

The Phase Rule: Beyond Myopia to Understanding

R. Stephen Berry and Boris M. Smirnov

Abstract The Gibbs phase rule relating the number of degrees of freedom f of a system to the number of components c and the number of coexisting phases p is a central, universally used relation, expressed by what is probably the simplest formula in the natural sciences, $f = c - p + 2$. Research into the behavior of small systems, notably atomic clusters, has shown in recent years that the phase rule is not as all-encompassing as is often assumed. Small systems can show coexistence of two or more phases in thermodynamic equilibrium over *bands* of temperature and pressure (with no other forces acting on them). The basis of this apparent violation of the phase rule, seeming almost like violation of a scientific law, is in reality entirely understandable, consistent with the laws of thermodynamics, and even allows one to estimate the upper size limit of any particular system for which such apparent violation could be observed.

Keywords Coexistence · Mass action · Phase equilibrium · Phase rule

Contents

1	Introduction and Background	222
2	Observations of Apparent Violations	222
3	Fundamental Explanation	223
3.1	Local Stability of Phases	223
3.2	The Coexistence of Phases	224

R.S. Berry (✉)

Department of Chemistry, The University of Chicago, Chicago, IL 60637, USA

e-mail: berry@uchicago.edu

B.M. Smirnov

Institute for High Temperatures of Russian Academy of Sciences, 127412 Moscow, Russia

e-mail: bmsmirnov@gmail.com

4	Observability of Coexistence in Apparent Violation of the Phase Rule	226
5	Phase Diagrams for Small Systems	227
6	Conclusions and Summary Remarks	228
	References	228

1 Introduction and Background

In 1876, Gibbs [1] introduced the relation between the number of controllable degrees of freedom, f , the number of chemically distinguishable components, c , in a system, and the number of phases, p , coexisting in thermodynamic equilibrium. It is this rule, $f = c - p + 2$, that tells us that we can vary the temperature and pressure of a liquid composed of a single substance, but that if we require that liquid to be in equilibrium with a solid of the same substance, we can vary only the temperature or the pressure, and that the value of the other variable will be fixed by our physical requirement. The rule of course applies just when the only control variables we have are temperature and pressure; the number “2” changes if we add such variables as applied electric or magnetic fields. The work was virtually unknown until Roozeboom recognized and publicized it in 1887 [2]. The concept became much more widely known when Wilhelm Ostwald translated and published Gibbs’ papers in German in 1892. It became a central focus of some physical chemists, notably Wilder Bancroft, soon after the turn of the twentieth century [3].

The history of the phase rule is the subject of a review written for its centenary [4]; that history need not concern our discussion here because it reviews the applications and the controversies over the derivation of the rule. Here, we concern ourselves with the *deviations and apparent violations* of the phase rule, a topic essentially untouched until computer simulations of small systems suggested that the phase rule might, after all, not be so universal.

2 Observations of Apparent Violations

The first indications that certain systems might violate the phase rule came from computer simulations of small clusters of atoms. A number of studies revealed clearly defined solid-like and liquid-like forms [5–14]. These embraced both molecular dynamics and Monte Carlo simulations, and explored a variety of clusters. These included several based on atomic models with interparticle Lennard-Jones forces, which mimic rare gas clusters rather well. There were also models of alkali halide clusters. Hence, the existence of solid and liquid forms for such small systems seemed not only plausible but general, not restricted to any one kind of system. Shortly after these studies appeared, another, of a 55-atom cluster with Lennard-Jones interparticle forces, showed not only solid and liquid forms but also a form in which the surface of the cluster (with icosahedral structure) is liquid

and the core is solid [15]. Simulations of metal clusters followed, and also showed solid and liquid forms [16, 17].

Experimental evidence for solid and liquid forms of clusters began to appear soon thereafter [18–21]. At that time, simply demonstrating the existence of the two phase-like forms for small clusters was a primary goal. Later, as we shall note shortly, more detailed and precise information came from experiments.

One particularly remarkable phenomenon seemed to appear in the early simulations. The observation of solid and liquid phases led Briant and Burton to make the tentative suggestion that even the small Lennard-Jones clusters that they were simulating could have a first-order phase transition [10]. However, Hill had argued in his monograph [22] that very small systems would have to have smooth passage between phases. But the simulations appeared to indicate that, at the molecular level in time evolution and in spatial structure, the solid and liquid forms are clearly distinguishable, with no indication, at least for many kinds and sizes of clusters, that there is no smooth, intermediate kind of behavior. Some of the simulations suggested that perhaps they showed dynamical coexistence of solid and liquid phases at more than one energy or temperature, with the simulated pressure constant, typically at zero.

In short, small clusters showed tantalizing kinds of behavior that did not seem to fit conventional concepts of how such systems should behave. This was a puzzle awaiting explanation.

3 Fundamental Explanation

3.1 *Local Stability of Phases*

The first step toward that explanation came from a rationalization of the simulation results of Briant and Burton [10] and of Eters and Kaelberer [11–13]. In this approach [23], one constructs quantum-mechanical models of the densities of states of rigid and highly nonrigid forms, i.e., solid and liquid forms, of the cluster of interest, and, from these, one can estimate partition functions and free energies of the two forms. The model uses a harmonic, rigid rotor model for the rigid form and an Einstein vibrator model for the nonrigid case. This is enough to specify the symmetry groups for the two limits, and hence to construct the correlation diagram between the two. (For N particles, the solid form has symmetry $SO(3) \times SO(3) \times U(3N - 6)$ and the liquid form has symmetry $U(3N - 3)$.) In this model, the parameter that varies from the solid extreme to the liquid extreme is simply some unspecified measure of the degree of nonrigidity. Such measures are available, but were not needed for the analysis of [23]. From the correlation diagrams and densities of states, it is straightforward to show that at low energies or temperatures, where only the low-lying levels are populated, only the rigid form has a minimum in its Helmholtz or Gibbs free energy. At sufficiently high

temperatures, only the nonrigid or liquid form has a minimum in its free energy. That is, at sufficiently low temperatures, only the solid is stable, and at sufficiently high temperatures, only the liquid is stable. However, there is an intermediate range of temperatures in which, according to this model, *both the rigid solid and the nonrigid liquid have local minima in their free energies.*

Within this model, there is a lowest temperature at which there is a minimum in the free energy in the region of the nonrigid limit, and another temperature that is the highest for which there is a free energy minimum in the region of the rigid, solid form.

From this point, several fruitful directions open. This result has told us that there is a distinct lower bound of temperature below which no liquid form is locally stable, and, correspondingly, a distinct upper bound of temperature above which no solid form is locally stable. These temperatures presumably depend on pressure, but that aspect has yet to be investigated. Whether those sharp bounds can be observed in real experiments is not clear, because of the question of whether fluctuations would hide or appear to smooth them. We refer to the lower limit of stability of the liquid as the *freezing limit* and the upper limit of stability of the solid as the *melting limit*.

3.2 The Coexistence of Phases

The next question is a very important one, perhaps the one most central to the issue of the relation between the phase rule and the behavior of small systems. This is the question of how the equilibrium between solid and liquid forms of small systems can be described in a manner that links that behavior to the behavior of macroscopic systems. It is, in fact, straightforward to address in traditional, classical thermodynamic terms.

We begin by writing the chemical *equilibrium constant* K_{eq} for the equilibrium between solid and liquid, $K_{\text{eq}} = [\text{liquid}]/[\text{solid}]$. This quantity is determined by the difference between the free energies ΔF of the two forms, $K_{\text{eq}} = \exp[-\Delta F/k_{\text{B}}T]$, where k_{B} is the Boltzmann constant and T is the absolute temperature. But ΔF is the difference in the chemical potentials $\Delta\mu$, multiplied by N , the number of particles in each system. We know that the traditional condition for equilibrium of two phases, e.g., solid and liquid, is the equality of the free energies or chemical potentials of the two forms. This is what sets the conditions for the coexistence curves required by the phase rule.

Now suppose we are dealing with a small but macroscopic system of, say, 10^{20} particles, much less than a millimole. And suppose the system is not quite at the traditional equilibrium point of $\Delta\mu = 0$; suppose that there is a deviation of $\pm 10^{-10}$ from the exact equality in $\Delta\mu/k_{\text{B}}T$. This means that the *exponent* determining the equilibrium constant is $\pm 10^{10}$, so $K_{\text{eq}} = \exp[\pm 10^{10}]$. This tells us something we already knew from the phase rule, that even at such tiny deviations from exact

equality of the chemical potentials, the amount of the unflavored phase, whichever it is, is present in unobservable tiny amounts.

But now let us apply this reasoning to small systems. Suppose we take $N = 20$; then if, for example, $\Delta\mu/k_B T = \pm 0.01$, then $K_{eq} = \exp[\pm 0.2]$, i.e., 1.22 or 0.81. In words, this means that the unflavored or minority form, whether solid or liquid, is present in almost as large an amount as the more stable form. *This way of approaching solid–liquid equilibria shows that the phase rule is strictly a consequence of large numbers, and that its range of validity is that of macroscopic systems.*

That range is something we examine later.

Because the equilibrium constant K_{eq} ranges from zero, when the system is all solid, to infinity, when it is all liquid, (strictly, one should include vaporization, neglected here) it is convenient to introduce another related function, a ratio we call D (for distribution), which contains the same information but ranges from -1 to $+1$: $D = (K_{eq} - 1)/(K_{eq} + 1)$. This allows us to portray graphically the behavior of a system in terms of the amount of each of two phases as a function of temperature. This is done in Fig. 1, for a small system (a), a mid-size system (b), and a large but not truly macroscopic system (c). However, even case (c) in this figure does not

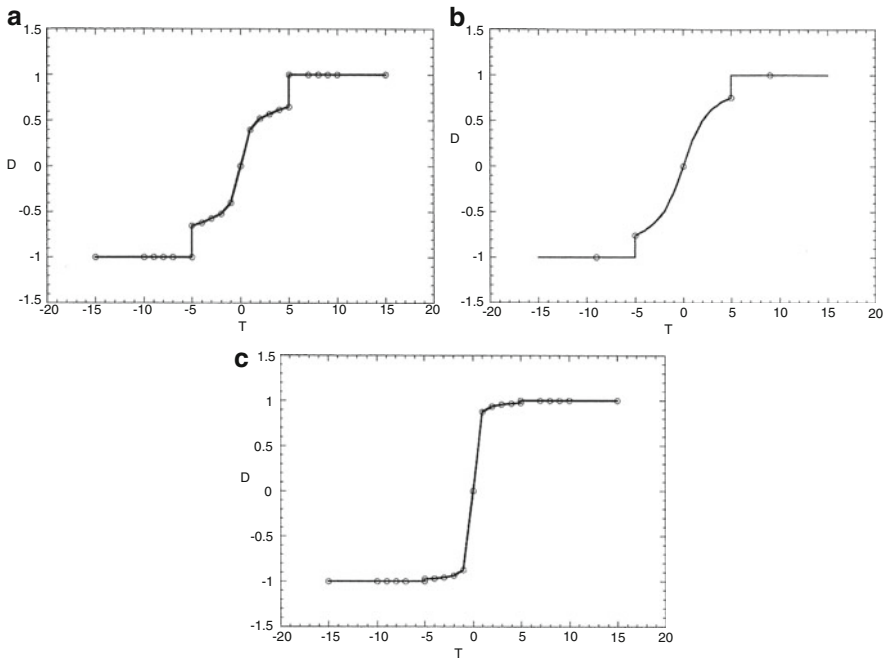


Fig. 1 Schematic behavior of two-phase systems as functions of temperature T . (a) A small system, with relatively large jumps in the distribution D at the freezing and melting limits, and a gradual change in D between these two limits, (b) a somewhat larger system, with smaller jumps and a steeper variation in D with temperature, and (c) a still larger system, for which the jumps in D are not discernible and its variation from -1 to $+1$ is very steep

make the change of D from -1 to $+1$ nearly as sharp as it is for any truly macroscopic system; for any system large enough to be visible to the naked eye, the change would be via a vertical line.

One important point that is virtually hidden in the argument showing that, for a small system, two phases can coexist over a range of temperature and pressure is this: the argument applies equally to more than two phases! So long as their free energies have sufficiently similar values, observable amounts of several phases can indeed coexist.

One particularly vivid example is the cluster of 55 atoms bound by Lennard-Jones forces, effectively Ar_{55} or by metallic binding forces. Besides its solid and liquid phases, this cluster (and others of similar size and constitution) can exhibit a *surface-melted phase* [15, 24, 25]. Strictly, as animations show, the term “surface melting” is not really accurate. In the phase that shows a high mobility of the atoms in the surface layer, the actual motion of almost all of those atoms is a large-amplitude, very anharmonic vibration, while at least one atom is promoted to move rather freely around the shell as a “floater”; the average is about one surface atom in 15 becomes a floater. The floater atom exchanges positions occasionally with an atom in the outer shell, so that, over time, all the surface atoms are, at some time, a floater. This process allows all the surface atoms to permute their positions and eventually to occupy all the surface sites – as one would expect of a liquid.

4 Observability of Coexistence in Apparent Violation of the Phase Rule

The next question is “How large can a cluster be and exhibit a band of coexisting solid and liquid phases?” What determines the size at which the phase rule begins to apply? The answer to this question lies in the property that most distinguishes solids and liquids, the difference in the entropies of the two forms [26, 27]. The approach requires an arbitrary decision as to what constitutes “observable” and the choice made in these references is that the minority phase should be present in an amount at least 10% as much as the majority phase. That means that one looks for conditions for which $0.1 < K_{\text{eq}} < 10$. This, in turn, means that $-2.3 < N\Delta\mu/k_{\text{B}}T < 2.3$. If we set the observable conditions at 1% instead of 10%, these values double. Of course $N\Delta\mu$ or ΔF is zero at the point of equal probability of the two phases, which we shall call, according to tradition, the melting point, where $\Delta E_m = T\Delta S_m$; here, the subscript m indicates the value at that melting point. We can assume that the energy and entropy changes, ΔE_m and ΔS_m in the vicinity of the melting point are very close to their values at the melting point. If we express the free energy change in units of $k_{\text{B}}T$, then $\Delta F = \Delta E/T - \Delta S$. From this, we can define the range of observability for the free energy as δF , the temperature derivative of ΔF evaluated at the melting point, or $\Delta E_m\delta T/T_m^2$. But since $\Delta E_m/T_m = \Delta S_m$, we can write the simple approximate expression $\delta T/T_m = \delta F/\Delta S$. If we set

observability limits at 10% of the minority species, then δF is 4.6, or essentially 5, so we can make the estimate the observability as $\delta T \sim 5T_m/\Delta S$.

Computer-based and model-based estimates of the entropy changes for rare gas clusters allow us to approximate the entropy change per particle, and infer that for a 100-atom cluster of argon atoms, with a criterion of 10%, $\delta T \approx 0.1$ K, and with a criterion of 1%, this extends to about 0.4 K [26]. Metal clusters have much smaller entropy changes when they melt, because of their much higher densities of states in their solid phases, so their ranges of observability of coexistence extend to several hundred particles. Such coexistence ranges have been observed in experiments, e.g., with sodium clusters [28, 29]. The first to be observed exhibiting such coexistence was Na_{139}^+ .

5 Phase Diagrams for Small Systems

Because the solid and liquid phases of small systems can coexist over ranges of temperature and pressure and in varying ratios, a phase diagram for such a system requires one coordinate more than the traditional two-dimensional p - V plot. It is convenient to use the distribution D to define that third coordinate, so that when $D = -1$, the system is entirely solid and when $D = +1$, it is entirely liquid. Figure 2 shows two examples of such expanded phase diagrams: in (a), we have a macroscopic system, exhibiting a discontinuous jump of D between these two values at the melting point. Example (b) is a schematic representation of a small system, for which there are discontinuities at the freezing and melting limits of temperature, but a continuous variation of D between these two discontinuities. In the former, because there are no intermediate values of D , the third coordinate is superfluous,

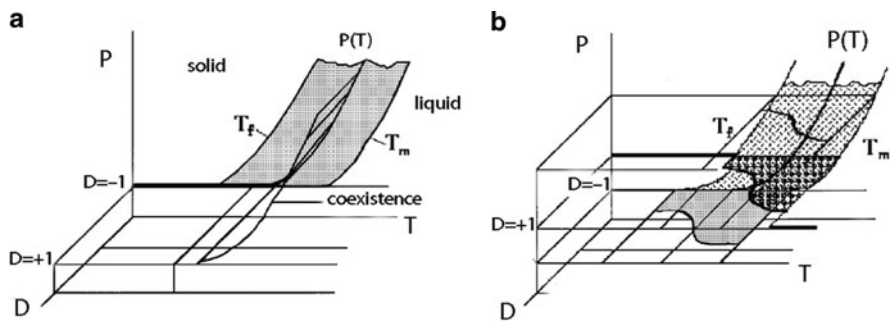


Fig. 2 Schematic phase diagrams for solid–liquid equilibria in three dimensions, pressure p , temperature T , and distribution D ; (a) for a macroscopic system, in which the change from solid to liquid is discontinuous at the melting point $T_m(p)$ and D changes from -1 to $+1$; (b) for a small system, for which D changes discontinuously between -1 and an intermediate value, and again between a much higher intermediate value and $+1$, but varies continuously between those two intermediate values

but in the latter, we need the third variable D in order to show the composition of the system in the coexistence range.

6 Conclusions and Summary Remarks

This review has shown that the Gibbs phase rule, powerful as it is, is a consequence of the large numbers of atomic particles in all macroscopic systems. Small systems, of order hundreds of particles or fewer, violate that rule in that multiple phases can coexist in thermodynamic equilibrium over ranges of temperature and pressure, and still behave in a way consistent with the basic laws of thermodynamics. If we treat phase equilibria just as we treat equilibria of chemical isomers and reacting species, we find that for large systems, the free energy change in the vicinity of the traditional melting point is so sharp that the unflavored phase can be present only in unobservable small quantities at temperatures deviating only tiny amounts from that melting point. However, for small systems, the free energy difference between phases can be small enough that observable amounts of the unfavored minority phase can easily exist under attainable conditions away from the melting point or, more generally, the traditional coexistence curve. Moreover the same reasoning shows that multiple phases may coexist within such ranges. It is possible to estimate from the entropy change between phases what the maximum size is for such coexistence to be detectable. One can construct phase diagrams for the coexistence of two phases that show the relative amounts of each phase as a function of temperature and pressure.

While Carl Johan Ballhausen worked throughout his scientific career on properties of atomic and molecular systems, especially complex ions, he was always fully aware of the thermodynamic aspects of these systems. He also was always determined to “get things right,” whether they fit conventional notions or not. It is in that spirit that this work addresses one of those conventional, accepted concepts and, by examining its realm of validity, strives to “get things right.”

Acknowledgments R.S.B. wishes to acknowledge the hospitality of the Aspen Center for Physics, which enabled the writing of this chapter. He also wishes to recognize with deep appreciation and very dear memories the companionship and stimulation of Carl Johan Ballhausen, from the time they were first together in the research group of William Moffitt at Harvard, in the 1950s until his passing.

References

1. Gibbs JW (1876) *Trans Conn Acad* 3:108; (1878) *Trans Conn Acad* 3:343
2. Roozeboom HW (1887) *Rec Trav Pays-Bas Belg* 6:262
3. Servos JW (1990) *Physical chemistry from Ostwald to Pauling*, Chapter 3. Princeton University Press, Princeton, p 100

4. Daub EE (1976) *J Chem Ed* 53:747
5. Cotterill RMJ, Damgaard-Kristensen W, Martin JW, Peterson LB, Jensen EJ (1973) *Comput Phys Comm* 5:28
6. Cotterill RMJ (1975) *Phil Mag* 32:1283
7. Lee JK, Barker JA, Abraham FF (1973) *J Chem Phys* 58:3166
8. McGinty DJ (1973) *J Chem Phys* 58:4733
9. Damgaard-Kristensen W, Jensen EJ, Cotterill RMJ (1974) *J Chem Phys* 6:4161
10. Briant CL, Burton JJ (1975) *J Chem Phys* 63:2045
11. Eters RD, Kaelberer JB (1975) *Phys Rev A* 11:1068
12. Eters RD, Kaelberer JB (1977) *J Chem Phys* 66:5112
13. Kaelberer JB, Eters RD (1977) *J Chem Phys* 66:3233
14. Amini M, Fincham D, Hockney RW (1979) *J Phys C* 12:4707
15. Nauchitel VV, Pertsin AJ (1980) *Mol Phys* 40:1341
16. Garzon IL, Jellinek J (1991) *Z Phys D* 20:235
17. Jellinek J, Garzon IL (1991) *Z Phys D* 20:239
18. Farges J, de Feraudy MF, Raoult B, Torchet G (1983) *J Chem Phys* 84:3491
19. Bartell LS (1985) *Chem Rev* 86:492
20. Bartell LS, Harsanyi L, Valente EJ (1989) *J Phys Chem* 93:6201
21. Bartell LS, Dibble TS (1991) *J Phys Chem* 95:1159
22. Hill TL (1963) *The thermodynamics of small systems*, Part 1; (1964) Part II. Benjamin, New York
23. Berry RS, Jellinek J, Natanson G (1984) *Phys Rev A* 30:919
24. Cheng H-P, Berry RS (1992) *Phys Rev A* 45:7969
25. Kunz RE, Berry RS (1994) *Phys Rev E* 49:1895
26. Berry RS, Smirnov BM (2009) *Int J Mass Spectrom* 280:204
27. Berry RS (2010) In: Sattler K (ed) *Handbook of nanophysics*, V.1, Chapter 20. CRC Press, Taylor and Francis, Boca Raton, FL
28. Schmidt M, Kusche R, Kronmüller W, von Issendorf B, Haberland H (1997) *Phys Rev Lett* 79:99
29. Schmidt M, Donges J, Hippler T, Haberland H (2003) *Phys Rev Lett* 90:103401