The Diffusion Equation Does Not Imply Instantaneous Action at a Distance*

A. R. Jadhav† and S. R. Kajale
College of Engineering (COEP), University of Pune, India
(Dated: 18–21 December 2006)

There are three astounding matters related to the Fourier theory of heat conduction: (i) the prediction that heat spreads with infinite speed—i.e. instantaneous action-at-a-distance (IAD); (ii) the explanation that the cause for this IAD lies with the diffusion equation itself—not with the solution theory; and (iii) the current practice that refrains from mentioning anything at all about this issue in the textbooks. Our viewpoint here happens to be at odds with all the three. In this paper, we begin by outlining Fourier's analytical theory of heat and Einstein's stochastic theory of diffusion. We then show precisely how the implication of IAD arises in the Fourier theory. After looking into the physical bases of the mathematical constructs, we reject the idea of IAD as *non sequitur*. Finally, we raise a couple of novel questions, and in discussing them, also propose a discriminant to indicate whether a given theory carries IAD or not.

I. THE DIFFUSION EQUATION

Consider an infinitesimal cubic element within a solid that has internal temperature gradients. Energy balance on per unit time basis yields:

$$\overrightarrow{\nabla} \cdot \left(k \overrightarrow{\nabla} \theta \right) + \delta \dot{q} = \frac{\partial \theta}{\partial t} \quad (1)$$

where k is thermal conductivity, θ is temperature, $\dot{q}$ is heat power, ρ is mass density, C_P is mass-specific heat capacity at constant pressure and t is time.

Reading from left to right, the terms of Eq. 1 are: Fourier's law of heat conduction; the non-functional forms of heat power generated in the elemental cube; and the rate of increase in the internal energy of the element. Now, if thermal conductivity k remains constant at all temperatures and $\dot{q}$ can be neglected, we get Fourier's heat equation (c. 1807):

$$\nabla^2 \theta = \frac{1}{\kappa} \frac{\partial \theta}{\partial t} \quad (2)$$

where

$$\kappa = \frac{k}{\rho C_P} \quad (3)$$

In the formulation of Eqs. 1 and 2, two ideas have been used, both on a time-rate basis: (i) a flux-to-gradient functional relationship (Fourier's law of conduction) and (ii) a conservation principle (the first law of thermodynamics). If, instead, Fick's first law of mass transport and conservation of mass-flow are used, we get the diffusion equation (c. 1855). Equations of exactly the same form are also satisfied by vorticity in a viscous fluid put into motion starting from a state of rest, and by electric field vector in the propagation of long waves in a good conductor. Finally, Schrödinger's equation also is formally nothing but a diffusion equation.

II. THE CLASSICAL TECHNIQUE TO SOLVE THE DIFFUSION EQUATION

In this section we outline the solution technique first put forth by Fourier (c. 1807) and generalized by Ostrogradski (c. 1828). Except for passing remarks in [1] and [2], textbooks do not discuss IAD in the context of diffusion at all. As such, it is necessary to recapitulate the Fourier theory of diffusion here, with a view to show precisely how the idea of IAD becomes a part of it. We therefore note many points of physical basis and interpretation that are not found discussed even in physics texts.

A. Separating the Variables—in 3D Space

Following d'Alembert (c. 1749), assume that the solution to Eq. (2) can be expressed as:

$$\theta(x, y, z, t) = X(x)Y(y)Z(z)T(t) \quad (4)$$

Substitute Eq. 4 into 2 to get:

$$\frac{1}{X}\frac{d^2X}{dx^2} + \frac{1}{Y}\frac{d^2Y}{dy^2} + \frac{1}{Z}\frac{d^2Z}{dz^2} = \frac{1}{\kappa T}\frac{dT}{dt} \equiv \varsigma \quad (5)$$

where ς is the separation constant.

For nontrivial solutions ς is required to be negative, say $\varsigma = -\sigma^2$ where $\sigma \in \Re$.

The LHS is a sum of X-, Y- and Z-involving terms, and equals ς. Therefore, the three x-, y- and z-related Helmholtz equations should be coupled to each other. This requirement can be met if we express σ^2 as a sum of three constants α^2, β^2 and χ^2, one each for a space axis:

$$\alpha^2 + \beta^2 + \chi^2 = \sigma^2 \quad (6)$$

Now, using these constants, we can express the three coupled Helmholtz equations as:

$$\frac{d^2X}{dx^2} + \alpha^2 X = 0 \quad (7a)$$

$$\frac{d^2Y}{dy^2} + \beta^2 Y = 0 \quad (7b)$$

$$\frac{d^2Z}{dz^2} + \chi^2 Z = 0 \quad (7c)$$

As to the time-dependence, only a rearrangement of the terms in Eq. 5 is necessary:

$$\frac{dT}{dt} + \sigma^2 \kappa T = 0 \quad (8)$$

---

*Published in the *Proceedings of the 51st Congress of the Indian Society of Theoretical and Applied Mechanics (An International Meet)* held at AUCE, Visakhapatnam, India, during 18–21 December 2006, pp. 62–72. *Ed.* B. Pradhan. *Pub.* ISTAM, IIT Kharagpur, India.
†Electronic address: ToneBrush@vsnl.net

B. Recognizing the Form of the Solution

The space-dependent part, i.e. Eq. 7a for example, has general solution of the form:

$$X(x) = C_x \cos(\alpha x) + S_x \sin(\alpha x) \tag{9a}$$

where C_x and C_y are constants. In contrast, for the time-dependent part i.e. Eq. 8 we have:

$$T(t) = \Theta_0 \exp\left(-\sigma^2 \kappa t\right) \tag{10}$$

where Θ_0 is a constant. Now, substituting Eqs 9 and 10 into 4, and then using Eq. 6:

$$\begin{aligned} \theta(x, y, z, t) = \Theta_0 & \\ & \exp\left(-\alpha^2 \kappa t\right) \left[C_x \cos(\alpha x) + S_x \sin(\alpha x)\right] \\ & \exp\left(-\beta^2 \kappa t\right) \left[C_y \cos(\beta y) + S_y \sin(\beta y)\right] \\ & \exp\left(-\chi^2 \kappa t\right) \left[C_z \cos(\chi z) + S_z \sin(\chi z)\right] \end{aligned} \tag{11}$$

Eq. 11 clearly shows that the 3D solution is nothing but a simple product of the individual 1D solution terms despite the coupling mentioned above. Therefore, assuming certain care, we are free to discuss only 1D domains without any loss of generality. (Further, if α^2, β^2 and χ^2 are not all positive, sinh and cosh terms enter the form of the solution. However, ignoring such a possibility does not affect the generality of the conclusions we draw in this paper.)

Accordingly, assuming $\Theta_0 = 1$, the solution form in 1D is given as:

$$\theta(x, t) = \left[C_x \cos(\alpha x) + S_x \sin(\alpha x)\right] \exp\left(-\alpha^2 \kappa t\right) \tag{12}$$

Now, rather than rush to the usual next steps of stating the Fourier series and absorbing some particular boundary conditions in it, let us first make sure we understand the basic form of Eq. 12 itself.

First, consider the space-dependent part which has come straight from Eq. 9. It is obvious that α is the wavenumber, i.e. the spatial equivalent of angular frequency. (It is defined as $\alpha \equiv \left|\left(\frac{\partial \phi}{\partial t}\right)_t\right| = \frac{2\pi}{\lambda}$). Accordingly, the square bracket in Eq. 12 represents a spatial waveform—i.e. an undulation frozen in time. It consists of a superposition of cosine and sine components of the same wavelength $\lambda = \frac{2\pi}{\alpha}$ but different scaling coefficients. Thus the solution has a certain characteristic length associated with it.

Secondly, consider the time-dependent part of Eq. (12). It consists of an exponential decay term—not sinusoidal. Consequently, the time evolution consists of nothing but a uniform stretching down of the entire undulation. In particular, unlike the waves implied by the wave equation proper, the *overall* waveform here does not move axially. Instead, with time, it simply attenuates while staying in the same place. The waveform vanishes in the limit that $t \to \infty$.

Due to the periodicity of the sine and cosine functions in Eq. (9a), any of their higher harmonics also itself is a solution. Further, due to the linearity of the governing Helmholtz equation, a sum of any two scaled solutions also itself is a solution. Therefore, the most general form of the space-dependent part can be expected to involve an infinite sum. But how about the overall diffusion solution? It too can be expected to be an infinite sum. Why? The reason is separated variables. In Eq. (12), even though the decay constant includes a dependence on α, for any particular j, the role of $T(t)$ is limited to being a simple multiplying factor.

C. The Fourier Series, Periodicity and Diffusion

If the domain is finite, there is a natural bound on the longest wavelength that can at all fit into it and still satisfy the imposed boundary conditions. The solution then assumes the form of the Fourier series:

$$\theta(x, t) = \sum_{j=0}^{\infty} \exp\left(-\alpha_j^2 \kappa t\right) \left[C_{x_j} \cos(\alpha_j x) + S_{x_j} \sin(\alpha_j x)\right] \tag{13}$$

where the constants C_{x_j} and S_{x_j} are to be determined using the Euler's formulae (c. 1777):

$$C_{x_0} = \frac{1}{L} \int_{-L/2}^{L/2} X(x)\, dx \tag{14a}$$

$$C_{x_j} = \frac{2}{L} \int_{-L/2}^{L/2} X(x) \cos\left(j \frac{2\pi x}{L}\right) dx; \tag{14b}$$

$$j = 1, 2, 3, \ldots \tag{14c}$$

$$S_{x_j} = \frac{2}{L} \int_{-L/2}^{L/2} X(x) \sin\left(j \frac{2\pi x}{L}\right) dx;$$

$$j = 1, 2, 3, \ldots$$

Here, the desired solution is assumed to be periodic over $[-L/2, +L/2]$. Paradoxically enough, though θ has thus been "solved" for, the function $X(x)$ still remains an unknown. The circularity between the Eqs 13 and 14 is prevented if $X(x)$ is set equal to the initial temperature distribution.

For finite domains, the technique relies on imagined periodic replication of the spatial waveform on each side, "up to" infinity. But why extend the profile ad infinitum? Apparently, to avoid modulation arising due to edge effects. This is fine. But note that the addition of the profiles is necessitated by the solution technique—not by the physics of the problem statement.

So, a natural question to raise is: Doesn't the procedure of adding an infinity of profiles implicitly bring in some additional physical assumptions? On this point, no definitive answer has ever been put forth in particular terms.

Indeed, no reason has ever been offered as to why a solution technique which is based on periodic phenomena should at all be used in finding solution of an essentially nonperiodic problem—except for the pragmatic justification that "it seems to work." It is important to note here that the Fourier series or integral does not constitute a general solution to any differential equation. The Fourier representation is just a more easily workable approach to remodel the given problem into different terms.

D. Boundary Conditions and Evolution in Time

Experience tells that to maintain generality and robustness in the solution procedure, boundary conditions should be applied before initial values are.

The boundary conditions basically determine the particular values of α_j's.

To take a concrete example, if the temperature at two endpoints remains constant at all times, then $C_j = 0$ $\forall j \in \aleph$. In such a case, only the sine terms remain in the solution, and the values of the α_j's are determined by the width of the domain: $\alpha_j = \frac{2\pi j}{L}$.

The initial values basically help determine the values of the Fourier coefficients C_{x_j} and S_{x_j}.

Once the initial value profile is represented in a Fourier series, how does the representation evolve in time? The answer is that due to the orthogonality property, each spatial harmonic decays *completely independent* of all others, with a speed determined by α_j. We can even associate a separate half-life with each Fourier component: $T_{j_{\mathrm{half}}} \approx \frac{0.693}{\alpha_j^2 \kappa} \approx 0.0176 \frac{\lambda_j^2}{\kappa}$. Thus, spatial harmonics of smaller wavelengths have a comparatively shorter half-life and therefore decay faster.

E. Infinite Domains

If the domain is infinite and the initial value distribution periodic, obviously the solution procedure is what was spelt out above. If the domain is infinite and the initial value profile *nonperiodic*, the solution is given by a Fourier integral.

For some strange reason, in the extremely rare case when the IAD is at all discussed in the context of diffusion, an infinite domain is invariably found assumed. However, infinitude of domain is not necessary either to pose the question of IAD or to answer it. Infinite speeds can always be theorized to occur over finite domains. We, therefore, need not go into the more complicated details of Fourier integrals, transforms or convolutions.

Instead, we will now take a look at an entirely different and fresh approach to diffusion phenomena—the stochastic approach.

III. STOCHASTICS: EINSTEIN'S DERIVATION OF THE DIFFUSION EQUATION

The discussion in this section is mainly based on Einstein's paper [3]. In this paper, Einstein addresses the 1D case.

The total particle population is a finite number n. The continuous-time Brownian motion of the particles is theoretically studied at two instants $t = 0$ and $t = \tau$, and inferences drawn from certain collective characteristics of the particle motions existing at those two instants. The time interval τ is sufficiently large that none of the particles can continue its motion unaltered across that duration. In other words, the motions of the same particle observed at the two instants can be assumed to be *mutually independent* of each other. Yet, τ is "very small compared with observable time intervals." Thus, τ is small enough that Einstein can later introduce a Taylor series expansion to the first order in τ.

Why this peculiar nature of τ? Why not make it infinitesimal ($d\tau$) straight away? It seems that Einstein here is anticipating the fractal nature of the Brownian movement—and ingeniously circumscribing the attendant complexity. It is interesting to see a 1950's description of fractals that Joos & Freeman give without using the word fractal; pp. 599 in [5]. Einstein avoids analysis in terms of velocity because, as $\tau \to 0$, velocity can become indefinitely large due to abruptness and the continual change of directions. Instead, he chooses to analyze the motion in terms of *displacements*.

Accordingly, suppose we record the displacements ξ's undergone by a single particle over many successive time intervals of τ duration each. The record would be the same as that for the displacements undergone by a large number of particles over a single time interval τ. As the number of trials tends to ∞, the ξ values will not approach a unique number. As Einstein supposes, it will give rise to a probability density function (PDF), say $f(\xi)$. Thus, for example, the number of particles that undergo displacement magnitudes between ξ and $\xi + d\xi$ during the time interval τ can be given as $dn = nf(\xi)d\xi$.

Note for later reference that Einstein clearly mentions his anticipation that the actual PDF will have nonzero values only for very small values of ξ.

We are now interested in finding out how the local density at a fixed location x evolves with the passage of time. Assume that the density N at a point depends only on the position x of that point and time t, i.e. $N = N(x,t)$.

Consider an infinitesimal element of width dx situated at the position x. The number of particles found in that element at a later time $t = \tau$ can be obtained using the PDF $f(\xi)$:

$$N(x, t+\tau)\, dx = dx \int_{-\infty}^{\infty} N(x+\xi, t)\, f(\xi) d\xi \qquad (15)$$

In Eq. 15, note how the immediately past densities at all the distant points $x + \xi$ have been translated into the currently existing density at the local point. (The parallel to the convolution integral is obvious.) The aforementioned mapping occurs over the finite interval τ.

Now, on LHS, expand N in the powers of τ. On RHS, we can expand N in the powers of ξ because the $f(\xi)$ is negligible at large values of ξ. We then get:

$$\begin{aligned} N(x,t) &+ \tau \frac{\partial N(x,t)}{\partial \tau} + \dots \\ &= \int_{-\infty}^{+\infty} \begin{bmatrix} N(x,t) + \xi \frac{\partial N(x,t)}{\partial \xi} \\ + \frac{1}{2}\xi^2 \frac{\partial^2 N(x,t)}{\partial \xi^2} + \dots \end{bmatrix} f(\xi) d\xi \end{aligned} \qquad (16)$$

Now, consider RHS. As an axiom of probability theory, $\int_{-\infty}^{+\infty} f(\xi) d\xi = 1$. (It can be taken as an expression of conservation of particles.) Therefore, the first integral term on RHS evaluates to $N(x,t)$ which then cancels out with the first term on LHS.

For the second integral term on RHS, observe that at the location x, the derivatives $\frac{\partial N}{\partial \tau}$, $\frac{\partial N}{\partial \xi}$, etc. have definite values—they are neither singular nor dependent on ξ. So, such terms can be taken out of the integral sign. Now, the integral $\int_{-\infty}^{+\infty} \xi f(\xi) d\xi$ evaluates to zero because $f(\xi)$ is an even function whereas ξ obviously is odd. With these observations, Eq. 16 simplifies to:

$$\begin{aligned} \frac{\partial N}{\partial \tau} &= \frac{1}{2\tau} \frac{\partial^2 N}{\partial \xi^2} \int_{-\infty}^{+\infty} \xi^2 f(\xi) d\xi \\ &= \frac{\langle \xi^2 \rangle}{2\tau} \frac{\partial^2 N}{\partial \xi^2} \end{aligned} \qquad (17)$$

where $\int_{-\infty}^{+\infty} \xi^2 f(\xi) d\xi = \langle \xi^2 \rangle$

Now, it can be demonstrated [5] that the mean square of "displacement" $\langle \xi^2 \rangle$ is proportional to the time interval of observation τ so that $\frac{\langle \xi^2 \rangle}{2\tau}$ is a constant. Accordingly, we can say:

$$\frac{\partial^2 N}{\partial \xi^2} = \frac{1}{D} \frac{\partial N}{\partial t} \tag{18}$$

where $D = \frac{\langle \xi^2 \rangle}{2\tau}$.

Observe that even though $N = N(x,t)$ the diffusion law is formulated in terms of stochastic displacements ξ, not position x. Even though both carry dimensions of length, ξdoes not denote an ordinary variable. It does not stand for a unique displacement undergone by a single particle—no such a unique limit exists. ξ instead refers to the dummy stochastic variable as used in, e.g., Eq. 15.

Conceptually, the right procedure is to first fix the point which will act as a sink. Then, ξ refers to the relative distances over which other points along the x-axis would turn active as sources during the time interval τ. Though in mathematical literature it seems fashionable not to use physical terms like sources and sinks, such precisely is the meaning implied by the mathematics.

As another important point, the diffusion law is insensitive to the particular form of $f(\xi)$[6].

IV. INSTANTANEOUS ACTION-AT-A-DISTANCE IN THE FOURIER THEORY

To illustrate how the Fourier theory implies infinite velocity, Greenberg [1] takes the example of an infinite 1D rod with the Heaviside step function as the initial distribution. But, as noted above, we can always consider a finite domain instead.

Suppose the initial distribution in a rod of length is "triangular" as shown in Fig. 1.

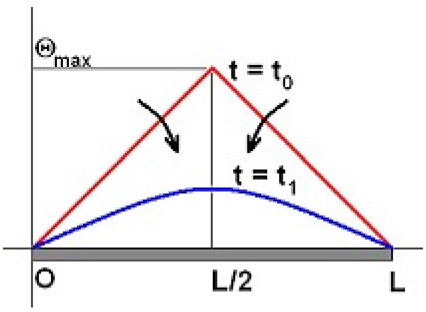

FIG. 1: Temperature evolution in a 1D rod of finite length. According to Fourier's theory, the sharp tip of an initially "triangular" distribution is lost in an infinitesimal time. This prediction implies IAD.

For finite domains, it is necessary to supply boundary conditions. Accordingly, suppose the rod is laterally insulated and both its endpoints are held at 0 temperature at all times. Let the initial temperature at the midpoint be $\theta_{\max}$. Considering the odd extension of the waveform, the Fourier representation is [4]:

$$\theta(x,t) = \theta_{\max} \frac{8}{\pi^2} \left\{ \begin{array}{c} \sin\left(\frac{\pi x}{L}\right) \left[\exp - \left(\frac{c\pi}{L}\right)^2 t\right] \\ -\frac{1}{3^2} \sin\left(3\frac{\pi x}{L}\right) \left[\exp - \left(3\frac{c\pi}{L}\right)^2 t\right] \\ + \dots \end{array} \right\} \tag{19}$$

The time evolution is shown schematically in Fig 1. Note that the sharp tip of the initial triangular profile is lost within an arbitrarily small time interval $\Delta t \to 0$. How come?

First, notice that near the sharp tip, the local radius of curvature approaches zero. Considering partial sums of the Fourier series, the sharp tip would be increasingly better represented in the limit that the wavenumber of the highest harmonic tends to zero. But, as the wavenumber approaches zero, so does the time taken by that harmonic component to decay down to vanishingly small magnitudes. (For example, the half-life of that harmonic approaches zero.)

Now, what does such a mathematics imply for the physics of the problem? Note that the local temperature never increases. Therefore, the heat from the midpoint has no other path to take but the one which goes over every part of the finite portion lying before the endpoint. But the midpoint loses the sharp tip in vanishingly small time. Therefore, heat must travel the entire distance $L/2$ in a vanishingly small time—i.e. with an infinite speed.

As another example, if a corner of a white blotting paper even momentarily touches a drop of blue ink, the entire blotting paper ought to turn blue in the same instant. The corner that touches the drop will stay darker than other portions. Yet, a *finite* portion of ink should travel all over the paper to turn it blue *everywhere*, and *in no time at all.*

Why does the Fourier theory predict this unexpected kind of behavior? The reason is found in the nature of the mathematical method employed for analysis—not in the physical observations of reality.

To see the mathematical method more closely, first consider the case that the initial spatial profile is a single sine wave. With infinitesimal passage of time, its time-dependent part $T(t)$ begins to attenuate exponentially.

At any instant, $T(t)$ *applies equally well at all the values of* x. This is an inexorable result of separating the desired solution as a product of space- and time-dependent parts, i.e., Eq. 4. Consequently, with time, *every* spatial part of a particular harmonic component gets stretched down *simultaneously.* The final solution consists of nothing but a linear superposition of all such components.

Further, inasmuch as boundary conditions determine eigenvalues, *primary* sources or sinks cannot be taken to exist everywhere. Thus, the theory provides no systematic means by which one can remove the simultaneity of decay at different points. Indeed, the famous orthogonality property ensures so. As a result, the final superposed solution shows the same characteristics. The implication of IAD is thus built into the very fabric of the mathematical theory.

The small footnote in [1] ascribes IAD to the diffusion equation. The reasons for this is not apparent to us. On the contrary, it is the Fourier technique which is inadequate to solve the equation.

V. REJECTING INSTANTANEOUS ACTION-AT-A-DISTANCE

As brought out in the earlier sections, IAD simply leads to contradictions. To reiterate, IAD comes about from a certain mathematical theory—not a set of physical observations or principles.

Now, the nature of the Fourier theory is such that it does not fundamentally solve the diffusion equation at all—it only remodels the problem in different terms. The rest of the "theory" consists of little more than fitting of curves to (theoretically posed) data points. But neither the terms nor the structure of the theory is fundamentally suited to diffusion [7]. Thus, IAD can be rejected as a *non sequitur* if one begins from the physical observations and principles.

Secondly, from a meta-reasoning viewpoint, the premise of IAD introduces the irremovable kind of incompleteness and redundancy into any physical theory that carries it.

To see the in-principle incompleteness, consider this question: "If heat is transported instantaneously, obviously the flux-gradient law is incomplete. So, what other mechanism is responsible for the gradual evolution of the temperature field?" Suppose the answer is X. Then, the same question can be repeated, now in relation to X. Infinite regress then follows.

To see how redundancy is brought in, consider the argument: "If heat can flow in no time, why think of any finite-time mechanism to account for any other aspects of diffusive transport?"

Thus, either ways, it is the explanation in de-finite terms which becomes redundant, useless, and finally cut off from the very theorization process. This may sound very far-fetched and not possible in the open environment of scientific research. But it isn't. For example, even though the idea of a finite support for the Fourier eigenfunctions is so simple to propose, it has been absent in the two centuries of the Fourier theory. (The recent wavelet theory does not fit the bill; see the next section.)

To conclude, the idea of IAD ought to be rejected out of the physical sciences precisely for the same reason that perpetual motion machines once were. Furthermore, even in the "abstract" mathematical sciences, IAD ought to be carefully identified as an inapplicable aspect of an inadequate theory.

VI. A COUPLE OF NOVEL QUESTIONS

There are some very interesting issues still left even after rejecting IAD. Consider the following:

Q1 *If IAD is implied by the Fourier theory, how come it doesn't come up in Einstein's theory? Or does it?*

Q2 *Since IAD seems to be built into the very fabric of the Fourier theory, rather than attempt modifying it, would it be possible to fundamentally extend the theory in such a way that IAD is no longer implied by it?*

Before we begin answering the above novel questions, note the idea of the support of a function f. It is defined [8] as the smallest closed set $S = suppf \subset X$ such that the values of the function f are zero everywhere on the complement $X\ S$. In other words, S is the closure of the set of all points $x \in X$ for which $f(x) \neq 0$.

The idea of support is what we mean when we talk of the "width" of a probability distribution.

To answer Q1, first consider an infinite domain.

The time interval τ is known to be finite. Therefore, so long as the PDF $f(\xi)$ has a finite width, the particle cannot attain an infinite speed. Yet, the derivation of Eq. 19 shows that it is insensitive to the particular form of $f(\xi)$. In particular, Eq. 19 is insensitive to the *width* of $f(\xi)$. The diffusion law can, in principle, be obeyed irrespective of whether the displacement ξ is finite or infinite.

Einstein did anticipate the actual PDF to drop to zero in a short distance. Physically, for the Brownian movement of a large particle surrounded on all sides by fluid molecules, it is difficult to think of a PDF having infinite tails. In any case, IAD has never been suspected to occur in the Brownian movement. Yet, nothing exists in the mathematics of the theory to prevent IAD from taking a hold in it. Indeed, all prominent PDFs do have infinite tails.

Therefore, introducing the "spooky action at a distance" in one of Einstein's own theories is very easily possible!

Next, consider the Brownian movement in a *finite* domain. Now the width of the PDF can be either smaller than or equal to width of the domain. If the two widths are equal, IAD can neither be affirmed nor denied. However, if the width of the PDF is smaller than that of the domain, then IAD can no longer be maintained.

With these observations, we now have a discriminant to tell if IAD is present in the theory or not. Note that for IAD to come about, it is irrelevant whether the diffusing matter is discrete or continuum. It is also irrelevant whether the initial value is spread for a finite or infinite extent. What matters is the width of the PDF.

Thus, an interesting question to raise here is: why must the support of the binomial distribution expand as the number of trials tends to infinity—why can't a finite but smooth curve be its limiting case.

Now, to answer Q2, if the stochastic theory can have IAD removed, then the same should also be possible in the analytical theories. Let us try to project this idea in conceptual terms. In case of a point source emitting a unit Dirac pulse the support of the density function spreads instantaneously due to the wavenumber-dependent decay of each Fourier component. Now, what needs to be done if such a spread is to be restricted to *finite* extents at all times?

In answer: The solution would have to be represented as an evolution consisting of a set of Fourier series/integral each of which is defined over the instantaneously existing finite support. The wavelength of the fundamental harmonic of each instantaneous Fourier series/integral will equal the extent of the support existing at that instant. For a given initial value distribution, the flux (i.e. energy) conservation principle should help determine the instantaneous extent of supports. Thus, it might be possible to combine the Fourier theoretical ideas and finite speed of propagation.

Note that the finitude of support is the important issue here—not orthogonality or "eigen-ness" of decomposition.

In this context, a few comments on the recent wavelet theory [9] are in order. The wavelet theory uses flexible window widths and leads to a "local" Fourier analysis [10]. Yet, notice that the support of the analyzing function is not always compact; e.g. the Meyer wavelet [11]. Popular literature notwithstanding, the wavelet theory is not, strictly speaking, a generalization of the Fourier theory. The two theories are analogous in many ways—that's all. It would be worthwhile to ponder why FFT has complexity of $O(n \log_2 n)$ whereas DWT has only $O(n)$. The wavelet concept of "scale" does not exactly generalize the Fourier concept of the frequency of an eigenfunction. It is not at all clear how effective the wavelet decomposition would be in representing the diffusion process. In the recent two decades of research on wavelets, none seems to have used them for modeling diffusion.

Overall, it appears that the wavelet theoretical ideas may not be of much help in removing IAD from the Fourier theory. On the contrary, wavelets may imply IAD in their own way, and if so, the theory may even benefit from the present ideas on how to detect and remove IAD.

To conclude this section, we express the hope that mathematicians would look into the many interesting issues that arise in pursuing the above two questions.

VII. FINALE

In this paper, we re-examined the issue of IAD in the pre-EPR context of diffusion, and reached certain novel explanations and results. We began by looking into the details of two distinct approaches: (i) Fourier's analytical theory (c. 1807), and (ii) Einstein's stochastic theory (c. 1905). We showed that IAD arises in the Fourier theory only due to its mathematical method and structure. In particular, each spatial harmonic component of the Fourier series simultaneously suffers exponential decay at all points of space. This fact naturally leads to IAD. Thus, the idea of IAD gains weight not because some "counter-intuitive" physical facts directly support it but because explanations have not been forthcoming as to how Fourier's is a fundamentally ill-fitting theory to model phenomena like diffusion. We also looked into some meta-reasoning and concluded that IAD should be rejected out of the physical theory. Thereafter, we posed a couple of novel questions to ascertain the conditions that might introduce IAD in, say, Einstein's own stochastic theory. In answer, we noted the role of the support (or width) of $f(\xi)$ as the discriminant. Finally, we briefly discussed what it would take to remove IAD from the Fourier theory, and why the answer does not seem to lie in wavelets (c. 1984).

[1] Greenberg, M. D. (1998) *Advanced Engineering Mathematics, 2nd ed.*, Pearson Education (Singapore), New Delhi

[2] Shercliff, J. A. (1977) *Vector Fields*, Cambridge University Press, Cambridge, UK

[3] Einstein, A. (1905) "On the movement of small particles suspended in a stationary liquid demanded by the molecular-kinetic theory of heat," *Ann. d. Phys.*, vol. 17, pp. 549–560

[4] Kreyszig, E. (2004) *Advanced Engineering Mathematics, 8th ed.*, John Wiley & Sons (Asia), Singapore

[5] Joos, G. and Freeman, I. (1958) *Theoretical Physics, 3rd ed.*, Blackie & Son, Glasgow, UK; reprint (1986) Dover, Mineola, NY, USA

[6] Huang, K. (2001) *Introduction to Statistical Physics*, Taylor & Francis, London, UK

[7] Nelson, E. (2001) *Dynamical Theories of Brownian Motion, 2nd ed.*, Princeton University Press, Princeton, NJ, USA; Internet URL: `http://www.math.princeton.edu/~nelson/books.html`

[8] Solomentsev, E. D. (2002) "Support of a function" in *Encyclopaedia of Mathematics*, *ed.* Hazewinkel, M., Springer-Verlag, Berlin, Germany; Internet URL: `http://eom.springer.de/S/s091280.htm`

[9] Mackenzie, D. (2001) *Wavelets: Seeing the Forest and the Trees*, NAS, Washington, DC, USA

[10] Kaiser, G. (1978) *Local Fourier analysis and synthesis*, research proposal no. MCS-7822673, NSF, USA

[11] Graps, A. L. (1995) "An introduction to wavelets," *IEEE Comp. Sc. & Engg.*, Vol. 2, No. 2, pp. 50–61