{5}$$

so the entropy is a concave function of its extensive variables. In the energy representation $U(S, V, N)$ is convex, the sense reversing because solving $S(E, V, N)$ for E inverts an increasing function.

The distinction between (5) and a local curvature test matters more than their similarity suggests. A negative second derivative at a point says only that the graph lies below its tangent nearby. A function can satisfy that throughout a neighbourhood and still lie below a chord joining two distant points, and then the uniform state described by the intermediate X has less entropy than a mixture of the two end states. That gap is what metastability amounts to, and Section 6 takes it up. The logical relation is one-way: (5) implies the local tests, and the local tests do not imply (5).

3.1. Homogeneity Supplies a Null Direction

There is a trap in applying a definiteness test to the full extensive-variable Hessian. Write $X = (E, V, N)^{\mathsf{T}}$ and let H_S be the matrix of second derivatives of S with respect to these variables.

Because S is homogeneous of degree one, each first derivative $\partial S/\partial X_i$ is homogeneous of degree zero, and differentiating that property along the scaling direction gives $\mathsf{H}_S X = 0$. The Hessian annihilates X itself. This is not instability: it says that doubling the system doubles its entropy, so the entropy does not curve in that direction at all.

Key Point

Stability requires H_S to be negative *semi*definite, with the scaling direction X as its only null direction. Strict negative definiteness of the full 3×3 matrix is never satisfiable, and $\det \mathsf{H}_S = 0$ is not a signal of marginal stability. The usable statement is that H_S is negative definite on any complement of the scaling direction — equivalently, that the two-variable Hessian of the entropy per particle $s(e, v)$ is negative definite.

The remedy is to remove the redundancy at the start: either work at fixed N with the pair (E, V), or use densities and study $s(e, v)$. Since $U(S, V, N) = N\, u(S/N, V/N)$, the Hessian of U with respect to (S, V) at fixed N is $1/N$ times the Hessian of $u(s, v)$, so the two routes give the same conditions. We take the fixed-N route and restore composition in Section 5.

3.2. Which Function Is Extremal

Laboratory constraints are usually not isolation: a thermostat fixes the temperature, a loaded piston the pressure, a particle reservoir the chemical potential. Each case follows from the entropy maximum without a new principle. Let the system exchange energy, volume, and particles with a reservoir whose T_0, P_0, and μ_0 do not change. The reservoir loses what the system gains, so $dS_R = -(dU + P_0\, dV - \mu_0\, dN)/T_0$ with unsubscripted symbols referring to the system, and adding the system's own entropy change,

$$dS_{\text{total}} = dS - \frac{dU + P_0\, dV - \mu_0\, dN}{T_0} = -\frac{1}{T_0}\, d\mathcal{A}, \qquad \mathcal{A} \equiv U - T_0 S + P_0 V - \mu_0 N. \tag{6}$$

Maximising the entropy of the isolated composite is minimising the *availability* $\mathcal{A}$ of the system, a function of the system's own variables in which the reservoir's parameters appear as constants. Its first variation reproduces $T = T_0$, $P = P_0$, $\mu = \mu_0$, and only at that point does $\mathcal{A}$ coincide, up to constants, with the Helmholtz or Gibbs free energy. The extremal function is therefore S_{tot} for an isolated composite, $U - T_0 S$ at fixed T_0, V, N, $U - T_0 S + P_0 V$ when the pressure is also fixed, and $\mathcal{A}$ when particles are exchanged.

For stability, the useful consequence is immediate. The reservoir's parameters are constants, so the terms they multiply are linear in the system's variables and contribute no curvature:

$$\delta^2 \mathcal{A} = \delta^2 U. \tag{7}$$

Every set of constraints leads back to the same object. What changes is not the curvature but which variations are admissible.

4 The Energy Hessian and the Response Functions

We will repeatedly need the sign of a quadratic form $Q(\delta x) = \delta x^{\mathsf{T}} \mathsf{H}\, \delta x$ with H real and symmetric. Completing the square on the first variable,

$$Q = \mathsf{H}_{11} \left(\delta x_1 + \frac{\mathsf{H}_{12}}{\mathsf{H}_{11}} \delta x_2 \right)^2 + \frac{\det \mathsf{H}}{\mathsf{H}_{11}} (\delta x_2)^2, \qquad \det \mathsf{H} = \mathsf{H}_{11}\mathsf{H}_{22} - \mathsf{H}_{12}^2, \tag{8}$$

so $Q > 0$ for all nonzero δx if and only if $\mathsf{H}_{11} > 0$ and $\det \mathsf{H} > 0$. Two features matter later. The diagonal entries alone decide nothing: a matrix with both diagonal entries positive is indefinite whenever H_{12}^2 exceeds their product. And the coefficient $\det \mathsf{H}/\mathsf{H}_{11}$ is the curvature that remains after δx_1 has been adjusted to its most favourable value, which is the form Le Châtelier's principle takes below.

Work at fixed N with the variations $(\delta S, \delta V)$ of a single body in contact with reservoirs, so that stability requires $\delta^2 U > 0$. From $dU = T\,dS - P\,dV$ the entries are second derivatives of $U(S,V)$, each a response function:

$$\mathsf{H}_{SS} = \left(\frac{\partial T}{\partial S}\right)_V = \frac{T}{C_V}, \qquad \mathsf{H}_{VV} = -\left(\frac{\partial P}{\partial V}\right)_S = \frac{1}{V\kappa_S}, \qquad \mathsf{H}_{SV} = \left(\frac{\partial T}{\partial V}\right)_S = -\frac{T\alpha}{C_V\kappa_T}. \tag{9}$$

The third uses the Maxwell relation belonging to $U(S,V)$,

$$\left(\frac{\partial T}{\partial V}\right)_S = -\left(\frac{\partial P}{\partial S}\right)_V = -\left(\frac{\partial P}{\partial T}\right)_V \left(\frac{\partial T}{\partial S}\right)_V = -\frac{\alpha}{\kappa_T}\frac{T}{C_V}.$$

The determinant looks unpromising, since it mixes an adiabatic with an isothermal compressibility. But with $C_P - C_V = TV\alpha^2/\kappa_T$ and $\kappa_S = \kappa_T C_V/C_P$, both established in the response-function unit, the first term expands as

$$\frac{T}{C_V V\kappa_S} = \frac{TC_P}{C_V^2 V\kappa_T} = \frac{T}{C_V V\kappa_T} + \frac{T^2\alpha^2}{C_V^2\kappa_T^2}, \tag{10}$$

whose second piece is exactly H_{SV}^2. Subtracting it leaves

$$\det \mathsf{H} = \frac{T}{C_V\, V\kappa_T}. \tag{11}$$

The adiabatic compressibility has cancelled. The identity is algebraic, so it holds wherever the response functions are defined and $C_V \neq 0$; it does not presuppose stability, and both sides change sign together when $\kappa_T < 0$. Applying (8) with $T > 0$ gives the criteria.

Local Stability of a One-Component Fluid

For a single homogeneous phase at fixed particle number, with a twice-differentiable fundamental relation and $T > 0$, strict local stability against variations of entropy and volume holds if and only if

$$C_V > 0 \qquad \text{and} \qquad \kappa_T > 0.$$

These are sufficient for a strict local minimum of the availability, and therefore for stability against small disturbances. They do not establish that the phase has the lowest availability among all states available to the system.

Two inequalities follow. Since $C_P - C_V = TV\alpha^2/\kappa_T$ and $\kappa_T - \kappa_S = TV\alpha^2/C_P$, positivity of C_V and κ_T gives

$$C_P \;\geq\; C_V > 0, \qquad \kappa_T \;\geq\; \kappa_S > 0, \tag{12}$$

with equality when $\alpha = 0$. The expansion coefficient α itself is unrestricted in sign, since it enters only through α^2: liquid water at atmospheric pressure between 0 and about 4 °C contracts on heating and is perfectly stable. It is worth noticing which quantities a stability argument constrains and which it merely happens to involve.

4.1. Le Châtelier's Principle and Relaxed Constraints

The verbal form of Le Châtelier's principle — that a system responds to an imposed change in a way that moderates it — is the completion of the square in (8). Writing the differentials of the intensive variables in terms of the Hessian, $\delta T = \mathsf{H}_{SS}\delta S + \mathsf{H}_{SV}\delta V$ and $\delta(-P) = \mathsf{H}_{SV}\delta S + \mathsf{H}_{VV}\delta V$, a clamped volume gives $\delta T = \mathsf{H}_{SS}\delta S$. If instead the pressure is held fixed, the second equation forces $\delta V = -(\mathsf{H}_{SV}/\mathsf{H}_{VV})\delta S$, and substituting into the first,

$$\left(\frac{\partial T}{\partial S}\right)_P = \mathsf{H}_{SS} - \frac{\mathsf{H}_{SV}^2}{\mathsf{H}_{VV}} = \frac{\det \mathsf{H}}{\mathsf{H}_{VV}} \;\le\; \mathsf{H}_{SS} = \left(\frac{\partial T}{\partial S}\right)_V , \tag{13}$$

the inequality holding because $\mathsf{H}_{VV} > 0$ in a stable phase. Since $(\partial T/\partial S)_V = T/C_V$ and $(\partial T/\partial S)_P = T/C_P$, this is $C_P \ge C_V$; exchanging the roles of the variables reproduces $\kappa_T \ge \kappa_S$. The two inequalities in (12) are one statement seen twice.

Key Point

Releasing a constraint cannot decrease a susceptibility. When a second variable is free to move, it moves so as to reduce the change in the conjugate field, and a given change in that field produces a larger change in the primary variable. The size of the effect is $\mathsf{H}_{SV}^2/\mathsf{H}_{VV}$, which vanishes when the variables are uncoupled.

Two qualifications. The inequality relates two equilibrium susceptibilities of the same state and is not a dynamical principle. And it presumes a stable state: if $\mathsf{H}_{VV} < 0$ the sign in (13) reverses, and the verbal form fails along with the stability it assumed. The related Le Châtelier–Braun principle, concerning the sign of the secondary response, needs a separate argument of the same kind; Callen gives it in Chapter 8.

5 Coupled Variations

5.1. A Wall That Moves and Conducts

Replace the rigid wall with one that is diathermal and free to slide while remaining impermeable, so that E_{tot} and $V_{\mathrm{tot}} = V_1 + V_2$ are fixed and the admissible variations form a two-dimensional family. Using $dS = dE/T + (P/T)dV$ for each body,

$$\delta S_{\mathrm{tot}} = \left(\frac{1}{T_1} - \frac{1}{T_2}\right)\delta E_1 + \left(\frac{P_1}{T_1} - \frac{P_2}{T_2}\right)\delta V_1, \tag{14}$$

which vanishes for every admissible pair only if $T_1 = T_2 \equiv T$ and $P_1 = P_2 \equiv P$. The bodies must agree in temperature and in pressure, and that is all stationarity establishes.

For the second variation we could differentiate (14) again, but the second derivatives of $S(E, V)$ are awkward objects. It is shorter to expand each body's energy instead, taking $(\delta S_1, \delta V_1)$ rather than $(\delta E_1, \delta V_1)$ as the independent parameters — a change of variables that is invertible because $T > 0$. Writing

$$\delta E_1 = T\,\delta S_1 - P\,\delta V_1 + \tfrac{1}{2}\,\delta x_1^{\mathsf{T}}\mathsf{H}^{(1)}\delta x_1 + O(\delta^3), \qquad \delta x_1 = (\delta S_1, \delta V_1)^{\mathsf{T}}, \tag{15}$$

with $\mathsf{H}^{(1)}$ the Hessian of (9) for body 1, and similarly for body 2, we add the two expansions and impose $\delta E_1 + \delta E_2 = 0$ and $\delta V_2 = -\delta V_1$. The pressure terms cancel against each other, leaving

$$0 = T\,(\delta S_1 + \delta S_2) + \tfrac{1}{2}\left[\delta x_1^{\mathsf{T}}\mathsf{H}^{(1)}\delta x_1 + \delta x_2^{\mathsf{T}}\mathsf{H}^{(2)}\delta x_2\right]. \tag{16}$$

At first order this says $\delta S_2 = -\delta S_1$, and with $\delta V_2 = -\delta V_1$ exactly, δx_2 may be replaced by $-\delta x_1$ inside the Hessian terms, which are already second order. Solving (16) for the entropy change,

$$\delta S_{\text{tot}} = -\frac{1}{2T}\,\delta x_1^{\mathsf{T}}\left[\mathsf{H}^{(1)} + \mathsf{H}^{(2)}\right]\delta x_1. \tag{17}$$

The first-order change has cancelled identically, so this is the whole change to second order, and the stationary point is a maximum exactly when the *sum* of the two energy Hessians is positive definite. Both earlier results are contained in it. Clamping the wall sets $\delta V_1 = 0$, and with $\mathsf{H}^{(i)}_{SS} = T/C_V^{(i)}$ this is the condition $1/C_V^{(1)} + 1/C_V^{(2)} > 0$. Making body 2 a reservoir sends all of its Hessian entries to zero — infinite heat capacity, infinite $V\kappa_S$ — leaving positive definiteness of $\mathsf{H}^{(1)}$ alone, which is where $C_V > 0$ and $\kappa_T > 0$ came from.

5.2. Why the Diagonal Entries Are Not Enough

Testing the two variations separately probes H_{SS}, thermal stability at fixed volume, and H_{VV}, mechanical stability at fixed entropy. Both can be positive while the matrix is indefinite, and by (8) the disturbance that finds the instability is a *correlated* one in which entropy and volume move together in a particular ratio. Neither axis-aligned test explores that direction. This is the structural counterpart of Le Châtelier's principle: enlarging the space of admissible variations can only make stability harder to satisfy, since the minimum of a quadratic form over a larger space cannot exceed its minimum over a subspace. Every constraint released — unclamping a wall, allowing a reaction to proceed, permitting two components to unmix — is a potential source of instability the constrained problem could not reveal.

The situation is not hypothetical. Take a van der Waals fluid,

$$P = \frac{Nk_BT}{V - Nb} - \frac{aN^2}{V^2}, \tag{18}$$

with constant $C_V = \frac{3}{2}Nk_B$, held by a thermostat and a pressure reservoir. Using (9) with $\alpha/\kappa_T = (\partial P/\partial T)_V$, and obtaining H_{VV} from $(\det \mathsf{H} + \mathsf{H}_{SV}^2)/\mathsf{H}_{SS}$ together with (11) and $1/(V\kappa_T) = -(\partial P/\partial V)_{T,N}$, the entries follow from (18) alone:

$$\mathsf{H}_{SS} = \frac{T}{C_V}, \quad \mathsf{H}_{SV} = -\frac{Nk_BT}{C_V(V - Nb)}, \quad \mathsf{H}_{VV} = \frac{Nk_BT}{(V - Nb)^2}\left(1 + \frac{Nk_B}{C_V}\right) - \frac{2aN^2}{V^3}. \tag{19}$$

Work in units with $N = k_B = 1$, $a = 3$, and $b = 1/3$, for which $v_c = 1$, $P_c = 1$, and $T_c = 8a/(27b) = 8/3$. At $T = 2.4 = 0.9\,T_c$ and $v = 1 = v_c$,

$$\mathsf{H} = \begin{pmatrix} 1.6 & -2.4 \\ -2.4 & 3.0 \end{pmatrix}, \qquad \det \mathsf{H} = -0.96, \qquad \text{eigenvalues } 4.8 \text{ and } -0.2. \tag{20}$$

Both diagonal entries are positive: the fluid is stable against heating at fixed volume, since $C_V = 3/2 > 0$, and against compression at fixed entropy, since $\kappa_S = 1/(V\mathsf{H}_{VV}) = 1/3 > 0$. It is nevertheless unstable, and the eigenvector belonging to the negative eigenvalue gives the direction, $\delta x \propto (4, 3)$, a fluctuation that raises the entropy and expands the volume together. The response function that has gone negative is $\kappa_T = -5/3$, consistent with $\det \mathsf{H} = T/(C_V V \kappa_T) < 0$, and $C_P = -15/2$ with it.

Key Point

Checking $C_V > 0$ and $\kappa_S > 0$ is not a stability test. The pair that decides local stability at fixed N is $C_V > 0$ and $\kappa_T > 0$, because those are the diagonal entry and the determinant of the same matrix. A state can satisfy both diagonal conditions and still be unstable against a correlated entropy–volume fluctuation.

5.3. Composition as a Third Variation

Letting the wall pass particles adds δN_1 and requires μ to match as well. For a single component the new condition is not new. At fixed T and V the Gibbs–Duhem relation gives $d\mu = v\,dP$ along an isotherm with $v = V/N$, so $(\partial\mu/\partial N)_{T,V} = v\,(\partial P/\partial N)_{T,V}$; and since the pressure depends on T and the density $n = N/V$ only, $(\partial P/\partial N)_{T,V} = (1/V)(\partial P/\partial n)_T$ while $(\partial P/\partial V)_{T,N} = -(n/V)(\partial P/\partial n)_T = -1/(V\kappa_T)$. Eliminating $(\partial P/\partial n)_T$,

$$\left(\frac{\partial\mu}{\partial N}\right)_{T,V} = \frac{V}{N^2\kappa_T}, \tag{21}$$

so diffusive stability restates $\kappa_T \geq 0$: in a one-component fluid, changing N at fixed V and changing V at fixed N are both changes of density. Holding T and P instead makes $G = N\mu(T,P)$ linear in N, so $(\partial\mu/\partial N)_{T,P} = 0$ identically. That is the homogeneity zero mode in its most visible form, not marginal stability and not a criterion awaiting application.

Composition becomes independent when there is more than one species. For a binary mixture at fixed T and P with mole fraction x of species 1, let $g(T,P,x)$ be the Gibbs energy per particle. From $g = x\mu_1 + (1-x)\mu_2$, $(\partial g/\partial x)_{T,P} = \mu_1 - \mu_2$, and Gibbs–Duhem in the form $x\,d\mu_1 + (1-x)\,d\mu_2 = 0$,

$$\left(\frac{\partial^2 g}{\partial x^2}\right)_{T,P} = \frac{1}{1-x}\left(\frac{\partial\mu_1}{\partial x}\right)_{T,P}, \tag{22}$$

so with $0 < x < 1$ local stability against a small transfer of composition is $(\partial^2 g/\partial x^2)_{T,P} \geq 0$: the Gibbs energy must be convex in composition. A genuine condition survives here because working per particle has already removed the scaling direction. The regular-solution model, with an ideal entropy of mixing and an energy of mixing proportional to the product of the concentrations,

$$\frac{\Delta g_{\mathrm{mix}}}{k_B T} = x\ln x + (1-x)\ln(1-x) + \chi\,x(1-x), \qquad \chi \equiv \frac{w}{k_B T}, \tag{23}$$

makes the competition explicit: mixing always gains entropy, and for $w > 0$ it costs energy. The logarithms take dimensionless arguments and χ is dimensionless. Differentiating twice gives $(k_BT)^{-1}(\partial^2\Delta g_{\mathrm{mix}}/\partial x^2)_{T,P} = 1/[x(1-x)] - 2\chi$, positive for all x when $\chi < 2$ and negative on an interval about $x = 1/2$ when $\chi > 2$. The homogeneous mixture is locally unstable where $x(1-x) > 1/(2\chi)$, and since $\chi \propto 1/T$ the instability appears on cooling, first at the critical point $\chi = 2$, $x = 1/2$. At $\chi = 2.5$ the boundaries are $x_\pm = \frac{1}{2}(1 \pm \sqrt{1 - 2/\chi}) = 0.276$ and 0.724.

The same completion of the square applies when the relaxing variable is a composition. If a mixture can react, with ξ the extent of reaction and $A = -(\partial G/\partial\xi)_{T,P}$ the affinity, then $C_P^{\mathrm{eq}} = C_P^{\mathrm{froz}} + T(\partial A/\partial T)^2_{P,\xi}/(-\partial A/\partial\xi)_{T,P} \geq C_P^{\mathrm{froz}}$, the denominator being positive in a state stable with respect to ξ. Heat supplied at constant pressure is partly absorbed by shifting the reaction, so it raises the temperature less — and the inequality fails exactly where stability with respect to ξ fails.

6 Locally Stable, Metastable, and Unstable

Every test so far has been local. Follow a single isotherm of the van der Waals fluid below its critical temperature: in reduced variables (18) becomes

$$\tilde{P} = \frac{8\tilde{T}}{3\tilde{v} - 1} - \frac{3}{\tilde{v}^2}, \tag{24}$$

independent of a and b. Local mechanical stability requires $(\partial \tilde{P}/\partial \tilde{v})_{\tilde{T}} < 0$, and setting that derivative to zero gives the boundary of the region, the *spinodal*:

$$\tilde{T}_{\mathrm{sp}} = \frac{(3\tilde{v} - 1)^2}{4\tilde{v}^3}, \qquad \tilde{P}_{\mathrm{sp}} = \frac{3\tilde{v} - 2}{\tilde{v}^3}. \tag{25}$$

Both reduce to 1 at $\tilde{v} = 1$, so the spinodal meets the critical point, as it must: the critical isotherm is the one whose unstable interval has shrunk to nothing. At $\tilde{T} = 0.9$ the spinodal volumes are $\tilde{v}_- = 0.719$ and $\tilde{v}_+ = 1.529$, at reduced pressures 0.420 and 0.724 (Figure 2).

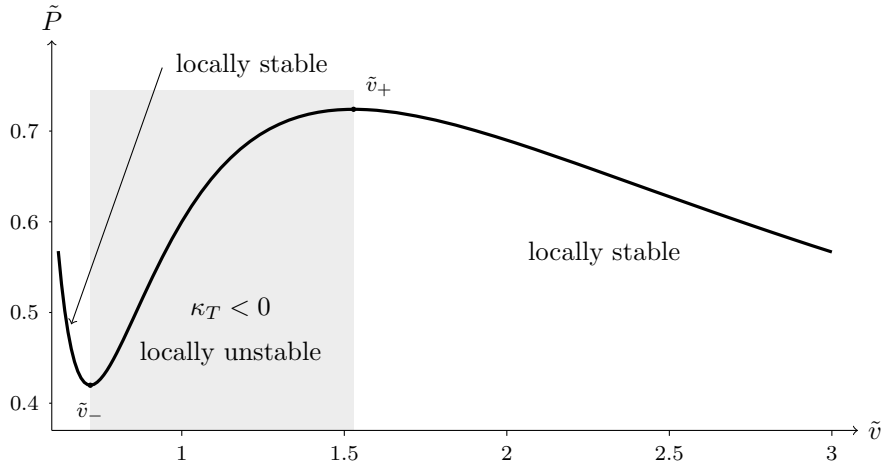

Figure 2: The reduced van der Waals isotherm (24) at $\tilde{T} = 0.9$. Between the spinodal volumes $\tilde{v}_\pm$ of (25) the isotherm rises with volume, so $\kappa_T < 0$ and the homogeneous fluid fails the local test. On both branches outside the shaded interval the fluid passes every local test, which does not settle whether it is the state of lowest Gibbs energy; those branches carry both genuinely stable and metastable states. The coexistence volumes lie inside the outer branches and are not drawn; locating them is the subject of the next unit.

Outside the shaded interval the fluid satisfies $C_V > 0$ and $\kappa_T > 0$, so small disturbances decay and the state is a local minimum of the availability. A homogeneous state at a volume just outside $\tilde{v}_+$ can nevertheless have a higher Gibbs energy than a mixture of a dense and a dilute phase occupying the same total volume, without any small disturbance revealing it. Such a state is *metastable*: stable against everything nearby, unstable against a comparison with a state far away. The regular solution of (23) shows the same three regions in a form that is easier to picture, since Δg_{mix} can be plotted directly: between $x_\pm$ the curve is concave and the mixture is locally unstable, while outside them it is convex but part of it still lies above the straight line joining two points of the curve, so that a mixture can lower its free energy by separating into the two phases at the points of tangency.

Key Point

Three statements about a candidate state should be kept apart. That its first variation vanishes makes it an equilibrium candidate. That the second variation is positive definite on the admissible space makes it locally stable, which covers both genuine stability and metastability. That no competing state has a lower value of the relevant potential makes it the equilibrium state. Each is strictly stronger than the one before, and a local test cannot deliver the third.

Metastable states are not a bookkeeping curiosity. Liquids can be superheated well past their boiling point and supercooled well below their freezing point, and vapours compressed past their condensation pressure, all reproducibly when nucleation sites are scarce; Debenedetti's monograph surveys the range. Deciding which of the locally stable states is the equilibrium one requires a comparison with competing states, which is a global construction rather than a derivative. The two equivalent constructions — a common tangent on the free-energy curve, and the equal-area rule on the isotherm — belong to the next unit.

7 Curvature Is Not Kinetics

A curvature tells us whether a disturbance grows. It does not tell us how fast or by what mechanism, and the two regions fail in different ways for reasons that are dynamical rather than thermodynamic. Inside the spinodal an arbitrarily small long-wavelength fluctuation lowers the free energy, so separation begins with no barrier to cross; what sets the rate and the length scale that appears first is the cost of the gradients such a fluctuation creates. Cahn and Hilliard modelled that cost with a term $\kappa(\nabla c)^2$ in the free-energy density, which penalises short wavelengths and, combined with a diffusive equation of motion, gives a growth rate peaking at an intermediate wavelength. Neither κ nor the mobility appears anywhere in the equilibrium fundamental relation.

Outside the spinodal but inside the coexistence region every small fluctuation decays instead. Separation must begin with a droplet of the other phase large enough that its bulk free-energy gain exceeds its surface cost, so the state waits for a rare fluctuation to supply one, and the waiting time depends exponentially on that barrier. Metastable lifetimes therefore range from microseconds to geological times with no corresponding change in the thermodynamic functions. One qualification: the spinodal is a sharp curve within a mean-field free energy such as (23), in which each region feels an average environment. Once fluctuations of finite range are included the change from barrier-crossing to continuous growth is gradual and the position of a sharp spinodal becomes model-dependent, the sharp version being a limiting concept for interactions of long but finite range.

8 Optional: When the Criteria Do Not Apply

This section may be skipped on a first reading. Nothing in the main argument depends on it, and the results quoted only mark the boundary of the preceding criteria.

The subdivision argument of Section 2 used the strongest of the standing assumptions: that the entropy of a body is the sum of the entropies of its parts. That is an excellent approximation when interactions are short-ranged, so that the interaction energy between two halves of a body scales with the area between them rather than the volume. For a pair potential falling off as $r^{-\alpha}$ in d

dimensions, the energy of a uniform system diverges with its size when $\alpha \le d$ and the system is nonadditive. Gravity with $\alpha = 1$, $d = 3$ is the standard example; unscreened Coulomb systems and two-dimensional vortices behave similarly. Campa, Dauxois, and Ruffo review the consequences.

The simplest quantitative statement comes from the virial theorem. For a bound self-gravitating system of point masses in a steady state, with no external pressure on its boundary, $2\langle K\rangle + \langle W\rangle = 0$, so $E = -\langle K\rangle$; defining a kinetic temperature by $\langle K\rangle = \frac{3}{2}Nk_BT$,

$$E = -\tfrac{3}{2}Nk_BT \qquad \Longrightarrow \qquad C = \frac{dE}{dT} = -\tfrac{3}{2}Nk_B < 0. \tag{26}$$

Losing energy makes such a system hotter. The attached conditions matter: a bound, virialised, steady state with no boundary term, and a temperature defined kinetically rather than through $\partial S/\partial E$. Antonov, and then Lynden-Bell and Wood, showed that for an isothermal sphere confined to a box larger than a critical radius no state of maximal entropy exists at all, the resulting gravothermal catastrophe being driven by exactly the mechanism of Figure 1, with a core of negative heat capacity in contact with a halo.

Equation (26) sits uneasily beside a familiar identity. Differentiating $\langle E\rangle = -\partial \ln Z/\partial\beta$ in the canonical ensemble gives

$$C_V = \frac{\mathrm{Var}(E)}{k_BT^2} \;\ge\; 0, \tag{27}$$

so a canonical heat capacity is never negative. There is no conflict: the statements concern different ensembles. Equation (27) describes a system held by a thermostat, and a system with a negative microcanonical heat capacity simply cannot be held at fixed temperature by one — the criterion $1/C_V^{(1)} + 1/C_V^{(2)} > 0$ fails in the reservoir limit, which is what Figure 1 showed. For additive systems the ensembles agree in the thermodynamic limit and the distinction never surfaces; for nonadditive ones they need not agree, and the microcanonical description is the appropriate one when the system is genuinely isolated.

The lesson concerns the scope of a theorem rather than its correctness. A negative heat capacity does not refute Section 2; it identifies a system violating the additivity that argument assumed. Schmidt and co-workers measured a negative microcanonical heat capacity for a cluster of 147 sodium atoms near melting, where the failure is not long-range forces but small size: at that size the surface contribution to the free energy is not negligible beside the bulk, so the parts are not additive either. Equation (27) also fixes the size of ordinary equilibrium fluctuations: with C_V proportional to N, the relative energy fluctuation $\sqrt{\mathrm{Var}(E)}/\langle E\rangle$ scales as $N^{-1/2}$ — a scaling that assumes an additive system away from a critical point, where C_V and κ_T remain finite.

9 Summary and the Next Question

Table 2 collects the criteria with the assumptions each carries and the variation space each applies to. Read on its own the table should not be able to mislead: a row is a statement about small disturbances within a named space, and none of the local rows settles the last one.

We now know when a homogeneous state is locally stable, and that passing every local test is not the same as being the equilibrium state. Both worked examples show a region where a uniform state is locally stable and yet a nonuniform one has lower free energy. That leaves a specific question for the next unit: given a free energy that is not convex, which pair of phases does the system separate into, and what fixes their densities or compositions? Answering it turns the common-tangent picture into the equality of chemical potentials and pressures between coexisting phases, and from there into the Clapeyron equation and the phase diagram.

Situation	Condition	Assumptions beyond differentiability
Two bodies, energy only	$1/C_V^{(1)} + 1/C_V^{(2)} > 0$	additive entropies; $T > 0$
Homogeneous body, energy only	$C_V > 0$	additivity, extensivity, one phase
Body with reservoirs, fixed N	$C_V > 0$ and $\kappa_T > 0$	one phase; $T > 0$; local only
Two bodies, energy and volume	$\mathsf{H}^{(1)} + \mathsf{H}^{(2)}$ positive definite	as above, on that variation space
Binary mixture at fixed T, P	$(\partial^2 g/\partial x^2)_{T,P} \geq 0$	one phase; per-particle variables
Global equilibrium	lowest potential among all states	requires comparison, not curvature

Table 2: Local stability criteria and the conditions they carry. Each row is valid only for the variation space named in it; enlarging that space can only add conditions.

A Practice Problems

Problem 1 (a body with a negative heat capacity). Body A has $T_A(E_A) = T_* \exp[-(E_A - E_*)/\Lambda]$ with $\Lambda > 0$ over the range of interest; body B has a constant heat capacity $C_B > 0$. They exchange only energy. (a) Show that $C_A = -\Lambda/T_A$ and find $S_A(E_A)$ up to a constant. (b) Find the condition on C_B for the state with $T_A = T_B = T_*$ to be a local entropy maximum. (c) Body A is placed instead in contact with a large reservoir at T_*. What happens, and what does the calculation not tell you? (d) Explain why (b) does not contradict the requirement $C_V > 0$ of Section 2.

Problem 2 (the response-function criteria). (a) Starting from (9) and the identities $C_P - C_V = TV\alpha^2/\kappa_T$, $\kappa_T/\kappa_S = C_P/C_V$, verify $\det \mathsf{H} = T/(C_V V \kappa_T)$. (b) Deduce $C_V > 0$ and $\kappa_T > 0$, stating where $T > 0$ was used. (c) Show that $C_P \geq C_V$ and $\kappa_T \geq \kappa_S$ follow, and identify where equality holds. (d) Show that α is unconstrained in sign, and give a stable substance with $\alpha < 0$.

Problem 3 (the spinodal, and a coupled instability). (a) Derive (25) and verify that both expressions reduce to 1 at $\tilde{v} = 1$. (b) Show that C_V of a van der Waals fluid is independent of volume, and explain why the instability therefore cannot appear as a negative C_V. (c) Using (19) with $N = k_B = 1$, $a = 3$, $b = 1/3$, $C_V = 3/2$ and $T = 2.4$, evaluate H at $v = 1.3$; confirm that both diagonal entries are positive while $\det \mathsf{H} < 0$, and find the eigenvalues and the unstable eigenvector. (d) Compute κ_T and C_P there, check the sign against $\det \mathsf{H} = T/(C_V v \kappa_T)$, and say which rows of Table 2 this state satisfies and which it violates.

Problem 4 (a regular solution). For (23): (a) show that the mixture is locally stable at all compositions when $\chi < 2$, and find the critical composition at $\chi = 2$; (b) for $\chi = 2.5$, find the spinodal compositions; (c) classify $x = 0.5$ and $x = 0.20$ at $\chi = 2.5$, and say what further information would decide whether the locally stable one is the equilibrium state; (d) describe how the unstable interval changes as T falls at fixed $w > 0$, and identify one assumption of the model that makes this temperature dependence simpler than that of a real alloy.

B Solution Sketches

Problem 1. (a) $dT_A/dE_A = -T_A/\Lambda$, so $C_A = -\Lambda/T_A < 0$. Integrating $dS_A/dE_A = 1/T_A$ gives $S_A = (\Lambda/T_*)\, e^{(E_A - E_*)/\Lambda}$ up to a constant, which is convex in E_A. (b) At T_*, $C_A = -\Lambda/T_*$, so $1/C_A + 1/C_B > 0$ requires $C_B < \Lambda/T_* = |C_A|$; Figure 1 plots this with $\Lambda = T_* = 1$. (c) A reservoir is the limit $C_B \to \infty$: the criterion fails and the stationary point is an entropy minimum.

Any transfer then continues in the same direction, with A heating further as it loses energy. The second variation locates the runaway but sets no timescale. (d) Section 2 also assumed additivity, extensivity, and homogeneity. Body A has a convex entropy and is not an additive body of that kind, so it lies outside the scope of that conclusion rather than contradicting it.

Problem 2. (a) Substitute $\kappa_S = \kappa_T C_V/C_P$ into $T/(C_V V \kappa_S)$, eliminate C_P, and subtract H_{SV}^2. (b) With $T > 0$, $\mathsf{H}_{SS} = T/C_V > 0$ forces $C_V > 0$, and then $\det \mathsf{H} > 0$ forces $\kappa_T > 0$; with $T < 0$ both reverse. (c) $C_P = C_V + TV\alpha^2/\kappa_T \geq C_V$ and $\kappa_T = \kappa_S + TV\alpha^2/C_P \geq \kappa_S$, with equality where $\alpha = 0$ — for water near 4 °C at atmospheric pressure. (d) Only α^2 appears; liquid water below about 4 °C has $\alpha < 0$ and is stable.

Problem 3. (a) $(\partial\tilde{P}/\partial\tilde{v})_{\tilde{T}} = -24\tilde{T}/(3\tilde{v}-1)^2 + 6/\tilde{v}^3 = 0$ gives the first relation; substituting into (24) gives the second. Both equal 1 at $\tilde{v} = 1$. (b) $(\partial C_V/\partial V)_{T,N} = T(\partial^2 P/\partial T^2)_{V,N} = 0$ because (18) is linear in T at fixed V, so C_V keeps its ideal-gas value everywhere and the instability must show up in the determinant, through κ_T. (c) $\mathsf{H}_{SS} = 1.600$, $\mathsf{H}_{SV} = -1.655$, $\mathsf{H}_{VV} = 1.550$, $\det \mathsf{H} = -0.260$; eigenvalues 3.230 and -0.081, with unstable eigenvector $(\delta S, \delta V) \propto (0.702, 0.713)$ — the region takes up entropy and expands together, while a pure δS or a pure δV raises the availability. (d) $\kappa_T = -4.730$, $C_P = -14.29$, and $2.4/(1.5 \times 1.3 \times (-4.730)) = -0.260$, matching the determinant. The state satisfies $C_V > 0$ (and $\kappa_S = 1/(v\mathsf{H}_{VV}) = 0.496 > 0$) but violates the third row through $\kappa_T < 0$, which is exactly why the two diagonal response functions are not the right test.

Problem 4. (a) $g'' > 0$ for all x requires $2\chi < \min_x[1/x(1-x)] = 4$, so $\chi < 2$; at $\chi = 2$ the first marginal composition is $x = 1/2$. (b) $x(1-x) = 1/(2\chi) = 0.2$ gives $x_\pm = \frac{1}{2}(1 \pm \sqrt{0.2}) = 0.2764$ and 0.7236. (c) $x = 0.5$ has $g''/k_BT = 4 - 5 = -1 < 0$, locally unstable; $x = 0.20$ has $g''/k_BT = 1/0.16 - 5 = 1.25 > 0$, locally stable. Deciding whether $x = 0.20$ is the equilibrium state needs its free energy compared with a two-phase mixture of the same overall composition — the common-tangent construction, a global comparison rather than a derivative. (d) Lowering T raises $\chi = w/k_BT$ and widens the interval about $x = 1/2$, the instability first appearing at $T = w/2k_B$. The model takes w temperature independent and the mixing entropy ideal, attributing the whole temperature dependence to the $1/T$ in χ; in a real alloy, lattice relaxation and vibrational entropy make w effectively temperature dependent.

References

[1] H. B. Callen, *Thermodynamics and an Introduction to Thermostatistics*, 2nd ed. (Wiley, New York, 1985), Ch. 8. Stability criteria and the Le Châtelier–Braun principle.

[2] P. G. Debenedetti, *Metastable Liquids: Concepts and Principles* (Princeton University Press, Princeton, 1996). Stability theory, the experimental range of superheated and supercooled states, and nucleation.

[3] J. W. Cahn and J. E. Hilliard, "Free energy of a nonuniform system. I. Interfacial free energy," J. Chem. Phys. **28**, 258–267 (1958). Source of the gradient term in Section 7.

[4] V. A. Antonov, Vest. Leningrad Univ. **7**, 135 (1962); D. Lynden-Bell and R. Wood, "The gravo-thermal catastrophe in isothermal spheres and the onset of red-giant structure for stellar systems," Mon. Not. R. Astron. Soc. **138**, 495–525 (1968).

[5] W. Thirring, "Systems with negative specific heat," Z. Phys. **235**, 339–352 (1970); T. Padmanabhan, "Statistical mechanics of gravitating systems," Phys. Rep. **188**, 285–362 (1990).

[6] A. Campa, T. Dauxois, and S. Ruffo, "Statistical mechanics and dynamics of solvable models with long-range interactions," Phys. Rep. **480**, 57–159 (2009).

[7] M. Schmidt, R. Kusche, T. Hippler, J. Donges, W. Kronmüller, B. von Issendorff, and H. Haberland, "Negative heat capacity for a cluster of 147 sodium atoms," Phys. Rev. Lett. **86**, 1191–1194 (2001).