Nonstandard existence proofs for reaction diffusion equations

Connor Olson, Marshall Mueller, and Sigurd B. Angenent

Abstract. 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.

CONTENTS

1. Introduction 1
2. Distribution solutions 4
3. The Cauchy problem for the heat equation 6
4. The Cauchy problem for a Reaction Diffusion Equation 11
References 16

1. Introduction

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

\[
\frac{\partial u}{\partial t} = D \frac{\partial^2 u}{\partial x^2} + f(u(x,t)), \quad x \in \mathbb{R}, t \geq 0
\]

with prescribed initial condition

\[
u(x,0) = u_0(x).
\]

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 [9], [2]) in which the reaction term is given by \(f(u) = u - u^2\).

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 [5] or [12]). As various techniques for constructing solutions are known, including the use of finite difference approximations to construct solutions (see [6, Ch.7.2]), our main goal is not to give another existence proof. Instead, we were inspired by several introductory texts on nonstandard analysis (notably, Keisler’s undergraduate calculus text [8], Goldblatt’s more advanced introduction to the hyperreals [4], Nelson’s “radically elementary approach” to probability [11], as well as Terry Tao’s blog post [14]), and wanted to see what some standard existence proofs would look like in the language of nonstandard analysis.

In [8] Keisler presents a proof of Peano’s existence theorem for solutions to ordinary differential equations

\[ \frac{dx}{dt} = f(t, x(t)), \quad x(0) = x_0, \]

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

\[ \frac{x_{i+1, \Delta t} - x_{i, \Delta t}}{\Delta t} = f(i\Delta t, x_{i, \Delta t}), \quad i = 0, 1, 2, \ldots \]

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 of 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’s text [8] 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 \( \Delta t \) 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.

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 makes the point in his blog post [14]), 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

\[
\frac{u(x, t + \Delta t) - u(x, t)}{\Delta t} = D \frac{u(x + \Delta x, t) - 2u(x, t) + u(x - \Delta x, t)}{(\Delta x)^2} + f(u(x, t)).
\]

This kind of discretization is very common in numerical analysis (see [10], or [13]).

For given initial data (but no boundary data) one can use this difference equation to inductively compute the values of \((x, t)\) for all \((x, t)\) in a triangular grid (see Figure 1).

Given a solution of the difference equation one can define a generalized function, or distribution

\[
\langle U, \varphi \rangle \overset{\text{def}}{=} \sum_t \sum_x U(x, t) \varphi(x, t).
\]

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 \(L^\infty\)) by letting \(\Delta x\) and \(\Delta t\) be infinitesimal positive hyperreals, and by taking the standard part of the expression on the right in (5). In both the standard and nonstandard setting this approach works for the linear heat equation, i.e. 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 point-wise 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 Ascoli-Arzela and extract a convergent subsequence. However, since the domain \(\mathbb{R}^2\) is not compact, Ascoli-Arzela 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. 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 (e.g., a very incomplete list:[7, 8, 4, 11, 1, 14]).

There are a few different approaches to using the hyperreals. Keisler, in his undergraduate nonstandard calculus text [8] gives 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 [11].

Our point of view in this paper is that of internal set theory as explained in Goldblatt’s book [4] (see also Keisler’s “instructor’s guide” [7] to his calculus text). Goldblatt explains the construction of the hyperreals using non principal ultra filters (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.

Acknowledgement. SBA 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.

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’s book [3].

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

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

for all so-called “test functions” $\varphi$. 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 $C^\infty_c(\mathbb{R}^2)$, or sometimes by $\mathcal{D}(\mathbb{R}^2)$, is an infinite dimensional vector space. By definition, a distribution is any linear functional $T : C^\infty_c(\mathbb{R}^2) \to \mathbb{R}$. The most common notation for the value of a distribution $T$ applied to a test function $\varphi$ is $\langle T, \varphi \rangle$. For instance, if $f : \mathbb{R}^2 \to \mathbb{R}$ is a continuous function, then the equation (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

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

The full definition of a distribution $T$ includes the requirement that the value $\langle T, \varphi \rangle$ depend continuously on the test function $\varphi$. To state this continuity condition precisely one must introduce a notion of convergence in the space of test functions $C^\infty_c(\mathbb{R}^2)$. We refer the reader to [3] 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

$$(7) \quad |\langle T, \varphi \rangle| \leq C \int \int_{\mathbb{R}^2} |\varphi(x, y)| \, dx \, dy$$

holds for all test functions $\varphi$. Alternatively, if a constant $C$ exists such that

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

holds for all $\varphi \in C^\infty_c(\mathbb{R}^2)$, 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 $Ndx$ and $Mdy$ are unlimited. Consider the rectangular grid

$$(9) \quad G = \{(kdx, ldy) \in \ast \mathbb{R}^2 \mid k, l \in \ast \mathbb{N}, |k| \leq N, |l| \leq M\}$$

From the point of view of nonstandard analysis and internal set theory, $G$ is a hyperfinite set, and for any internal function $f : G \to \ast \mathbb{R}$ 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

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

where $x \approx y$ 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 $L, L' \in \ast \mathbb{N}$ for which $Ldx \leq \ell < (L + 1)dx$ and $L'dy \leq \ell < (L' + 1)dy$. Then the nonstandard definition of the Riemann integral implies

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

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

$$\sum_{k=-L}^{L} \sum_{l=-L'}^{L'} g(kdx, ldy) \, dx \, dy = \sum_{(x, y) \in G} g(x, y) \, dx \, dy.$$
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 $|f(x,y)| \leq C$ for all $(x,y) \in G$. Then the expression

$$
⟨Tf, ϕ⟩ \triangleq \text{St} \left( \sum_{(x,y) \in G} f(x,y)ϕ(x,y) \, dx \, dy \right)
$$

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 $|f(x,y)| \leq C$ for all $(x,y)$ we have

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

Hence the sum in the definition (10) of $⟨Tf, ϕ⟩$ is a limited hyperreal, whose standard part is a well defined real number which satisfies

$$
|⟨Tf, ϕ⟩| \leq C \int_{\mathbb{R}^2} |ϕ(x,y)| \, dx \, dy.
$$

Therefore $T_f$ is a well defined distribution.

Let $⟨f, ϕ⟩$ be defined as in (6). Fix $ϕ$. We then have

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

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 $u(x,0) = u_0(x)$.

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

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

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

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 $ϕ \in C_c^\infty(\Omega)$. To explain this in more
detail, recall that $\delta$ is Dirac’s delta function, so that the action of the right hand side in (11) on a test function is

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

The definition of distributional derivative [3, ch. 9] says that the action of the left hand side in (11) is given by

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

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 $t \geq 0$ (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

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

for all test functions $\varphi \in C_0^\infty(\mathbb{R}^2)$. 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:

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

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

See for example, [10, chapter 1]. With this notation

$$D_t^2 U(x,t) = D_x^+ D_x^- U(x,t) = \frac{U(x + dx,t) - 2U(x,t) + U(x - dx,t)}{(dx)^2}$$  

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.

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

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

$$\begin{align*}
U(x,t + dt) &= \alpha U(x - dx,t) + (1 - 2\alpha)U(x,t) + \alpha U(x + dx,t)
\end{align*}$$  

where

$$\alpha \equiv \frac{dt}{(dx)^2}.$$
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.

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

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

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

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

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

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

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

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 it in slightly greater generality than needed in this section. The lemma explains why we need the condition $\alpha = dt/(dx)^2 \leq 1/4$, 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 $W : G_C \to \ast \mathbb{R}$ satisfy $W(x,t) \geq 0$ for all $(x,t) \in G_C$, and suppose that for some nonnegative $m \in \mathbb{R}$ one has

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

at all $(x,t) \in G_C$. For each $t = ndt$ with $0 \leq n \leq N$ consider

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

If $0 \leq \alpha \leq 1/2$ then

$$w(t) \leq e^{nt} w(0).$$
Taking the maximum over $x$ we then have for $t \in G_C$ that
$$W(x, t + dt) = αW(x + dx, t) + (1 - 2α)W(x, t) + αW(x - dx, t) + dt \cdot W(x, t)$$
$$≤ (α + (1 - 2α) + α + m dt)w(t)$$
$$= (1 + m dt)w(t).$$

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

### 3.4. Proof of Theorem 3.3.1

The relation (14) which defines $U(x, t)$ implies that $W(x, t) \overset{\text{def}}{=} |U(x, t)|$ satisfies
$$W(x, t + dt) = |αU(x - dx, t) + (1 - 2α)U(x, t) + αU(x + dx, t)|$$
$$≤ αW(x - dx, t) + (1 - 2α)W(x, t) + αW(x + dx, t),$$
where we have used $α ≥ 0$ and $1 - 2α ≥ 0$.

Since the initial condition is bounded by $|U(x, 0)| = |u_0(x)| ≤ M$ the Gronwall type lemma 3.3.2 implies that $|U(x, t)| ≤ M$ 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
$$\langle u_t - u_{xx}, φ \rangle = \langle u_0(x)δ(t), φ \rangle = \int_{\mathbb{R}} u_0(x)φ(x, 0) \, dx.$$ 

First, we see from the definition of distributional derivative that
$$\langle u_t - u_{xx}, φ \rangle = -\langle u, φ_t + φ_{xx} \rangle.$$ 

We then have from the definition of $u$ that
$$-\langle u, φ_t + φ_{xx} \rangle ≈ \sum_{(x,t) \in G_C} -U(x, t)(φ_t(x, t) + φ_{xx}(x, t)) \, dx \, dt \overset{\text{def}}{=} T.$$ 

Using Taylor’s formula we replace the partial derivatives of the test function with its corresponding finite differences, i.e. we write $φ_t(x, t) = D_t^+ φ(x, t) + ε_t(x, t)$ and $φ_{xx}(x, t) = D_x^2 φ(x, t) + ε_{xx}(x, t)$, where $ε_t, ε_{xx} : G_C \rightarrow ^*\mathbb{R}$ are the infinitesimal error terms in the Taylor expansion. Substituting these gives us
$$T = \sum_{G_C} -U(x, t)(D_t^+ φ(x, t) + D_x^2 φ(x, t) + ε_t + ε_{xx}) \, dx \, dt.$$ 

We can split this sum into three parts, $T = T_1 + T_2 + T_3$, with
$$T_1 = \sum_{G_C} -U(x, t)D_t^+ φ(x, t) \, dx \, dt$$
$$T_2 = \sum_{G_C} -U(x, t)D_x^2 φ(x, t) \, dx \, dt$$
$$T_3 = \sum_{G_C} -U(x, t)(ε_t + ε_{xx}) \, dx \, dt.$$ 

We will first handle the error term, $T_3$. Since the test function $φ$ has compact support, there exists a real $ℓ > 0$ such that $φ = 0$ outside the rectangle $Ω = [-ℓ, ℓ] \times [-ℓ, ℓ]$. The errors in the Taylor expansion therefore also vanish outside of $Ω$ so that we can write $T_3$ as
$$T_3 = \int_{Ω \cap G_C} -U(x, t)(ε_t + ε_{xx}) \, dx \, dt.$$
The key to estimating this sum is that we can estimate all the errors \( \varepsilon_1(x, t) \) 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 \( |\varepsilon_1| : G_C \to ^*\mathbb{R} \) attains its largest value at one of the \( (x, t) \in G_C \), say at \( (x_1, t_1) \). Then \( |\varepsilon_1(x, t)| \leq |\varepsilon_1(x_1, t_1)| \) 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

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

Then both \( \varepsilon_1 \) and \( \varepsilon_2 \) are positive infinitesimals for which

\[
|\varepsilon_1(x, t)| \leq \varepsilon_1, \quad |\varepsilon_{xx}(x, t)| \leq \varepsilon_2
\]

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

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

\[
|T_3| \leq \sum_{G_C \cap \Omega} 2M\varepsilon \, dx \, dt \leq 2M\varepsilon \, dx \, dt \frac{\ell}{2dx} \frac{\ell}{dt} = M\ell^2\varepsilon,
\]

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 \( \varphi \), 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

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

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

\[
T_2 = \sum_{l=0}^{L} \sum_{k=-K-1}^{K+1} -U(k \, dx, l \, dt) D_x^+ \varphi(k \, dx, l \, dt) \, dt \, dx
\]

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

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 \( T_3 \approx 0 \),

\[
T_1 + T_2 + T_3 \approx T_1 + T_2 = \sum_{k=-K}^{K} U(k \, dx, 0) \varphi(k \, dx, 0) \, dx
\]

\[
+ \sum_{k=-K}^{K} \sum_{l=0}^{L} \varphi(k \, dx, l \, dt) \left( D_x^+ U(k \, dx, l \, dt) - D_x^2 U(k \, dx, l \, dt) \right) \, dx \, dt.
\]
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
\[
\text{St}(T) = \text{St} \left( \sum_{k=-K}^{K} u_0(k dx)\varphi(k dx, 0) dx \right) = \int_{-\ell}^{\ell} u_0(x)\varphi(x, 0) dx = \langle u_0(x)\delta(t), \varphi \rangle.
\]
This completes the proof that \( \langle u_t - u_{xx}, \varphi \rangle = \langle u_0(x)\delta(t), \varphi \rangle \) for all test functions \( \varphi \), 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
\[
\sum_{|k| \leq K} f(k dx) dx \approx \int_{-\ell}^{\ell} f(x) dx
\]
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
\[
\begin{aligned}
&\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + f(u(x,t)), \\
&(x \in \mathbb{R}, t > 0)
\end{aligned}
\]
with initial condition (1b). It is known that one cannot expect solutions to exist for all time \( t > 0 \) without imposing some growth conditions on the nonlinearity \( f(u) \). We will assume that \( f \) is a Lipschitz continuous function, i.e. that for some positive real \( K_1 \) one has
\[
\forall u, v \in \mathbb{R} : \quad |f(u) - f(v)| \leq K_1 |u - v|.
\]
This implies that \( f(u) \) grows at most linearly in \( u \):
\[
\forall u \in \mathbb{R} : \quad |f(u)| \leq K_0 + K_1 |u|
\]
where \( K_0 \overset{\text{def}}{=} |f(0)| \).

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.
\[
\forall x, y \in \mathbb{R} : \quad |u_0(x) - u_0(y)| \leq L|x - y|.
\]
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

\[
\int_{\mathbb{R} \times [0, \infty)} \left\{ u(x, t)(-\varphi_{xx} - \varphi_t) - f(u(x, t))\varphi \right\} dx dt = \int_{\mathbb{R}} u_0(x)\varphi(x, 0) dx,
\]

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

**Theorem 4.1.1.** If $f$ is Lipschitz continuous as in (18), and if the initial function $u_0$ is bounded by $\forall x \in \mathbb{R} : |u_0(x)| \leq M,$

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 §3.3 with infinitesimal mesh sizes $dx, dt > 0$, and consider the finite difference scheme

\[
D_t^+ U(x, t) = D_x^2 U(x, t) + f(U(x, t)).
\]

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

\[
U(x, t + dt) = \alpha U(x + dx, t) + (1 - 2\alpha)U(x, t) + \alpha U(x - dx, t) + dt f(U(x, t)),
\]

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 establish a bound for $|U(x, t)|$.

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

\[
|U(x, t)| \leq e^{K_1 t} M + K_0 K_1 (e^{K_1 t} - 1).
\]

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

\[
|U(x, t + dt)| \leq \alpha |U(x + dx, t)| + (1 - 2\alpha) |U(x, t)| + \alpha |U(x - dx, t)| + dt (K_0 + K_1 |U(x, t)|).
\]
In terms of $M(t) = \max_x |U(x,t)|$ this implies
\[
M(t + dt) \leq M(t) + dt(K_0 + K_1 M(t)) = (1 + K_1 dt)M(t) + K_0 dt.
\]
Setting $t = ndt$ we see that this is an inequality of the form $M_n \leq a M_{n-1} + b$ with $M_n = M(ndt)$. By induction this implies that
\[
M(t) = M(ndt) \leq (1 + K_1 dt)^n M(0) + \frac{(1 + K_1 dt)^n - 1}{1 + K_1 dt - 1} K_0 dt
\]
where $L$ is the Lipschitz constant for the initial function $u_0$, as in (20).

\[
\text{Proof. Let}
\]
\[
V(x,t) \overset{\text{def}}{=} \frac{U(x+dx,t) - U(x,t)}{dx} = D_x^+ U(x,t).
\]
Applying $D_x^+$ to both sides of the equation (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 $D_t^+ f(U)$, we have
\[
|D_t^+ f(U)| = \left| \frac{f(U(x+dx,t)) - f(U(x,t))}{dx} \right| \leq \frac{K_1 |U(x+dx,t) - U(x,t)|}{dx} = K_1 |V(x,t)|,
\]
so that
\[
|V(x,t + dt)| \leq \alpha |V(x + dx,t)| + (1 - 2\alpha)|V(x,t)| + \alpha |V(x - dx,t)| + K_1 dt |V(x,t)|.
\]
Using Gronwall’s Inequality on $\max_x V$, we get the inequality
\[
\max_x |V(x,t)| \leq e^{K_1 t} \max_x |V(x,0)|.
\]
The initial condition $u_0$ satisfies $|u_0(x) - u_0(x')| \leq L|x - x'|$ for all $x, x' \in \mathbb{R}$, and therefore the extension of $u_0$ to the hyperreals satisfies this same inequality. Therefore $|V(x,0)| \leq L$ for all grid points $(x,0)$, and thus we have $|V(x,t)| \leq Le^{K_1 t}$. This implies (26).
4.5. Hölder continuity in time of the approximate solution.

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

\[
|U(x_0, t_1) - U(x_0, t_0)| \leq C\sqrt{t_1 - t_0}
\]

where \( C \) is a constant that only depends on \( \tilde{t}, K_0, K_1, L, \) and \( M. \)

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

\[
|U(x, t)| \leq A_0 \overset{\text{def}}{=} e^{K_0 i M} + \frac{K_1}{K_0} \left( e^{K_0 i} - 1 \right),
\]

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

\[
\forall (x, t) \in G_C \text{ with } t \leq \tilde{t} : |f(U(x, t))| \leq A.
\]

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

\[
\bar{U}(x, t) \overset{\text{def}}{=} U(x_0, t_0) + a(t - t_0) + \frac{b}{2} (x - x_0)^2 + c.
\]

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

\[
D_t^+ \bar{U} - D_x^2 \bar{U} \geq A + 1,
\]

and

\[
\bar{U}(x, 0) > U(x, 0) \text{ for all } (x, 0) \in G_C.
\]

A direct computation shows that \( D_t^+ \bar{U} - D_x^2 \bar{U} = a - b, \) so for 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 \( \bar{L} \overset{\text{def}}{=} e^{K_0 i} L: \)

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

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

Next, we apply a maximum principle argument to compare \( U \) and \( \bar{U} \). Consider \( W(x, t) = U(x, t) - \bar{U}(x, t) \). 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 \( U(x_0, t) < \bar{U}(x_0, t) \) 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 \( t_1 - t_0 = \sqrt{t_1 - t_0} \sqrt{t_1 - t_0} \leq \sqrt{t_1 - t_0} \) we get

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

\[
|U(x, t) - U(y, s)| \leq |U(x, t) - U(y, t)| + |U(y, t) - U(y, s)| 
\leq L|x - y| + C \sqrt{|t - s|}.
\]

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 \((\hat{x}, \hat{t}) \in G_C\) with \(x \approx \hat{x}\) and \(t \approx \hat{t}\), and then define \(u(x, t) = \text{St}(U(\hat{x}, \hat{t}))\). The continuity property (31) of the approximate solution \(U\) implies that the value of \(\text{St}(U(\hat{x}, \hat{t}))\) does not depend on how we chose the grid point \((\hat{x}, \hat{t})\), for if \((\hat{x}, \hat{t}) \in G_C\) also satisfied \(\hat{x} \approx x\), \(\hat{t} \approx t\), then \(\hat{x} \approx \hat{x}\) and \(\hat{t} \approx \hat{t}\), so that \(U(\hat{x}, \hat{t}) \approx U(\hat{x}, \hat{t})\). 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 \(x' = \text{St}(x)\) and \(t' = \text{St}(t)\) are well defined real numbers. By continuity of both \(u\) we have \(u(x, t) \approx u(x', t')\). By 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 \(\varphi \in C_0^\infty(\mathbb{R}^2)\).

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

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

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

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

We now intend to replace \(u\) by \(U\), and the derivatives of \(\varphi\) 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

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

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 $\varepsilon$, is again infinitesimal. Therefore we have

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

for some infinitesimal $\varepsilon > 0$.

The remainder of the argument is very similar to our proof in § 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 $\varphi$ by finite differences of $\varphi$ in (32), then (33) implies that we only make an infinitesimal error on both sides. We therefore only have to prove

$$\sum_{(x, t) \in G_{C \ell}} \left\{U(x, t)(-D_x^2 \varphi - D_{xx}^+ \varphi_t) - f(u(x, t))\varphi\right\} dx \, dt \approx \sum_{(x, 0) \in G_{C \ell}} u_0(x)\varphi(x, 0) \, dx.$$

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

References

[1] Sergio Albeverio, Raphael Høegh-Krohn, Jens Erik Fenstad, and Tom Lindstrom. Nonstandard methods in stochastic analysis and mathematical physics, volume 122 of Pure and Applied Mathematics. Academic Press, Inc., Orlando, FL, 1986.

[2] R. A. Fisher. The wave of advance of advantageous genes. Annals of Eugenics, 7(4):355–369, 1937.

[3] Gerald B. Folland. Real analysis. Pure and Applied Mathematics (New York). John Wiley & Sons, Inc., New York, second edition, 1999. Modern techniques and their applications, A Wiley-Interscience Publication.

[4] Robert Goldblatt. Lectures on the hyperreals, volume 188 of Graduate Texts in Mathematics. Springer-Verlag, New York, 1998. An introduction to nonstandard analysis.

[5] D. Henry. Geometric theory of semilinear parabolic equations. Lecture notes in mathematics. Springer-Verlag, 1981.

[6] F. John. Partial Differential Equations. Applied Mathematical Sciences. Springer New York, 1991.

[7] H.J. Keisler. Foundations of Infinitesimal Calculus. on-line at https://www.math.wisc.edu/~keisler/foundations.html, 2007.

[8] H.J. Keisler. Elementary Calculus: An Infinitesimal Approach. Dover Books on Mathematics Series. Dover Publications, 2012.

[9] A. Kolmogorov, I. Petrovskii, and N. Piskunov. A study of the diffusion equation with increase in the amount of substance, and its application to a biological problem. In V. M. Tikhomirov, editor, Selected Works of A. N. Kolmogorov I, page 248–270. Kluwer, 1991. ISBN 90-277-2796-1. Translated by V. M. Volosov from Bull. Moscow Univ., Math. Mech. 1, 1–25, 1937.

[10] Randall J. LeVeque. Finite difference methods for ordinary and partial differential equations. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2007. Steady-state and time-dependent problems.

[11] Edward Nelson. Radically elementary probability theory, volume 117 of Annals of Mathematics Studies. Princeton University Press, Princeton, NJ, 1987.
[12] A. Pazy. *Semigroups of Linear Operators and Applications to Partial Differential Equations*. Applied Mathematical Sciences. Springer New York, 2012.

[13] W.H. Press, S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery. *Numerical Recipes 3rd Edition: The Art of Scientific Computing*. Cambridge University Press, 2007.

[14] Terence Tao. Ultrafilters, nonstandard analysis, and epsilon management. https://terrytao.wordpress.com/2007/06/25/ultrafilters-nonstandard-analysis-and-epsilon-management/ Accessed: 2018-07-19.

C.O.: Department of Mathematics, University of Wisconsin – Madison

M.M.: Department of Mathematics, Tufts University – 503 Boston Avenue, Medford, MA 02155

S.B.A.: Department of Mathematics, University of Wisconsin – Madison