involve
a journal of mathematics

Nonstandard existence proofs for reaction diffusion equations

Connor Olson, Marshall Mueller and Sigurd B. Angenent

msp

2019 vol. 12, no. 6

INVOLVE 12:6 (2019)
dx.doi.org/10.2140/involve.2019.12.1015

Nonstandard existence proofs for reaction diffusion equations

Connor Olson, Marshall Mueller and Sigurd B. Angenent

(Communicated by Suzanne Lenhart)

To the memory of Terry Millar

We give an existence proof for distribution solutions to a scalar reaction diffusion equation, with the aim of illustrating both the differences and the common ingredients of the nonstandard and standard approaches. In particular, our proof shows how the operation of taking the standard part of a nonstandard real number can replace several different compactness theorems, such as Ascoli's theorem and the Banach–Alaoglu theorem on weak*-compactness of the unit ball in the dual of a Banach space.

1. Introduction

1.1. ***Reaction diffusion equations.*** We consider the Cauchy problem for scalar reaction diffusion equations of the form

$$\frac{\partial u}{\partial t} = D\frac{\partial^2 u}{\partial x^2} + f(u(x,t)), \quad x \in \mathbb{R},\ t \geq 0, \tag{1a}$$

with prescribed initial condition

$$u(x,0) = u_0(x). \tag{1b}$$

In the setting of reaction diffusion equations the function $u(x,t)$ represents the density at location $x \in \mathbb{R}$ and time $t \geq 0$ of some substance which diffuses, and simultaneously grows or decays due to chemical reaction, biological mutation, or some other process. The term $D\partial^2 u/\partial x^2$ in the PDE (1a) accounts for the change in u due to diffusion, while the nonlinear term $f(u)$ accounts for the reaction rates. The prototypical example of such a reaction diffusion equation is the Fisher–KPP equation, see [Kolmogorov et al. 1937; Fisher 1937], in which the reaction term is given by $f(u) = u - u^2$.

MSC2010: 26E35, 35K57.
Keywords: nonstandard analysis, partial differential equations, reaction diffusion equations.

The Cauchy problem for the reaction diffusion equation (1a) is to find a function $u : \mathbb{R} \times [0, \infty) \to \mathbb{R}$ that satisfies the partial differential equation (1a) as well as the initial condition (1b). This is a classical problem, and the existence of such solutions is well known; see for example [Henry 1981; Pazy 1983]. As various techniques for constructing solutions are known, including the use of finite difference approximations to construct solutions (see [John 1982, Chapter 7.2]), our main goal is not to give another existence proof. Instead, we were inspired by several introductory texts on nonstandard analysis (notably, the undergraduate calculus text [Keisler 1976a], the more advanced introduction to the hyperreals [Goldblatt 1998], the "radically elementary approach" to probability in [Nelson 1987], as well as the blog post [Tao 2007]) and wanted to see what some standard existence proofs would look like in the language of nonstandard analysis.

Keisler [1976a] presented a proof of Peano's existence theorem for solutions to ordinary differential equations

$$\frac{dx}{dt} = f(t, x(t)), \quad x(0) = x_0, \tag{2}$$

using nonstandard analysis. One possible standard proof of Peano's theorem proceeds by constructing the numerical approximation to the solution by solving Euler's method for any small step size $\Delta t > 0$; i.e., one defines numbers $x_{i,\Delta t}$ by setting $x_{0,\Delta t} = x_0$ and then inductively solving

$$\frac{x_{i+1,\Delta t} - x_{i,\Delta t}}{\Delta t} = f(i\Delta t, x_{i,\Delta t}), \quad i = 0, 1, 2, \ldots. \tag{3}$$

The function $x_{\Delta t} : [0, \infty) \to \mathbb{R}$ obtained by linearly interpolating between the values $x_{\Delta t}(i\Delta t) = x_{i,\Delta t}$ is Euler's numerical approximation to the solution of the differential equation (2). The standard analysis proof of Peano's existence theorem then uses Ascoli's compactness theorem to extract a sequence of step sizes $\Delta t_n \to 0$ such that the approximate solutions $x_{\Delta t_n}(t)$ converge uniformly to some function $\tilde{x} : [0, \infty) \to \mathbb{R}$ and concludes by showing that the limit $\tilde{x}$ is a solution to the differential equation (2).

The nonstandard proof in [Keisler 1976a] follows the same outline, but one notable feature of this proof is that instead of using Ascoli's theorem, one "simply" chooses the step size Δt to be a positive infinitesimal number. The approximate solution then takes values in the hyperreals, and instead of applying a compactness theorem (Ascoli's in this case), one "takes the standard part" of the nonstandard approximate solution. The proof is then completed by showing that the function that is obtained actually satisfies the differential equation.

The standard and nonstandard proofs have some common ingredients. In both proofs one must find suitable estimates for the approximate solutions $x_{i,\Delta t}$, where the estimates should not depend on the step size Δt. Namely, the approximate

solutions $x_{\Delta t}$ should be uniformly bounded, and they should be uniformly Lipschitz continuous, i.e., $|x_{\Delta t}(t) - x_{\Delta t}(s)| \le L|t-s|$ for all $t, s \in [0,\infty)$, $\Delta t > 0$. In the standard proof these estimates allow one to use Ascoli's theorem; in the nonstandard proof they guarantee that the standard part of the approximating solution with infinitesimally small Δt still defines a continuous function on the standard reals.

There appear to be two main differences between the standard and nonstandard proofs. The first, very obviously, is that the nonstandard setting allows one to speak rigorously of infinitely small numbers, and thereby avoid the need to consider limits of sequences. The second difference is that the process of "taking the standard part" of a hyperreal number acts as a replacement for one compactness theorem or another: in the nonstandard proof of Peano's theorem one avoids Ascoli's theorem by taking standard parts. This too is probably well known in some circles (Terry Tao [2007] made the point in a blog post), but is not as obviously stated in the nonstandard analysis texts we have seen.

In this paper we intend to further illustrate this point by proving an existence theorem for weak or distributional solutions of certain partial differential equations that is analogous to the proof of Peano's theorem sketched above (see Section 2.1 below for a very short summary of the theory of distributions). Thus, to "solve" the reaction diffusion equation (1a) we choose space and time steps $\Delta x > 0$ and $\Delta t > 0$, and discretize the PDE by replacing the second derivative with a second difference quotient and the time derivative with a forward difference quotient, resulting in a finite difference equation,

$$\frac{u(x,t+\Delta t)-u(x,t)}{\Delta t} = D\,\frac{u(x+\Delta x,t)-2u(x,t)+u(x-\Delta x,t)}{(\Delta x)^2} + f(u(x,t)). \quad (4)$$

This kind of discretization is very common in numerical analysis;[1] see [LeVeque 2007; Press et al. 2007]. For given initial data (but no boundary data) one can use this difference equation to inductively compute the values of $u(x,t)$ for all (x,t) in a triangular grid (see Figure 1).

Given a solution $U(x,t)$ of the difference equation (4), one can define a generalized function, or distribution,

$$\langle U, \varphi\rangle \stackrel{\text{def}}{=} \sum_t \sum_x U(x,t)\,\varphi(x,t)\,\Delta x\,\Delta t. \quad (5)$$

[1] As one of the reviewers pointed out, we have chosen the simplest among the many other difference schemes that approximate the reaction diffusion equation (1a). For many other schemes the arguments in this paper could probably be adapted, although for implicit schemes one would have to arbitrarily select boundary values and observe that these will only have an infinitesimal effect on the solution in bounded regions of the form $|x|, |t| \le R$ for any standard $R \in \mathbb{R}$. We leave it to the interested reader to pursue these questions.

In a standard existence proof of weak solutions to the equation one would now use a compactness theorem to extract a sequence $(\Delta x_i, \Delta t_i) \to (0,0)$ for which the corresponding distributions U_i converge in the sense of distributions, and then show that the limiting distribution satisfies the PDE (1a). The compactness theorem that is required in this proof is the Banach-Alaoglu theorem about weak*-compactness in duals of Banach spaces (in our case, $L^\infty(\mathbb{R}^2)$, which is the Banach space dual of $L^1(\mathbb{R}^2)$).

The nonstandard proof, which we give in this paper, avoids the compactness theorem (or notions of Lebesgue integration required to define L^∞) by letting Δx and Δt be infinitesimal positive hyperreals, and by taking the standard part of the expression on the right in (5). In both the standard and nonstandard settings this approach works for the linear heat equation, that is, in the case where the reaction term $f(u)$ is absent (i.e., $f(u) \equiv 0$). The nonlinear case is a bit more complicated because there is no adequate definition of $f(u)$ when u is a distribution rather than a pointwise-defined function. In both the standard and nonstandard proofs we overcome this by proving that the approximating functions are Hölder continuous, so that $f(u(x,t))$ can be defined. In the standard proof this again allows one to use Arzelà–Ascoli and extract a convergent subsequence. However, since the domain $\mathbb{R}^2$ is not compact, Arzelà–Ascoli cannot be applied directly, and the standard proof therefore requires one to apply the compactness theorem on an increasing sequence of compact subsets $K_i \subset \mathbb{R}^2$, after which Cantor's diagonal trick must be invoked to get a sequence of functions that converges uniformly on every compact subset of $\mathbb{R}^2$. As we show below, these issues do not come up in the nonstandard proof.

1.2. ***Comments on nonstandard analysis.*** In nonstandard analysis one exploits the existence of an ordered field ${}^*\mathbb{R}$ called the hyperreal numbers, which contains the standard real numbers $\mathbb{R}$, but also contains infinitesimally small numbers, i.e., numbers $x \in {}^*\mathbb{R}$ with $x \neq 0$ that violate the Archimedean axiom by satisfying $n|x| < 1$ for all standard integers $n \in \mathbb{N}$. When two hyperreals $x, y \in {}^*\mathbb{R}$ differ by an infinitesimal, one writes $x \approx y$. For each hyperreal number $x \in {}^*\mathbb{R}$ there is a unique standard real number $\mathrm{St}(x)$, called the standard part of x, such that $x \approx \mathrm{St}(x)$. Beyond this simple description of the hyperreals we will not even try to give an exposition of nonstandard analysis in this paper and instead refer the reader to the many texts that have been written on the subject; see, e.g., a very incomplete list: [Keisler 1976a; 1976b; Goldblatt 1998; Nelson 1987; Albeverio et al. 1986; Tao 2007].

There are a few different approaches to using the hyperreals. Keisler [1976a] gave an axiomatic description of the hyperreals and their relation with the standard reals. In this approach, functions that are defined for standard reals automatically extend to the hyperreals, according to the *transfer principle.* A different approach

that also begins with an axiomatic description of the hyperreals can be found in Nelson's "radically elementary" treatment of probability theory [1987].

Our point of view in this paper is that of internal set theory as explained in [Goldblatt 1998] (see also the "instructor's guide" [Keisler 1976b] to his calculus text). Goldblatt explains the construction of the hyperreals using nonprincipal ultrafilters (which can be thought of as analogous to the construction of real numbers as equivalence classes of Cauchy sequences of rational numbers). He then extends this construction and defines internal sets, internal functions, etc.

2. Distribution solutions

2.1. ***The definition of distributions.*** We recall the definition of a "generalized function," i.e., of a distribution, which can be found in many textbooks on real analysis, such as [Folland 1999].

A real-valued function f on $\mathbb{R}^2$ is traditionally defined by specifying its values $f(x, y)$ at each point $(x, y) \in \mathbb{R}^2$. In the theory of distributions a generalized function f is defined by specifying its weighted averages

$$\langle f, \varphi \rangle = \int_{\mathbb{R}^2} f(x, y)\, \varphi(x, y)\, dx\, dy \tag{6}$$

for all so-called "test functions" φ. A test function is any function $\varphi : \mathbb{R}^2 \to \mathbb{R}$ that is infinitely often differentiable, and which vanishes outside a sufficiently large ball $B_R = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 < R\}$ whose radius R is allowed to depend on the particular test function. The set of all test functions, which is denoted by $C_c^\infty(\mathbb{R}^2)$, or sometimes by $\mathcal{D}(\mathbb{R}^2)$, is an infinite-dimensional vector space. *By definition,* a distribution is any linear functional $T : C_c^\infty(\mathbb{R}^2) \to \mathbb{R}$. The most common notation for the value of a distribution T applied to a test function φ is $\langle T, \varphi \rangle$. For instance, if $f : \mathbb{R}^2 \to \mathbb{R}$ is a continuous function, then (6) defines f as a distribution. The canonical example of a distribution that does not correspond to a function f is the Dirac delta function, which is defined by

$$\langle \delta, \varphi \rangle \stackrel{\text{def}}{=} \varphi(0, 0).$$

The full definition of a distribution T includes the requirement that the value $\langle T, \varphi \rangle$ depend continuously on the test function φ. To state this continuity condition precisely one must introduce a notion of convergence in the space of test functions $C_c^\infty(\mathbb{R}^2)$. We refer the reader to [Folland 1999] for the details, and merely observe that a sufficient condition for a linear functional $\varphi \mapsto \langle T, \varphi \rangle$ to be a distribution is that there exist a constant C such that

$$|\langle T, \varphi \rangle| \le C \iint_{\mathbb{R}^2} |\varphi(x, y)|\, dx\, dy \tag{7}$$

holds for all test functions φ. Alternatively, if a constant C exists such that

$$|\langle T, \varphi\rangle| \le C \sup_{(x,y)\in\mathbb{R}^2} |\varphi(x, y)| \tag{8}$$

holds for all $\varphi \in C_c^\infty(\mathbb{R}^2)$, then T also satisfies the definition of a distribution. The conditions (7) and (8) are not equivalent: either one of these implies that T is a distribution.

2.2. ***Distributions defined by nonstandard functions on a grid.*** Let dx, dy be two positive infinitesimal hyperreal numbers, and let N, M be two positive hyperintegers such that $N\,dx$ and $M\,dy$ are unlimited. Consider the rectangular grid

$$G = \{(k\,dx, l\,dy) \in {}^*\mathbb{R}^2 \mid k, l \in {}^*\mathbb{N},\ |k| \le N,\ |l| \le M\}. \tag{9}$$

From the point of view of nonstandard analysis and internal set theory, G is a hyperfinite set, and for any internal function $f : G \to {}^*\mathbb{R}$ there is an $(x, y) \in G$ for which $f(x, y)$ is maximal.

Lemma 2.2.1. *If $g : \mathbb{R}^2 \to \mathbb{R}$ is a continuous function with compact support, then*

$$\int_{\mathbb{R}^2} g(x, y)\,dx\,dy \approx \sum_{(x,y)\in G} g(x, y)\,dx\,dy.$$

Recall that $x \approx y$ means that $x - y$ is infinitesimal.

Proof. The statement of the lemma is very close to the nonstandard definition of the Riemann integral of a continuous function, the only difference being that we are integrating over the unbounded domain $\mathbb{R}^2$ rather than a compact rectangle $[-\ell, \ell] \times [-\ell, \ell] \subset \mathbb{R}^2$. Since the function g has compact support, there is a real $\ell > 0$ such that $g(x, y) = 0$ outside the square $[-\ell, \ell] \times [-\ell, \ell]$. By definition we then have

$$\int_{\mathbb{R}^2} g(x, y)\,dx\,dy = \int_{-\ell}^{\ell}\int_{-\ell}^{\ell} g(x, y)\,dx\,dy.$$

Choose hyperintegers $L, L' \in {}^*\mathbb{N}$ for which

$$L\,dx \le \ell < (L + 1)\,dx \quad \text{and} \quad L'\,dy \le \ell < (L' + 1)\,dy.$$

Then the nonstandard definition of the Riemann integral implies

$$\int_{-\ell}^{\ell}\int_{-\ell}^{\ell} g(x, y)\,dx\,dy \approx \sum_{k=-L}^{L}\sum_{l=-L'}^{L'} g(k\,dx, l\,dy)\,dx\,dy.$$

Finally, if $(x, y) \in G$ then $g(x, y) = 0$ unless $|x| \le \ell$ and $|y| \le \ell$, so that

$$\sum_{k=-L}^{L}\sum_{l=-L'}^{L'} g(k\,dx, l\,dy)\,dx\,dy = \sum_{(x,y)\in G} g(x, y)\,dx\,dy. \qquad \square$$

Lemma 2.2.2. *Suppose that $f : G \to {}^*\mathbb{R}$ is a hyperreal-valued function which is bounded, in the sense that there exists a limited $C > 0$ such that $|f(x, y)| \le C$ for all $(x, y) \in G$. Then the expression*

$$\langle T_f, \varphi \rangle \stackrel{\text{def}}{=} \mathrm{St}\left(\sum_{(x,y)\in G} f(x, y)\, \varphi(x, y)\, dx\, dy \right) \tag{10}$$

defines a distribution on $\mathbb{R}^2$.

If the function f is the nonstandard extension of a (standard) continuous function $f : \mathbb{R}^2 \to \mathbb{R}$, then the distribution T_f coincides with the distribution defined by (6).

Proof. We first verify that the distribution is well-defined. Since $|f(x, y)| \le C$ for all (x, y) we have

$$\left| \sum_{(x,y)\in G} f(x, y)\, \varphi(x, y)\, dx\, dy \right| \le C \sum_{(x,y)\in G} |\varphi(x, y)|\, dx\, dy \approx C \int_{\mathbb{R}^2} |\varphi(x, y)|\, dx\, dy.$$

Hence the sum in the definition (10) of $\langle T_f, \varphi \rangle$ is a limited hyperreal, whose standard part is a well-defined real number which satisfies

$$|\langle T_f, \varphi \rangle| \le C \int_{\mathbb{R}^2} |\varphi(x, y)|\, dx\, dy.$$

Therefore T_f is a well-defined distribution.

Let $\langle f, \varphi \rangle$ be defined as in (6). Fix φ. We then have

$$\int_{\mathbb{R}^2} f(x, y)\varphi(x, y)\, dx\, dy \approx \sum_{(x,y)\in G} f(x, y)\varphi(x, y)\, dx\, dy,$$

which implies that the distribution defined in (6) coincides with T_f. □

3. The Cauchy problem for the heat equation

In this section we recall the definition of distribution solutions to the Cauchy problem for the heat equation and show how, by solving the finite difference approximation to the heat equation on a hyperfinite grid, one can construct a distribution solution to the Cauchy problem.

3.1. ***Formulation in terms of distributions.*** We consider the Cauchy problem for the linear heat equation $u_t = u_{xx}$ with bounded and continuous initial data $u(x, 0) = u_0(x)$. Without losing generality we may assume that the diffusion coefficient D is 1, e.g., by nondimensionalizing space and time and introducing $\tau = Dt$ as the new time variable.

Definition 3.1.1. A distribution u on $\mathbb{R}^2$ is a solution to the heat equation $u_t = u_{xx}$ with initial data u_0 if u satisfies

$$u_t - u_{xx} = u_0(x)\,\delta(t), \quad x \in \mathbb{R},\ t \in \mathbb{R}, \tag{11}$$

in the sense of distributions, and if $u = 0$ for $t \le 0$.

Equality in the sense of distributions in (11) means that both sides of the equation are to be interpreted as distributions, and that they should yield the same result when evaluated on any test function $\varphi \in C_c^\infty(\Omega)$. To explain this in more detail, recall that δ is Dirac's delta function, so that the action of the right-hand side in (11) on a test function is

$$\langle u_0(x)\,\delta(t), \varphi\rangle \stackrel{\text{def}}{=} \int_{\mathbb{R}} u_0(x)\,\varphi(x,0)\,dx.$$

The definition of distributional derivative [Folland 1999, Chapter 9] says that the action of the left-hand side in (11) is given by

$$\langle u_t - u_{xx}, \varphi\rangle = \langle u, -\varphi_t - \varphi_{xx}\rangle.$$

If the distribution u is given by a function $u : \mathbb{R}^2 \to \mathbb{R}$ which vanishes for $t < 0$ and is continuous for $t \ge 0$ (so that it has a simple jump discontinuity at $t = 0$) then we get

$$\langle u_t - u_{xx}, \varphi\rangle = \int_{\mathbb{R}}\int_0^\infty u(x,t)\{-\varphi_t - \varphi_{xx}\}\,dt\,dx.$$

A piecewise continuous function u therefore satisfies (11) in the sense of distributions if

$$\int_{\mathbb{R}} u_0(x)\varphi(x,0)\,dt + \int_{\mathbb{R}}\int_0^\infty u(x,t)\{\varphi_t + \varphi_{xx}\}\,dt\,dx = 0 \tag{12}$$

for all test functions $\varphi \in C_c^\infty(\mathbb{R}^2)$. This is one form of the classical definition of a weak solution to the Cauchy problem.

3.2. ***The finite difference equation.*** To construct a distribution solution to (11) we introduce a grid with spacing dx and dt, and replace the differential equation by the simplest finite difference scheme that appears in numerical analysis. If u is the solution to the differential equation, then we write U for the approximating solution to the finite difference equation, using the following common notation for finite differences:

$$D_x^+ U(x,t) = \frac{U(x+dx,t) - U(x,t)}{dx},$$

$$D_x^- U(x,t) = \frac{U(x,t) - U(x-dx,t)}{dx}$$

$$D_t^+ U(x,t) = \frac{U(x,t+dt) - U(x,t)}{dt}.$$

See for example, [LeVeque 2007, Chapter 1]. With this notation

$$D_x^2 U(x,t) \stackrel{\text{def}}{=} D_x^+ D_x^- U(x,t) = \frac{U(x+dx,t) - 2U(x,t) + U(x-dx,t)}{(dx)^2}.$$

The operators D_x^+, D_x^-, and D_t^+ all commute. A finite difference equation corresponding to the heat equation $u_t = u_{xx}$ is then $D_t^+ U = D_x^2 U$, i.e.,

$$\frac{U(x,t+dt) - U(x,t)}{dt} = \frac{U(x+dx,t) - 2U(x,t) + U(x-dx,t)}{(dx)^2}. \tag{13}$$

We can solve this algebraic equation for $U(x,t+dt)$, resulting in

$$U(x,t+dt) = \alpha U(x-dx,t) + (1-2\alpha)U(x,t) + \alpha U(x+dx,t), \tag{14}$$

where

$$\alpha \stackrel{\text{def}}{=} \frac{dt}{(dx)^2}.$$

3.3. ***The approximate solution.*** Let $N \in {}^*\mathbb{N}$ be an unlimited hyperfinite integer, and assume that dx and dt are positive infinitesimals. Assume moreover that N is so large that both $N\,dt$ and $N\,dx$ are unlimited hyperreals. We then consider the hyperfinite grid

$$G_C = \{(m\,dx, n\,dt) \mid m, n \in {}^*\mathbb{N}, |m| + n \le N\}.$$

See Figure 1. The initial function $u_0 : \mathbb{R} \to \mathbb{R}$ extends to an internal function $u_0 : {}^*\mathbb{R} \to {}^*\mathbb{R}$. By assumption there is a $C \in \mathbb{R}$ such that $|u_0(x)| \le C$ for all $x \in \mathbb{R}$, so this also holds for all $x \in {}^*\mathbb{R}$.

We define $U : G_C \to {}^*\mathbb{R}$ by requiring:

- $U(x,0) = u_0(x)$ for all x with $(x,0) \in G_C$, i.e., for all $x = k\,dx$ with $k \in \{-N, \ldots, +N\}$.
- U satisfies (4) or, equivalently, (14) at all $(x,t) = (m\,dx, n\,dt) \in G_C$ with $|m| + n < N$.

Theorem 3.3.1. *Let $U : G_C \to {}^*\mathbb{R}$ be the hyperreal solution of the finite difference scheme (13) with initial values $U(x,0) = u_0(x)$, and suppose that $\alpha \le \frac{1}{2}$. Then the expression*

$$\langle u, \varphi \rangle \stackrel{\text{def}}{=} \text{St}\left(\sum U(x,t)\varphi(x,t)\,dx\,dt \right), \quad \varphi \in C_c^\infty(\mathbb{R}^2), \tag{15}$$

defines a distribution on $\mathbb{R}^2$ that satisfies (11).

To show that this expression does indeed define a distribution we must show that the $U(x,t)$ are bounded by a standard real number. This follows from a discrete version of the maximum principle, which we will again use in Section 4, so we state

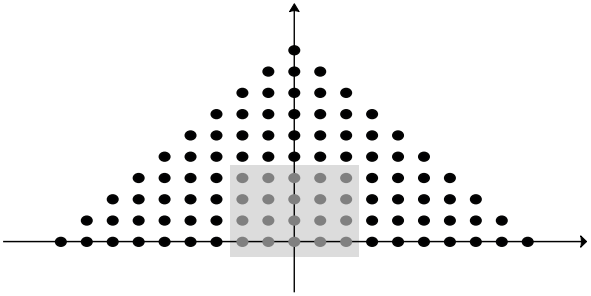

Figure 1. The triangular grid G_C; if $U(x,t)$ is known at all grid points at the bottom of the triangle, then the finite difference equation (13) uniquely determines $U(x,t)$ at all other grid points.

it in slightly greater generality than needed in this section. The lemma explains why we need the condition $\alpha = dt/(dx)^2 \le \frac{1}{2}$ and is well known in numerical analysis as a necessary condition for stability of the finite difference scheme.

Lemma 3.3.2 (Gronwall-type estimate). *Let $W : G_C \to {}^*\mathbb{R}$ satisfy $W(x,t) \ge 0$ for all $(x,t) \in G_C$, and suppose that for some nonnegative $m \in \mathbb{R}$ one has*

$$W(x,t+dt) \le \alpha W(x+dx,t) + (1-2\alpha)W(x,t) + \alpha W(x-dx,t) + dt\, m W(x,t)$$

at all $(x,t) \in G_C$. For each $t = n\,dt$ with $0 \le n \le N$ consider[2]

$$w(t) \stackrel{\text{def}}{=} \max_x W(x,t).$$

If $0 \le \alpha \le \frac{1}{2}$ then

$$w(t) \le e^{mt} w(0).$$

Proof. (Compare [John 1982, §7.2, Lemma I].) The assumption on α implies that $\alpha \ge 0$ and $1 - 2\alpha \ge 0$. Hence for all x with $(x, t+dt) \in G_C$ we have

$$\begin{aligned} W(x,t+dt) &= \alpha W(x+dx,t) + (1-2\alpha)W(x,t) + \alpha W(x-dx,t) + dt\, m W(x,t) \\ &\le (\alpha + (1-2\alpha) + \alpha + m\,dt) w(t) \\ &= (1 + m\,dt) w(t). \end{aligned}$$

Taking the maximum over x we see that $w(t+dt) \le (1+m\,dt)w(t) \le e^{m\,dt} w(t)$. By induction we then have for $t = n\,dt$ that $w(t) \le (e^{m\,dt})^n w(0) = e^{mt} w(0)$. □

[2] For any given $t = n\Delta t$ there are infinitely many hyperreal numbers $W(x,t)$, so the standard analyst may be surprised to see "max" instead of "sup" in the definition of $w(t)$. However, in the internal-set-theory interpretation, the set of numbers $\{W(x,t) : (x,t) \in G_C\} = \{W(m\Delta x, n\Delta t) : |m| \le N - n\}$ is a hyperfinite internal set of real numbers, indexed by $m \in \{0, \pm 1, \pm 2, \ldots, \pm(N-n)\}$. Therefore one of the numbers $W(m\Delta x, n\Delta t)$ is the largest, so that the maximum is a well-defined hyperreal number.

3.4. ***Proof of Theorem 3.3.1.*** The relation (14) which defines $U(x,t)$ implies that $W(x,t) \stackrel{\text{def}}{=} |U(x,t)|$ satisfies

$$\begin{aligned} W(x,t+dt) &= |\alpha U(x-dx,t) + (1-2\alpha)U(x,t) + \alpha U(x+dx,t)| \\ &\le \alpha W(x-dx,t) + (1-2\alpha)W(x,t) + \alpha W(x+dx,t), \end{aligned}$$

where we have used $\alpha \ge 0$ and $1-2\alpha \ge 0$.

Since the initial condition is bounded by $|U(x,0)| = |u_0(x)| \le M$, the Gronwall-type Lemma 3.3.2 implies that $|U(x,t)| \le M$ for all $(x,t) \in G_C$. According to Lemma 2.2.2 this implies that the expression (15) does define a distribution u on $\mathbb{R}^2$.

We want to prove that u satisfies the heat equation in the sense of distributions; i.e., we want to show for any test function φ that

$$\langle u_t - u_{xx}, \varphi \rangle = \langle u_0(x)\,\delta(t), \varphi \rangle = \int_{\mathbb{R}} u_0(x)\,\varphi(x,0)\,dx.$$

First, we see from the definition of distributional derivative that

$$\langle u_t - u_{xx}, \varphi \rangle = -\langle u, \varphi_t + \varphi_{xx} \rangle.$$

We then have from the definition of u that

$$-\langle u, \varphi_t + \varphi_{xx} \rangle \approx \sum_{(x,t)\in G_C} -U(x,t)(\varphi_t(x,t) + \varphi_{xx}(x,t))\,dx\,dt \stackrel{\text{def}}{=} T.$$

Using Taylor's formula we replace the partial derivatives of the test function with its corresponding finite differences; i.e., we write $\varphi_t(x,t) = D_t^+\varphi(x,t) + \varepsilon_t(x,t)$ and $\varphi_{xx}(x,t) = D_x^2\varphi(x,t) + \varepsilon_{xx}(x,t)$, where $\varepsilon_t, \varepsilon_{xx} : G_C \to {}^*\mathbb{R}$ are the infinitesimal error terms in the Taylor expansion. Substituting these gives us

$$T = \sum_{G_C} -U(x,t)(D_t^+\varphi(x,t) + D_x^2\varphi(x,t) + \varepsilon_t + \varepsilon_{xx})\,dx\,dt.$$

We can split this sum into three parts, $T = T_1 + T_2 + T_3$, with

$$\begin{aligned} T_1 &= \sum_{G_C} -U(x,t) D_t^+\varphi(x,t)\,dx\,dt, \\ T_2 &= \sum_{G_C} -U(x,t) D_x^2\varphi(x,t)\,dx\,dt, \\ T_3 &= \sum_{G_C} -U(x,t)(\varepsilon_t + \varepsilon_{xx})\,dx\,dt. \end{aligned}$$

We will first handle the error term, T_3. Since the test function φ has compact support, there exists a real $\ell > 0$ such that $\varphi = 0$ outside the rectangle $\Omega = [-\ell, \ell] \times [-\ell, \ell]$. The errors in the Taylor expansion therefore also vanish outside of Ω so that we

can write T_3 as

$$T_3 = \sum_{\Omega \cap G_C} -U(x,t)(\varepsilon_t + \varepsilon_{xx})\,dx\,dt.$$

The key to estimating this sum is that we can estimate all the errors $\varepsilon_t(x,t)$ and $\varepsilon_{xx}(x,t)$ by one fixed infinitesimal $\varepsilon > 0$ that does not depend on (x,t). Indeed, the grid G_C is a hyperfinite internal set, and therefore any internal function such as $|\varepsilon_t| : G_C \to {}^*\mathbb{R}$ attains its largest value at one of the $(x,t) \in G_C$, say at (x_1,t_1). Then $|\varepsilon_t(x,t)| \le |\varepsilon_t(x_1,t_1)|$ for all $(x,t) \in G_C$. Similarly, there is an $(x_2,t_2) \in G_C$ that maximizes $|\varepsilon_{xx}(x,t)|$. Now define

$$\varepsilon_1 = |\varepsilon_t(x_1,t_1)|, \quad \varepsilon_2 = |\varepsilon_{xx}(x_2,t_2)|.$$

Then both ε_1 and ε_2 are positive infinitesimals for which

$$|\varepsilon_t(x,t)| \le \varepsilon_1, \quad |\varepsilon_{xx}(x,t)| \le \varepsilon_2$$

hold at all grid points $(x,t) \in G_C$.

If we let $\varepsilon = \max\{\varepsilon_1, \varepsilon_2\}$, then $|T_3| \le \sum_{G_C \cap \Omega} 2U(x,t)\,\varepsilon\,dx\,dt$. By the construction of U, we have $|U(x,t)| \le M$ for all $(x,t) \in G_C$, so we get

$$|T_3| \le \sum_{G_C \cap \Omega} 2M\varepsilon\,dx\,dt \le 2M\varepsilon\,dx\,dt\frac{\ell}{2\,dx}\frac{\ell}{dt} = M\ell^2\varepsilon,$$

which is infinitesimal, so T_3 is infinitesimal.

From the definition we have $T_1 = \sum_{G_C} -U(x,t)(\varphi(x,t+dt) - \varphi(x,t))\,dx$. Using the compact support of φ, we can then rewrite this sum as

$$T_1 = -\sum_{k=-K}^{K}\sum_{l=0}^{L+1} U(k\,dx, l\,dt)\{\varphi(k\,dx,(l+1)\,dt) - \varphi(k\,dx, l\,dt)\}\,dx,$$

where $K\,dx \approx \ell$ and $L\,dt \approx \ell$.

Applying summation by parts to this sum we then get

$$T_1 = \sum_{k=-K}^{K}\left\{U(k\,dx,0)\,\varphi(k\,dx,0) + \sum_{l=0}^{L}\varphi(k\,dx,(l+1)\,dt)\,D_t^+U(k\,dx,l\,dt)\,dt\right\}dx.$$

Next, for T_2, we can split the sum into two parts.

$$T_2 = \sum_{l=0}^{L}\sum_{k=-K-1}^{K+1} -U(k\,dx,l\,dt)D_x^+\varphi(k\,dx,l\,dt)\,dt\,dx + \sum_{l=0}^{L}\sum_{k=-K-1}^{K+1} U(k\,dx,l\,dt)D_x^-\varphi(k\,dx,l\,dt)\,dt\,dx. \quad (16)$$

Applying summation by parts again to both sums we get

$$T_2 = -\sum_{l=0}^{L}\sum_{k=-K}^{K} \varphi(k\,dx, l\,dt)\, D_x^2 U(k\,dx, l\,dt)\, dx\, dt.$$

Putting the terms T_1, T_2, T_3 all together, we have, because $T_3 \approx 0$,

$$\begin{aligned} T_1+T_2+T_3 &\approx T_1+T_2 \\ &= \sum_{k=-K}^{K} U(k\,dx, 0)\varphi(k\,dx, 0)\, dx \\ &\quad + \sum_{k=-K}^{K}\sum_{l=0}^{L} \varphi(k\,dx, l\,dt) \\ &\qquad \times (D_t^+ U(k\,dx, l\,dt) - D_x^2 U(k\,dx, l\,dt))\, dx\, dt. \end{aligned} \tag{17}$$

Since U satisfies the difference equation $D_t^+ U = D_x^2 U$ at all grid points this reduces to

$$T = T_1+T_2+T_3 \approx \sum_{k=-K}^{K} U(k\,dx, 0)\, \varphi(k\,dx, 0)\, dx \approx \sum_{k=-K}^{K} u_0(k\,dx)\, \varphi(k\,dx, 0)\, dx.$$

Taking the standard part we get the distribution

$$\mathrm{St}(T) = \mathrm{St}\left(\sum_{k=-K}^{K} u_0(k\,dx)\, \varphi(k\,dx, 0)\, dx \right) = \int_{-\ell}^{\ell} u_0(x)\, \varphi(x, 0)\, dx = \langle u_0(x)\, \delta(t), \varphi \rangle.$$

This completes the proof that $\langle u_t - u_{xx}, \varphi \rangle = \langle u_0(x)\delta(t), \varphi \rangle$ for all test functions φ, and thus that u is a distributional solution of (11).

3.5. ***Comments on the proof.*** In our construction of solutions to the linear heat equation we completely avoided estimating derivatives of the approximate solution U. The only estimate we used was that the approximate solution $U(x, t)$ has the same upper bound as the given initial function u_0.

We assumed that the initial function u_0 is continuous. The one place in the proof where we needed this assumption was at the end, when we used the fact that for continuous functions $f : \mathbb{R} \to \mathbb{R}$ one has

$$\sum_{|k|\le K} f(k\,dx)\, dx \approx \int_{-\ell}^{\ell} f(x)\, dx$$

and applied this to the function $f(x) = u_0(x)\, \varphi(x, 0)$.

4. The Cauchy problem for a reaction diffusion equation

We consider the reaction diffusion equation

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + f(u(x, t)), \quad x \in \mathbb{R},\ t > 0, \tag{1a}$$

with initial condition (1b). It is known that one cannot expect solutions to exist for all times $t > 0$ without imposing some growth conditions on the nonlinearity $f(u)$. We will assume that f is a Lipschitz continuous function;[3] i.e., for some positive real K_1 one has

$$\text{for all } u, v \in \mathbb{R}, \quad |f(u) - f(v)| \le K_1 |u - v|. \tag{18}$$

This implies that $f(u)$ grows at most linearly in u:

$$\text{for all } u \in \mathbb{R}, \quad |f(u)| \le K_0 + K_1 |u|, \tag{19}$$

where $K_0 \stackrel{\text{def}}{=} |f(0)|$.

In contrast to the linear heat equation, (1a) contains the nonlinear term $f(u)$, which is meaningless if u is an arbitrary distribution. One can follow the same procedure as in the previous section; i.e., one can replace the differential equation by a finite difference scheme on the hyperfinite grid G_C and construct an approximating solution $U : G_C \to {}^*\mathbb{R}$. After establishing suitable bounds one can then show that by taking standard parts as in Lemma 2.2.2, both $U(x,t)$ and $f(U(x,t))$ define distributions u and F on $\mathbb{R}^2$. The problem is to give a meaning to the claim that "$F = f(u)$", because u is merely a distribution and can therefore not be substituted in a nonlinear function. In this section we show how to overcome this problem by adding the assumption that the initial function is Lipschitz continuous, i.e.,

$$\text{for all } x, y \in \mathbb{R}, \quad |u_0(x) - u_0(y)| \le L|x - y| \tag{20}$$

for some real $L > 0$, and showing that the standard part of the approximating solution U is a continuous function on $\mathbb{R} \times [0, \infty)$. The substitution $f(U(x,t))$ is then well-defined and we can verify that the continuous standard function corresponding to U is a distributional solution of the reaction diffusion equation (1a).

4.1. ***Weak solutions to the reaction diffusion equation.*** Rather than writing the initial value problem in the distributional form $u_t - u_{xx} - f(u) = u_0(x)\,\delta(t)$, we use the integral version (12) of the definition of weak solution. Thus we define a *weak solution* to (1a), (1b) to be a continuous function $u : \mathbb{R} \times [0, \infty) \to \mathbb{R}$ that satisfies

$$\iint_{\mathbb{R}\times[0,\infty)} \{u(x,t)(-\varphi_{xx} - \varphi_t) - f(u(x,t))\,\varphi\}\,dx\,dt = \int_{\mathbb{R}} u_0(x)\,\varphi(x,0)\,dx \tag{21}$$

for all test functions $\varphi \in C_c^\infty(\mathbb{R}^2)$.

[3] The assumption that f be globally Lipschitz continuous rules out the Fisher–Kolmogorov nonlinearity $f(u) = u - u^2$. However, for that particular nonlinearity the only solutions that are relevant to the interpretation of u as an allele ratio are those with $0 \le u \le 1$. If $f(u) = u - u^2$, then a quick look at our finite difference scheme (22) shows that for initial data that satisfy $0 \le u(x,0) \le 1$ the approximate solution to the difference equation also satisfies $0 \le U(x,t) \le 1$, provided $\alpha < \frac{1}{2}$, so that the subsequent arguments still apply.

Theorem 4.1.1. *If f is Lipschitz continuous as in (18), and if the initial function u_0 is bounded by*

$$\text{for all } x \in \mathbb{R}, \quad |u_0(x)| \le M$$

for some positive real M, and if u_0 also is Lipschitz continuous, as in (20), then the reaction diffusion equation (1a), (1b) has a weak solution.

4.2. ***Definition of the approximate solution.*** To construct the solution we consider the grid G_C as defined in Section 3.3 with infinitesimal mesh sizes $dx, dt > 0$, and consider the finite difference scheme

$$D_t^+ U(x,t) = D_x^2 U(x,t) + f(U(x,t)). \tag{22}$$

Solving for $U(x, t+dt)$ we get

$$U(x,t+dt) = \alpha U(x+dx,t) + (1-2\alpha)U(x,t) + \alpha U(x-dx,t) + dt\, f(U(x,t)), \tag{23}$$

where, as before, $\alpha = dt/(dx)^2$. We extend the continuous function u_0 to an internal function $u_0 : {}^*\mathbb{R} \to {}^*\mathbb{R}$, and specify the initial conditions $U(x,0) = u_0(x)$ for $x = m\,dx$, $m = -N, \ldots, +N$. The finite difference equation (23) then determines $U(x,t)$ for all $(x,t) \in G_C$.

We now establish a number of a priori estimates for the approximate solution U that will let us verify that its standard part is well-defined and that it is a weak solution of the initial value problem.

4.3. ***Boundedness of the approximate solution.*** First we bound $|U(x,t)|$.

Lemma 4.3.1. *For all $(x,t) \in G_C$ we have*

$$|U(x,t)| \le e^{K_1 t} M + \frac{K_0}{K_1}(e^{K_1 t} - 1). \tag{24}$$

Proof. Using (19), i.e., $|f(u)| \le K_0 + K_1|u|$, we get

$$\begin{aligned} |U(x,t+dt)| \le{}& \alpha|U(x+dx,t)| + (1-2\alpha)|U(x,t)| \\ &+ \alpha|U(x-dx,t)| + dt(K_0 + K_1|U(x,t)|). \end{aligned} \tag{25}$$

In terms of $M(t) = \max_x |U(x,t)|$ this implies

$$M(t+dt) \le M(t) + dt(K_0 + K_1 M(t)) = (1 + K_1\, dt)M(t) + K_0\, dt.$$

Setting $t = n\,dt$ we see that this is an inequality of the form $M_n \le a M_{n-1} + b$, with $M_n = M(n\,dt)$. By induction this implies

$$\begin{aligned} M(t) = M(n\,dt) &\le (1 + K_1\, dt)^n M(0) + \frac{(1 + K_1\, dt)^n - 1}{1 + K_1\, dt - 1} K_0\, dt \\ &\le e^{K_1 t} M(0) + \frac{K_0}{K_1}(e^{K_1 t} - 1). \end{aligned}$$

Since $M(0) = M$, this proves (24). □

4.4. ***Lipschitz continuity in space of the approximate solution.*** We now show that $U(x,t)$ is Lipschitz continuous in the space variable.

Lemma 4.4.1. *For any two points* $(x,t), (x',t) \in G_C$ *we have*

$$|U(x,t) - U(x',t)| \le Le^{K_1 t}|x - x'|, \tag{26}$$

where L *is the Lipschitz constant for the initial function* u_0*, as in* (20)*.*

Proof. Let

$$V(x,t) \stackrel{\text{def}}{=} \frac{U(x+dx,t) - U(x,t)}{dx} = D_x^+ U(x,t).$$

Applying D_x^+ to both sides of (4) for U, and using the definition of V and commutativity of the difference quotient operators, we find

$$D_t^+ V = D_x^2 V + D_x^+ f(U).$$

Solving for $V(x, t+dt)$ we find

$$V(x,t+dt) = \alpha V(x+dx,t) + (1-2\alpha)V(x,t) + \alpha V(x-dx,t) + D_x^+ f(U).$$

Examining $D_x^+ f(U)$, we have

$$\begin{aligned} |D_x^+ f(U)| &= \frac{|f(U(x+dx,t)) - f(U(x,t))|}{dx} \\ &\le \frac{K_1\,|U(x+dx,t) - U(x,t)|}{dx} = K_1 |V(x,t)|, \end{aligned}$$

so that

$$|V(x,t+dt)| \le \alpha|V(x+dx,t)| + (1-2\alpha)|V(x,t)| + \alpha|V(x-dx,t)| + K_1\,dt|V(x,t)|.$$

Using Gronwall's inequality on $\max_x V$, we get the inequality

$$\max_x |V(x,t)| \le e^{K_1 t} \max_x |V(x,0)|.$$

The initial condition u_0 satisfies $|u_0(x) - u_0(x')| \le L|x - x'|$ for all $x, x' \in \mathbb{R}$, and therefore the extension of u_0 to the hyperreals satisfies this same inequality. Hence $|V(x,0)| \le L$ for all grid points $(x,0)$, and thus we have $|V(x,t)| \le Le^{K_1 t}$. This implies (26). □

4.5. ***Hölder continuity in time of the approximate solution.***

Lemma 4.5.1. *Given any real* $\bar{t} > 0$*, for any two grid points* $(x_0, t_0), (x_0, t_1) \in G_C$ *with* $0 \le t_0 \le t_1 \le \bar{t}$*, we have*

$$|U(x_0,t_1) - U(x_0,t_0)| \le C\sqrt{t_1 - t_0}, \tag{27}$$

where C *is a constant that only depends on* $\bar{t}$*,* K_0*,* K_1*,* L*, and* M*.*

Proof. We begin by observing that $f(U(x,t))$ is bounded on the time interval we are considering. Indeed, for $t \le \bar{t}$ we have shown for all $(x,t) \in G_C$ that

$$|U(x,t)| \le A_0 \stackrel{\text{def}}{=} e^{K_1 \bar{t}} M + \frac{K_1}{K_0}(e^{K_1 \bar{t}} - 1),$$

while the Lipschitz condition for f implies $|f(U(x,t))| \le K_0 + K_1|U| \le K_0 + K_1 A_0$. So if we set $A = K_0 + K_1 A_0$, then we have

$$\text{for all } (x,t) \in G_C \text{ with } t \le \bar{t}, \quad |f(U(x,t))| \le A. \tag{28}$$

Next, we construct a family of upper barriers for U using parabolas. In particular, for any real $a, b, c > 0$ we consider

$$\overline{U}(x,t) \stackrel{\text{def}}{=} U(x_0,t_0) + a(t-t_0) + \frac{b}{2}(x-x_0)^2 + c.$$

For any $b > 0$ we will find $a, c > 0$ so that $\overline{U}$ is an upper barrier, in the sense that

$$D_t^+ \overline{U} - D_x^2 \overline{U} \ge A + 1, \tag{29}$$

$$\overline{U}(x,0) > U(x,0) \quad \text{for all } (x,0) \in G_C. \tag{30}$$

A direct computation shows that $D_t^+ \overline{U} - D_x^2 \overline{U} = a - b$, so for a given b we choose $a = b + A + 1$ and (29) will hold.

To satisfy (30) we use (26), i.e., that $U(x,t)$ is Lipschitz continuous with Lipschitz constant $\bar{L} \stackrel{\text{def}}{=} e^{K_1 \bar{t}} L$:

$$U(x,t) \le U(x_0,t_0) + \bar{L}|x - x_0| \le U(x_0,t_0) + \frac{2}{b}(x-x_0)^2 + \frac{\bar{L}^2}{2b}.$$

If we choose $c > \bar{L}^2/(2b)$, e.g., $c = \bar{L}^2/b$, then our upper barrier $\overline{U}$ also satisfies (30).

Next, we apply a maximum principle argument to compare U and $\overline{U}$. Consider $W(x,t) = U(x,t) - \overline{U}(x,t)$. Then we have shown that $W(x,0) < 0$ for all x, and $D_t^+ W - D_x^2 W < 0$, which implies

$$W(x, t+dt) < \alpha W(x-dx,t) + (1-2\alpha) W(x,t) + \alpha W(x+dx,t)$$

for all (x,t) for which $(x \pm dx, t) \in G_C$. By induction we get $W(x,t) < 0$ for all $(x,t) \in G_C$. In particular $U(x_0,t) < \overline{U}(x_0,t)$ for all $t > t_0$; i.e., we have shown

$$U(x_0,t_1) < U(x_0,t_0) + (b + A + 1)(t_1 - t_0) + \frac{\bar{L}^2}{b}.$$

This upper bound holds for any choice of $b > 0$. To get the best upper bound we minimize the right-hand side over all $b > 0$. The best bound appears when $b = \bar{L}/\sqrt{t_1 - t_0}$. After some algebra one then finds

$$U(x_0,t_1) - U(x_0,t_0) < (A+1)(t_1 - t_0) + 2\bar{L}\sqrt{t_1 - t_0}.$$

Finally, using $t_1 - t_0 = \sqrt{t_1 - t_0}\sqrt{t_1 - t_0} \le \sqrt{\bar{t}}\sqrt{t_1 - t_0}$ we get

$$U(x_0, t_1) - U(x_0, t_0) < ((A+1)\sqrt{\bar{t}} + 2\bar{L})\sqrt{t_1 - t_0}.$$

This proves the upper bound in (27). To get the analogous lower bound one changes the signs of the coefficients a, b, c, which will turn $\overline{U}$ into a lower barrier. After working through the details, one finds the appropriate lower bound. □

Now that we have Lipschitz in space and Hölder in time, we also have for any pair of points $(x, t), (y, s) \in G_C$ that

$$\begin{aligned} |U(x,t) - U(y,s)| &\le |U(x,t) - U(y,t)| + |U(y,t) - U(y,s)| \\ &\le L|x - y| + C\sqrt{|t - s|}. \end{aligned} \tag{31}$$

4.6. ***Definition of the weak solution.*** So far we have been establishing estimates for the solution U to the finite difference scheme. It is worth pointing out that a standard existence proof would have required exactly the same estimates. At this point, however, the standard and nonstandard proofs diverge.

For $(x, t) \in \mathbb{R} \times [0, \infty)$ we choose $(\tilde{x}, \tilde{t}) \in G_C$ with $x \approx \tilde{x}$ and $t \approx \tilde{t}$, and then define $u(x, t) = \mathrm{St}(U(\tilde{x}, \tilde{t}))$. The continuity property (31) of the approximate solution U implies that the value of $\mathrm{St}(U(\tilde{x}, \tilde{t}))$ does not depend on how we chose the grid point $(\tilde{x}, \tilde{t})$, for if $(\hat{x}, \hat{t}) \in G_C$ also satisfied $\hat{x} \approx x$, $\hat{t} \approx t$, then $\tilde{x} \approx \hat{x}$ and $\tilde{t} \approx \hat{t}$, so that $U(\tilde{x}, \tilde{t}) \approx U(\hat{x}, \hat{t})$. It follows directly that the function $u : \mathbb{R} \times [0, \infty) \to \mathbb{R}$ is well-defined, that it satisfies the continuity condition (31), and that it satisfies the same bounds as in (24).

By the transfer principle, the standard function u extends in a unique way to an internal function ${}^*\mathbb{R} \times {}^*[0, \infty) \to {}^*\mathbb{R}$. It is common practice to abuse notation and use the same symbol u for the extension. The extended function satisfies the same continuity condition (31).

Lemma 4.6.1. *If $(x, t) \in G_C$ is limited, then $u(x, t) \approx U(x, t)$.*

Proof. If (x, t) is limited, then $x' = \mathrm{St}(x)$ and $t' = \mathrm{St}(t)$ are well-defined real numbers. By continuity of both u we have $u(x, t) \approx u(x', t')$. By the definition of u it follows from $x' \approx x$ and $t' \approx t$ that $u(x', t') \approx U(x, t)$. Combined we get $u(x, t) \approx U(x, t)$. □

4.7. ***Proof that u is a weak solution.*** We will now show that u is a weak solution whose existence is claimed in Theorem 4.1.1; i.e., we verify that u satisfies (21) for any test function $\varphi \in C_c^\infty(\mathbb{R}^2)$.

Since φ has compact support, there is a positive real ℓ such that $\varphi(x, t) = 0$ outside the square $[-\ell, \ell] \times [-\ell, \ell]$. We therefore have to verify

$$\int_0^\ell \int_{-\ell}^{\ell} \{u(x,t)(-\varphi_{xx} - \varphi_t) - f(u(x,t))\varphi\}\, dx\, dt = \int_{-\ell}^{\ell} u_0(x)\varphi(x, 0)\, dx.$$

Since the integrands are continuous functions we only make an infinitesimal error when we replace the two Riemann integrals by Riemann sums over the part of the grid G_C that lies within the square $[-\ell, \ell] \times [-\ell, \ell]$. Thus we must prove

$$\sum_{(x,t)\in G_C} \{u(x,t)(-\varphi_{xx}-\varphi_t)-f(u(x,t))\varphi\}\,dx\,dt \approx \sum_{(x,0)\in G_C} u_0(x)\varphi(x,0)\,dx. \quad (32)$$

We now intend to replace u by U, and the derivatives of φ by the corresponding finite differences. In doing so we make errors that we must estimate. Let $G_{C\ell} = G_C \cap [-\ell, \ell]^2$, so that the only nonzero terms in the two sums come from terms evaluated at points in $G_{C\ell}$. The intersection of internal sets is again internal, so the set $G_{C\ell}$ is internal and hyperfinite.

For each $(x, t) \in G_{C\ell}$ the quantities

$$|u(x,t) - U(x,t)|, \quad |\varphi_t(x,t) - D_t^+\varphi(x,t)|, \quad \text{and} \quad |\varphi_{xx}(x,t) - D_x^2\varphi(x,t)|$$

are infinitesimal. Since they are defined by internal functions, one of the numbers in the hyperfinite set

$$\{|u(x,t) - U(x,t)|, |\varphi_t(x,t) - D_t^+\varphi(x,t)|, |\varphi_{xx}(x,t) - D_x^2\varphi(x,t)| : (x,t) \in G_{C\ell}\}$$

is the largest. This number, which we call ε, is again infinitesimal. Therefore we have

$$\max_{G_{C\ell}}\{|u(x,t)-U(x,t)|, |\varphi_t(x,t)-D_t^+\varphi(x,t)|, |\varphi_{xx}(x,t)-D_x^2\varphi(x,t)|\} \le \varepsilon \quad (33)$$

for some infinitesimal $\varepsilon > 0$.

The remainder of the argument is very similar to our proof in Section 3.4 that the distribution u defined was a distribution solution to the linear heat equation. Namely, if we replace u by U and derivatives of φ by finite differences of φ in (32), then (33) implies that we only make an infinitesimal error on both sides. We therefore only have to prove

$$\sum_{(x,t)\in G_C} \{U(x,t)(-D_x^2\varphi - D_t^+\varphi_t) - f(u(x,t))\varphi\}\,dx\,dt \approx \sum_{(x,0)\in G_C} u_0(x)\varphi(x,0)\,dx.$$

This follows after applying summation by parts, and using the finite difference equation (22) satisfied by U. This completes the existence proof.

Acknowledgement

Angenent has enjoyed and benefited from conversations about nonstandard analysis with members of the logic group at Madison, and would in particular like to thank Terry Millar for his relentless advocacy of matters both infinitesimal and unlimited.

References

[Albeverio et al. 1986] S. Albeverio, R. Høegh-Krohn, J. E. Fenstad, and T. Lindstrøm, *Nonstandard methods in stochastic analysis and mathematical physics*, Pure and Appl. Math. **122**, Academic Press, Orlando, FL, 1986. MR Zbl

[Fisher 1937] R. A. Fisher, "The wave of advance of advantageous genes", *Ann. Eugenics* **7**:4 (1937), 355–369. Zbl

[Folland 1999] G. B. Folland, *Real analysis: modern techniques and their applications*, 2nd ed., Wiley, New York, 1999. MR Zbl

[Goldblatt 1998] R. Goldblatt, *Lectures on the hyperreals*, Graduate Texts in Math. **188**, Springer, 1998. MR Zbl

[Henry 1981] D. Henry, *Geometric theory of semilinear parabolic equations*, Lecture Notes in Math. **840**, Springer, 1981. MR Zbl

[John 1982] F. John, *Partial differential equations*, 4th ed., Appl. Math. Sciences **1**, Springer, 1982. MR Zbl

[Keisler 1976a] H. J. Keisler, *Elementary calculus: an infinitesimal approach*, Prindle, Weber & Schmidt, Boston, 1976. Zbl

[Keisler 1976b] H. J. Keisler, *Foundations of infinitesimal calculus*, Prindle, Weber & Schmidt, Boston, 1976.

[Kolmogorov et al. 1937] A. N. Kolmogorov, I. G. Petrovsky, and N. S. Piskunov, "A study of the diffusion equation with increase in the amount of substance, and its application to a biological problem", *Vestnik Moskov. Univ. Mat. Meh.* **1**:6 (1937), 1–26. In Russian; translated in *Selected works of A. N. Kolmogorov, Volume I* (1991), Springer, 242–270.

[LeVeque 2007] R. J. LeVeque, *Finite difference methods for ordinary and partial differential equations*, Soc. Indust. Appl. Math., Philadelphia, 2007. MR Zbl

[Nelson 1987] E. Nelson, *Radically elementary probability theory*, Ann. Math. Studies **117**, Princeton Univ. Press, 1987. MR Zbl

[Pazy 1983] A. Pazy, *Semigroups of linear operators and applications to partial differential equations*, Appl. Math. Sciences **44**, Springer, 1983. MR Zbl

[Press et al. 2007] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical recipes: the art of scientific computing*, 3rd ed., Cambridge Univ. Press, 2007. MR Zbl

[Tao 2007] T. Tao, "Ultrafilters, nonstandard analysis, and epsilon management", blog post, 2007, available at https://tinyurl.com/ultrafilt.

Received: 2018-09-19 Revised: 2019-03-28 Accepted: 2019-04-02

cjo5325@psu.edu *Department of Mathematics, University of Wisconsin, Madison, WI, United States*

marshall.mueller@tufts.edu *Department of Mathematics, Tufts University, Medford, MA, United States*

angenent@wisc.edu *Department of Mathematics, University of Wisconsin, Madison, WI, United States*

involve

msp.org/involve

INVOLVE YOUR STUDENTS IN RESEARCH

Involve showcases and encourages high-quality mathematical research involving students from all academic levels. The editorial board consists of mathematical scientists committed to nurturing student participation in research. Bridging the gap between the extremes of purely undergraduate research journals and mainstream research journals, *Involve* provides a venue to mathematicians wishing to encourage the creative involvement of students.

See inside back cover or msp.org/involve for submission instructions. The subscription price for 2019 is US $195/year for the electronic version, and $260/year (+$35, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues and changes of subscriber address should be sent to MSP.

Involve (ISSN 1944-4184 electronic, 1944-4176 printed) at Mathematical Sciences Publishers, 798 Evans Hall #3840, c/o University of California, Berkeley, CA 94720-3840, is published continuously online. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices.

Involve peer review and production are managed by EditFlow® from Mathematical Sciences Publishers.

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

involve
2019 vol. 12 no. 6