Alleged acausality of the diffusion equations: A reply

Peter Kostädt* and Mario Liu†
Theoretische Physik, Universität Hannover, 30167 Hannover, Germany

(Received 26 February 2001; published 19 September 2001)

Two facts are pointed out: (1) Only when the diffusion equation is inadmissibly applied, outside its defined range of validity, does it lead to acausal predictions; (2) although there are many instances in which usually diffusive hydrodynamic variables propagate wave-like in condensed systems, none of them are connected with this appearance of acausality.

DOI: 10.1103/PhysRevD.64.088504 PACS number(s): 04.40.Nr, 05.70.Ln, 47.75.+f, 95.30.Lz

There are three issues plaguing relativistic hydrodynamics: causality, stability, and uniqueness. In Ref. [1], the paper criticized by the preceding Comment [2], the focus is on the stability. More specifically, by examining the mathematical properties of the relativistic diffusion equation, we found that the Hiscock-Lindblom instability emerges from an ill-posed initial-value problem, which does not meet the mathematical conditions required for realistic physical problems. Therefore, we concluded that this instability cannot serve as an argument against the relativistic hydrodynamics of Landau and Lifshitz.

In the Introduction, when embedding our issue of stability in the context of published works, we did also comment on causality, by quoting results from Geroch [3] and Lindblom [4]. It is this parenthetical remark that has provoked the writing of [2], which deals exclusively with causality. We need to make this point very clear at the beginning, because we are afraid that this Comment, especially its Abstract, may be rather misleading in this respect. In fact, only the very diligent reader who reads all footnotes and does not miss Ref. [13] will realize that it is not the purpose of [2] to challenge our results. Rather, it is (we quote) "*only to focus on the unfortunate quoted phrase which is actually irrelevant to the stability proof.*"

If we did not have severe problems with some statements of [2], we would have ended our reply here. As things stand, however, rebuttals are called for.

We start by outlining our point of view on causality. Strictly speaking, the diffusion equation implies signals with infinite velocity, or horizontal "world lines." While quite generally unphysical, this defect is aggravated by relativity: When viewed from a different frame, the world lines tilt, implying signals that go backward in time.

To repair this, extended thermodynamic theories were put forward, which start from the hydrodynamic theories but include additional dynamic variables. The resultant larger set of coefficients can be chosen such that all the differential equations are hyperbolic, ensuring causality. The price for this nice feature is a rather more complicated theory, and the difficulty of finding a universally valid and accepted set of additional variables, rendering the varying results *ad hoc*.

On the other hand, we may also take a perspective view, and understand that the diffusion equation is not an exact mathematical statement. Rather, it is an approximate description with a clearly outlined range of validity. Two of the more important constraints are the hydrodynamic low frequency regime and an inaccuracy of the variables on the order of the respective thermal fluctuations. Taking these into account, and considering only amplitudes of the variables that are above a minimal threshold, the signal velocity can be shown never to exceed that of the constituent particles [5], excluding any acausal consequences.

For instance, given the initial condition of a temperature peak, the diffusion equation predicts a broadening of the peak that is exponentially decaying in space. So the fast, acausal velocity is confined to exponentially small temperature signals—much smaller than any hydrodynamic theory is meant to account for. Excluding all signals below a minimal threshold, given by, say, one-hundredth of the fluctuation inherent in the temperature, we find the spread of the remaining signals—the physically significant ones—to be universally slow, much slower not only than the light velocity, but also than the velocity of the particles transmitting the heat. We therefore conclude that acausality is a result of pushing the diffusion equation beyond its range of validity.

If one still chooses to free the equation of motion for the temperature of any appearance of acausality, one needs to execute the extension such that the resultant equation (i) agrees with the heat diffusion equation inside its range of validity, where it is one of the best experimentally validated equations, yet (ii) remains explicitly causal outside it, where signals are much smaller than thermal fluctuations. In light of these two constraints, it is hardly surprising that Geroch and Lindblom found, in those cases they considered, that "*the complicated dynamical structure which ensures causality is unobservable.*"

The above statements do not at all imply that we believe that temperature must obey the parabolic diffusion equation in all existing condensed systems—as may be construed from reading [2]—we know it does not. This sentence only means that the above addressed, purely formal acausality is not enough reason to abandon the diffusion equation. And if we do abandon it for no other reason than to ensure causality, the additional and complicated structure is unobservable—in order not to disrupt the agreement with experiments.

*Email address: kostaedt@itp.uni-hannover.de
†Email address: liu@itp.uni-hannover.de

 ©2001 The American Physical Society

Spontaneously broken symmetries are the main reasons for the diffusion equation to be invalid in condensed systems, frequently turning a given diffusion equation into a wave equation. None of these instances is in any way related to the formal acausality or infinite velocity discussed above. In fact, when this happens, the respective variable propagates faster, not more slowly.

Generally speaking, simple fluids with no broken symmetries are accounted for by all its five conserved variables: the densities of energy, mass, and momentum, or, equivalently, temperature T, pressure P, longitudinal velocity $v_\parallel$, and the two components of transverse velocity $v_\perp$. The two variables P and $v_\parallel$ combine to sustain sound waves, while T and $v_\perp$ obey diffusion equations.

Circumstances change when there are broken symmetries [6]. Broken translational symmetry in crystals [7–10] (and liquid crystals [11,12]) gives rise to the displacement U_j as additional variables, which combine with $v_\perp$ to sustain shear waves. In other words, the two $v_\perp$ obey wave equations in crystals rather than diffusion equations. Similarly, the broken phase symmetry in superfluids, both ^{4}He [13,14] and ^{3}He [15–18], gives rise to the phase φ as an additional variable, which combines with the temperature to sustain the so called second sound. So temperature obeys a wave and not a diffusion equation here. These are indeed systems in which some of the diffusion equations are not always valid. Yet $v_\perp$ goes on obeying the diffusion equation in superfluids, just as T does in crystals. Also noteworthy is the fact mentioned above, that T propagates faster, not more slowly, in superfluids; the same applies to $v_\perp$ in crystals.

Another known reason that renders the diffusion equation invalid is long lived yet nonhydrodynamic degrees of freedom, with a finite relaxation time τ. Their existence frequently cause the system to mimic broken symmetry behavior in the high frequency regime $\omega\tau \gg 1$. In the present two cases, we have second sound (i.e., temperature wave) in nearly perfect crystals at low temperatures [19], and shear waves as a result of viscoelasticity in non-Newtonian polymer solutions [20]. One must not confuse these two cases: While second sound is a true hydrodynamic mode in superfluids, it possesses a (frequency) gap (for vanishing wave vector) in crystals: $\omega(q \to 0) = i/\tau$. This is a qualitative difference that the authors of [2] do not seem to have grasped.

Generally speaking, the presence of broken symmetries and long-lived degrees of freedom does modify and complicate the hydrodynamic equations of a condensed system. And it is well conceivable that some parabolic equations become hyperbolic hereby, as in the above cases, or in the many examples given in [2]. The crucial point, however, remains that the modification of the characteristics happens coincidentally—not because the parabolic equations were acausal, or the hyperbolic ones in any other sense preferable, as implied over and over again in [2]. There is no doubt that second sound exists in superfluids and crystals. It is the connection of this fact to the formal acausality of the parabolic diffusion equation that is misguided.

Equally disturbing is the great emphasis put in [2] on the difference between the particles making up the fluid and those transporting the heat. While it is commonplace that these two may, but need not, be the same, there is simply no important difference between the two in the present context: Phonons transmit heat in ^{4}He liquid at low temperatures, a task that in ^{3}He liquid is mainly accomplished by the ^{3}He atoms themselves. Yet both have wave equations for the temperature in their superfluid phase, and diffusion equation in their normal phase [15]. The authors of [2] further assert that the relaxation time τ is frequently confused with the mean collision time t_c, as a consequence of which their approach and results are not appreciated. We do not share their observation: In hydrodynamic theories and irreversible thermodynamic considerations of dense systems, the usual discussion is always about τ alone, without the possibility of confusion, as t_c is not a well-defined and readily available quantity. This is in clear contrast to kinetic theories more appropriate for dilute systems.

We conclude that the preceding Comment has nothing to do with the paper being commented upon, that formal acausality is not a valid reason to abandon the diffusion equation, and that division of all physics into bad, parabolic and good, hyperbolic versions is an endeavor devoid of any experimental support and rational justification.

We summarize and comment upon some of the results of our original paper [1], although this is unrelated to the preceding Comment [2]. Our main result follows: The initial-value problem for the relativistic diffusion equation is well posed only if the initial data are provided for the proper time slice, $\tau = \text{const}$, which is the characteristics of the differential equation; it is ill posed if one takes any boosted time slice, $t = \text{const}$, none of which is a characteristics. One needs to be careful drawing physical conclusions from this mathematical result. We do not think it means that the relativistic diffusion equation is defunct. As there is a medium in any macroscopic theory, the existence of which clearly breaks Lorentz invariance, the capability of accommodating arbitrary initial data on boosted time slices is not a *sine qua non* for such a theory to be healthy and acceptable.

We all know that in nonrelativistic, coarse-grained, macroscopic theories the initial-value problems of spatial coordinates are similarly ill posed, even in the rest frame. (In fact, the relativistic, temporal ill-posedness is nothing but the Lorentz-transformed spatial one; see [1].) No one considers this fact remotely unhealthy, as one may simply stick to those initial spatial data that do not lead to obviously nonphysical results, such as field variations on scales smaller than the descriptive grains, or with an amplitude either too large or too small, either unbound at infinity or comparable to fluctuations [21]. We expect this pragmatic operative prescription to be just as helpful in the relativistic context—it certainly suffices to rule out the Hiscock-Lindblom instability, which is both unbound for $t \to \infty$ and varies on microscopically small time scales. We therefore conclude: *The relativistic diffusion equation does possess an initial-value formulation, although the initial values (if provided on a noncharacteristic line such as a boosted time slice) are not completely arbitrary*.

In our paper, we have explicitly omitted the case of non-uniform motion, in which the medium does not possess a unique rest frame. (This omission is relevant if the velocity differences are relativistically large. Because the term ''velocity'' implies the mean velocity of volume elements in macroscopic, coarse-grained theories, large differences could only happen in astronomically large objects.) Nevertheless, we believe that the above operational prescription to disregard any obviously nonphysical results should also be useful here.

[1] P. Kostädt and M. Liu, Phys. Rev. D **62**, 023003 (2000).
[2] L. Herrera and D. Pavon, preceding Comment, Phys. Rev. D **64**, 088503 (2001).
[3] R. Geroch, J. Math. Phys. **36**, 4226 (1995).
[4] L. Lindblom, Ann. Phys. (N.Y.) **247**, 1 (1996).
[5] S. R. de Groot, W. A. van Leeuwen, and Ch. G. van Weert, *Relativistic Kinetic Theory* (North-Holland, Amsterdam, 1980), Ch. VI, Sec. 1e.
[6] D. Forster, *Hydrodynamic Fluctuations, Broken Symmetry and Correlation Functions* (Benjamin, New York, 1975); I. E. Dyaloshinskii and G. E. Volovik, Ann. Phys. (N.Y.) **125**, 67 (1980).
[7] L. D. Landau and E. M. Lifshitz, *Theory of Elasticity* (Pergamon, Oxford, 1986).
[8] P. M. Chaikin and T. C. Lubensky, *Principles of Condensed Matter Physics* (Cambridge University Press, Cambridge, England, 1995).
[9] P. C. Martin, O. Parodi, and P. S. Pershan, Phys. Rev. A **6**, 2401 (1972).
[10] P. D. Fleming III and C. Cohen, Phys. Rev. B **13**, 500 (1976).
[11] P. G. de Gennes, *The Physics of Liquid Crystals* (Clarendon, Oxford, 1974).
[12] H. Pleiner and H. R. Brand, in *Pattern Formation in Liquid Crystals*, edited by A. Buka and L. Kramer (Springer, New York, 1996).
[13] L. D. Landau and E. M. Lifshitz, *Fluid Mechanics* (Pergamon, Oxford, 1987).
[14] I. M. Khalatnikov, *Introduction to the Theory of Superfluidity* (Bejamin, New York, 1965).
[15] D. Vollhardt and P. Wölfle, *The Superfluid Phases of* ^3He (Taylor and Frances, London, 1990).
[16] T. L. Ho and N. D. Mermin, Phys. Rev. Lett. **44**, 330 (1980).
[17] C.-R. Hu and W. M. Saslow, Phys. Rev. Lett. **38**, 605 (1977).
[18] M. Liu and M. C. Cross, Phys. Rev. Lett. **41**, 250 (1978); **43**, 296 (1979); M. Liu, *ibid.* **43**, 1740 (1979).
[19] M. Liu, Phys. Rev. B **18**, 1165 (1978), and references therein.
[20] H. Temmen, H. Pleiner, M. Liu, and H. R. Brand, Phys. Rev. Lett. **84**, 3228 (2000), and references therein.
[21] A good example is the so-called sq mode in ^{4}He, but not the same mode in ^{3}He, cf. M. Grabinski and M. Liu, Phys. Rev. Lett. **58**, 800 (1987); J. Low Temp. Phys. **73**, 79 (1988); Y. Sun, P. Wölfle, and S. Yip, Phys. Rev. Lett., **63**, 1613 (1989); P. Kostädt and M. Liu, *ibid.* **71**, 3513 (1993).