LBL-22420
Preprint

Lawrence Berkeley Laboratory

UNIVERSITY OF CALIFORNIA

Physics Division

Mathematics Department

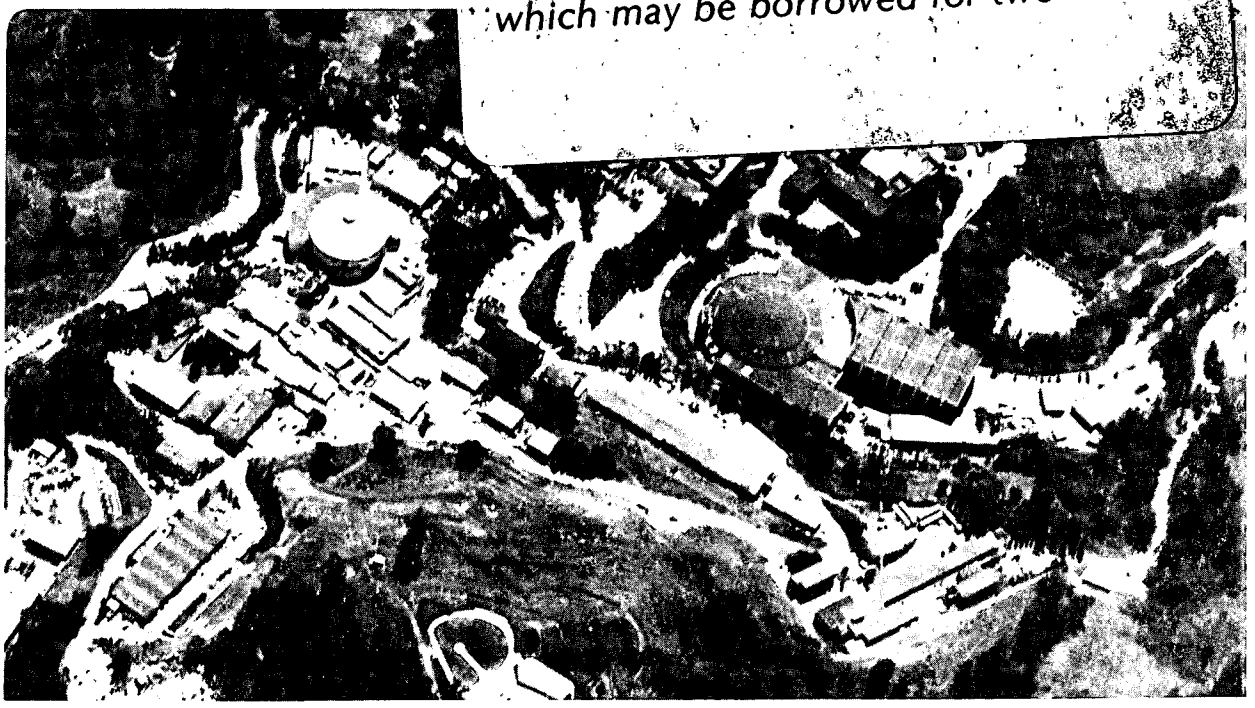
To be submitted for publication

CONVERGENCE OF A RANDOM PARTICLE METHOD TO SOLUTIONS OF THE KOLMOGOROV EQUATION $u_t = \nu u_{xx} + u(1-u)$

E.G. Puckett

April 1987

TWO-WEEK LOAN COPY
This is a Library Circulating Copy
which may be borrowed for two weeks.



LBL-22420 c.2

Prepared for the U.S. Department of Energy under Contract DE-AC03-76SF00098

DISCLAIMER

This document was prepared as an account of work sponsored by the United States Government. While this document is believed to contain correct information, neither the United States Government nor any agency thereof, nor the Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or the Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof or the Regents of the University of California.

LBL-22420

CONVERGENCE OF A RANDOM PARTICLE METHOD TO SOLUTIONS OF THE KOLMOGOROV EQUATION $u_t = \nu u_{xx} + u(1-u)$[†]

Elbridge Gerry Puckett[‡]

Department of Mathematics
and
Lawrence Berkeley Laboratory
University of California
Berkeley, California 94720

April 1987

[†]Supported in part by the Applied Mathematical Sciences Subprogram of the Office of Energy Research, U.S. Department of Energy under contract DE-AC03-76SF00098.

[‡]Current address: Lawrence Livermore National Laboratory, Livermore, CA 94550.

Convergence of a Random Particle Method to Solutions of the Kolmogorov Equation $u_t = \nu\, u_{xx} + u\,(1-u)$

Elbridge Gerry Puckett †

U. C. Berkeley Mathematics Department ‡
and Lawrence Berkeley Laboratory
Berkeley, California 94720

ABSTRACT

We study a random particle method for solving the reaction-diffusion equation $u_t = \nu\, u_{xx} + f\,(u)$ which is a one dimensional analogue of the random vortex method in fluid mechanics. Our method is a fractional step method in which $u_t = \nu\, u_{xx}$ is solved by random walking the particles while $u_t = f\,(u)$ is solved by Euler's method. For the case when $f\,(u) = u\,(1-u)$, i.e. the Kolmogorov equation, we prove that for $\Delta t = O\,(\sqrt[4]{N}^{-1})$ the numerical method converges like $\ln N \cdot \sqrt[4]{N}^{-1}$ uniformly as the diffusion coefficient ν tends to 0. Thus, travelling waves with arbitrarily steep wavefronts may be modeled without an increase in the computational cost. We also present numerical results that include experiments with second order time discretization and second order operator splitting.

Keywords : Kolmogorov equation, particle method, random vortex method, random walk

AMS Subject Classification : 35, 60, 65, 76

† This work was supported in part by the Applied Mathematical Sciences Subprogram of the Office of Energy Research, U.S. Department of Energy under contract DE-AC03-76SF00098 and is part of the author's PhD dissertation in the U. C. Berkeley Mathematics Department.

‡ Current address: Lawrence Livermore National Laboratory, Livermore, CA 94550

1. Introduction

In this paper we study a random particle method, due to Chorin [15] for approximating solutions of the one dimensional *reaction-diffusion equation,*

$$u_t = \nu\, u_{xx} + f\,(u) \tag{1.1a}$$

$$u\,(x\,,0) = u^0(x\,) \tag{1.1b}$$

where the *forcing function,* $f\,(u\,)$, satisfies

$$f\,(0) = f\,(1) = 0\,, \tag{1.2a}$$

$$f\,(u\,) > 0 \qquad \text{for} \quad 0 < u < 1\,, \tag{1.2b}$$

$$f\,'(u\,) \leq 1 \qquad \text{for} \quad 0 \leq u \leq 1. \tag{1.2c}$$

We call this method the *random gradient method.* Algorithms based on this method have been used by Sherman and Peskin to solve Nagumo's equation [50] and the Hodgkin-Huxley Equations [51]. We prove the convergence of the random gradient method to solutions of the *Kolmogorov equation,*

$$u_t = \nu\, u_{xx} + u\,(1-u\,) \tag{1.3a}$$

$$u\,(x\,,0) = u^0(x\,) \tag{1.3b}$$

subject to the constraints

$$0 \leq u^0(x\,) \leq 1\,, \tag{1.3c}$$

$$\lim_{x \to -\infty} u^0(x\,) = 1\,, \tag{1.3d}$$

$$\lim_{x \to +\infty} u^0(x) = 0 \,. \tag{1.3e}$$

Our work follows that of Roberts [47] who proved the convergence of a random particle method to Burgers' equation,

$$u_t + uu_x = \nu u_{xx} \,.$$

Related theoretical work includes Brenier [9], Cottet and Gallic [20], Hald [32, 33], Raviart [44, 45], and Rosen [48]. A review of particle methods which use random walks to model diffusion may be found in Ghoniem and Sherman [26].

Our interest in the random gradient method is primarily motivated by the fact that it is a one dimensional analogue of Chorin's random vortex method [13] for approximating solutions of the Navier-Stokes equations in two and three dimensions. It is hoped that a thorough examination of the errors obtained when using the random gradient method will yield a greater understanding of the error inherent in using the random vortex method, particularly the error due to the use of random walks to simulate diffusion. In order to motivate the subsequent discussion it seems appropriate to list here the most important characteristics that these two methods have in common.

i) Both are ***particle methods***, with the particles representing point concentrations of some derivative of the solution. (Gradients of u with respect to x in the case of the random gradient method, vorticity in the case of the random vortex method.)

ii) Both are ***splitting*** or ***fractional step*** methods. That is, the equation to be solved is split into two evolution equations, each of which is solved separately. This process is coupled by using the solution obtained after solving one of the evolution equations as the initial data for the other.

iii) In both methods one of the fractional steps is the ***heat equation***,

$$u_t = \nu \, \Delta u \, . \tag{1.4}$$

In each instance the numerical solution to the heat equation is obtained by random walking the particles.

iv) Finally, in both instances the second of the fractional steps is a nonlinear evolution equation. In the case of the random gradient method this is simply the ordinary differential equation (ODE)

$$u_t = f\,(u\,) \tag{1.5}$$

whereas for the random vortex method it is the Euler equations.

Similar analogies may be drawn between the present method and Chorin's vortex sheet method [14] for approximating solutions of the Prandtl boundary layer equations.

Numerical estimates of the convergence rate for the random vortex method were made by Roberts [46] while a convergence proof for the method in the absence of boundaries may be found in Goodman [27]. Theoretical work on the vortex method solution of the Euler equations include Anderson and Greengard [3], Beale and Majda [5, 6], Cottet [19], Greengard [28], Hald and Del Prete [30], and Hald [31, 34].

It should be noted that in our treatment of the random gradient method particles are not permitted to divide in two when their strengths surpass some critical value as was originally proposed by Chorin. This greatly simplifies the convergence proof. In fact, it is interesting to note that creation of particles is not necessary for convergence, at least when the approximation to the initial data is taken to be monotonically decreasing as we do here. There are, however, several good arguments for why the algorithm with particle creation should be more accurate. Furthermore, the difficulties which arise in trying to prove convergence for the algorithm with particle creation are very similar to those which arise when trying to prove convergence of the

random vortex method when boundaries are present. In this case particle creation corresponds to the creation of vorticity, an important phenomenon in fluid flow. Hald [33] has successfully attacked this problem for a one dimensional diffusion equation with thermal convection.

The main result of this paper may be stated as follows.

THEOREM 1.1 Assume $\nu \leq 1$. Fix $T > 0$ and choose a time step $\Delta t > 0$ such that $T = k\,\Delta t$ for some integer k. Let $u(x,T)$ be the solution at time T of equation (1.3) with initial data u^0 and let $\tilde{u}^k(x)$ be the corresponding 'computed solution' with initial data $\tilde{u}^0$. Denote the number of particles used to generate $\tilde{u}^k$ by N and assume that $\Delta t = O(\sqrt[4]{N}^{-1})$. Then there exist positive constants C_1 and C_2, independent of ν, Δt, and N, such that

$$\mathrm{E}\ \| u(T) - \tilde{u}^k \|_{L^1} \leq (1+T)\left[e^T \| u^0 - \tilde{u}^0 \|_{L^1} + C_1 \sqrt{\nu}\,\Delta t + C_2 \frac{\ln N}{\sqrt[4]{N}} \right] \tag{1.6}$$

and

$$var\left(\| u(T) - \tilde{u}^k \|_{L^1} \right) \leq (1+T)\left[e^T \| u^0 - \tilde{u}^0 \|_{L^1} + C_1 \sqrt{\nu}\,\Delta t + C_2 \frac{\ln N}{\sqrt[4]{N}} \right]^2. \tag{1.7}$$

Here E denotes expected value and *var* the variance.

In order to prove this theorem several assumptions regarding u^0 and $\tilde{u}^0$ have been made. In addition to satisfying the constraints (1.3c-e) it has been assumed that u^0 is continuously differentiable on $\mathbb{R}$ and $u^0_x \in L^1(\mathbb{R}) \cap L^\infty(\mathbb{R})$. The approximate initial data, $\tilde{u}^0$, is a step function approximation to u^0 and is required to be monotonically decreasing. This is perhaps the only 'unnatural' restriction on the initial data. We discuss this issue at the beginning of §2.3 and in somewhat greater depth in the remark after the proof of Theorem 4.1.

One of the most important consequences of Theorem 1.1 is that the error is independent of the diffusion coefficient, or *viscosity*, ν. Thus, solutions with arbitrarily steep wavefronts may

be modeled without any increase in the computational cost. In contrast, a finite difference calculation would require $\Delta x << O(\sqrt{\nu})$ to accurately resolve wavefronts with viscosity ν. In this regard see Sherman and Peskin [51].

This 'favorable' dependence on ν is what makes random walk methods competitive. In the absence of some variance reduction technique (for example see Chang [11]) the random walk algorithm will converge at a rate of $O(\sqrt{N}^{-1})$. This is much slower than, say, a finite difference method designed to solve the same problem. However, with a finite difference method one must take $\Delta x << O(\sqrt{\nu})$ in order to accurately resolve the effects due to viscosity whereas a random walk method automatically concentrates the computational elements in regions of interest. Furthermore, the random walk introduces no 'numerical diffusion'. (For a discussion of numerical diffusion see Sod [53].) Unfortunately, the price one pays for being able to compute at small ν is a convergence rate of $\sqrt{N}^{-1}$.

Our approach is to prove the theorem for $\nu = 1$ and then use a simple scaling argument to demonstrate the validity of the result for arbitrary $\nu \leq 1$. The second term in the sum on the right hand side of (1.6), $C_1 \sqrt{\nu}\, \Delta t$, results from bounding the error due to exact operator splitting. That is, the error that results when (1.4) and (1.5) are solved exactly. We remark that in [7] Beale and Majda proved that exact operator splitting for the Navier-Stokes equations is $O(\nu\, \Delta t)$ in $L^2(\mathbb{R}^n)$ for $n = 2,3$.

The details of the random gradient method are developed in chapter 2; beginning with some notation (§2.1) and followed by the algorithm itself (§2.2). In §2.3 the class, $\mathbf{S}$, of permissible starting approximations, $\tilde{u}^0$, is defined and several preliminary lemmas are proved. Chapter 3 is devoted to outlining the proof of Theorem 1.1. Most of the error analysis is written in the language of solution operators. This notation is introduced in §3.1 and §3.2. A brief account of the proof may be found in §3.3, together with a description of how the details are divided among the chapters 4 through 7. In chapter 8 we put the various parts together and

prove the theorem for $\nu = 1$. Then, in §8.3, we remove the restriction $\nu = 1$ and prove Theorem 1.1 for arbitrary $\nu \leq 1$.

2. A Description of the Random Gradient Method

We begin this chapter with the introduction of some notation and a detailed description of the algorithm itself. This is followed by a discussion of the difficulties that are encountered for non-monotonic initial data and the proof of several basic facts that hold for monotonic initial data.

2.1 Step Function Notation We will denote the numerical approximation of some function, intended to be obtained on a computer, by the symbol ' $\tilde{}$ '. Thus, $\tilde{u}(x,t)$ denotes an approximation to the solution, $u(x,t)$, of equation (1.1a,b). We will use the term step function to refer to any piecewise constant function of $x \in \mathbb{R}$ that has a finite number of discontinuities.

In the random gradient method $\tilde{u}$ is a step function approximation to u. Consequently, knowledge of the position of each discontinuity and of the amount of each jump is all that is required in order to know $\tilde{u}$. It is convenient to think of $\tilde{u}$ at a given time t as being represented by N particles, each particle having associated with it a position on the x-axis and a *strength* or *weight*; the particle's position being a point at which $\tilde{u}$ is discontinuous and its' strength being the amount by which $\tilde{u}$ changes at that point. (The terms weight and strength will be used interchangeably.) The position of the *ith* particle at time $t = j\Delta t$ will be denoted by X_i^j and its weight by w_i^j. It is often desirable to write the computed solution after j time steps as $\tilde{u}^j(x)$ instead of $\tilde{u}(x, j\Delta t)$. Thus, the computed solution can be written in the following way:

$$\tilde{u}^j(x) = \sum_{i=1}^{N} H(X_i^j - x)\, w_i^j \tag{2.1}$$

where $H(x)$ is the Heaviside function

$$H(x) = \begin{cases} 0, & x < 0, \\ 1, & x \geq 0. \end{cases} \tag{2.2}$$

It will often be convenient to assume that the particles have been labeled so that for each j,

$$X_1^j \leq X_2^j \leq \ldots \leq X_N^j . \tag{2.3}$$

Note that this may require a relabeling of the particles at each time step, since random walking the particles can result in a different ordering of the particle positions. This is simply a notational convenience and has no effect on the actual details of the convergence proof.[1] Henceforth, we will assume that (2.3) holds, usually without comment.

Let $\tilde{u}_i^j = \tilde{u}^j(X_i^j)$ denote the value of $\tilde{u}^j$ at the ith particle position. For future reference we note that, by (2.1) and (2.2),

$$\tilde{u}_i^j = \sum_{r=1}^{N} H(X_r^j - X_i^j)\, w_r^j = \sum_{r \geq i} w_r^j \tag{2.4}$$

and, consequently, the strength of the ith particle is given by

$$w_i^j = \sum_{r \geq i} w_r^j - \sum_{r \geq i+1} w_r^j = \tilde{u}_i^j - \tilde{u}_{i+1}^j .$$

The variable N will always be used to denote the number of particles present in the flow; with N being fixed for a given run of the numerical method.

2.2 The Algorithm We begin the random gradient method by determining a step function approximation, $\tilde{u}^0$, to the exact initial data u^0. Given the computed solution, $\tilde{u}^j$, at time $j\Delta t$ the solution at time $(j+1)\Delta t$ is obtained in two distinct steps.

Step I: The first step is the numerical solution of $u_t = f(u)$. For fixed x this is an ODE in t

[1] This is, however, a reflection of the need to keep the particle positions sorted if one wishes to recover the function values $\tilde{u}(x)$ efficiently. See the paragraph following equation (2.7) for further details.

with initial data $\tilde{u}^j(x)$. The solution of this equation can easily be obtained using any explicit ODE solver. In the convergence proof that follows we will assume that Euler's method is used. It should be noted, however, that the analysis (Chapter 6) carries through for higher order Runge-Kutta methods as well. In addition, there are many cases in which $u_t = f(u)$ may be solved exactly. We note two such instances that may be of interest:

1) For equation (1.3a,b) we have $u_t = u(1-u)$ which is the case studied in this paper.

2) If $f(u) = -v(x)u$ (and so f is also a function of x), then equation (1.1a,b) is a one dimensional Schrodinger equation written in imaginary time, $u_t = u_{xx} - v(x)u$. This is a one dimensional version of the case considered by Alder *et al* in [1] and [2], although there they solve the ODE by creating and annihilating particles with fixed strengths rather than by altering the weights.

When the solution of the ODE is obtained using Euler's method the value of the intermediate solution, $\tilde{v}^{j+1}$, at the point x is given by

$$\tilde{v}^{j+1}(x) = \tilde{u}^j(x) + \Delta t\, f(\tilde{u}^j(x)). \tag{2.5}$$

Here Δt is the time step and the variable $\tilde{v}$ has been used to denote the solution after one half of a two part fractional step method. Since $\tilde{u}^j$ is a step function, so is $\tilde{v}^{j+1}$; the height of the step above the point x having been increased or decreased by the amount $\Delta t\, f(\tilde{u}^j(x))$. This is equivalent to altering the weights w_i^j so that the new weights, w_i^{j+1}, satisfy

$$\tilde{v}^{j+1}(x) = \sum_{i=1}^{N} H(X_i^j - x)\, w_i^{j+1}. \tag{2.6}$$

A simple formula can be derived for the w_i^{j+1}. Assume that the particles have been labeled so that (2.3) holds and define $\tilde{v}_i^{j+1} = \tilde{v}^{j+1}(X_i^j)$. Then, letting $\tilde{u}_{N+1}^j = 0$, we see that for each $i = 1, \ldots, N$,

$$
\begin{aligned}
w_i^{j+1} &= \tilde{v}_i^{j+1} - \tilde{v}_{i+1}^{j+1} \\
&= \tilde{u}_i^j - \tilde{u}_{i+1}^j + \Delta t \left[f\left(\tilde{u}_i^j\right) - f\left(\tilde{u}_{i+1}^j\right) \right] \qquad (2.7) \\
&= w_i^j + \Delta t \left[f\left(\tilde{u}_i^j\right) - f\left(\tilde{u}_{i+1}^j\right) \right].
\end{aligned}
$$

This is the formula that is used in practice to determine the new weights. It is apparent from (2.4) that the N new weights can be calculated in $O(N)$ operations if the particle positions are stored in the computer so that (2.3) holds. This necessitates sorting the particles at each time step, a task which can, at best, only be done in $O(N \log N)$ operations. This is the only aspect of the algorithm which is not $O(N)$. While it may be possible to design a sort which, on the average, takes $O(N)$ operations (for example, by using the fact that the particles at the previous time step were already sorted), in actual practice one tends to run out of storage before time becomes a crucial factor.

Boundary Conditions: Note that $\tilde{v}^{j+1}$ automatically satisfies the boundary condition (1.3e) since $H(X_i^j - x) = 0$ for all $x > X_i^j$. Furthermore, by summing over the w_i^j and using (1.2a) it is easy to show that $\sum w_i^j = 1 \Rightarrow \sum w_i^{j+1} = 1$. Since step II does not alter the particle strengths (as will be seen below), it follows that if $\sum w_i^0 = 1$, then the sum of the particle strengths is a conserved quantity in the random gradient method. In other words, $\tilde{u}^j$ satisfies the boundary condition (1.3d) at each time step if $\tilde{u}^0$ does initially.

Step II: It remains to solve the heat equation $u_t = \nu\, u_{xx}$ with initial data $\tilde{v}^{j+1}$. First select N random numbers $\eta_1, \eta_2, \ldots, \eta_N$ from a Gaussian distribution with mean 0 and variance $2\nu\Delta t$. (With regards to generating normally distributed random numbers on a computer see [36, 38, 40].) The position of the *ith* particle, X_i^j, is then altered by the amount η_i to obtain $X_i^{j+1} = X_i^j + \eta_i$. Thus,

$$\tilde{u}^{j+1}(x) = \sum_{i=1}^{N} \mathrm{H}(\mathrm{X}_i^j + \eta_i - x)\, w_i^{j+1} = \sum_{i=1}^{N} \mathrm{H}(\mathrm{X}_i^{j+1} - x)\, w_i^{j+1}. \tag{2.8}$$

It should be noted that there is an error introduced into the computation which is due to our inability to generate truly random numbers (whatever that means! ([40])) on a computer. We will neglect this source of error and assume that the $\eta_1, \ldots, \eta_N$ really are independent and Gaussian distributed. Since in practice this error is of the same order as computer roundoff error, our omission will be of no more consequence than the standard practice of ignoring the effect of roundoff error in a convergence proof.

2.3 Restriction to Monotonic Initial Data In order to prove the convergence of this method to solutions of (1.3a-e) we have found it necessary to assume that the initial approximation, $\tilde{u}^0$, is monotonic. This is due to the fact that if one allows particle weights with different signs, then some realizations of the $\eta_1, \ldots, \eta_N$ will result in $\tilde{u}^{j+1}(x) < 0$ for certain x, in spite of the fact that $\tilde{u}^j$ (and hence $\tilde{v}^{j+1}$) may satisfy $0 \leq \tilde{u}^j \leq 1$ everywhere. (See fig. 3 of [32].) Not only are such negative solutions unphysical (in Chapter 5 we will see that solutions of (1.3a-e) always lie in [0,1]) but, as we shall demonstrate in Chapter 4, solutions of $u_t = u(1-u)$ with negative initial data blow up in finite time. This can lead to particle strengths which increase without bound, further degrading the numerical solution. In [32] Hald encountered precisely the same problem and also found it necessary to assume that the initial data is monotonic. Given these considerations we start by defining the class S of 'good' starting approximations, $\tilde{u}^0$.

DEFINITION 2.1 Let S be the class of all monotonically decreasing step functions, $\tilde{u}$, which satisfy $\tilde{u}(-\infty) = 1$ and $\tilde{u}(\infty) = 0$. Thus, $\tilde{u} \in \mathrm{S}$ if and only if $\tilde{u}$ can be written in the form

$$\tilde{u}(x) = \sum_{i=1}^{N} H(\mathrm{X}_i - x)\, w_i \tag{2.9}$$

where the weights $w_1, \ldots, w_N$ satisfy

$$0 < w_i \leq 1, \tag{2.10a}$$

$$\sum_{i=1}^{N} w_i = 1. \tag{2.10b}$$

Next, we will show that the random gradient method maps the class $\mathbf{S}$ into itself, thereby avoiding the difficulties outlined above.

LEMMA 2.2 Fix $\Delta t \leq 1$ and assume that $\tilde{u}^j \in \mathbf{S}$. Let $\tilde{v}^{j+1}$ and $\tilde{u}^{j+1}$ be given by (2.6) and (2.8) respectively. Then $\tilde{v}^{j+1} \in \mathbf{S}$ and $\tilde{u}^{j+1} \in \mathbf{S}$.

Assumption: Here and for the remainder of this paper f will always be taken to be the function $f(u) = u(1-u)$. While many of the theorems that follow continue to hold for more general f, this assumption greatly simplifies several of the proofs.

Proof: To see that $\tilde{v}^{j+1} \in \mathbf{S}$, use (2.7) to write

$$\begin{aligned}
\sum_{i=1}^{N} w_i^{j+1} &= \sum_{i=1}^{N} w_i^j + \Delta t \sum_{i=1}^{N} \left(f(\tilde{u}_i^j) - f(\tilde{u}_{i+1}^j) \right) \\
&= 1 + \Delta t \left(f(\tilde{u}_0^j) - f(\tilde{u}_N^j) \right) \\
&= 1.
\end{aligned}$$

Here we have used $\tilde{u}_0^j = 1$, $\tilde{u}_N^j = 0$ and $f(1) = f(0) = 0$. Furthermore,

$$\begin{aligned}
w_i^{j+1} &= w_i^j + \Delta t \left(f(\tilde{u}_i^j) - f(\tilde{u}_{i+1}^j) \right) \\
&= w_i^j \left[1 + \Delta t \left(1 - (\tilde{u}_i^j + \tilde{u}_{i+1}^j) \right) \right] \qquad (2.11) \\
&> 0
\end{aligned}$$

since $w_i^j\ [1 + \Delta t\ (1 - (\tilde{u}_i^j + \tilde{u}_{i+1}^j))]$ is the product of two positive quantities. This can be seen as follows. By assumption, $w_i^j > 0$. Since $0 \leq \tilde{u}_i^j \leq 1$ for all i and since $\tilde{u}_i^j = 1$ only if $i = 0$, we have

$$-1 \ < \ 1 - (\tilde{u}_i^j + \tilde{u}_{i+1}^j) \ \leq \ 1. \tag{2.12}$$

Hence, for $0 < \Delta t \leq 1$,

$$1 + \Delta t\ (1 - (\tilde{u}_i^j + \tilde{u}_{i+1}^j)) > 0.$$

Finally, $w_i^{j+1} > 0$ and $\sum_{i=1}^{N} w_i^{j+1} = 1$ together imply $w_i^{j+1} \leq 1$ for all i. Thus, $\tilde{v}^{j+1} \in \mathbf{S}$ as claimed.

By using the fact that $\tilde{v}^{j+1} \in \mathbf{S}$ and noting that an alteration of the particle positions has no effect on the weights it follows that $\tilde{u}^{j+1} \in \mathbf{S}$ as well.

It follows immediately that if $\tilde{u}^0 \in \mathbf{S}$, then at all subsequent time steps, $\tilde{v}^j, \tilde{u}^j \in \mathbf{S}$. Although this is a trivial consequence of Lemma 2.2 we state it here for future reference.

COROLLARY 2.3 Fix $\Delta t \leq 1$ and assume that $\tilde{u}^0 \in \mathbf{S}$. Then for all $j = 1, 2, \ldots,$ we have $\tilde{v}^j \in \mathbf{S}$ and $\tilde{u}^j \in \mathbf{S}$.

Assumption: From now on, it will be assumed that the time step, Δt, always satisfies $\Delta t \leq 1$.

One final fact will be established in this section; a bound on the particle strengths, w_i^j. By (2.10a) and Corollary 2.3, $w_i^j \leq 1$ for all i and all j. This is not enough, however, for one needs to know that $N\, w_i^j = O(1)$ as $N \to \infty$. If the strengths are initially chosen so that $w_i^0 = O(N^{-1})$, then this a consequence of the following lemma.

LEMMA 2.4 For $\tilde{u}^0 \in \mathbf{S}$ let $\tilde{u}^j = \sum \mathrm{H}\,(\mathrm{X}_i^j - x)\, w_i^j$ be the computed solution at time $T = j\,\Delta t$ which has been derived from $\tilde{u}^0$ by the random gradient method. Then for all $i = 1, \ldots, N$ the particle strengths, w_i^j, satisfy

$$w_i^j \leq e^T\, w_i^0. \tag{2.13}$$

Proof: It is evident from (2.11) and (2.12) that

$$\begin{aligned} w_i^{j+1} &= w_i^j\, [1 + \Delta t\, (1 - (\tilde{u}_i^j + \tilde{u}_{i+1}^j))] \\ &\leq w_i^j\, (1 + \Delta t\,). \end{aligned}$$

The inequality in (2.13) follows immediately.

3. Solution Operator Notation and an Outline of the Error Analysis

The primary purpose of this chapter is to develop a notation with which to discuss the error. We shall then present an outline of the convergence proof. We begin by assuming that $\nu = 1$. This makes the exposition simpler. We will remove this restriction at the end of Chapter 8 and show, by a scaling argument, that Theorem 1.1 holds for arbitrary $\nu \leq 1$.

3.1 The Exact Solution Operators The *exact solution operator* for the Kolmogorov equation, F_t, is defined by

$$F_t\, u^0(x) = u\,(x\,,t\,)$$

where $u\,(x\,,t\,)$ is the solution to (1.2a,b) at time t. In other words, F_t is a one parameter family of maps, with parameter t, which takes functions on $\mathbb{R}$ to functions on $\mathbb{R}$ such that the initial value u^0 is mapped onto the corresponding solution of the Kolmogorov equation at time t. Note that if $t = j\,\Delta t$, then

$$u\,(x\,,t\,) = F_{\Delta t}^{\,j}\, u^0.$$

The superscript j has been used here to indicate the *jth* power of the operator $F_{\Delta t}$, *i.e.* $F_{\Delta t}$ composed with itself j times.

The *reaction operator* R_t and the *diffusion operator* D_t are defined similarly. Thus, $R_t\, u^0$ is the solution at time t to the *reaction equation* (which is an ODE for fixed x) with initial data u^0,

$$u_t = u\,(1-u\,) \tag{3.1a}$$

$$u\,(x\,,0) = u^0(x\,), \tag{3.1b}$$

and $D_t u^0$ is the solution at time t to the *heat equation* with initial data u^0,

$$u_t = u_{xx} \tag{3.2a}$$

$$u(x,0) = u^0(x). \tag{3.2b}$$

3.2 The Approximate Solution Operators We now define approximations, $\widetilde{R}_{\Delta t}$ and $\widetilde{D}_{\Delta t}$, to the operators $R_{\Delta t}$ and $D_{\Delta t}$. Let $u(x)$ be an arbitrary piecewise continuous function. The *approximate reaction operator*, $\widetilde{R}_{\Delta t}$, is defined by

$$\widetilde{R}_{\Delta t}\, u(x) = u(x) + \Delta t\ u(x)(1 - u(x)). \tag{3.3}$$

In other words, for each fixed $x \in \mathbb{R}$, $\widetilde{R}_{\Delta t}\, u(x)$ is simply the Euler's method approximation, after one time step, to the solution of the reaction equation (3.1a,b) with initial data $u(x)$.

To define the *approximate diffusion operator*, $\widetilde{D}_{\Delta t}$, let $\tilde{u}$ be an arbitrary step function of the form,

$$\tilde{u}(x) = \sum_{i=1}^{N} H(X_i - x)\, \omega_i\,. \tag{3.4}$$

Then $\widetilde{D}_{\Delta t}\, \tilde{u}$ is given by

$$\widetilde{D}_{\Delta t}\, \tilde{u}(x) = \sum_{i=1}^{N} H(X_i + \eta_i - x)\, \omega_i \tag{3.5}$$

where $\eta_1, \eta_2, \ldots, \eta_N$ are N independent random numbers chosen from a Gaussian distribution with mean 0 and variance $2\Delta t$. In the notation of Chapter 2,

$$\tilde{v}^{j+1} = \widetilde{R}_{\Delta t}\, \tilde{u}^j$$

$$\tilde{u}^{j+1} = \widetilde{D}_{\Delta t}\, \tilde{v}^{j+1}.$$

Thus, the computed solution after j time steps, $\tilde{u}^j$, may be written in terms of the initial data, $\tilde{u}^0$, in the following way:

$$\tilde{u}^j = (\widetilde{D}_{\Delta t}\, \widetilde{R}_{\Delta t})^j\, \tilde{u}^0.$$

For future reference we note that, as a consequence of Lemma 2.2, the operators $\widetilde{R}_{\Delta t}$ and $\widetilde{D}_{\Delta t}$ map the space $\mathbf{S}$ into itself.

3.3 An Introduction to the Error Analysis We now present an outline of the convergence proof. Our approach has been greatly influenced by the work of Roberts [47].

As above, let $\tilde{u}^0 \in \mathbf{S}$ be a step function approximation to the initial data u^0. The L^1 difference, at time $T = k\,\Delta t$, between the exact solution, u, of (1.2a,b) and the approximate solution, $\tilde{u}^k$, is

$$\| u(\,\cdot\,, T) - \tilde{u}^k(\,\cdot\,) \|_1 = \| F_{\Delta t}^k\, u^0 - (\widetilde{D}_{\Delta t}\, \widetilde{R}_{\Delta t})^k\, \tilde{u}^0 \|_1.$$

This error may be divided into three distinct components,

$$\| F_{\Delta t}^k\, u^0 - (\widetilde{D}_{\Delta t}\, \widetilde{R}_{\Delta t})^k\, \tilde{u}^0 \|_1 \le \| F_{\Delta t}^k\, u^0 - (D_{\Delta t}\, R_{\Delta t})^k\, u^0 \|_1 +$$

$$\| (D_{\Delta t}\, R_{\Delta t})^k\, u^0 - (D_{\Delta t}\, R_{\Delta t})^k\, \tilde{u}^0 \|_1 + \| (D_{\Delta t}\, R_{\Delta t})^k\, \tilde{u}^0 - (\widetilde{D}_{\Delta t}\, \widetilde{R}_{\Delta t})^k\, \tilde{u}^0 \|_1. \tag{3.6}$$

The first term on the right is called the *splitting error*. It is the error due to the fractional step or the error due to *exact operator splitting*. In Chapter 5 this error is shown to be $O(\Delta t)$,

$$\| F_{\Delta t}^k\, u^0 - (D_{\Delta t}\, R_{\Delta t})^k\, u^0 \|_1 \le C\, \Delta t. \tag{3.7}$$

(For the remainder of this chapter C will denote a generic constant, independent of both N and Δt .) Our proof of (3.7) is modeled after Roberts' use of techniques from the theory of first order quasilinear partial differential equations (PDE) (Kruzkov [41]) to prove a similar fact for Burgers' equation.

The second term on the right is the error due to our approximation of the initial data, u^0, by the step function, $\tilde{u}^0$. In Chapter 4 it is shown that the operators R_t and D_t are stable in the L^1 norm. This implies

$$\| (D_{\Delta t} R_{\Delta t})^k u^0 - (D_{\Delta t} R_{\Delta t})^k \tilde{u}^0 \|_1 \leq e^T \| u^0 - \tilde{u}^0 \|_1.$$

The third term on the right is the error due to the numerical approximation of the solutions to equations (3.1a,b) and (3.2a,b). That is, it is the error that results from approximating the exact operators $R_{\Delta t}$ and $D_{\Delta t}$ by the approximate operators $\widetilde{R}_{\Delta t}$ and $\widetilde{D}_{\Delta t}$. Since the effect of the operator $\widetilde{D}_{\Delta t}$ is non-deterministic, the bound on this error takes the form

$$P\left(\| (D_{\Delta t} R_{\Delta t})^k \tilde{u}^0 - (\widetilde{D}_{\Delta t} \widetilde{R}_{\Delta t})^k \tilde{u}^0 \|_1 \geq \gamma C \frac{\ln N}{\sqrt[4]{N}} \right) \leq 7 T N^{1 - \frac{9}{4}\gamma} \tag{3.8}$$

where $\gamma \geq 1$ is an arbitrary real number. The approach here is to break the error into $2k$ pieces,

$$\| (D_{\Delta t} R_{\Delta t})^k \tilde{u}^0 - (\widetilde{D}_{\Delta t} \widetilde{R}_{\Delta t})^k \tilde{u}^0 \|_1 \leq e^T \sum_{j=0}^{k-1} \left(\| R_{\Delta t} \tilde{u}^j - \widetilde{R}_{\Delta t} \tilde{u}^j \|_1 + \| D_{\Delta t} \tilde{v}^j - \widetilde{D}_{\Delta t} \tilde{v}^j \|_1 \right). \tag{3.9}$$

In Chapter 6 we show that

$$P\left(\| R_{\Delta t} \tilde{u}^j - \widetilde{R}_{\Delta t} \tilde{u}^j \|_1 \geq \gamma C \sqrt{\ln N} (\Delta t)^2 \right) \leq N^{1 - \frac{9}{4}\gamma^2}. \tag{3.10}$$

This estimate follows from the fact that Euler's method has local truncation error $O((\Delta t)^2)$ and

that

$$R_t \tilde{u}^j(x) = \widetilde{R}_{\Delta t}\, \tilde{u}^j(x) \text{ for all } |x| > \max_i |X_i^j| .$$

Chapter 7 is devoted to proving that

$$P\left(\| D_{\Delta t}\, \tilde{v}^j - \widetilde{D}_{\Delta t}\, \tilde{v}^j \|_1 \geq \gamma C \frac{\ln N}{\sqrt{N}} \right) \leq 4\, N^{1-\frac{2}{4}\gamma} \tag{3.14}$$

The proof of this inequality relies heavily on the pointwise estimate

$$P\left(| D_{\Delta t}\, \tilde{v}^j(x) - \widetilde{D}_{\Delta t}\, \tilde{v}^j(x) | \geq C\, \alpha \right) \leq e^{-2N\alpha^2}$$

where $\alpha > 0$ is an arbitrary real number. This is a simple consequence of an inequality for sums of bounded, independent random variables due to Hoeffding [37]. Such exponential inequalities are the key to establishing an estimate that is $o(1)$ after summing $k = O(\Delta t^{-1})$ times in (3.9). If we now set $\Delta t = \sqrt[4]{N}^{-1}$ and use (3.10) and (3.14) in (3.9) we obtain (3.8).

4. The Exact Solution Operators R_t and D_t

The purpose of this chapter is to develop some of the basic properties of the operators R_t and D_t. The principal result here is that both operators are stable in the L^1 norm. In §4.3 this fact is used to examine the propagation of the error which is induced by approximating the initial data with a step function.

4.1 The Exact Reaction Operator, R_t It is a simple matter to check that the function defined by

$$R_t u^0(x) = \frac{u^0(x)\, e^t}{1 + (e^t - 1)\, u^0(x)} \tag{4.1}$$

is a solution of the reaction equation (3.1a,b). Observe that equation (3.1a,b) is an ODE in t; the variable x is simply a parameter which selects the particular initial value, $u^0(x)$, to be used. The L^1 stability of R_t is a simple consequence of having an exact expression for $R_t u^0$.

LEMMA 4.1 (L^1 **Stability of** R_t) Let u and v be measurable functions on $\mathbb{R}$ such that $0 \le u, v \le 1$ and $\| u - v \|_1 < \infty$. Then for any time $t > 0$,

$$\| R_t u - R_t v \|_1 \le e^t \| u - v \|_1.$$

Proof: First note that $u \ge 0$ implies $1 + (e^t - 1)\, u \ge 1$ and similarly for v. Thus,

$$| R_t u - R_t v | = \left| \frac{u e^t}{1 + (e^t - 1)\, u} - \frac{v e^t}{1 + (e^t - 1)\, v} \right|$$

$$= \frac{| u - v |\, e^t}{(1 + (e^t - 1)\, u)(1 + (e^t - 1)\, v)}$$

$$\leq \mid u - v \mid e^t$$

since the denominator is the product of two terms, each of which is greater than one. This implies

$$\| R_t u - R_t v \|_1 = \int_{-\infty}^{\infty} \mid R_t u(x) - R_t v(x) \mid dx \leq e^t \| u - v \|_1.$$

Remark: In order to establish a bound of the form

$$\| R_t u - R_t v \|_1 \leq C(t) \| u - v \|_1 \tag{4.2}$$

that is valid for all time $t \geq 0$ it is necessary that $u, v \geq 0$. This is due to the fact that, if $u < 0$, then $R_t u$ blows up in finite time. To be more specific, equation (4.1) implies

$$R_t u(x) \rightarrow -\infty \quad as \quad t \rightarrow \ln(1 - u^{-1}(x))$$

whenever $u(x) < 0$.

In the course of proving that the error due to exact operator splitting is small (Chapter 5) it will be necessary to know how the reaction operator, R_t, affects the gradient of a C^1 function. By differentiating (4.1) we find

$$(R_t u)_x = \frac{u_x e^t}{(1 + (e^t - 1) u)^2}.$$

This immediately implies

LEMMA 4.2 Let $u \in C^1(\mathbb{R})$ and assume that $0 \leq u \leq 1$. Then for any time $t > 0$,

$$\| u_x \|_1 < \infty \quad \Rightarrow \quad \| (R_t u)_x \|_1 \leq e^t \| u_x \|_1$$

$$\| u_x \|_\infty < \infty \quad \Rightarrow \quad \| (R_t u)_x \|_\infty \leq e^t \| u_x \|_\infty.$$

4.2 The Exact Diffusion Operator, $D_{\Delta t}$ Define the *heat kernel*, $G(x,t)$, by

$$G(x,t) = \frac{e^{\frac{-x^2}{4t}}}{\sqrt{4\pi t}}. \tag{4.3}$$

Occasionally, when there is no possibility of confusion, we will write $G_t(x)$ instead of $G(x,t)$. It is a well known fact that the solution of the heat equation (3.2a,b) is given by

$$u(x,t) = (G_t * u^0)(x)$$

where * denotes convolution,

$$(G_t * u^0)(x) = \int_{-\infty}^{\infty} G_t(x-y) u^0(y) dy \,.$$

(For example, see John [39], p. 209.) Thus, $D_{\Delta t} u^0 = G_{\Delta t} * u^0$. A basic result from the theory of PDE is that the diffusion operator, D_t, maps $L^p(\mathbb{R})$ onto $L^p(\mathbb{R})$ for $1 \leq p \leq \infty$ (Folland [24]). We state and prove this fact for $p = 1$ and $p = \infty$ since it will be needed in the sequel.

LEMMA 4.3 Let u be any measurable function of $x \in \mathbb{R}$. Then for any time $t > 0$,

$$\| u \|_1 < \infty \quad \Rightarrow \quad \| D_t u \|_1 \leq e^t \| u \|_1$$

$$\| u \|_\infty < \infty \quad \Rightarrow \quad \| D_t u \|_\infty \leq e^t \| u \|_\infty.$$

Proof: Since for any $t > 0$,

$$\int_{-\infty}^{\infty} | G_t(x) | \, dx = \int_{-\infty}^{\infty} \frac{e^{\frac{-x^2}{4t}}}{\sqrt{4\pi t}} dx = 1$$

it follows that

$$\| D_t u(x) \|_1 = \int_{-\infty}^{\infty} | \int_{-\infty}^{\infty} G_t(x-y) \, u(y) \, dy | \, dx$$

$$\leq \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} | G_t(y) | \, | u(x-y) | \, dy \, dx$$

$$= \int_{-\infty}^{\infty} | G_t(y) | \int_{-\infty}^{\infty} | u(x-y) | \, dx \, dy$$

$$= \int_{-\infty}^{\infty} | G_t(y) | \, dy \; \| u \|_1$$

$$= \| u \|_1 .$$

The proof of the second inequality is similar.

Remark: For any bounded differentiable function u on $\mathbb{R}$ which satisfies $u_x \in L^1$ we have $(G_t * u)_x = G_t * u_x$. Consequently, it follows from Lemma 4.3 that

$$\| (D_t u)_x \|_1 \leq \| u_x \|_1 .$$

A similar result holds in in the sup norm. These bounds will be needed several times in Chapter 5.

The L^1 stability of the diffusion operator, D_t, is an immediate consequence Lemma 4.3 and the fact that $D_t u - D_t v = D_t(u - v)$.

COROLLARY 4.4 (L^1 **Stability of** D_t) For any bounded measurable functions u, v defined on $\mathbb{R}$ such that $\| u - v \|_1 < \infty$ and any time $t > 0$ we have

$$\| D_t u - D_t v \|_1 \leq \| u - v \|_1.$$

4.3 The Error Due to Approximating the Initial Data The L^1 stability of the operators $R_{\Delta t}$ and $D_{\Delta t}$ allows us to analyze the error that occurs as a result of approximating the initial data, u^0, with a step function, $\tilde{u}^0$. This is accomplished by examining the second term on the right in (3.8),

$$\| (D_{\Delta t} R_{\Delta t})^k u^0 - (D_{\Delta t} R_{\Delta t})^k \tilde{u}^0 \|_1.$$

This expression will be bounded in terms of the initial error, $\| u^0 - \tilde{u}^0 \|_1$, by repeatedly applying Corollary 4.4 to functions of the form $u = R_{\Delta t}(D_{\Delta t} R_{\Delta t})^j u^0$, $v = R_{\Delta t}(D_{\Delta t} R_{\Delta t})^j \tilde{u}^0$ and Lemma 4.1 to functions of the form $u = (D_{\Delta t} R_{\Delta t})^j u^0$, $v = (D_{\Delta t} R_{\Delta t})^j \tilde{u}^0$. Since the hypotheses of Lemma 4.1 require that $0 \leq u, v \leq 1$ the following lemma will be needed.

LEMMA 4.5 Let u be a measurable function on the real line such that $0 \leq u \leq 1$. Then for all time $t > 0$ and all $x \in \mathbb{R}$, $0 \leq R_t u(x) \leq 1$ and $0 \leq D_t u(x) \leq 1$.

Proof: Rewrite (4.1) to obtain

$$R_t u = \frac{u}{(1-u)e^{-t} + u}.$$

For $0 \leq u \leq 1$ both the numerator and the denominator are non-negative and hence, $R_t u \geq 0$. Similarly, $u \leq 1$ implies $u \leq (1-u)e^{-t} + u$ whereby $R_t u \leq 1$.

For u non-negative,

$$(D_t u)(x) = \int_{-\infty}^{\infty} G_t(x-y)u(y)dy$$

is the integral of a non-negative quantity and therefore, $D_t u \geq 0$. If $0 \leq u(x) \leq 1$, then $|u(x)| \leq 1$ implying

$$|D_t u(x)| = \left| \int_{-\infty}^{\infty} G_t(x-y)u(y)dy \right| \leq \int_{-\infty}^{\infty} G_t(x-y)|u(y)|dy \leq \int_{-\infty}^{\infty} G_t(x-y)dy = 1.$$

Thus, $D_t u(x) \leq 1$.

The main result of this chapter now follows easily.

THEOREM 4.6 Let u and v be bounded measurable functions defined on $\mathbb{R}$ satisfying $0 \leq u, v \leq 1$. Then for all $\Delta t > 0$,

$$\|(D_{\Delta t} R_{\Delta t})^k u - (D_{\Delta t} R_{\Delta t})^k v\|_1 \leq e^{k\Delta t} \|u - v\|_1. \tag{4.4}$$

Remark: Thus, if $u = u^0$ and $v = \tilde{u}^0$, where $\tilde{u}^0$ is a step function approximation to u^0, then Theorem 4.6 implies that after k time steps the error due to this approximation is no more than a constant times the initial error. Furthermore, the constant, $e^{k\Delta t}$, depends only on the time, $T = k\Delta t$, at which this error is measured.

Proof: The proof is by induction on k. The case $k = 0$ is automatic. If (4.4) holds for k then Lemma 4.1 and Corollary 4.4 imply that

$$\|(D_{\Delta t} R_{\Delta t})^{k+1} u - (D_{\Delta t} R_{\Delta t})^{k+1} v\|_1 = \|D_{\Delta t} R_{\Delta t} (D_{\Delta t} R_{\Delta t})^k u - D_{\Delta t} R_{\Delta t} (D_{\Delta t} R_{\Delta t})^k v\|_1$$

$$\leq \| R_{\Delta t} (D_{\Delta t} R_{\Delta t})^k u - R_{\Delta t} (D_{\Delta t} R_{\Delta t})^k v \|_1$$

$$\leq e^{\Delta t} \| (D_{\Delta t} R_{\Delta t})^k u - (D_{\Delta t} R_{\Delta t})^k v \|_1$$

$$\leq e^{(k+1)\Delta t} \| u - v \|_1.$$

It should be noted that in order to apply Lemma 4.1 in the induction step (to the second line) it is first necessary to ascertain that $0 \leq (D_{\Delta t} R_{\Delta t})^k u \leq 1$ and similarly for $(D_{\Delta t} R_{\Delta t})^k v$. This is a consequence of Lemma 4.5.

5. The Error Due to Exact Operator Splitting

This chapter is devoted to proving

$$\| F_{\Delta t}^{k} u^0 - (D_{\Delta t} R_{\Delta t})^k u^0 \|_1 \leq C \Delta t \tag{5.1}$$

and hence, that the error due to exact operator splitting goes to 0 like Δt. The key idea, (Roberts, [47]), is that the function, w, defined by

$$w(x,t) = F_t u^0(x) - D_t R_t u^0(x)$$

is a solution of

$$w_t = w_{xx} + a(x,t)w + b(x,t) \tag{5.2a}$$

$$w(x,0) = w^0(x) \tag{5.2b}$$

with $w^0(x) \equiv 0$ and $\| b(\cdot,t) \|_1 = O(t)$. In Theorem 5.4 we show that there exists a constant $A > 0$ such that solutions of (5.2a,b) satisfy the inequality

$$\| w(\cdot,T) \|_1 \leq e^{AT} \| w^0 \|_1 + e^{AT} \int_0^T \| b(\cdot,t) \|_1 \, dt \tag{5.3}$$

at all times $T > 0$. Thus, by setting $T = \Delta t$, it follows that the splitting error after one time step is $O(\Delta t^2)$,

$$\| F_{\Delta t} u^0 - D_{\Delta t} R_{\Delta t} u^0 \|_1 = \| w(\cdot,\Delta t) \|_1 \leq C (\Delta t)^2.$$

This inequality, together with the L^1 stability of the operator $F_{\Delta t}$ (Lemma 5.6), yields (5.1).

The techniques employed in the proof of (5.3) come from Kruzkov [41]. They have been used by Roberts [47] to prove the convergence of a particle method with random walks to solutions of Burgers' equation. Similar results in the L^2 norm have been obtained by Beale and Majda [7] for viscous splitting of the Navier-Stokes equations.

Roberts' work on exact operator splitting (his Chapter 2) and the material here represent two special cases of the following more general result. Under the appropriate assumptions on the coefficients $a(x,t)$, $b(x,t)$ and $c(x,t)$ and the function $w(x,t)$ solutions of the differential equation (DE)

$$w_t = w_{xx} + aw_x + bw + c$$

$$w(x,0) = w^0(x)$$

satisfy the L^1 inequality

$$\| w(\cdot,T)\|_1 \leq e^{AT} \| w^0 \|_1 + e^{AT} \int_0^T \| b(\cdot,t)\|_1 \, dt.$$

This is a special case of results from [41].

Chapter 5 is divided into four sections. In §5.1 we begin with a maximum principle for solutions of the Kolmogorov equation. This is followed by a proof of the existence and uniqueness of such solutions. The purpose of §5.2 is to prove Theorem 5.4: Under suitable conditions, solutions of (5.2a,b) satisfy the inequality in (5.3). In §5.3 it is shown that for solutions, u, of the Kolmogorov equation $u_x^0 \in L^\infty$ implies $u_x \in L^\infty$ at any time $t > 0$. This is then used to prove that operator $F_{\Delta t}$ is stable in in the L^1 norm. These results are brought together in §5.4 to show that the splitting error after one time step is $O(\Delta t^2)$ and from this, that the inequality in (5.1) holds.

5.1 Existence, Uniqueness, and a Maximum Principle Solutions of the nonlinear reaction-diffusion equation (1.1a-e) satisfy a maximum principle in much the same way as do solutions of linear parabolic DE's (Bramson [8]). Here we confine ourselves to proving a maximum principle for solutions of the Kolmogorov equation (1.3a,b). (See also Aronson and Wienberger [4], Fife and Mcleod [23], and Smoller [52].)

LEMMA 5.1 (Maximum Principle) Let u and v be bounded solutions of (1.3a,b) in the strip $\Omega = \mathbb{R}\times[0,T]$ with initial data u^0 and v^0 respectively. Suppose that $v^0(x) \leq u^0(x)$ for all $x \in \mathbb{R}$. Then for all $(x,t) \in \Omega$,

$$v(x,t) \leq u(x,t).$$

Remark: It follows from Lemma 5.1 that, if u is a *bounded* solution to (1.3a,b) in Ω such that the initial data satisfies $0 \leq u^0(x) \leq 1$ for all $x \in \mathbb{R}$, then $0 \leq u(x,t) \leq 1$ in Ω.

Proof: Let $w = u - v$. Then $w(x,0) \geq 0$ and

$$w_t = w_{xx} + aw$$

for $a = 1 - (u+v)$. Since a and w are bounded in Ω, it follows from the maximum principle for linear parabolic PDE's (Friedman [25], p. 43) that $w(x,t) \geq 0$ in Ω.

With the aid of a fixed point theorem one can demonstrate that for some $\delta > 0$, there exists a solution to (1.3a,b) on the interval $[0,\delta]$. (For example, see Theorem 11.12 of Smoller [52], p. 115.) But, in order to extend the solution beyond this interval, one must have some knowledge of how large the solution is on $[0,\delta]$. The maximum principle provides us with this knowledge. Hence, as a consequence of the maximum principle, it is possible to show that for $u^0 \in [0,1]$ there exists a unique bounded solution to (1.3a,b) for all time $t > 0$.

Theorem 5.2 (Existence and Uniqueness) Let $u^0 \in C(\mathbb{R})$ and assume that $0 \leq u^0 \leq 1$. Then for all $T > 0$ there exists a unique solution, u, to the Kolmogorov equation (1.3a,b) on the strip $\Omega = \mathbb{R} \times [0,T]$.

Proof: First note that there exists $\delta > 0$ such that for all functions u^0 which satisfy $\| u^0 \|_\infty \leq 1$ one can find a solution $u(x,t)$ to (1.3a,b) on $\mathbb{R} \times [0,\delta]$. This is a consequence of Theorem 11.12 of Smoller [52], p. 115 (where the Banach space is $X = C(\mathbb{R}) \cap L^\infty(\mathbb{R})$ with the sup norm). By the remark after Lemma 5.1 it follows that $\| u(\cdot,\delta) \|_\infty \leq 1$. Hence, there exists a solution to (1.3a) on $\mathbb{R} \times [\delta, 2\delta]$ with initial data $u(x,\delta)$. By taking successive intervals, each of length δ, one can 'glue' together a solution on $[0,T]$ for any $T > 0$. This is the content of Theorem 11.14 of [52]. Uniqueness follows from a Gronwall inequality as, for example, in Theorem 11.13 of [52], p. 116.

Remark: Actually, Smoller's theorem only implies the existence of $u \in C([0,\delta],X)$ (continuous functions from $[0,\delta]$ to X) satisfying

$$u(x,t) = \int_{-\infty}^{\infty} G(x-y,t)u^0(y)\,dy + \int_0^t \int_{-\infty}^{\infty} G(x-y,t-s) f(u(y,s))\,dy\,ds$$

where $f(u) = u(1-u)$ and $G(x,t)$ is the heat kernel (4.3). However, due to the well known properties of the heat kernel (John [39], p. 209), the right hand side of this expression is a C^∞ function on $\mathbb{R} \times (0,\delta)$. It follows that u is in fact a 'classical' solution of (1.3a,b). (For further discussion see the remarks at the top of p. 115 of Smoller [52].) Here and in the sequel 'solution' will always be taken to mean a classical solution.

5.2 An L^1 Bound on Solutions of $w_t = w_{xx} + aw + b$ The sole task of this section is to prove Theorem 5.4: that, under the proper conditions, solutions of (5.2a,b) satisfy the inequality in (5.3). This is the cornerstone of the analysis which ultimately leads to the convergence of exact operator splitting. An essential ingredient in the proof of Theorem 5.4 is a detailed knowledge of the behavior of solutions to the DE (5.4a,b) below, particularly the fact that for continuous initial data with compact support these solutions decay exponentially fast as $|x| \to \infty$. This is Lemma 5.3. A more general result may be found in Lemma 4 of Kruzkov [41], p. 232.

LEMMA 5.3 Fix $T > 0$ and let $\Omega = \mathbb{R} \times [0,T]$. Suppose that $a(x,t) \in C^1(\Omega)$ and that a and a_x are bounded in Ω. Let $q^0(x)$ be a continuous function of $x \in \mathbb{R}$ such that for some $r > 0$, $q^0(x) = 0$ when $|x| \geq r$. Then the following statements hold.

i) There exists a bounded solution, $q(x,t)$, in Ω to the problem

$$L(q) = q_{xx} + aq - q_t = 0, \tag{5.4a}$$

$$q(x,0) = q^0(x). \tag{5.4b}$$

ii) Let $C_0 = \sup_{\mathbb{R}} |q^0(x)|$ and $A = \sup_{\Omega} \{0, a(x,t)\}$. Then for all $(x,t) \in \Omega$

$$|q(x,t)| \leq C_0 e^{At}. \tag{5.5}$$

iii) Let $C_1 = e^{(1+A)T + \sqrt{1+r^2}}$. Then, for all $(x,t) \in \Omega$,

$$|q(x,t)| \leq C_1 e^{-\sqrt{1+x^2}}. \tag{5.6}$$

In other words, $q \to 0$ exponentially fast as $|x| \to \infty$ and the rate of decrease is uniform for $t \in [0,T]$.

Proof: *i)* Let $\Gamma(x,t;y,\tau)$ be a fundamental solution of the parabolic DE $L(q)=0$ (Friedman [25], Chapter 1). The assumptions on the coefficient a and the initial data q^0 guarantee that there exists a solution of the form

$$q(x,t)=\int_{-\infty}^{\infty}\Gamma(x,t;y,0)q^0(y)dy$$

to the Cauchy problem (5.4a,b). For example, this is a consequence of Theorem 12 (p. 25) in Friedman [25]. Furthermore, the proof of Friedman's theorem shows that if the function q^0 (Friedman's ϕ) is assumed to be bounded (instead of exponentially bounded), then the solution $q(x,t)$ is bounded in Ω. (This follows from equation (6.12) on p. 24 of [25].)

ii) The knowledge that q is bounded allows us to use the maximum principal for linear parabolic PDE's to compute the bound (5.5). Consider the function $Q(x,t)=C_0\,e^{At}$. It is easy to check that $Q(x,0)\geq q(x,0)$ for all $x\in\mathbb{R}$ and $L(Q)=(a(x,t)-A)\,Q\leq 0$ on $\Omega_0=\mathbb{R}\times(0,T]$. Consequently, the function $w=Q-q$ satisfies the differential inequality $L(w)\leq 0$ subject to $w(x,0)\geq 0$. From *i)* we know that q, and hence w, is bounded on Ω. Hence, it follows from Theorem 9 in Chapter 2 (p. 43) of Friedman [25] that $w(x,t)\geq 0$ for all $(x,t)\in\Omega$. Thus,

$$q(x,t)\leq Q(x,t)=C_0\,e^{At}.$$

The reverse inequality,

$$q(x,t)\geq -Q(x,t)=-C_0\,e^{At},$$

can be shown in a similar manner. The inequality in (5.5) follows immediately.

iii) By setting $Q(x,t)=C_0\,e^{(1+A)t+\sqrt{1+r^2}-\sqrt{1+x^2}}$ one can obtain (5.6) using the same argument as in *ii)* above.

The stage is now set to prove Theorem 5.4. Following Roberts ([47], p. 25), we multiply both sides of (5.2a) by some carefully chosen function, $q(x,t)$, integrate over $\Omega = \mathbb{R}\times[0,T]$, and then use integration by parts to push the derivatives from w onto q. If q is chosen properly, the resulting expression yields (5.3) fairly easily.

THEOREM 5.4 Fix $T > 0$ and let $a(x,t)$ and $b(x,t)$ be bounded, continuous functions on the strip $\Omega = \mathbb{R}\times[0,T]$ such that $a \in C^1(\mathbb{R})$, a_x is bounded in Ω, and $b \in L^1(\Omega)$. Suppose that $w(x,t)$ is a solution of (5.2a,b) in $\mathbb{R}\times(0,T]$ with $w^0 \in L^1(\mathbb{R})$, and assume that w and w_x are bounded in Ω. Then w satisfies the inequality in (5.3) for $A = \sup_{\Omega} \{0, a(x,t)\}$.

Proof: Let $q(x,t)$ be a function defined on Ω. Multiplying (5.2a) by q and integrating over Ω yields

$$\int_0^T \int_{-\infty}^{\infty} w_t\, q\, dx\, dt = \int_0^T \int_{-\infty}^{\infty} (w_{xx}\, q + awq + bq)\, dx\, dt. \tag{5.7}$$

For motivation, assume that q and q_x decay exponentially fast as $|x| \to \infty$. Then integration by parts yields

$$\int_{-\infty}^{\infty} wq\,\big|_{t=T}\, dx = \int_{-\infty}^{\infty} w^0 q\,\big|_{t=0}\, dx + \int_0^T \int_{-\infty}^{\infty} w\, L^*(q)\, dx\, dt + \int_0^T \int_{-\infty}^{\infty} b\, q\, dx\, dt \tag{5.8}$$

where $L^*(q) = q_{xx} + aq + q_t$. This suggests the following choice for q: Let q be a solution of

$$L^*(q) = 0 \tag{5.9a}$$

$$q(x,T) = q^0(x) \tag{5.9b}$$

(note that we are solving backwards in time) where the initial data, q^0, has yet to be specified. Hence, the second term on the right hand side of (5.8) disappears.

The substitution $t \to t-T$ transforms the problem (5.9,b) into one of the form (5.4a,b). Assume, for the moment, that q^0 is continuous and has compact support. Then Lemma 5.3 guarantees the existence of a solution, q, to (5.9,b) and a constant $C > 0$ such that $|q(x,t)| \leq C\, e^{-\sqrt{1+x^2}}$ for all $(x,t) \in \Omega$.

The exponential decay of q in Ω enables us to apply integration by parts once to the term $\int_0^T \int_{-\infty}^{\infty} w_{xx}\, q\; dx\, dt$ in (5.7) to obtain $-\int_0^T \int_{-\infty}^{\infty} w_x\, q_x\, dx\, dt$. However, there is no assurance that q_x behaves nicely enough at infinity to integrate by parts again. To rectify this problem consider the *cutoff function* $\eta(x) \in C_0^\infty(\mathbb{R})$ which satisfies $0 \leq \eta(x) \leq 1$ for all x, $\eta(x) = 1$ for $|x| \leq 1$, and $\eta(x) = 0$ for $|x| \geq 2$. (Concerning the existence of such functions see Folland [24], p. 18.) Then for any $r > 0$ the function $\eta^r(x) = \eta(\frac{x}{r})$ satisfies $0 \leq \eta^r(x) \leq 1$ for all x, $\eta^r(x) = 1$ for $|x| \leq r$, and $\eta^r(x) = \eta^r_x(x) = 0$ for $|x| \geq 2r$. Consequently, the function $q\,\eta^r$ and its first derivative $(q\,\eta^r)_x$ have compact support. Thus, after multiplying (5.2a) by $q(x,t)\,\eta^r(x)$ and integrating over Ω, it is now permissible to integrate by parts twice. Remembering that $L^*(q) = 0$ we find

$$\int_{-\infty}^{\infty} wq\,\big|_{t=T}\,\eta^r\, dx = \int_{-\infty}^{\infty} wq\,\big|_{t=0}\,\eta^r\, dx + \int_0^T \int_{-\infty}^{\infty} bq\,\eta^r\, dx\, dt \tag{5.10}$$

$$+ \int_0^T \int_{-\infty}^{\infty} (wq\,\eta^r_{xx} + 2wq_x\,\eta^r_x)\, dx\, dt.$$

In the last term on the right the integrand has compact support and hence integration by parts can be used to push the x-derivative from q_x onto $\eta^r_x\, w$. Since $\eta^r_x = \eta^r_{xx} = 0$ for $|x| \leq r$, $(w\,\eta^r_{xx} + 2w_x\,\eta^r_x)$ is bounded uniformly in t, and $|q(x,t)| \leq Ce^{-\sqrt{1+x^2}}$, this implies

$$\int_0^T \int_{-\infty}^{\infty} (wq\,\eta^r_{xx} + 2wq_x\,\eta^r_x)\, dx\, dt = -\int_0^T \int_{|x|>r} q\,(w\,\eta^r_{xx} + 2w_x\,\eta^r_x)\, dx\, dt \;\to 0 \;\; as \;\; r \to \infty.$$

Now if we assume that q^0 has compact support, then $w(x,T)\,q^0(x) \in L^1(\mathbb{R})$. This, together with the fact that $q\,\eta^r \to q$ as $r \to \infty$, allows us to apply the Lebesque Dominated Convergence Theorem to (5.10) to obtain

$$\int_{-\infty}^{\infty} w(x,T)q^0(x)dx = \int_{-\infty}^{\infty} w^0(x)q(x,0)dx + \int_0^T \int_{-\infty}^{\infty} q(x,t)b(x,t)dx\,dt \tag{5.11}$$

in the limit as $r \to \infty$.

All that remains is to choose the initial data, q^0, so that q^0 is continuous and has compact support. To this end introduce $\phi(x) \in C_0^\infty(\mathbb{R})$ which satisfies $\phi(x) \geq 0$ all x, $\phi(x) = 0$ for $|x| \geq 1$, and $\int_{-\infty}^{\infty} \phi(x)\,dx = 1$. For $\epsilon > 0$ define $\phi^\epsilon(x) = \epsilon^{-1}\phi(\frac{x}{\epsilon})$. The class $\{\phi^\epsilon\}_{\epsilon>0}$ is known as an *approximation to the identity* (Folland [24], p. 14) or as *mean functions* (Kruzkov [41], p. 221). Fix $R > 0$ and let

$$h(x) = \begin{cases} sign\,(w(x,T)), & |x| \leq R, \\ 0, & |x| > R. \end{cases}$$

The function $h^\epsilon(x) = (\phi^\epsilon * h)(x)$ is smooth, has compact support, and satisfies $sup\ |h^\epsilon(x)| \leq 1$. Hence, by Lemma 5.3, (5.9,b) can be solved with initial data $q^0(x) = h^\epsilon(x)$. Denote this solution by $q^\epsilon(x,t)$. Replacing q by q^ϵ in (5.11) and taking the absolute value yields

$$\left| \int_{-\infty}^{\infty} w(x,T)h^\epsilon(x)dx \right| \leq \int_{-\infty}^{\infty} |w^0(x)|\,|q^\epsilon(x,0)|\,dx + \int_0^T \int_{-\infty}^{\infty} |b(x,t)|\,|q^\epsilon(x,t)|\,dx\,dt$$

$$\leq e^{AT} \int_{-\infty}^{\infty} |w^0(x)|\,dx + e^{AT} \int_0^T \int_{-\infty}^{\infty} |b(x,t)|\,dx\,dt$$

since, by (5.5), $|q^\epsilon(x,t)| \leq e^{AT}$ for all $(x,t) \in \Omega$. Note that $h^\epsilon = (\phi^\epsilon * h) \to h$ as $\epsilon \to 0$ in

the L^1 norm. (For example, see Theorem 0.13 (p. 16) of Folland [24].) Therefore, since $w(x,T)\,h(x) \in L^1(\mathbb{R})$, one can apply the Lebesque Dominated Convergence theorem to obtain

$$\int_{-R}^{R} |\, w(x,T)\,|\, dx \;\leq\; e^{AT} \int_{-\infty}^{\infty} |\, w^0(x)\,|\, dx \;+\; e^{AT} \int_0^T \int_{-\infty}^{\infty} |\, b(x,t)\,|\, dx\, dt.$$

Letting $R \to \infty$ and using Fatou's lemma yields (5.3).

5.3 L^1 Stability of Solutions to the Kolmogorov Equation The next step in the proof of (5.1) is to show that solutions, u, of the Kolmogorov equation (1.3a,b) are stable in the L^1 norm and that $u_x^0 \in L^\infty(\mathbb{R})$ implies $u_x \in L^\infty(\mathbb{R})$ at all subsequent times $t > 0$. Similar statements have already been proved for solutions of the reaction equation (Lemmas 4.1 and 4.2) and the heat equation (Lemma 4.3 and Corollary 4.4).

We begin with the L^∞ bound on the gradient, u_x. This bound serves two purposes. First, it will be used in Lemma 5.6 to show that the difference between two solutions of (1.3a,b) satisfies the hypotheses of Theorem 5.4. Assuming that their difference at time $t = 0$ is in L^1, this leads to the L^1 stability of solutions of the Kolmogorov equation. Later, in Theorem 5.8, it will be used to show that the function $w(x,t) = F_{\Delta t}\, u^0 - D_{\Delta t}\, R_{\Delta t}\, u^0$ satisfies the hypotheses of Theorem 5.4 and hence, that the splitting error after one time step is small.

LEMMA 5.5 Let $u(x,t)$ be a solution of the Kolmogorov equation (1.3a,b) with initial data $u^0 \in C^1(\mathbb{R})$. Assume that $0 \leq u^0(x) \leq 1$ for all $x \in \mathbb{R}$ and that $\| u_x^0 \|_\infty < \infty$. Then for any time $t \geq 0$,

$$\| u_x(t,\cdot) \|_\infty \leq e^t \;\| u_x^0 \|_\infty. \tag{5.12}$$

Remark: This bound is essentially a corollary to Lemma 3.2 of Bramson [8] which asserts that,

if one considers only bounded solutions of (1.3a,b), then the reaction-diffusion equation (1.1a) with $f' \leq 1$ (and hence the Kolmogorov equation) is stable in the L^∞ norm. This is a consequence of the maximum principle for linear parabolic PDE's.

Proof: Fix $r \in \mathbb{R}$ and note that $v(x,t) = u(x+r,t)$ is a solution of (1.3a) with initial data $v^0(x) = u^0(x+r)$. By the mean value theorem

$$| u^0(x) - v^0(x) | = | u^0(x) - u^0(x+r) | \leq \| u^0_x \|_\infty | r | .$$

Now use Lemma 3.2 (p. 38) of [8] (with $\epsilon = | r | \| u^0_x \|_\infty$) to obtain

$$| u(x,t) - v(x,t) | \leq e^t \| u^0_x \|_\infty | r |$$

for all $x \in \mathbb{R}$ and $t \geq 0$. Setting $y = x + r$ yields

$$| u(x,t) - u(y,t) | \leq e^t \| u^0_x \|_\infty | x - y |$$

for all $x \in \mathbb{R}$ and $t \geq 0$. But this holds for all $y \in \mathbb{R}$ as well since r was arbitrary. The inequality in (5.12) follows immediately.

By restricting ourselves to initial data, u^0, which have bounded first derivative, we will now use Theorem 5.4 and Lemma 5.5 to prove the L^1 stability of solutions to (1.3a,b). Besides providing us with some insight into how solutions of the Kolmogorov equation behave, L^1 stability is needed in order to derive (5.1) from a bound on the splitting error after one time step. It is interesting to note that Theorem 5.4, which is central to the proof of this bound on the splitting error, also has L^1 stability of the Kolmogorov equation as a consequence.

LEMMA 5.6 (L^1 Stability) Let u and v be solutions to the Kolmogorov equation (1.3a,b) with initial data u^0, $v^0 \in C^1(\mathbb{R})$. Assume that $0 \leq u^0, v^0 \leq 1$ with $\| u^0 - v^0 \|_1 < \infty$ and

that the derivatives u_x^0, v_x^0 are bounded on $\mathbb{R}$. Then for any time $t \geq 0$,

$$\| u(\cdot,t) - v(\cdot,t) \|_1 \leq e^t \ \| u^0 - v^0 \|_1.$$

Proof: Let $w = u - v$. Then w satisfies the equation

$$w_t = w_{xx} + a(x,t)w$$

with $a = 1 - (u+v)$. By the maximum principle, Lemma 5.1, $0 \leq u, v \leq 1$ and therefore $|a| \leq 1$. Since, by assumption, u_x^0 and v_x^0 are bounded on $\mathbb{R}$, Lemma 5.5 guarantees that u_x and v_x are bounded on the strip $\mathbb{R}\times[0,t]$ for any $t > 0$. Thus, it follows from Theorem 5.4 that

$$\| w(\cdot,t) \|_1 \leq e^t \ \| w(\cdot,0) \|_1.$$

5.4 Convergence of Exact Operator Splitting We are now ready to prove (5.1). This is accomplished in two steps. The first step, Theorem 5.8, consists of using Theorem 5.4 to show that the L^1 norm of the function $w(x,t) = F_t u^0 - D_t R_t u^0$ is $O(t^2)$. If one regards $D_{\Delta t} R_{\Delta t} u^0$ as a numerical approximation to $F_{\Delta t} u^0$ after one time step of length Δt, then this is simply the statement that the local truncation error is of order $(\Delta t)^2$. In other words, our numerical scheme (approximating $F_{\Delta t} u^0$ by $D_{\Delta t} R_{\Delta t} u^0$) is consistent.

THEOREM 5.8 Let $u(x,t)$ be a solution of the Kolmogorov equation (1.3a,b) with initial data $u^0 \in C^1(\mathbb{R})$. Assume that $0 \leq u^0 \leq 1$ and that $u_x^0 \in L^1(\mathbb{R}) \cap L^\infty(\mathbb{R})$. Then

$$\| F_{\Delta t} u^0 - D_{\Delta t} R_{\Delta t} u^0 \|_1 \leq \mathrm{C} (\Delta t)^2 \tag{5.13}$$

where C is given by

$$C = e^{3\Delta t}\left\{e^{\Delta t}\,\|u_x^0\|_\infty + 4\frac{\sqrt{2\Delta t}}{\sqrt{\pi}}\right\}\|u_x^0\|_1. \tag{5.14}$$

Proof: As indicated at the beginning of this chapter, the idea is to show that the function

$$w(x,t) = F_t\,u^0(x) - D_t R_t\,u^0(x)$$

satisfies a DE of the form (5.2a,b) on $\Omega = \mathbb{R}\times[0,\Delta t]$ with initial data $w^0(x) \equiv 0$ and where the coefficient $b(x,t)$ in (5.2a) satisfies $\|b\|_{L^1(\Omega)} = O(\Delta t^2)$. Then (5.13) is a consequence of Theorem 5.4.

Let $u(x,t) = F_t\,u^0(x)$ be the solution to the Kolmogorov equation (1.3a,b) with initial data u^0 and let $v(x,t) = R_t\,u^0(x)$ denote the solution to the reaction equation (3.1a,b) with initial data u^0. Then

$$w(x,t) = u(x,t) - (G*v)(x,t)$$

where $G(x,t)$ is the Gaussian kernel (4.5) and $*$ denotes convolution. By differentiating w with respect to t and using the fact that $v_t = v(1-v)$ and $G_t = G_{xx}$ it is not difficult to show that w satisfies equation (5.2a,b) for $a = 1-(u + G*v)$ and $b = G*v^2 - (G*v)^2$.

It follows from Lemma 5.1 that $\|u(\cdot,t)\|_\infty \le 1$ for any time $t \ge 0$. Furthermore, by Lemma 5.5, $\|u_x(\cdot,t)\|_\infty \le e^t\,\|u_x^0\|_\infty$. Identical estimates hold for v and $G*v$. For, by Lemma 4.5, $\|v(\cdot,t)\|_\infty \le 1$ and hence, $\|(G*v)\|_\infty \le 1$ for all $t \ge 0$. (In what follows, we will often suppress mention of the variable t.) By Lemma 4.2, $\|v_x\|_\infty \le e^t\,\|u_x^0\|_\infty$ and, upon writing $(G*v)_x = (G*v_x)$, one finds that $\|(G*v)_x\|_\infty \le e^t\,\|u_x^0\|_\infty$. Thus, a, a_x, b, w and w_x are bounded continuous functions on Ω.

It remains to show that $\| b \|_{L^1(\Omega)} = O(\Delta t^2)$. To this end Theorem 5.4 will once again be used, this time applied to the function $b = G * v^2 - (G * v)^2$. Differentiating b with respect to t and remembering that $v_t = v(1 - v)$ and $G_t = G_{xx}$ one finds that

$$b_t = (G * v^2)_{xx} + G * 2vv_t - 2(G * v)[(G * v)_{xx} + G * v_t]$$

$$= b_{xx} + 2G * v(G * v)_{xx} + 2((G * v)_x)^2 + 2 * G * v^2 - 2G * v^3$$

$$-2(G * v)(G * v)_{xx} 2(G * v)(G * v) + 2(G * v)(G * v^2)$$

$$= b_{xx} + 2[(G * v_x)^2 + G * v^2 - G * v^3 - (G * v)^2 + (G * v)(G * v^2)]$$

$$= b_{xx} + 2b + 2[(G * v_x)^2 + (G * v)(G * v^2) - G * v^3].$$

Hence, b satisfies the DE

$$b_t = b_{xx} + 2b + c(x, t) \tag{5.15a}$$

$$b(x, 0) = 0 \tag{5.15b}$$

with $c = 2[(G * v_x)^2 + (G * v)(G * v^2) - G * v^3]$. Noting that v and v_x are bounded and continuous in Ω it follows that b, b_x, and c are as well.

For the remainder of this proof, we will assume that $0 \leq t \leq \Delta t$. We claim that

$$\int_0^t \int_{-\infty}^{\infty} | c(x, s) | \, dx \, ds \leq 2Ct e^{-2t}$$

where C is given by (5.14). This follows from $\| c(\cdot, t) \|_1 \leq 2Ce^{-2t}$ whenever $t \leq \Delta t$, a fact which we shall now show. We can use Lemmas 4.2 and 4.3 to obtain

$$\|(G*v_x)^2\|_1 \le \|G*v_x\|_\infty \|G*v_x\|_1 \le \|v_x\|_\infty \|v_x\|_1 \le e^{2t} \|u_x^0\|_\infty \|u_x^0\|_1.$$

To estimate the remaining portion of c recall that $0 \le v \le 1$ and write

$$\|(G*v)(G*v^2)-(G*v^3)\|_1 \le \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} G(x-y)G(x-z)\,|\,v(z)v^2(y)-v^3(z)\,|\,dy\,dz\,dx$$

$$\le \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} G(x-y)\,G(x-z)\,v(z)\,(v(y)+v(z))\;|\,v(y)-v(z)\,|\,dy\,dz\,dx$$

$$= 2\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} G(x-y)\,G(x-z)\;\left|\int_0^1 v_x(z+\theta(y-z))\,d\theta\right|\;|\,y-z\,|\,dy\,dz\,dx$$

$$\le 2\,\|v_x(\cdot,t)\|_1 \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} G(y)G(z)\,|\,y-z\,|\,dy\,dz$$

$$\le 4\frac{\sqrt{2t}}{\sqrt{\pi}}e^t\,\|u_x^0\|_1$$

where we have used

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} G(y)G(z)\,|\,y-z\,|\,dy\,dz = \frac{\sqrt{8t}}{\sqrt{\pi}}.$$

(To see this, rewrite the integral in terms of polar coordinates and perform the necessary integration.) Thus,

$$\|c(\cdot,t)\|_1 \le 2\{\|(G*v_x)^2\|_1 + \|(G*v_x)(G*v_x^2)-(G*v_x^3)\|_1\}$$

$$\le 2\{e^{2t}\,\|u_x^0\|_1\,\|u_x^0\|_\infty + \frac{4\sqrt{2t}}{\sqrt{\pi}}\,e^t\,\|u_x^0\|_1\}$$

$$= 2e^t\,\{e^t\,\|u_x^0\|_\infty + \frac{4\sqrt{2t}}{\sqrt{\pi}}\,e^t\}\,\|u_x^0\|_1$$

and hence, for $t \in [0,\Delta t]$,

$$\int_0^t \int_{-\infty}^{\infty} |\, c\,(x\,,s)\,|\; dx\; ds \;=\; \int_0^t \|\, c\,(\,\cdot\,,s\,)\,\|_1\; ds \;\leq\; 2\mathrm{C}\,t\; e^{-2t}$$

as claimed.

Since $b\,(x\,,t\,)$ is a solution of (5.15a,b) with initial data $b\,(x\,,0) \equiv 0$, it follows from Theorem 5.4 that $\|\, b\,(\,\cdot\,,t\,)\,\|_1 \leq 2\mathrm{C}\,t$. This fact allows us to apply Theorem 5.4 once more, this time to the function $w\,(x\,,t\,)$ on $\Omega = \mathbb{R}\times[0,\Delta t\,]$, to obtain

$$\|\, w\,(\,\cdot\,,\Delta t\,)\,\|_1 \leq \mathrm{C}\,(\Delta t\,)^2.$$

This is (5.13).

Continuing to think of $D_{\Delta t}\, R_{\Delta t}\, u^0$ as a numerical approximation to $F_{\Delta t}\, u^0$ we may now use consistency (Theorem 5.8) together with stability (Theorem 5.6) to prove that for fixed time $T = k\,\Delta t$, $(D_{\Delta t}\, R_{\Delta t}\,)^k\, u^0 \rightarrow F^k_{\Delta t}\, u^0$ as $\Delta t \rightarrow 0$. The proof proceeds precisely as in the case of Euler's method; at a given time step, j, use stability and consistency to reduce the error at time $j\,\Delta t$ into the error at time $(j-1)\Delta t$ plus a term of order $(\Delta t\,)^2$. Thus, the error at time $T = k\,\Delta t$ is the sum of k terms, each of order $(\Delta t\,)^2$ plus the error due to the initial approximation (which in our case is 0). The only difference here is that, in order to use the stability and consistency results, it is first necessary to show that the functions they are being applied to satisfy the appropriate hypotheses. It is at this point that much of our previous effort towards obtaining bounds on $F_t\, u^0$ and $D_t\, R_t\, u^0$ and the gradients of these functions is rewarded.

THEOREM 5.9 Let u be a solution of the Kolmogorov equation (1.3a,b) with initial data $u^0 \in \mathrm{C}^1(\mathbb{R})$ such that $0 \leq u^0 \leq 1$ and $u^0_x \in L^1(\mathbb{R}) \cap L^\infty(\mathbb{R})$. Then

$$\| F_{\Delta t}^k u^0 - (D_{\Delta t} R_{\Delta t})^k u^0 \|_1 \leq C \Delta t$$

where, if we let $T = k \Delta t$, the constant C is given by

$$C = T e^{3T} \{ e^T \| u_x^0 \|_\infty + \frac{4\sqrt{2\Delta t}}{\sqrt{\pi}} \} \| u_x^0 \|_1. \tag{5.16}$$

Proof: For $j = 1, \ldots, n$ let $u^j = F_{\Delta t}^j u^0$ be the exact solution to the Kolmogorov equation (1.3a,b) at time $t = j \Delta t$. By the maximum principle, Lemma 5.1, $0 \leq u^j \leq 1$ for each j and, since $u_x^0 \in L^\infty(\mathbb{R})$, Lemma 5.5 implies that that $(u^j)_x \in L^\infty(\mathbb{R})$. Similarly, let $v^j = (D_{\Delta t} R_{\Delta t})^j u^0$ denote the function obtained after j time steps from u^0 via exact operator splitting. Repeated application of Lemma 4.5 yields $0 \leq v^j \leq 1$. Furthermore, Lemmas 4.2 and 4.3 together with $(D_{\Delta t} R_{\Delta t} v^j)_x = (G_{\Delta t} * (R_{\Delta t} v^j))_x = G_{\Delta t} * (R_{\Delta t} v^j)_x$ imply that $\| (v^j)_x \|_\infty \leq e^{j \Delta t} \| u_x^0 \|_\infty$ and $\| (v^j)_x \|_1 \leq e^{j \Delta t} \| u_x^0 \|_1$. Hence, $(v^j)_x \in L^1(\mathbb{R}) \cap L^\infty(\mathbb{R})$. We can now use the stability of $F_{\Delta t}$ (one must verify, by induction, that $(u^{j-1} - v^{j-1}) \in L^1$) and Theorem 5.8 to obtain

$$\| F_{\Delta t}^j u^0 - (D_{\Delta t} R_{\Delta t})^j u^0 \|_1 \leq \| F_{\Delta t} u^{j-1} - F_{\Delta t} v^{j-1} \|_1 + \| F_{\Delta t} v^{j-1} - D_{\Delta t} R_{\Delta t} v^{j-1} \|_1$$

$$\leq e^{\Delta t} \| u^{j-1} - v^{j-1} \|_1 + C_{j-1} (\Delta t)^2$$

where the constant, C_{j-1}, is given by (5.14) with u^0 replaced by v^{j-1}. Hence,

$$\| F_{\Delta t}^k u^0 - (D_{\Delta t} R_{\Delta t})^k u^0 \|_1 \leq \sum_{j=0}^{k-1} e^{(k-1-j)\Delta t} C_j \, (\Delta t)^2 \leq \max_{j<k} C_j \; T e^T \Delta t .$$

Now use the L^1 and L^∞ bounds on $(v^j)_x$ found above to obtain

$$\max_{j<k} C_j = \max_{j<k} e^{3\Delta t} \{ e^{\Delta t} \| (v^j)_x \|_\infty + \frac{4\sqrt{2\Delta t}}{\sqrt{\pi}} \} \| (v^j)_x \|_1$$

$$= \max_{j<k} e^{(j+3)\Delta t} \left\{ e^{(j+1)\Delta t} \| u_x^0 \|_\infty + \frac{4\sqrt{2\Delta t}}{\sqrt{\pi}} \right\} \| u_x^0 \|_1$$

$$\leq e^{3T} \left\{ e^{T} \| u_x^0 \|_\infty + \frac{4\sqrt{2\Delta t}}{\sqrt{\pi}} \right\} \| u_x^0 \|_1 .$$

The theorem follows immediately.

6. The Error Due to the Approximate Reaction Operator $\widetilde{R}_{\Delta t}$

This goal of this chapter is to prove that if $\tilde{u}^j = (\widetilde{D}_{\Delta t}\,\widetilde{R}_{\Delta t})^j\,\tilde{u}^0$ is the computed solution after j time steps, then for all real $\gamma \geq 1$

$$P\left(\,\|R_{\Delta t}\,\tilde{u}^j - \widetilde{R}_{\Delta t}\,\tilde{u}^j\,\|_1 \;\geq\; \gamma C\sqrt{\ln N}\;(\Delta t)^2\,\right) \leq N^{1-\frac{9}{4}\gamma^2} \tag{6.1}$$

where the constant, C, depends only on the initial data, $\tilde{u}^0$, and the time $t = j\,\Delta t$. The proof of this inequality is based on the following two points:

1) Given any $L > 0$ such that $X_i^j \in (-L,L)$ for all i, then, by (2.1),

$$x < -L \;\Rightarrow\; \tilde{u}^j(x) = 1 \quad \text{and} \quad x > L \;\Rightarrow\; \tilde{u}^j(x) = 0.$$

This implies $R_{\Delta t}\,\tilde{u}(x) = \widetilde{R}_{\Delta t}\,\tilde{u}(x)$ for $|\,x\,| > L$ and thus, the L^1 estimate of the error is reduced to an estimate over the interval (-L,L),

$$\|R_{\Delta t}\,\tilde{u}^j - \widetilde{R}_{\Delta t}\,\tilde{u}^j\,\|_1 = \int_{-L}^{L} |\,R_{\Delta t}\,\tilde{u}(x) - \widetilde{R}_{\Delta t}\,\tilde{u}(x)\,|\;dx \;=\; \|R_{\Delta t}\,\tilde{u}^j - \widetilde{R}_{\Delta t}\,\tilde{u}^j\,\|_{(-L,L)}.$$

2) For fixed x, $\widetilde{R}_{\Delta t}$ is simply Euler's method for approximating the solution of an ODE and hence, the local truncation error is known to be $O(\Delta t^2)$. This fact can be exploited to obtain a bound for $|\,R_{\Delta t}\,\tilde{u}(x) - \widetilde{R}_{\Delta t}\,\tilde{u}(x)\,|$ which is uniform in x.

Together, 1) and 2) imply

$$\|R_{\Delta t}\,\tilde{u}^j - \widetilde{R}_{\Delta t}\,\tilde{u}^j\,\|_1 \;=\; \|R_{\Delta t}\,\tilde{u}^j - \widetilde{R}_{\Delta t}\,\tilde{u}^j\,\|_{(-L,L)} \;\leq\; const.\;\; 2\,L\,(\Delta t)^2.$$

In general, however, the size of the interval (-L,L) cannot be given a deterministic bound. For the particle positions, X_i^j, are random variables yet L has been chosen so that $|\,X_i^j\,| < L$ for all i. Consequently, the most that one can hope for is to find the probability that this

interval is a given size. This is accomplished by examining the movement of the particles, a task to which we now turn.

6.1 A Probabilistic Bound on the Particle Positions Recall that the position of the *ith* particle at the *jth* time step is denoted by X_i^j and that X_i^j is obtained from X_i^{j-1} by adding a normally distributed random variable, η_i^j, with mean 0 and variance $2\Delta t$ to X_i^{j-1}. Hence, $X_i^j = X_i^{j-1} + \eta_i^j$. The movement of the particles is thus governed by the distribution of the η_i^j. This distribution is given by

$$P(\eta_i^j < x) = \frac{1}{\sqrt{4\pi\Delta t}} \int_{-\infty}^{x} e^{-\frac{s^2}{4\Delta t}}\, dy = \phi\left(\frac{x}{\sqrt{2\Delta t}}\right)$$

where ϕ is the probability distribution function for a Gaussian distribution with mean 0 and variance 1,

$$\phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{s^2}{2}}\, ds\,. \tag{6.2}$$

LEMMA 6.1 Let $K > 0$ be chosen so that $X_i^0 \in (-K,K)$ for all i. Then for all $\alpha > 0$,

$$P(X_i^j < -K - \alpha) < \phi\left(\frac{-\alpha}{\sqrt{2j\,\Delta t}}\right), \tag{6.3a}$$

$$P(X_i^j > K + \alpha) < \phi\left(\frac{-\alpha}{\sqrt{2j\,\Delta t}}\right). \tag{6.3b}$$

Proof: First write the position of the *ith* particle, X_i^j, in terms of its initial position, X_i^0, and the subsequent random walks, $\eta_i^1, \ldots, \eta_i^j$,

$$X_i^j = X_i^{j-1} + \eta_i^j$$

$$= X_i^{j-2} + \eta_i^{j-1} + \eta_i^j$$

$$= X_i^0 + \eta_i^1 + \ \ldots \ + \eta_i^j$$

$$= X_i^0 + Y_i \, .$$

where $Y_i = \eta_i^1 + \ \ldots \ + \eta_i^j$. Since the the sum of j independent normally distributed random variables each with mean 0 and variance $2\Delta t$ is itself a normally distributed random variable with mean 0 and variance $2j\,\Delta t$ it follows that the probability distribution of Y_i is given by

$$P(Y_i < x) = \phi(\frac{x}{\sqrt{2j\,\Delta t}}).$$

By assumption, $-K < X_i^0 < K$ for all i, whereby

$$P(X_i^j < -K - \alpha) = P(X_i^0 + Y_i < -K - \alpha)$$

$$= P(Y_i < -K - X_i^0 - \alpha)$$

$$\leq P(Y_i < -\alpha)$$

$$= \phi(\frac{-\alpha}{\sqrt{2k\,\Delta t}}).$$

The inequality (6.3b) is proved in a similar fashion.

It is well known that $\phi(x)$ decreases at an exponential rate as $x \to -\infty$. This fact allows us to compute a bound on $\phi(x)$ and hence, on the probability that the particles lie outside a given interval.

LEMMA 6.2 For any $x < 0$,

$$\phi(x) \leq \frac{-1}{x\sqrt{2\pi}} e^{-\frac{x^2}{2}}. \tag{6.4}$$

Remark: This is the first of a sequence of bounds for ϕ that can be derived by considering an asymptotic expansion of $\phi(x)$ in powers of $\frac{1}{x}$. See Feller [22], p. 175 for further details.

Proof: Note that $s \leq x < 0$ implies $\frac{s}{x} \geq 1$. Therefore,

$$\phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{s^2}{2}} ds \leq \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \frac{s}{x} e^{-\frac{s^2}{2}} ds = \frac{-1}{x\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

as claimed.

6.2 A Probability Inequality for the Error Due to $\widetilde{R}_{\Delta t}$ We are now prepared to prove a probability inequality for the error, $\| R_{\Delta t}\, \tilde{u}^j - \widetilde{R}_{\Delta t}\, \tilde{u}^j \|_1$. To begin we will establish an inequality which depends on the parameter α which, as in Lemma 6.1, corresponds to the amount the particles may have strayed from their initial positions. Then, in the subsequent corollary, we show how to choose α to obtain the inequality stated at the beginning of the chapter, a form which will be more convenient for later use.

THEOREM 6.3 Let $\tilde{u}^0 \in \mathbf{S}$, let $\mathrm{K} > 0$ be such that $\mathrm{X}_i^0 \in (-\mathrm{K},\mathrm{K})$ for all i and let $\tilde{u}^j(x) = (\widetilde{D}_{\Delta t}\, \widetilde{R}_{\Delta t})^j\, \tilde{u}^0$ be the step function obtained from $\tilde{u}^0$ via the random gradient method at time $T = j\,\Delta t$. Then for all $\alpha > 0$,

$$P\left(\| R_{\Delta t}\, \tilde{u}^j - \widetilde{R}_{\Delta t}\, \tilde{u}^j \|_{L^1} > \frac{\sqrt{3}}{9}(\mathrm{K}+\alpha)(\Delta t)^2 \right) \leq \frac{N\sqrt{4T}}{\alpha\sqrt{\pi}}\, e^{\frac{-\alpha^2}{4T}}. \tag{6.5}$$

Proof: Set $\mathrm{L} = \mathrm{K} + \alpha$. As indicated in the introduction to this chapter, we prove (6.5) by bounding $\| R_{\Delta t}\, \tilde{u}^j - \widetilde{R}_{\Delta t}\, \tilde{u}^j \|_{L^1}$ under the assumption that at time $T = j\,\Delta t$ the particles all

lie entirely within the interval (-L,L) and then estimating the probability that this assumption is true. Therefore, assume that

$$-\mathrm{L} < \mathrm{X}_i^j < \mathrm{L}, \qquad i = 1, \ldots, N. \tag{6.6}$$

By Corollary 2.3, $\tilde{u}^j \in \mathbf{S}$, whereby the w_i^j satisfy $\sum w_i^j = 1$. Hence, $\tilde{u}^j(x) = 1$ for $x < -\mathrm{L}$ and $\tilde{u}^j(x) = 0$ for $x > \mathrm{L}$. Now use (4.1) to write

$$R_{\Delta t}\, \tilde{u}^j(x) = \frac{\tilde{u}^j(x)\, e^{\Delta t}}{1 - \tilde{u}^j(x) + \tilde{u}^j(x)\, e^{\Delta t}} = \begin{cases} 1, & \text{if } \tilde{u}^j(x) = 1, \\ 0, & \text{if } \tilde{u}^j(x) = 0. \end{cases}$$

In other words,

$$R_{\Delta t}\, \tilde{u}^j(x) = \begin{cases} 1, & \text{if } x < -\mathrm{L}, \\ 0, & \text{if } x > \mathrm{L}. \end{cases}$$

Similarly, since

$$\widetilde{R}_{\Delta t}\, \tilde{u}^j(x) = \sum_{i=1}^{N} H(\mathrm{X}_i^j - x)\, w_i^{j+1}$$

where the w_i^{j+1} are given by (2.4), and since, by Lemma 2.2, the w_i^{j+1} satisfy $\sum w_i^{j+1} = 1$, it is clear that $\widetilde{R}_{\Delta t}\, \tilde{u}^j$ satisfies

$$\widetilde{R}_{\Delta t}\, \tilde{u}^j(x) = \begin{cases} 1, & \text{if } x < -\mathrm{L}, \\ 0, & \text{if } x > \mathrm{L}. \end{cases}$$

Consequently,

$$\| R_{\Delta t}\, \tilde{u}^j - \widetilde{R}_{\Delta t}\, \tilde{u}^j \|_1 = \| R_{\Delta t}\, \tilde{u}^j - \widetilde{R}_{\Delta t}\, \tilde{u}^j \|_{(-\mathrm{L},\mathrm{L})}.$$

The next step consists of placing a pointwise (in x) bound on the difference $R_{\Delta t}\,\tilde{u}^j(x) - \widetilde{R}_{\Delta t}\,\tilde{u}^j(x)$ and then using this bound to estimate the L^1 integral over (–L,L). Fix x and consider $R_t\tilde{u}^j(x)$ as a function of t alone, say $g(t) = R_t\tilde{u}^j(x)$. The approximation $\widetilde{R}_{\Delta t}\,\tilde{u}^j(x)$ is simply Euler's method used to approximate the solution, g, of the ODE $g' = g(1-g)$ with initial data $g_0 = \tilde{u}^j(x)$ (see (2.2)). In other words,

$$\widetilde{R}_{\Delta t}\,\tilde{u}^j(x) = \tilde{u}^j(x) + \Delta t\;\tilde{u}^j(x)\,(1-\tilde{u}^j(x))$$

$$= g_0 + \Delta t\;g_0(1-g_0).$$

It is well known that the error in this approximation after one time step, called the *local truncation error* or *local discretization error*, is bounded by

$$|\,R_{\Delta t}\,\tilde{u}^j(x) - \widetilde{R}_{\Delta t}\,\tilde{u}^j(x)\,| < \max_{0\le\xi\le\Delta t}\frac{|\,g''(\xi)\,|}{2}(\Delta t)^2. \tag{6.7}$$

(See Conte & de Boor [18], p. 359.) In order to bound g'' first note that since g satisfies $g' = g(1-g)$,

$$g'' = g' - 2gg' = 2g^3 - 3g^2 + g\,.$$

Corollary 2.3 implies $0 \le \tilde{u}^j(x) \le 1$ for all $x \in \mathbb{R}$ and consequently, by Lemma 4.5, $g(t) = R_t\tilde{u}^j(x)$ satisfies $0 \le g(t) \le 1$ for all t. Thus,

$$\max_{\xi\in[0,\Delta t]}|\,g''(\xi)\,| = \max_{0\le g\le 1}|\,2g^3 - 3g^2 + g\,| \le \frac{\sqrt{3}}{18}.$$

Substituting this estimate into (6.7) yields the pointwise bound

$$|\,R_{\Delta t}\,\tilde{u}^j(x) - \widetilde{R}_{\Delta t}\,\tilde{u}^j(x)\,| \le \frac{\sqrt{3}}{18}(\Delta t)^2$$

which must hold for any $x \in \mathbb{R}$ since x was arbitrary. Now integrate over $x \in$ (-L,L) to obtain

$$\| R_{\Delta t}\, \tilde{u}^j - \widetilde{R}_{\Delta t}\, \tilde{u}^j \|_{L^1} = \int_{-L}^{L} | R_{\Delta t}\, \tilde{u}^j(x) - \widetilde{R}_{\Delta t}\, \tilde{u}^j(x) |\, dx$$

$$\leq \frac{\sqrt{3}}{18} (\Delta t)^2 \int_{-L}^{L} dx$$

$$= \frac{\sqrt{3}}{9} \mathrm{L}\, (\Delta t)^2.$$

This estimate is valid as long as the assumption (6.6) holds. In other words, the probability that the L^1 error exceeds this estimate is bounded by the probability that at least one of the particles lies outside (-L,L). Thus, by Lemmas 6.1 and 6.2,

$$P(\, \| R_{\Delta t}\, \tilde{u}^j - \widetilde{R}_{\Delta t}\, \tilde{u}^j \|_{L^1} > 6\mathrm{L}(\Delta t)^2 \,) \leq P(\, \exists\, i : \mathrm{X}_i^j \notin (-\mathrm{L},\mathrm{L}) \,)$$

$$\leq \sum_{i=1}^{N} [\, P(\mathrm{X}_i^j < -\mathrm{L}) + P(\mathrm{X}_i^j > \mathrm{L}) \,]$$

$$= 2N\, \phi(\frac{-\alpha}{\sqrt{2j\,\Delta t}})$$

$$\leq \frac{N\sqrt{4T}}{\alpha\sqrt{\pi}}\, e^{\frac{-\alpha^2}{4T}}.$$

The main result of this chapter, equation (6.1), follows from Theorem 6.3 immediately upon setting $\alpha = 3\gamma\sqrt{T}\,\sqrt{\ln N}$ for any real $\gamma \geq 1$.

COROLLARY 6.4 Let $\tilde{u}^0 \in \mathbf{S}$ be generated by N particles and let $\mathrm{K} > 0$ be chosen so that the particle positions, X_i^0, all lie in the interval (-K,K). Let $T = j\,\Delta t$ denote the time and let

$$\mathrm{C} = \frac{\sqrt{3}}{9} (\mathrm{K} + 3\sqrt{T}\,). \tag{6.8}$$

Then for all real $\gamma \geq 1$,

$$P\left(\| R_{\Delta t}\, \tilde{u}^j - \widetilde{R}_{\Delta t}\, \tilde{u}^j \|_1 \;\geq\; \gamma C \sqrt{\ln N}\, (\Delta t)^2 \right) \leq N^{1 - \frac{9}{4}\gamma^2}.$$

7. The L^1 Convergence of the Approximate Diffusion Operator, $\widetilde{D}_{\Delta t}$

In this chapter we prove that if $\tilde{u}^0 \in \mathbf{S}$ and if $\tilde{v}^j = \widetilde{R}_{\Delta t}(\widetilde{D}_{\Delta t}\widetilde{R}_{\Delta t})^{j-1}\tilde{u}^0$, then for all real $\gamma \geq 1$,

$$P\Big(\| D_{\Delta t}\tilde{v}^j - \widetilde{D}_{\Delta t}\tilde{v}^j \|_1 \geq \gamma C \frac{\ln N}{\sqrt{N}}\Big) \leq 4\, N^{1-\frac{2}{4}\gamma}$$

where C depends only on the initial data, $\tilde{u}^0$, and the time, $t = j\,\Delta t$. Setting $\gamma = 1$, it follows that given any $\epsilon > 0$ we can find $N_0 = N_0(\epsilon)$ such that for all $N \geq N_0$,

$$P\Big(\| D_{\Delta t}\tilde{v}^j - \widetilde{D}_{\Delta t}\tilde{v}^j \|_1 \geq C\frac{\ln N}{\sqrt{N}}\Big) \leq \epsilon$$

or, equivalently,

$$P\Big(\| D_{\Delta t}\tilde{v}^j - \widetilde{D}_{\Delta t}\tilde{v}^j \|_1 < C\frac{\ln N}{\sqrt{N}}\Big) > 1-\epsilon.$$

Thus, by using sufficiently many particles, one can guarantee that the error due to approximating $D_{\Delta t}$ by $\widetilde{D}_{\Delta t}$ is small with arbitrarily high probability. It is in this sense that the approximate diffusion operator, $\widetilde{D}_{\Delta t}$, converges to the exact diffusion operator, $D_{\Delta t}$. Hald has proven similar results in the L^2 norm ([33]) and the sup norm ([29]).

There are two fundamental facts about using random walks to solve the heat equation (3.2a,b) that should be noted here. First, the rate of convergence does not depend on the time step, Δt, and hence is exclusively a function of the number of particles, N. Second, the rate of convergence is $O(\ln N/\sqrt{N})$. (Actually, the rate of convergence is probably $O(1/\sqrt{N})$, with the factor $\ln N$ simply being a spurious term introduced by the analysis.) Therefore, since all of the other sources of error behave like $O(\Delta t)$, and since it is considerably cheaper to halve the time step than to quadruple the number of particles, this quickly becomes the dominant source of error. This is an essential feature of numerical methods which use random walks to model

diffusion processes.

The same problem occurs when one uses random sampling to approximate an integral as in Monte Carlo integration ([21, 36, 49]). There have been many suggestions for improving the rate of convergence for Monte Carlo methods. For example, one can use equidistributed sequences instead of randomly chosen numbers to sample the integral ([17, 35, 36]) or one can use some *a priori* knowledge of the solution to reduce the variance of the error ([1, 2, 10, 12, 43, 49]). However, in spite of a considerable amount of work improvement of the convergence rate continues to be difficult.

This chapter is divided into four sections. In §7.1 we begin by introducing some notation. Following this, in §7.2, we present the basic ideas. In the third section, §7.3, we prove that $\widetilde{D}_{\Delta t}$ converges to $D_{\Delta t}$ pointwise: for all $x \in \mathbb{R}$ and all $\alpha > 0$,

$$P(\ |\ D_{\Delta t}\, \tilde{v}^{\,j}(x) - \widetilde{D}_{\Delta t}\, \tilde{v}^{\,j}(x)\ |\ \ \geq\ \alpha N\, \overline{w}\ )\ \leq\ 2e^{-2N\alpha^2}$$

where $\overline{w}$ is an upper bound on the particle strengths. Finally, in §7.4, we use this pointwise bound, together with the fact that the pointwise error decreases exponentially fast as $|\ x\ | \rightarrow \infty$, to prove the error estimate in the L^1 norm.

The results in this chapter are based on the work of Roberts. Most of the reasoning is identical to the arguments in Chapter 4 of [47]. The main difference between Roberts' convergence proof for the approximate diffusion operator and our own is that in his numerical method the particle strengths remain constant for all time, whereas in the random gradient method, the particle strengths are random variables. This difference manifests itself primarily in the proof of Lemma 7.4 where, in order to establish a pointwise bound on the difference between $D_{\Delta t}\, \tilde{v}^{\,k}$ and $\widetilde{D}_{\Delta t}\, \tilde{v}^{\,k}$, it is first necessary to bound the particle strengths. See the remarks preceding Lemma 7.4 for further details.

7.1 The Underlying Probability Space, Ω In Chapter 8 we will show that

$$P\Big(\| F^k_{\Delta t}\, u^0 - (\widetilde{D}_{\Delta t}\, \widetilde{R}_{\Delta t})^k\, \tilde{u}^0 \|_1 \geq \gamma C \frac{\ln N}{\sqrt[4]{N}} \Big) \;\leq\; 5\, T\, N^{\frac{5}{4} - \frac{9}{4}\gamma} \tag{7.1}$$

where C is a constant, $\gamma \geq 1$ is an arbitrary real number, and $T = k\,\Delta t$. Implicit in this statement is the existence of a probability space (Ω, Σ, μ) upon which the error, $\| F^k_{\Delta t}\, u^0 - (\widetilde{D}_{\Delta t}\, \widetilde{R}_{\Delta t})^k\, \tilde{u}^0 \|_1$, is a random variable that satisfies (7.1). There is a simple one to one correspondence between elements of Ω and a given run of the random gradient method: each $\omega \in \Omega$ corresponds to one realization of the random walks, $\omega = (\eta^1_1, \ldots, \eta^1_N, \ldots, \eta^k_1, \ldots, \eta^k_N)$. Most authors generally find it unnecessary to develop the underlying probability space (for example: Hald [32, 33], Goodman [27]). However, since several of the probability estimates below are estimates over subspaces of Ω rather than Ω itself (each subspace corresponds to one time step), we choose to go into greater detail here.

Let $\mathbb{R}^1$ denote the real line, B^1 the standard Euclidean Borel field on $\mathbb{R}^1$, and μ^1 Gaussian measure on $\mathbb{R}^1$,

$$\mu^1(S) = \frac{1}{\sqrt{4\pi\Delta t}} \int_S e^{\frac{-y^2}{4\Delta t}}\, dy \qquad \forall\, S \in B^1.$$

(Note that μ^1 depends on the time step, Δt.) Let $(\Omega, \Sigma, \mu) = (\prod_{j=1}^{k} \Omega_j, \prod_{j=1}^{k} \Sigma_j, \prod_{j=1}^{k} \mu_j)$ where the *jth* component space, $(\Omega_j, \Sigma_j, \mu_j) = (\mathbb{R}^N, B^N, \mu^N)$, is simply N-dimensional Euclidean space, $\mathbb{R}^N$, with the standard Euclidean Borel field, $B^N = \prod_{i=1}^{N} B^1$, and N-dimensional Gaussian measure $\mu^N = \prod_{i=1}^{N} \mu^1$,

$$\mu^N(S) = \mu^1(S_1) \cdot \mu^1(S_2) \cdot \;\ldots\; \cdot \mu^1(S_N)$$

for all $S = S_1 \times S_2 \times \ldots \times S_N \in B^N$. Here N is the number of particles and k is the time step at which we wish to examine the error; both N and k are fixed.

The component space $(\Omega_j, \Sigma_j, \mu_j)$ has been chosen so that there is a one to one correspondence between an element $\omega_j \in \Omega_j$ and the N random numbers used at the jth time step to random walk the particles, $\eta_1^j, \ldots, \eta_N^j$. In what follows, ω_j will be used to denote a random variable on Ω_j whereas $(\eta_1^j, \ldots, \eta_N^j)$ is (ideally) a specific realization of that variable. Sometimes, however, it is convenient to ignore this distinction and regard the vector $(\eta_1^j, \ldots, \eta_N^j)$ as a random variable on the subspace Ω_j. In practice it should be apparent from the context whether $(\eta_1^j, \ldots, \eta_N^j)$ denotes the random variable, ω_j, or a specific realization of ω_j.

7.2 A Brief Outline of the Argument In order to understand our motivation for constructing the probability space Ω it is helpful to know how we arrive at (7.1). Recall that the error at time $t = k\,\Delta t$ can be bounded by

$$\| F_{\Delta t}^k u^0 - (\widetilde{D}_{\Delta t} \widetilde{R}_{\Delta t})^k \tilde{u}^0 \|_1 \leq \| F_{\Delta t}^k u^0 - (D_{\Delta t} R_{\Delta t})^k u^0 \|_1$$

$$+ \| (D_{\Delta t} R_{\Delta t})^k u^0 - (D_{\Delta t} R_{\Delta t})^k \tilde{u}^0 \|_1 + \| (D_{\Delta t} R_{\Delta t})^k \tilde{u}^0 - (\widetilde{D}_{\Delta t} \widetilde{R}_{\Delta t})^k \tilde{u}^0 \|_1.$$

The first two terms on the right, which were estimated in Chapters 4 and 5, are deterministic in the sense that they are unaffected by the choice of the random numbers, $\eta_1, \ldots, \eta_N$, at each time step. Hence, considered as random variables on Ω, these two terms are constants. The third term, however, is not constant on Ω, since it contains the error due to approximating $D_{\Delta t}$ by $\widetilde{D}_{\Delta t}$. Using the stability results from Chapter 4 (Lemma 4.1 and Corollary 4.4) we can split this term into two components,

$$\| (D_{\Delta t} R_{\Delta t})^k \tilde{u}^0 - (\widetilde{D}_{\Delta t} \widetilde{R}_{\Delta t})^k \tilde{u}^0 \|_1 \leq \| D_{\Delta t} R_{\Delta t} (D_{\Delta t} R_{\Delta t})^{k-1} \tilde{u}^0 - D_{\Delta t} R_{\Delta t} (\widetilde{D}_{\Delta t} \widetilde{R}_{\Delta t})^{k-1} \tilde{u}^0 \|_1$$

$$+ \ \| D_{\Delta t} R_{\Delta t} (\widetilde{D}_{\Delta t} \widetilde{R}_{\Delta t})^{k-1} \tilde{u}^0 - \widetilde{D}_{\Delta t} \widetilde{R}_{\Delta t} (\widetilde{D}_{\Delta t} \widetilde{R}_{\Delta t})^{k-1} \tilde{u}^0 \|_1$$

$$\leq \ e^{\Delta t} \ \| (D_{\Delta t} R_{\Delta t})^{k-1} \tilde{u}^0 - (\widetilde{D}_{\Delta t} \widetilde{R}_{\Delta t})^{k-1} \tilde{u}^0 \|_1$$

$$+ \ \| D_{\Delta t} R_{\Delta t} \tilde{u}^{k-1} - \widetilde{D}_{\Delta t} \widetilde{R}_{\Delta t} \tilde{u}^{k-1} \|_1 .$$

Proceeding inductively and using

$$\| D_{\Delta t} R_{\Delta t} \tilde{u}^{j-1} - \widetilde{D}_{\Delta t} \widetilde{R}_{\Delta t} \tilde{u}^{j-1} \|_1 \ \leq \ \| D_{\Delta t} R_{\Delta t} \tilde{u}^{j-1} - D_{\Delta t} \widetilde{R}_{\Delta t} \tilde{u}^{j-1} \|_1$$

$$+ \ \| D_{\Delta t} \widetilde{R}_{\Delta t} \tilde{u}^{j-1} - \widetilde{D}_{\Delta t} \widetilde{R}_{\Delta t} \tilde{u}^{j-1} \|_1$$

$$\leq \ \| R_{\Delta t} \tilde{u}^{j-1} - \widetilde{R}_{\Delta t} \tilde{u}^{j-1} \|_1 \ + \ \| D_{\Delta t} \tilde{v}^{j} - \widetilde{D}_{\Delta t} \tilde{v}^{j} \|_1 ,$$

we find

$$\| (D_{\Delta t} R_{\Delta t})^k \tilde{u}^0 - (\widetilde{D}_{\Delta t} \widetilde{R}_{\Delta t})^k \tilde{u}^0 \|_1 \ \leq \ \sum_{j=1}^{k} e^{(k-j)\Delta t} \ \| D_{\Delta t} R_{\Delta t} \tilde{u}^{j-1} - \widetilde{D}_{\Delta t} \widetilde{R}_{\Delta t} \tilde{u}^{j-1} \|_1 \tag{7.2}$$

$$\leq \ e^T \sum_{j=1}^{k} \left(\| R_{\Delta t} \tilde{u}^{j-1} - \widetilde{R}_{\Delta t} \tilde{u}^{j-1} \|_1 + \| D_{\Delta t} \tilde{v}^{j} - \widetilde{D}_{\Delta t} \tilde{v}^{j} \|_1 \right).$$

Therefore,

$$P(\| (D_{\Delta t} R_{\Delta t})^k \tilde{u}^0 - (\widetilde{D}_{\Delta t} \widetilde{R}_{\Delta t})^k \tilde{u}^0 \|_1 \ \geq \ \epsilon \ k \ e^T) \ \leq \ \sum_{j=1}^{k} P(\| R_{\Delta t} \tilde{u}^{j-1} - \widetilde{R}_{\Delta t} \tilde{u}^{j-1} \|_1 \ \geq \ \epsilon_1)$$

$$+ \sum_{j=1}^{k} P(\| D_{\Delta t} \tilde{v}^{j} - \widetilde{D}_{\Delta t} \tilde{v}^{j} \|_1 \ \geq \ \epsilon_2)$$

for all $\epsilon = \epsilon_1 + \epsilon_2 \in \mathbb{R}$.

The first k terms on the right hand side of this inequality can be estimated using the results of Chapter 6. Let us consider what is involved in estimating terms of the form $P(\| D_{\Delta t}\, \tilde{v}^{\,j} - \widetilde{D}_{\Delta t}\, \tilde{v}^{\,j} \|_1 \geq \epsilon_2)$. We first divide the real line into three pieces, $(-\infty,-L)$, $(-L,L)$, and (L,∞), where $L > 0$ is free to be chosen as we wish. Thus,

$$P(\| D_{\Delta t}\, \tilde{v}^{\,j} - \widetilde{D}_{\Delta t}\, \tilde{v}^{\,j} \|_1 \geq \epsilon_2) \leq P(\| D_{\Delta t}\, \tilde{v} - \widetilde{D}_{\Delta t}\, \tilde{v} \|_{(-\infty,-L)} \geq \epsilon_3)$$

$$+ P(\| D_{\Delta t}\, \tilde{v}^{\,j} - \widetilde{D}_{\Delta t}\, \tilde{v}^{\,j} \|_{(-L,L)} \geq \epsilon_4) + P(\| D_{\Delta t}\, \tilde{v}^{\,j} - \widetilde{D}_{\Delta t}\, \tilde{v}^{\,j} \|_{(L,\infty)} \geq \epsilon_3)$$

for any ϵ_3, ϵ_4 satisfying $\epsilon_2 = 2\epsilon_3 + \epsilon_4$. Here we have used $\| \ldots \|_{(-L,L)}$ to denote the L^1 norm on the interval $(-L,L)$ and similarly for $\| \ldots \|_{(-\infty,-L)}$ and $\| \ldots \|_{(L,\infty)}$.

The error over the tails, $(-\infty,-L)$ and (L,∞), can be estimated using ideas very similar to those used in Chapter 6. We first bound the error over $(-\infty,-L)$ (*resp.* (L,∞)) under the assumption that all of the particles lie in the interval $(-B,B) \subset (-L,L)$ at times $(j-1)\Delta t$ and $j\Delta t$. Lemmas 6.1 and 6.2 are then used to estimate the probability that this assumption holds.

The bound for the error over the interval $(-L,L)$ is a bit more involved. Observe that the error, $\| D_{\Delta t}\, \tilde{v}^{\,j} - \widetilde{D}_{\Delta t}\, \tilde{v}^{\,j} \|_1$, considered as a function of $\omega = (\omega_1, \ldots, \omega_k) \in \Omega$, does not depend on $\omega_{j+1}, \ldots, \omega_k$, *i.e.* on the random walks after the *jth* time step. Furthermore, $\tilde{v}^{\,j} = \widetilde{R}_{\Delta t}(\widetilde{D}_{\Delta t}\, \widetilde{R}_{\Delta t})^{j-1}\tilde{u}^{\,0}$ depends only on $\omega_j^* = (\omega_1, \ldots, \omega_{j-1})$, since it is also unaffected by the random walks taken at the *jth* time step. Our estimate proceeds in the following way. We first prove that for each possible realization of $\tilde{v}^{\,j} = \tilde{v}^{\,j}(\omega_j^*)$, (*i.e.* for each fixed $\omega_j^* \in \Omega_1 \times \ldots \times \Omega_{j-1}$),

$$P_{\Omega_j}(\| D_{\Delta t}\, \tilde{v}^{\,j} - \widetilde{D}_{\Delta t}\, \tilde{v}^{\,j} \|_{(-L,L)} \geq \epsilon_4) = \mu_j\, \{\omega_j \in \Omega_j : \| D_{\Delta t}\, \tilde{v}^{\,j} - \widetilde{D}_{\Delta t}\, \tilde{v}^{\,j} \|_{(-L,L)} \geq \epsilon_4\}$$

$$= \int_{\Omega_j} \mathrm{H}\,(\| D_{\Delta t}\, \tilde{v}^{\,j} - \widetilde{D}_{\Delta t}\, \tilde{v}^{\,j} \|_1 - \epsilon_4)\, d\mu_j$$

$$\leq \delta$$

where δ depends on N, Δt, $t = j\Delta t$, and the initial data, $\tilde{u}^0$, but not on ω_j^*. Here we have used P_{Ω_j} to denote probability with respect to the subspace Ω_j. Therefore, since the error does not depend on $\omega_{j+1}, \ldots, \omega_k$, we have

$$\begin{aligned} P(\| D_{\Delta t}\tilde{v}^j - \widetilde{D}_{\Delta t}\tilde{v}^j \|_1 \geq \epsilon_4) &= \int_{\Omega} \mathrm{H}\,(\| D_{\Delta t}\tilde{v}^j - \widetilde{D}_{\Delta t}\tilde{v}^j \|_1 - \epsilon_4)\, d\mu \\ &= \int_{\Omega_1} \cdots \int_{\Omega_j} \mathrm{H}(\| D_{\Delta t}\tilde{v}^j - \widetilde{D}_{\Delta t}\tilde{v}^j \|_1 - \epsilon_4)\, d\mu_j \ldots d\mu_1 \\ &\leq \delta \int_{\Omega_1} \cdots \int_{\Omega_{j-1}} d\mu_{j-1} \ldots d\mu_1 \\ &= \delta. \end{aligned}$$

Thus, the problem has been reduced to the following: Given any fixed $\tilde{v}^j = \tilde{v}^j(\omega_j^*) \in \mathbf{S}$, estimate the L^1 error on (-L,L) between the exact solution of the heat equation with initial data $\tilde{v}^j$, $D_{\Delta t}\tilde{v}^j$, and the random walk solution, $\widetilde{D}_{\Delta t}\tilde{v}^j$. As we shall demonstrate below, it is possible to estimate this error in terms of the pointwise error, $P_{\Omega_j}(\,|\,D_{\Delta t}\tilde{v}^j(x) - \widetilde{D}_{\Delta t}\tilde{v}^j(x)\,| \;\geq\; \alpha\,)$, at N points in (-L,L). We now turn to the task of establishing such pointwise estimates.

7.3 Pointwise Estimates In this section we investigate the size of the pointwise error $|\,D_{\Delta t}\tilde{v}^j(x) - \widetilde{D}_{\Delta t}\tilde{v}^j(x)\,|$. The principal result will be that

$$P_{\Omega_j}(\,|\,D_{\Delta t}\tilde{v}^j(x) - \widetilde{D}_{\Delta t}\tilde{v}^j(x)\,| \;\geq\; \alpha N\overline{w}\,) \;\leq\; 2\,e^{-2N\alpha^2}$$

holds uniformly for all $x \in \mathbb{R}$ where $\overline{w}$ is the maximum particle strength for the function $\tilde{v}^j$.

We wish to emphasize that this is an estimate over the subspace Ω_j. In other words, for fixed $\tilde{v}^j$, we examine the pointwise error $|D_{\Delta t}\tilde{v}^j(x) - \widetilde{D}_{\Delta t}\tilde{v}^j(x)|$ as a function of the random walks at the *jth* time step, $\eta_1^j, \ldots, \eta_N^j$. Furthermore, this estimate holds uniformly for all possible $\tilde{v}^j$ (and hence, for all $\omega_j^* \in \Omega_1 \times \ldots \times \Omega_{j-1}$).

Our first step will be to establish that for all $x \in \mathbb{R}$, $\mathrm{E}_{\Omega_j}[\widetilde{D}_{\Delta t}\tilde{v}^j(x)] = D_{\Delta t}\tilde{v}^j(x)$ where E_{Ω_j} denotes expected value with respect to $\eta_1^j, \ldots, \eta_N^j$. While it is also true that $\mathrm{E}_{\Omega}[\widetilde{D}_{\Delta t}\tilde{v}^j(x)] = D_{\Delta t}\tilde{v}^j(x)$ (in fact this is a consequence of the equality over Ω_j) at the moment we are only concerned with the error that is introduced at the *jth* time step.

LEMMA 7.1 Fix $\tilde{u}^0 \in \mathbf{S}$ and let $\tilde{v}^j = \widetilde{R}_{\Delta t}(\widetilde{D}_{\Delta t}\widetilde{R}_{\Delta t})^{j-1}\tilde{u}^0$. Then for any $x \in \mathbb{R}$,

$$\mathrm{E}_{\Omega_j}[\widetilde{D}_{\Delta t}\tilde{v}^j(x)] = D_{\Delta t}\tilde{v}^j(x).$$

Proof: First, suppose that $\tilde{v}^j(x) = \mathrm{H}(\mathrm{X} - x)$ for arbitrary $\mathrm{X} \in \mathbb{R}$. Then, by (3.5), $\widetilde{D}_{\Delta t}\tilde{v}^j(x) = \mathrm{H}(\mathrm{X} + \eta - x)$ where η is a Gaussian random variable with mean 0 and variance $2\Delta t$. The expected value of $\widetilde{D}_{\Delta t}\tilde{v}^j(x)$ is therefore

$$\mathrm{E}_{\Omega_j}[\widetilde{D}_{\Delta t}\tilde{v}^j(x)] = \frac{1}{\sqrt{4\pi\Delta t}}\int_{-\infty}^{\infty}\mathrm{H}(\mathrm{X} + \eta - x)\, e^{\frac{-\eta^2}{4\Delta t}}\, d\eta$$

$$= (G_{\Delta t} * \tilde{v}^j)(x)$$

$$= D_{\Delta t}\tilde{v}^j(x).$$

In general, $\tilde{v}^j$ is of the form

$$\tilde{v}^{\,j}(x) = \sum_{i=1}^{N} \mathrm{H}\,(\mathrm{X}_i - x)\, w_i\,.$$

(When no confusion is likely we shall drop the superscripts $j-1$ and j from X_i^{j-1}, $\eta_i^{\,j}$, and w_i^j.) In this case $\widetilde{D}_{\Delta t}\,\tilde{v}^{\,j}(x)$ is given by

$$\widetilde{D}_{\Delta t}\,\tilde{v}^{\,j}(x) = \sum_{i=1}^{N} \mathrm{H}\,(\mathrm{X}_i + \eta_i - x)\, w_i$$

where $\eta_1, \ldots, \eta_N$ are N independent Gaussian random variables with mean 0 and variance $2\Delta t$. Since both the expected value operator, E, and the diffusion operator, $D_{\Delta t}$, are linear, it now follows that

$$\begin{aligned}
\mathrm{E}_{\Omega_j}[\,\widetilde{D}_{\Delta t}\,\tilde{v}^{\,j}(x)\,] &= \sum_{i=1}^{N} \mathrm{E}\,[\,\mathrm{H}(\mathrm{X}_i + \eta_i - x)\,]\, w_i \\
&= \sum_{i=1}^{N} (D_{\Delta t}\,\mathrm{H}(\mathrm{X}_i + \eta_i - x)\,)\, w_i \\
&= D_{\Delta t}\,\tilde{v}^{\,j}(x).
\end{aligned}$$

Armed with the fact that $\mathrm{E}_{\Omega_j}[\,\widetilde{D}_{\Delta t}\,\tilde{v}^{\,j}(x)\,] = D_{\Delta t}\,\tilde{v}^{\,j}(x)$ we can now estimate the size of the error due to approximating $D_{\Delta t}\,\tilde{v}^{\,j}(x)$ by $\widetilde{D}_{\Delta t}\,\tilde{v}^{\,j}(x)$. The simplest approach is to use Chebychev's inequality ([16], p. 46),

$$P(\mathrm{Z} \geq s) \;\leq\; \frac{\mathrm{E}\,[\phi(\mathrm{Z})]}{\phi(s)}$$

where Z is a random variable, $s > 0$, $\phi(-s) = \phi(s)$, and ϕ is a strictly positive and increasing function on $(0,\infty)$. Setting $\mathrm{Z} = |\,D_{\Delta t}\,\tilde{v}^{\,j}(x) - \widetilde{D}_{\Delta t}\,\tilde{v}^{\,j}(x)\,|$ and $\phi(s) = s^2$ we find

$$P_{\Omega_j}(\,|\,D_{\Delta t}\,\tilde{v}^{\,j}(x)-\widetilde{D}_{\Delta t}\,\tilde{v}^{\,j}(x)\,|\;\geq s\;)\;\leq\;\frac{var\,(\widetilde{D}_{\Delta t}\,\tilde{v}^{\,j}(x)\,)}{s^{\,2}}\qquad\qquad\forall\,s\;>0$$

where

$$var\,(\widetilde{D}_{\Delta t}\,\tilde{v}^{\,j}(x)\,)\;=\;\mathrm{E}_{\Omega_j}[\,(\widetilde{D}_{\Delta t}\,\tilde{v}^{\,j}(x)-\mathrm{E}_{\Omega_j}[\,\widetilde{D}_{\Delta t}\,\tilde{v}^{\,j}(x)\,]\,)^2\,]$$

$$=\mathrm{E}_{\Omega_j}[\,(\widetilde{D}_{\Delta t}\,\tilde{v}^{\,j}(x)-D_{\Delta t}\,\tilde{v}^{\,j}(x)\,)^2\,]$$

is the variance with respect to the random walks at the jth time step, $\eta_1,\ldots,\eta_N$. In order to estimate the variance first note that $\mathrm{H}(\mathrm{X}_i+\eta_i-x)^2=\mathrm{H}(\mathrm{X}_i+\eta_i-x)$ and $0\;\leq\;\mathrm{H}(\mathrm{X}_i+\eta_i-x)\;\leq\;1$. Hence, since $\max\limits_{x\,\in\,[0,1]}x-x^2=\frac{1}{4}$,

$$var\,(\mathrm{H}(\mathrm{X}_i+\eta_i-x)\,)\;=\;\mathrm{E}\,[\,\mathrm{H}(\mathrm{X}_i+\eta_i-x)^2\,]-\mathrm{E}\,[\,\mathrm{H}(\mathrm{X}_i+\eta_i-x)\,]^2$$

$$=\;\mathrm{E}\,[\,\mathrm{H}(\mathrm{X}_i+\eta_i-x)\,]-\mathrm{E}\,[\,\mathrm{H}(\mathrm{X}_i+\eta_i-x)\,]^2$$

$$\leq\;\frac{1}{4}.$$

Furthermore, w_i is constant as a function of $\eta_1,\ldots,\eta_N$ and therefore, since the $\eta_1,\ldots,\eta_N$ are independent, $\mathrm{H}(\mathrm{X}_i+\eta_i-x)w_i$ is independent of $\mathrm{H}(\mathrm{X}_{i'}+\eta_{i'}-x)w_{i'}$ for $i\neq i'$. It follows that for each fixed $x\in\mathbb{R}$,

$$var\,(\widetilde{D}_{\Delta t}\,\tilde{v}^{\,j}(x)\,)\;=\;var\,(\sum_{i=1}^{N}\mathrm{H}(\mathrm{X}_i+\eta_i-x)w_i\;)$$

$$=\;\sum_{i=1}^{N}var\,(\mathrm{H}(\mathrm{X}_i+\eta_i-x)\,)w_i^2$$

$$\leq\;\frac{1}{4}N\max w_i^2.$$

Thus, if $\max w_i^2 \leq CN^{-2}$ (by Lemma 2.4 this is true (with $C = e^{2T}$) for any computed solution generated from initial data satisfying $w_i^0 = N^{-1}$), then for all $x \in \mathbb{R}$ and all $\alpha > 0$,

$$P_{\Omega_j}\left(\left| D_{\Delta t}\, \tilde{v}^j(x) - \tilde{D}_{\Delta t}\, \tilde{v}^j(x) \right| \geq \frac{\alpha}{\sqrt{N}} \right) \leq \frac{C}{4\alpha^2} \qquad \forall\, \alpha > 0.$$

Similar estimates have been derived by Hald [33] (see his Theorem 1 and the subsequent remark) and Roberts [47] (Theorem 4.3.1). Unfortunately, this inequality is too crude to use here, for we require an estimate which yields an $o(1)$ bound after adding the pointwise error at $O(N)$ points and then summing over $k = O(\Delta t^{-1})$ time steps. This problem has been considered by Roberts in [47]. His solution was to use the following inequality of exponential type due to Hoeffding [37]. Inequalities of this type may also be found in §19.1 of Loève [42].

LEMMA 7.2 (Hoeffding [37], p. 16) Let $Z_1, \ldots, Z_N$ be N independent random variables satisfying $0 \leq Z_i \leq 1$. Then for any $\alpha > 0$,

$$P\Big(\frac{1}{N}\sum_{i=1}^{N} Z_i - \frac{1}{N}\sum_{i=1}^{N} E\,[Z_i] \geq \alpha \Big) \leq e^{-2N\alpha^2}.$$

Applying Lemma 7.2 twice, once to the Z_i and once to the random variables $1 - Z_i$, and then using the fact that for all random variables X and Y

$$P(\,|X - Y|\, \geq \alpha) \leq P(X - Y \geq \alpha) + P(Y - X \geq \alpha),$$

we obtain the following more useful form of this inequality.

COROLLARY 7.3 Let $Z_1, \ldots, Z_N$ be N independent random variables satisfying $0 \leq Z_i \leq 1$. Then for any $\alpha > 0$,

$$P\Big(\left| \frac{1}{N}\sum_{i=1}^{N} Z_i - \frac{1}{N}\sum_{i=1}^{N} E\,[Z_i] \right| \geq \alpha \Big) \leq 2e^{-2N\alpha^2}.$$

By noting that for each fixed $x \in \mathbb{R}$, $Z_i = H(X_i + \eta_i - x)\, w_i\, \overline{w}^{-1}$ satisfies the hypotheses of Corollary 7.3 where $\overline{w} = \max w_i$, we can now derive an exponential bound on $P_{\Omega_j}(\,|\, D_{\Delta t}\, \tilde{v}^j(x) - \widetilde{D}_{\Delta t}\, \tilde{v}^j(x)\,| \geq \alpha N \overline{w}\,)$. As was mentioned in the introduction to this chapter our result differs slightly from Roberts ([47], Theorem 4.3.4) in that his particle strengths are constant over time whereas, in the random gradient method, they are altered at each time step. This is analogous to the difficulty one encounters when trying to generalize a convergence proof for the two dimensional vortex method to three dimensions. For, in the vortex method, particle strengths correspond to vorticity and vorticity is constant along particle paths in two dimensional inviscid flow but not in three dimensional inviscid flow ([3]). [1]

LEMMA 7.4 Let $\tilde{u}^0 \in \mathbf{S}$, let $\tilde{v}^j = \widetilde{R}_{\Delta t}(\widetilde{D}_{\Delta t}\, \widetilde{R}_{\Delta t})^{j-1}\tilde{u}^0$, and let $\overline{w} = \max_i w_i^j$. Then for any $x \in \mathbb{R}$ and $\alpha > 0$,

$$P_{\Omega_j}(\,|\, D_{\Delta t}\, \tilde{v}^j(x) - \widetilde{D}_{\Delta t}\, \tilde{v}^j(x)\,| \geq \alpha N \overline{w}\,) \leq 2e^{-2N\alpha^2}.$$

Remark : If $w_i^0 = O(N^{-1})$, then this estimate depends exclusively on the parameter α and the number of particles, N and the time $t = j\,\Delta t$. For, by Lemma 2.4, $N\overline{w}$ is $O(e^t)$ for any $\tilde{v}^j$ which has been generated by the random gradient method from initial data with particle strengths that are $O(N^{-1})$.

Proof:. By Corollary 2.3 $\tilde{v}^j \in \mathbf{S}$. Therefore, w_i satisfies $0 < w_i \leq 1$ for each i and hence, $0 < w_i \leq \overline{w}$. Thus, $0 < w_i\, \overline{w}^{-1} \leq 1$. Define

$$Z_i = H(X_i + \eta_i - x)\, w_i\, \overline{w}^{-1}.$$

[1]While it is doubtful that the work here will shed much light on the problem of proving convergence for vortex methods with variable particle strengths, one might achieve such a benefit by adapting Goodman's convergence proof for the two dimensional random vortex method [27] to a convergence proof for the random gradient method. One of the principal difficulties that would have to be overcome is how to apply the ideas from Goodman's proof to a method which has variable particle strengths. Since this is one of the problems that must be solved in order to prove the convergence of a three dimensional random vortex method, it's possible that one might gain some insight by solving this simpler problem first.

For each fixed $x \in \mathbb{R}$, the Z_i are N independent random variables satisfying $0 \leq Z_i \leq 1$. Furthermore,

$$\frac{1}{N\overline{w}} \widetilde{D}_{\Delta t}\, \tilde{v}^{\,j}(x) = \frac{1}{N} \sum_{i=1}^{N} Z_i$$

and hence, by Lemma 7.1,

$$\frac{1}{N\overline{w}} D_{\Delta t}\, \tilde{v}^{\,j}(x) = \frac{1}{N} \sum_{i=1}^{N} \mathrm{E}\,[Z_i].$$

It now follows from Corollary 7.3 that for any $x \in \mathbb{R}$,

$$\begin{aligned} P_{\Omega_j}\Big(\mid D_{\Delta t}\, \tilde{v}^{\,j}(x) - \widetilde{D}_{\Delta t}\, \tilde{v}^{\,j}(x) \mid > N\overline{w}\,\alpha \Big) &= P_{\Omega_j}\Big(\frac{1}{N\overline{w}} \mid D_{\Delta t}\, \tilde{v}^{\,j}(x) - \widetilde{D}_{\Delta t}\, \tilde{v}^{\,j}(x) \mid > \alpha \Big) \\ &= P_{\Omega_j}\Big(\mid \frac{1}{N}\sum_{i=1}^{N} \mathrm{E}\,[Z_i] - \frac{1}{N}\sum_{i=1}^{N} Z_i \mid > \alpha \Big) \\ &\leq 2\, e^{-2N\alpha^2}. \end{aligned}$$

7.4 The L^1 Convergence of $\widetilde{D}_{\Delta t}$ We will now use the pointwise estimates from §7.3 to derive an error bound in the L^1 norm. The proof is accomplished by dividing the real line into three intervals, $(-\infty,-L)$, $(-L,L)$, and (L,∞), and establishing bounds on each of these intervals separately. With the exception of the first result, Corollary 7.6, all bounds in this section are over the probability space Ω.

Our first task is to establish a bound for the L^1 error over the interval $(-L,L)$. To begin, we divide this interval into $N-1$ appropriately chosen subintervals, (a_{j-1}, a_j), Then we use the monotonicity[2] of the functions $D_{\Delta t}\, \tilde{v}^{\,j}$ and $\widetilde{D}_{\Delta t}\, \tilde{v}^{\,j}$, to demonstrate that, on each of these subin-

[2]As Roberts has pointed out ([47]), monotonicity is not necessary here.

tervals, $| D_{\Delta t}\, \tilde{v}^{\,j}(x) - \widetilde{D}_{\Delta t}\, \tilde{v}^{\,j}(x) |$ is no greater than $\sqrt{N}^{\,-1}$ plus the maximum value of the error at the endpoints of the subinterval. Hence, the L^1 error over (–L,L) is bounded by $2\mathrm{L}[\sqrt{N}^{\,-1}+\Theta]$, where Θ is the maximum value of the pointwise error at the a_j. Calculation of the L^1 error over (–L,L) thus reduces to estimating the size of N pointwise errors.

THEOREM 7.5 Let $\tilde{u}^0 \in \mathbf{S}$ and assume that $\tilde{u}^0$ is generated by N particles. Let $\tilde{v}^{\,j} = \widetilde{R}_{\Delta t}(\widetilde{D}_{\Delta t}\, \widetilde{R}_{\Delta t})^{j-1}\tilde{u}^0$ and set $\overline{w} = \max_i w_i^j$. Then for all real α, $\mathrm{L} > 0$,

$$P_{\Omega_j}\Big(\| D_{\Delta t}\, \tilde{v}^{\,j} - \widetilde{D}_{\Delta t}\, \tilde{v}^{\,j} \|_{(-\mathrm{L},\mathrm{L})} > 2\mathrm{L}[\frac{1}{\sqrt{N}} + \alpha N \overline{w}] \Big) \leq 2Ne^{-2N\alpha^2}. \tag{7.3}$$

Proof: For the duration of this proof we drop the superscript j from $\tilde{v}^{\,j}$ and write $\tilde{v}$. Observe that $D_{\Delta t}\, \tilde{v}$ and $\widetilde{D}_{\Delta t}\, \tilde{v}$ are monotonically decreasing functions bounded between 0 and 1. To see this, first note that by Corollary 2.3, $\widetilde{D}_{\Delta t}\, \tilde{v} \in \mathbf{S}$. Thus, $\widetilde{D}_{\Delta t}\, \tilde{v}$ is a monotone decreasing function of x satisfying $0 \leq \widetilde{D}_{\Delta t}\, \tilde{v} \leq 1$. Similarly, by writing $D_{\Delta t}\, \tilde{v}$ as the convolution of the heat kernel with $\tilde{v}$ and using the fact that $\tilde{v} \in \mathbf{S}$, it is an easy matter to show that $D_{\Delta t}\, \tilde{v}$ is monotone decreasing with $0 \leq D_{\Delta t}\, \tilde{v} \leq 1$.

Since $D_{\Delta t}\, \tilde{v}$ is monotone and $0 \leq D_{\Delta t}\, \tilde{v} \leq 1$ it follows that we can find a sequence $a_1, a_2, \ldots, a_N$ with $-\mathrm{L} = a_1 < a_2 < \ldots < a_N = \mathrm{L}$ such that

$$| D_{\Delta t}\, \tilde{v}(a_{j-1}) - D_{\Delta t}\, \tilde{v}(a_j) | < \frac{1}{\sqrt{N}} \qquad\qquad j = 2, \ldots, N.$$

(Of course we could use $(N-1)^{-1}$ here instead of $\sqrt{N}^{\,-1}$. But this turns out to be only a temporary gain, since this term will be added to terms of order $\sqrt{N}^{\,-1}$ below.) For each j let $\theta(a_j)$ denote the absolute value of the difference between $D_{\Delta t}\, \tilde{v}$ and $\widetilde{D}_{\Delta t}\, \tilde{v}$ at the point a_j,

$$\theta(a_j) = | D_{\Delta t}\, \tilde{v}(a_j) - \widetilde{D}_{\Delta t}\, \tilde{v}(a_j) | .$$

Now, since $D_{\Delta t}\,\tilde{v}$ and $\widetilde{D}_{\Delta t}\,\tilde{v}$ are monotone decreasing functions of x, it follows that for each $x \in (a_{j-1}, a_j)$,

$$\begin{aligned} D_{\Delta t}\,\tilde{v}(x) - \widetilde{D}_{\Delta t}\,\tilde{v}(x) &\le D_{\Delta t}\,\tilde{v}(a_{j-1}) - \widetilde{D}_{\Delta t}\,\tilde{v}(a_j) + D_{\Delta t}\,\tilde{v}(a_j) - D_{\Delta t}\,\tilde{v}(a_j) \\ &\le |\, D_{\Delta t}\,\tilde{v}(a_{j-1}) - D_{\Delta t}\,\tilde{v}(a_j)\,| \;+\; |\, D_{\Delta t}\,\tilde{v}(a_j) - \widetilde{D}_{\Delta t}\,\tilde{v}(a_j)\,| \\ &\le |\, D_{\Delta t}\,\tilde{v}(a_{j-1}) - D_{\Delta t}\,\tilde{v}(a_j)\,| \;+\; \theta(a_j) \\ &\le \frac{1}{\sqrt{N}} + \max\{\theta(a_{j-1}),\, \theta(a_j)\}. \end{aligned}$$

Similarly, one can show that

$$-\,(D_{\Delta t}\,\tilde{v}(x) - \widetilde{D}_{\Delta t}\,\tilde{v}(x)) \le \frac{1}{\sqrt{N}} + \max\{\theta(a_{j-1}),\, \theta(a_j)\}.$$

Hence, for $x \in (a_{j-1}, a_j)$,

$$|\, D_{\Delta t}\,\tilde{v}(x) - \widetilde{D}_{\Delta t}\,\tilde{v}(x)\,| \;\le\; \frac{1}{\sqrt{N}} + \max\{\theta(a_{j-1}),\, \theta(a_j)\}.$$

This yields the following estimate for the L^1 norm over the interval (-L,L):

$$\begin{aligned} \|\, D_{\Delta t}\,\tilde{v} - \widetilde{D}_{\Delta t}\,\tilde{v} \,\|_{(-L,L)} &= \sum_{j=2}^{N} \|\, D_{\Delta t}\,\tilde{v} - \widetilde{D}_{\Delta t}\,\tilde{v} \,\|_{(a_{j-1}, a_j)} \\ &\le \sum_{j=2}^{N} (a_j - a_{j-1})\Big[\frac{1}{\sqrt{N}} + \max\,\{\theta(a_{j-1}),\, \theta(a_j)\}\Big] \qquad (7.4) \\ &\le 2L\Big[\frac{1}{\sqrt{N}} + \max_j\, \theta(a_j)\Big]. \end{aligned}$$

The function $\Theta = \max_j\, \theta(a_j)$ is a random variable with which depends on $\eta_1, \ldots, \eta_N$. The pro-

bability that the error over the interval (-L,L) is greater than $2L[\sqrt{N}^{-1} + \alpha N\overline{w}]$ can be estimated in terms of the probability that $\Theta > \alpha N\overline{w}$. To see this, note that by (7.4),

$$\| D_{\Delta t}\,\tilde{v} - \widetilde{D}_{\Delta t}\,\tilde{v} \|_{(-L,L)} \geq 2L[\frac{1}{\sqrt{N}} + \alpha N\overline{w}] \Rightarrow 2L[\frac{1}{\sqrt{N}} + \Theta] \geq 2L[\frac{1}{\sqrt{N}} + \alpha N\overline{w}]$$

$$\Longleftrightarrow \Theta \geq \alpha N\overline{w}.$$

Therefore, since $\Theta \geq \alpha N\overline{w}$ implies $| D_{\Delta t}\,\tilde{v}(a_j) - \widetilde{D}_{\Delta t}\,\tilde{v}(a_j) | \geq \alpha N\overline{w}$ for some a_j, we can use Lemma 7.4 to obtain

$$P_{\Omega_j}\Big(\| D_{\Delta t}\,\tilde{v} - \widetilde{D}_{\Delta t}\,\tilde{v} \|_{(-L,L)} \geq 2L[\frac{1}{\sqrt{N}} + \alpha N\overline{w}] \Big) \leq P_{\Omega_j}\Big(\exists\, j : | D_{\Delta t}\,\tilde{v}(a_j) - \widetilde{D}_{\Delta t}\,\tilde{v}(a_j) | \geq \alpha N\overline{w} \Big)$$

$$\leq \sum_{j=1}^{N} P_{\Omega_j}\Big(| D_{\Delta t}\,\tilde{v}(a_j) - \widetilde{D}_{\Delta t}\,\tilde{v}(a_j) | \geq N\overline{w}\,\alpha \Big)$$

$$\leq 2Ne^{-2N\alpha^2}.$$

Theorem 7.5 holds for all possible $\tilde{v}^j$. In other words, if we regard $\tilde{v}^j$ as a function of $\omega_j^* = (\omega_1, \ldots, \omega_{j-1})$, then the bound in (7.3) holds for all $\omega_j^* \in \Omega_1 \times \ldots \times \Omega_{j-1}$. Therefore, using the argument outlined at the end of §7.2, we now obtain a bound over the probability space Ω.

COROLLARY 7.6 Let $\tilde{u}^0 \in \mathbf{S}$ and assume that $\tilde{u}^0$ is generated by N particles. Let $\tilde{v}^j = \widetilde{R}_{\Delta t}(\widetilde{D}_{\Delta t}\,\widetilde{R}_{\Delta t})^{j-1}\tilde{u}^0$ and set $\overline{w} = \max_i w_i^j$. Then for all real numbers α, L > 0,

$$P\Big(\| D_{\Delta t}\,\tilde{v}^j - \widetilde{D}_{\Delta t}\,\tilde{v}^j \|_{(-L,L)} > 2L[\frac{1}{\sqrt{N}} + \alpha N\overline{w}] \Big) \leq 2Ne^{-2N\alpha^2}. \tag{7.5}$$

The next step is to prove a probability inequality for the error in the L^1 norm over the tails, $(-\infty,-L)$ and (L,∞). Note that we are still free to choose L. Let K > 0 be given such

that at time $t = 0$ the particles all lie in (-K,K). Let $L = K + 2\beta$ where $\beta > 0$ is an arbitrary parameter. The idea, due to Roberts [47], is to estimate the error as a function of β under the assumption that at time $(j-1)\Delta t$ or $j\Delta t$ the particles remain in the interval $(-K-\beta, K+\beta)$. This reduces the problem to that of finding the probability that the particles are in this interval at the $(j-1)st$ and jth time steps, a problem which is easily solved using Lemmas 6.1 and 6.2.

THEOREM 7.7 Let $\tilde{u}^0 \in S$. Assume that $\tilde{u}^0$ is generated by N particles and that for some $K > 0$ all of these particles lie in the interval (-K,K). Let $\tilde{v}^j = \widetilde{R}_{\Delta t}(\widetilde{D}_{\Delta t}\widetilde{R}_{\Delta t})^{j-1}\tilde{u}^0$ and denote the time by $T = j\Delta t$. For any real number $\beta > 0$ set $L = K + 2\beta$. Then

$$P\Big(\| D_{\Delta t}\tilde{v}^j - \widetilde{D}_{\Delta t}\tilde{v}^j \|_{(-\infty,-L)} \geq \frac{\sqrt{\Delta t}}{\sqrt{\pi}} e^{\frac{-\beta^2}{4\Delta t}}\Big) \leq \frac{4N\sqrt{T}}{\beta\sqrt{\pi}} e^{\frac{-\beta^2}{4T}}, \tag{7.6a}$$

$$P\Big(\| D_{\Delta t}\tilde{v}^j - \widetilde{D}_{\Delta t}\tilde{v}^j \|_{(L,\infty)} \geq \frac{\sqrt{\Delta t}}{\sqrt{\pi}} e^{\frac{-\beta^2}{4\Delta t}}\Big) \leq \frac{4N\sqrt{T}}{\beta\sqrt{\pi}} e^{\frac{-\beta^2}{4T}}. \tag{7.6b}$$

Proof: We prove (7.6a); the proof of (7.6b) is similar. Let $B = K+\beta$. We begin by proving

$$\| D_{\Delta t}\tilde{v}^j - \widetilde{D}_{\Delta t}\tilde{v}^j \|_{(-\infty,-L)} \geq \frac{\sqrt{\Delta t}}{\sqrt{\pi}} e^{\frac{-\beta^2}{4\Delta t}} \Rightarrow \exists\, i \text{ such that } X_i^{j-1} \text{ or } X_i^j \notin (-B,B) \tag{7.7}$$

where X_i^{j-1} and X_i^j are the particle positions for $\tilde{v}^j$ and $\widetilde{D}_{\Delta t}\tilde{v}^j$ respectively. The probability that some X_i^{j-1} or $X_i^j \notin (-B,B)$ is then be estimated using Lemmas 6.1 and 6.2. The inequality in (7.6a) follows immediately. We prove the contrapositive of (7.7). Assume that

$$-B < X_i^{j-1}, X_i^j < B \qquad \forall\, i \in \{1, \ldots, N\}. \tag{7.8}$$

By the triangle inequality,

$$\| D_{\Delta t}\tilde{v}^j - \widetilde{D}_{\Delta t}\tilde{v}^j \|_{(-\infty,-L)} \leq \| D_{\Delta t}\tilde{v}^j - 1 \|_{(-\infty,-L)} + \| 1 - \widetilde{D}_{\Delta t}\tilde{v}^j \|_{(-\infty,-L)}.$$

Hence, it suffices to show

$$\| 1 - \widetilde{D}_{\Delta t}\, \tilde{v}^{\,j} \|_{(-\infty,-\mathrm{L})} = 0 \tag{7.9}$$

and

$$\| D_{\Delta t}\, \tilde{v}^{\,j} - 1 \|_{(-\infty,-\mathrm{L})} < \frac{\sqrt{\Delta t}}{\sqrt{\pi}} e^{\frac{-\beta^2}{4\Delta t}}. \tag{7.10}$$

The proof of (7.9) is easy. For, by (7.8), $x < -\mathrm{L} = -\mathrm{K}-2\beta < -\mathrm{B} = -\mathrm{K}-\beta$ implies $x < \mathrm{X}_i^j$ for all i. Therefore, for all $x < -\mathrm{L}$,

$$\widetilde{D}_{\Delta t}\, \tilde{v}^{\,j}(x) = \sum_{i=1}^{N} \mathrm{H}\,(\mathrm{X}_i^j - x)\, w_i^j = \sum_{i=1}^{N} w_i^j = 1$$

where we have used $\widetilde{D}_{\Delta t}\, \tilde{v}^{\,j} \in \mathbf{S}$ (Corollary 2.3). Therefore,

$$\| 1 - \widetilde{D}_{\Delta t}\, \tilde{v}^{\,j} \|_{(-\infty,-\mathrm{L})} = \int_{-\infty}^{-\mathrm{L}} | 1 - \widetilde{D}_{\Delta t}\, \tilde{v}^{\,j}(x) |\, dx = 0.$$

To prove (7.10) observe that (7.8) implies $-\mathrm{B} - x < \mathrm{X}_i^{j-1} - x$ for all $x \in \mathbb{R}$. Thus, since $D_{\Delta t}\,\mathrm{H}\,(x)$ is a monotonically increasing function of x that is bounded above by 1,

$$D_{\Delta t}\,\mathrm{H}\,(-\mathrm{B} - x) \le D_{\Delta t}\,\mathrm{H}\,(\mathrm{X}_i^{j-1} - x) \le 1, \qquad i = 1, \ldots, N.$$

Multiply each side of this inequality by w_i^j and, using the fact that $D_{\Delta t}$ is linear and $\sum w_i^j = 1$, sum over $i = 1, \ldots, N$ to obtain

$$D_{\Delta t}\,\mathrm{H}\,(-\mathrm{B} - x) \le D_{\Delta t}\, \tilde{v}^{\,j}(x) \le 1.$$

Hence, for all $x \in \mathbb{R}$,

$$0 \le 1 - D_{\Delta t}\, \tilde{v}^{\,j}(x) \le 1 - D_{\Delta t}\,\mathrm{H}\,(-\mathrm{B} - x).$$

Since $1 - D_{\Delta t}\,\mathrm{H}\,(-x) = D_{\Delta t}\,(1 - \mathrm{H}\,(-x))$ and $1 - \mathrm{H}\,(-x) = \mathrm{H}\,(x)$ (when $x \neq 0$), we can

integrate over $(-\infty,-L)$ to obtain

$$\| 1 - D_{\Delta t}\, \tilde{v}^{\,j} \|_{(-\infty,-L)} = \int_{-\infty}^{-L} | 1 - D_{\Delta t}\, \tilde{v}^{\,j}(x) | \, dx$$

$$\leq \int_{-\infty}^{-L} | 1 - D_{\Delta t}\, \mathrm{H}(-B - x) | \, dx$$

$$= \int_{-\infty}^{-\beta} | 1 - D_{\Delta t}\, \mathrm{H}(-s) | \, ds$$

$$= \int_{-\infty}^{-\beta} | D_{\Delta t}\, \mathrm{H}(s) | \, ds$$

$$= \| D_{\Delta t}\, \mathrm{H} \|_{(-\infty,-\beta)}.$$

In order to estimate this last quantity, recall the function ϕ defined by 6.2 and note that $D_{\Delta t}\, \mathrm{H}(x) = G_{\Delta t} * \mathrm{H}(x) = \phi(\frac{x}{\sqrt{2\Delta t}})$. Hence,

$$\| D_{\Delta t}\, \mathrm{H} \|_{(-\infty,-\beta)} = \int_{-\infty}^{-\beta} \int_{-\infty}^{\frac{x}{\sqrt{2\Delta t}}} \frac{e^{\frac{-s^2}{2}}}{\sqrt{2\pi}} \, ds \; dx$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{-\beta}{\sqrt{2\Delta t}}} \int_{s\sqrt{2\Delta t}}^{-\beta} e^{\frac{-s^2}{2}} \, dx \; ds$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{-\beta}{\sqrt{2\Delta t}}} (-\beta - s\sqrt{2\Delta t}\,) e^{\frac{-s^2}{2}} \, ds$$

$$< \frac{\sqrt{2\Delta t}}{\sqrt{2\pi}} \int_{-\infty}^{\frac{-\beta}{\sqrt{2\Delta t}}} -s e^{\frac{-s^2}{2}} \, ds$$

$$= \frac{\sqrt{\Delta t}}{\sqrt{\pi}} e^{\frac{-\beta^2}{4\Delta t}}.$$

Thus, we have shown

$$\| D_{\Delta t}\, \tilde{v}^{\,j} - 1 \|_{(-\infty,-L)} \leq \| D_{\Delta t} \mathrm{H} \|_{(-\infty,-\beta)} < \frac{\sqrt{\Delta t}}{\sqrt{\pi}} e^{\frac{-\beta^2}{4\Delta t}}.$$

This proves (7.10) and hence, (7.7). All that remains is to estimate the probability that the right hand side of (7.7) is true. By Lemmas 6.1 and 6.2,

$$P(\ \exists\, i\ : \mathrm{X}_i^{j-1} \text{ or } \mathrm{X}_i^{j} \notin (-\mathrm{B},\mathrm{B})\ ) \leq \sum_{i=1}^{N} [\ P(\mathrm{X}_i^{j-1} < -\mathrm{K}-\beta\ ) + P(\mathrm{X}_i^{j-1} > \mathrm{K}+\beta\ )\ ]$$

$$+ \sum_{i=1}^{N} [\ P(\mathrm{X}_i^{j} < -\mathrm{K}-\beta\ ) + P(\mathrm{X}_i^{j} > \mathrm{K}+\beta\ )\ ]$$

$$\leq 2N\, \phi(\frac{-\beta}{\sqrt{2(j-1)\Delta t}}) + 2N\, \phi(\frac{-\beta}{\sqrt{2j\,\Delta t}})$$

$$\leq \frac{4N\sqrt{j\,\Delta t}}{\beta\sqrt{\pi}} e^{\frac{-\beta^2}{4j\,\Delta t}}$$

where we have used the fact that $\phi(x)$ is an increasing function of x to bound $\phi(\frac{-\beta}{\sqrt{2(j-1)\Delta t}})$ by $\phi(\frac{-\beta}{\sqrt{2j\,\Delta t}})$.

Combining Corollary 7.6, Theorem 7.7, and an appropriate choice of the parameters α and β yields the main result of this chapter, a probability inequality for the error in the L^1 norm.

THEOREM 7.8 Let $\tilde{u}^0 \in \mathbf{S}$ be generated by $N \geq 3$ particles, each with weight $w_i^0 = N^{-1}$, and assume that for some $\mathrm{K} > 0$ all of the particles lie in the interval (-K,K). Let

$\tilde{v}^j = \widetilde{R}_{\Delta t}(\widetilde{D}_{\Delta t}\widetilde{R}_{\Delta t})^{j-1}\tilde{u}^0$ and denote the time by $T = j\,\Delta t$. Then for any real number $\gamma \geq 1$,

$$P(\ \| D_{\Delta t}\tilde{v}^j - \widetilde{D}_{\Delta t}\tilde{v}^j \|_{L^1} \ \geq\ \gamma\,\mathrm{C}\,\frac{\ln N}{\sqrt{N}}\ ) \ \leq\ 4N^{1-\frac{2}{4}\gamma} \tag{7.11}$$

where C is given by

$$\mathrm{C} = 2[\,(\mathrm{K} + 6\sqrt{T}\,)(1 + 2e^{T}) + \frac{\sqrt{\Delta t}}{\sqrt{\pi}}\,]. \tag{7.12}$$

Proof: Fix $\gamma \in \mathbb{R}$ with $\gamma \geq 1$ and let $\alpha = 3\sqrt{\gamma \ln N}\,/\sqrt{8N}$, $\beta = 3\sqrt{\gamma T \ln N}$, and $\mathrm{L} = \mathrm{K} + 2\beta$. Referring now to equation (7.5), we find

$$2Ne^{-2N\alpha^2} = 2Ne^{-\frac{2}{4}\gamma \ln N} = 2N^{1-\frac{2}{4}\gamma}.$$

Furthermore, since γ and $\ln N$ are both ≥ 1,

$$2\,\mathrm{L}\,[\,\frac{1}{\sqrt{N}} + \alpha N\overline{w}\,] \ =\ 2\,(\mathrm{K} + 6\sqrt{\gamma T}\ \sqrt{\ln N}\ )\,(\frac{1}{\sqrt{N}} + \frac{3\sqrt{\gamma}}{\sqrt{8}}\frac{\sqrt{\ln N}}{\sqrt{N}}N\overline{w}\ )$$

$$\leq\ 2\gamma\,(\mathrm{K} + 6\sqrt{T}\ )\,(1 + 2e^{T}\,)\frac{\ln N}{\sqrt{N}}$$

where we have used Lemma 2.4, together with $w_i^0 = N^{-1}$, to deduce that $N\overline{w} \leq e^{T}$. Consequently, (7.5) becomes

$$P\Big(\ \| D_{\Delta t}\tilde{v}^j - \widetilde{D}_{\Delta t}\tilde{v}^j \ \|_{(-\mathrm{L},\mathrm{L})} \geq 2\gamma(\mathrm{K} + 6\sqrt{T}\,)(1 + 2e^{T})\frac{\ln N}{\sqrt{N}}\ \Big) \ \leq\ 2N^{1-\frac{2}{4}\gamma}. \tag{7.13}$$

We can derive similar estimates for the error over the tails, $(-\infty,-\mathrm{L})$ and (L,∞). With our choice of α and β the right hand side of (7.6a,b) becomes

$$\frac{4N\sqrt{T}}{\beta\sqrt{\pi}}e^{-\frac{\beta^2}{4T}} = \frac{4N}{3\sqrt{\gamma\pi}}\frac{N^{-\frac{\theta}{4}\gamma}}{\sqrt{\ln N}} \leq N^{1-\frac{\theta}{4}\gamma}$$

since $4/3\sqrt{\pi} \leq 1$. To estimate the other quantity that appears in (7.6a,b) use $j \geq 1$ and $\gamma \geq 1$ to write

$$\frac{\sqrt{\Delta t}}{\sqrt{\pi}}e^{-\frac{\beta^2}{4\Delta t}} = \frac{\sqrt{\Delta t}}{\sqrt{\pi}}N^{-\frac{\theta}{4}j\gamma} \leq \gamma\frac{\sqrt{\Delta t}}{\sqrt{\pi}}\frac{1}{\sqrt{N}}.$$

Substituting these two inequalities into (7.6a,b) yields

$$P\Big(\| D_{\Delta t}\,\tilde{v}^{\,j} - \widetilde{D}_{\Delta t}\,\tilde{v}^{\,j} \|_{(-\infty,-L)} \geq \gamma\frac{\sqrt{\Delta t}}{\sqrt{\pi}}\frac{1}{\sqrt{N}}\Big) \leq N^{1-\frac{\theta}{4}\gamma}$$

$$P\Big(\| D_{\Delta t}\,\tilde{v}^{\,j} - \widetilde{D}_{\Delta t}\,\tilde{v}^{\,j} \|_{(L,\infty)} \geq \gamma\frac{\sqrt{\Delta t}}{\sqrt{\pi}}\frac{1}{\sqrt{N}}\Big) \leq N^{1-\frac{\theta}{4}\gamma}.$$

Combining these estimates with (7.13) above yields (7.11).

8. Convergence of the Method

We now use the results from the previous chapters to prove the convergence of the random gradient method. The main result of this chapter is the following. If the hypotheses listed in §8.1 hold, then for any real number $\gamma \geq 1$,

$$P\left(\| F_{\Delta t}^{k} u^0 - (\widetilde{D}_{\Delta t} \widetilde{R}_{\Delta t})^k \tilde{u}^0 \|_1 \geq \gamma [e^T \| u^0 - \tilde{u}^0 \|_1 + C_1 \Delta t + C_2 \frac{\ln N}{\sqrt[4]{N}}] \right) \leq 5\, T\, N^{\frac{5}{4} - \frac{9}{4}\gamma} \tag{8.1}$$

where the constants C_1 and C_2 depend only u^0, $\tilde{u}^0$, and the time, $T = k\,\Delta t$. The most important hypothesis here is that $\Delta t = C_0/\sqrt[4]{N}$ for some constant C_0. This has the effect of balancing the error due to the time step (temporal discretization) with the error due to the number of particles (spatial discretization).

This inequality is useful for a couple of reasons. On the one hand, it tells us that the probability of the error being greater than $\gamma [e^T \| u^0 - \tilde{u}^0 \|_1 + C_1 \Delta t + C_2 \frac{\ln N}{\sqrt[4]{N}}]$ decreases exponentially as a function of γ. This allows us to find bounds for the expected value and the variance of the error:

$$E\left[\| F_{\Delta t}^{k} u^0 - (\widetilde{D}_{\Delta t} \widetilde{R}_{\Delta t})^k \tilde{u}^0 \|_1 \right] \leq (1 + T) [e^T \| u^0 - \tilde{u}^0 \|_1 + C_1 \Delta t + C_2 \frac{\ln N}{\sqrt[4]{N}}],$$

$$var\left(\| F_{\Delta t}^{k} u^0 - (\widetilde{D}_{\Delta t} \widetilde{R}_{\Delta t})^k \tilde{u}^0 \|_1 \right) \leq (1 + T) [e^T \| u^0 - \tilde{u}^0 \|_1 + C_1 \Delta t + C_2 \frac{\ln N}{\sqrt[4]{N}}]^2.$$

Both of these estimates follow from the fact that for any random variable $Z \geq 0$ and any real number $a > 0$,

$$E[Z] \leq a\left(1 + \sum_{r=1}^{\infty} P(Z \geq ra)\right).$$

On the other hand, when $\gamma = 1$, the inequality in (8.1) implies that

$$P\left(\| F^k_{\Delta t} u^0 - (\widetilde{D}_{\Delta t} \widetilde{R}_{\Delta t})^k \tilde{u}^0 \|_1 < [e^T \| u^0 - \tilde{u}^0 \|_1 + C_1 \Delta t + C_2 \frac{\ln N}{\sqrt[4]{N}}] \right) \geq 1 - \frac{5T}{N}. \quad (8.2)$$

This estimate may be used to guide our choice of N. For example, say $T = 1$. Then, if we wish to obtain accuracy of order ϵ, it suffices to choose N so that

$$e^T \| u^0 - \tilde{u}^0 \|_1 + C_1 \frac{1}{\sqrt[4]{N}} + C_2 \frac{\ln N}{\sqrt[4]{N}} = \epsilon. \quad (8.3)$$

(More particles results in a better initial error.) Now note that the right hand side of (8.2) is an increasing function of N and that for $N = 1000$,

$$1 - 5T N^{-1} = \frac{995}{1000}.$$

Thus, if $N \geq 1000$ is chosen so that (8.3) holds, then the probability inequality in (8.2) assures us that better than 99% of the time

$$\| F^k_{\Delta t} u^0 - (\widetilde{D}_{\Delta t} \widetilde{R}_{\Delta t})^k \tilde{u}^0 \|_1 < \epsilon.$$

The estimate in (8.2) is probably not sharp. In fact, in Chapter 9 we present numerical experiments with an exact traveling wave solution for which the error tends to 0 like $1/\sqrt{N}$ instead of $\ln N / \sqrt[4]{N}$.

8.1. The Hypotheses Throughout this chapter let $T = k\,\Delta t$ denote the time at which we wish to compare the computed solution to the exact solution. We shall assume that the following hypotheses hold:

Hypothesis A_1: In addition to (1.3c-e) the exact initial data, u^0, satisfies $u^0 \in C^1(\mathbb{R})$, $0 \leq u^0(x) \leq 1$ for all $x \in \mathbb{R}$, and $u^0_x \in L^1(\mathbb{R}) \cap L^\infty(\mathbb{R})$.

Hypothesis A_2: The approximate initial data, $\tilde{u}^0$, satisfies $\tilde{u}^0 \in \mathbf{S}$ (see §2.3), $\tilde{u}^0$ is generated by $N \geq 10$ particles, and the initial weights satisfy $w_i^0 = N^{-1}$.

Hypothesis A_3: The computational parameters N and Δt have been chosen so that for some constant C_0 we have $\Delta t = C_0/\sqrt[4]{N}$.

We also assume that the constant $K > 0$ has been chosen so that the variation of $\tilde{u}^0$ lies in (-K,K),

$$|X_i^0| < K \qquad i = 1, \ldots, N. \tag{8.4}$$

8.2. A Bound on the Probability Distribution of the Error We prove the probability inequality stated at the beginning of this chapter in two steps. In the first step (Theorem 8.1) we use the estimates from Chapters 6 and 7 to establish a probability inequality for the error due to the approximate operators, $\|(D_{\Delta t} R_{\Delta t})^k \tilde{u}^0 - (\tilde{D}_{\Delta t} \tilde{R}_{\Delta t})^k \tilde{u}^0\|_1$. In the second step (Theorem 8.2) we use the (deterministic) bounds from Chapters 4 and 5 to control the remaining sources of error, $\|F_{\Delta t}^k u^0 - (D_{\Delta t} R_{\Delta t})^k u^0\|_1$ and $\|(D_{\Delta t} R_{\Delta t})^k u^0 - (D_{\Delta t} R_{\Delta t})^k \tilde{u}^0\|_1$.

THEOREM 8.1 Assume that hypotheses A_1, A_2, and A_3 hold. Let

$$C_2 = \{B_1 C_0^2 + B_2\} \frac{Te^T}{C_0} \tag{8.5}$$

where $B_1 = \frac{\sqrt{3}}{9}(K + 3\sqrt{T})$, $B_2 = 2\left[(K + 6\sqrt{T})(1 + 2e^T) + \frac{\sqrt{\Delta t}}{\sqrt{\pi}}\right]$, and C_0 is given by hypothesis A_3. Then for any real number $\gamma \geq 1$,

$$P\left(\|(D_{\Delta t} R_{\Delta t})^k \tilde{u}^0 - (\tilde{D}_{\Delta t} \tilde{R}_{\Delta t})^k \tilde{u}^0\|_1 \geq \gamma C_2 \frac{\ln N}{\sqrt[4]{N}} \right) \leq 5 T N^{1 - \frac{9}{4}\gamma}.$$

Proof: Let $Z_j = \|(D_{\Delta t} R_{\Delta t})^j \tilde{u}^0 - (\tilde{D}_{\Delta t} \tilde{R}_{\Delta t})^j \tilde{u}^0\|_1$, $V_j = \|R_{\Delta t} \tilde{u}^{j-1} - \tilde{R}_{\Delta t} \tilde{u}^{j-1}\|_1$, and

$W_j = \| D_{\Delta t}\, \tilde{v}^{\,j} - \widetilde{D}_{\Delta t}\, \tilde{v}^{\,j} \|_1$ where $\tilde{v}^{\,j}$ is as in Chapter 7. From equation (7.2) we have

$$Z_k \;\le\; e^{T} \sum_{j=1}^{k} (V_j \;+\; W_j\,).$$

We have $Z_0 = 0$ and thus, since $k = T\sqrt[4]{N}/C_0$ and $(\Delta t)^2 = C_0^2/\sqrt{N}$,

$$Z_k \;\ge \gamma\, C_2 \frac{\ln N}{\sqrt[4]{N}} \;\Rightarrow\; \sum_{j=1}^{k} (V_j + W_j\,) \ge \frac{\gamma C_2}{e^{T}} \frac{\ln N}{\sqrt[4]{N}}$$

$$\Rightarrow\; \exists\, j \text{ such that } (V_j + W_j\,) \;\ge\; \frac{\gamma C_2}{k e^{T}} \frac{\ln N}{\sqrt[4]{N}} \;=\; \frac{\gamma C_0 C_2}{T e^{T}} \frac{\ln N}{\sqrt{N}}$$

$$\Rightarrow\; \exists\, j \text{ such that } V_j \;\ge\; \gamma B_1 (\ln N)(\Delta t)^2 \text{ or } W_j \;\ge\; \gamma B_2 \frac{\ln N}{\sqrt{N}}.$$

We can now apply Corollary 6.4 and Theorem 7.8 to obtain

$$P\left(Z_k \;\ge\; \gamma C_2 \frac{\ln N}{\sqrt[4]{N}} \right) \le \sum_{j=1}^{k} P\left(V_j \ge \gamma\, B_1 \ln N (\Delta t)^2 \right) + \sum_{j=1}^{k} P\left(W_j \ge \gamma\, B_2 \frac{\ln N}{\sqrt{N}} \right)$$

$$\le\; k\; N^{1-\frac{9}{4}\gamma^2} \;+\; k\; 4\, N^{1-\frac{9}{4}\gamma}$$

$$\le\; 5k\, N^{1-\frac{9}{4}\gamma}$$

$$=\; 5\, T\, N^{\frac{5}{4}-\frac{9}{4}\gamma}.$$

Using Theorem 8.1 and the bounds from Theorems 4.6 and 5.9, we will now derive the essential result of this paper, a bound on the probability distribution of the error, $\| F_{\Delta t}^{k}\, u^0 - (\widetilde{D}_{\Delta t}\, \widetilde{R}_{\Delta t}^{\,-})^k\, \tilde{u}^{\,0} \|_1$.

THEOREM 8.2 Assume that hypotheses A_1, A_2, and A_3 hold. Let C_1 be given by (5.16) and

C_2 by (8.5). Then for all real $\gamma \geq 1$,

$$P\left(\| F^k_{\Delta t} u^0 - (\widetilde{D}_{\Delta t} \widetilde{R}_{\Delta t})^k \tilde{u}^0 \|_1 \geq \gamma \left[e^T \| u^0 - \tilde{u}^0 \|_1 + C_1 \Delta t + C_2 \frac{\ln N}{\sqrt[4]{N}} \right] \right) \leq 5\, T\, N^{\frac{5}{4} - \frac{9}{4}\gamma}.$$

Proof: Applying Theorems 4.6 and 5.9 to equation (3.6) we see that

$$\| F^k_{\Delta t} u^0 - (\widetilde{D}_{\Delta t} \widetilde{R}_{\Delta t})^k \tilde{u}^0 \|_1 \leq C_1 \Delta t + e^T \| u^0 - \tilde{u}^0 \|_1 + \| (D_{\Delta t} R_{\Delta t})^k \tilde{u}^0 - (\widetilde{D}_{\Delta t} \widetilde{R}_{\Delta t})^k \tilde{u}^0 \|_1.$$

Thus, since $\gamma \geq 1$,

$$\| F^k_{\Delta t} u^0 - (\widetilde{D}_{\Delta t} \widetilde{R}_{\Delta t})^k \tilde{u}^0 \|_1 \geq \gamma \left(e^T \| u^0 - \tilde{u}^0 \|_1 + C_1 \Delta t + C_2 \frac{\ln N}{\sqrt[4]{N}} \right)$$

implies

$$\| (D_{\Delta t} R_{\Delta t})^k \tilde{u}^0 - (\widetilde{D}_{\Delta t} \widetilde{R}_{\Delta t})^k \tilde{u}^0 \|_1 \geq \gamma\, C_2 \frac{\ln N}{\sqrt[4]{N}}.$$

It now follows from Theorem 8.1 that

$$\begin{aligned} P\left(\| F^k_{\Delta t} u^0 - (\widetilde{D}_{\Delta t} \widetilde{R}_{\Delta t})^k \tilde{u}^0 \|_1 \geq \gamma \left(e^T \| u^0 - \tilde{u}^0 \|_1 + C_1 \Delta t + C_2 \frac{\ln N}{\sqrt[4]{N}} \right) \right) \\ \leq P\left(\| (D_{\Delta t} R_{\Delta t})^k \tilde{u}^0 - (\widetilde{D}_{\Delta t} \widetilde{R}_{\Delta t})^k \tilde{u}^0 \|_1 \geq \gamma\, C_2 \frac{\ln N}{\sqrt[4]{N}} \right) \\ \leq 5\, T\, N^{\frac{5}{4} - \frac{9}{4}\gamma}. \end{aligned}$$

8.3. The Expected Value and Variance of the Error In addition to telling us how large the error is likely to be, the bound in Theorem 8.2 also allows us to estimate the expected value and the variance of the error. To do this we will need the following lemma, a variant of which may be found on page 41 of Chung [16].

LEMMA 8.3 Let $Z \geq 0$ be a random variable with $E[Z] < \infty$ and let $a > 0$ be an arbitrary real number. Then

$$E[Z] \leq a \left[1 + \sum_{r=1}^{\infty} P(Z \geq ra) \right].$$

Proof: Since $E[Z]$ is finite we can write

$$E[Z] \leq \sum_{r=0}^{\infty} (r+1)\, a\, P(ra \leq Z < (r+1)a)$$

$$= a \sum_{r=0}^{\infty} (r+1) \left[P(Z \geq ra) - P(Z \geq (r+1)a) \right]$$

$$= a \sum_{r=0}^{\infty} P(Z \geq ra)$$

$$\leq a \left[1 + \sum_{r=1}^{\infty} P(Z \geq ra) \right].$$

The last inequality follows from the fact that $P(Z \geq ra) = P(Z \geq 0) = 1$ when $r = 0$.

We will now derive a bound on the expected value of the error by setting $\gamma = 1, 2, \ldots$ in Theorem 8.2 and applying Lemma 8.3.

THEOREM 8.4: Assume that hypotheses A_1, A_2, and A_3 hold. Let C_1 be given by (5.16) and C_2 by (8.5). Then

$$E\left[\| F_{\Delta t}^{k} u^0 - (\widetilde{D}_{\Delta t} \widetilde{R}_{\Delta t})^k \tilde{u}^0 \|_1 \right] \leq (1 + T) \left[e^T \| u^0 - \tilde{u}^0 \|_1 + C_1 \Delta t + C_2 \frac{\ln N}{\sqrt[4]{N}} \right]. \quad (8.6)$$

Proof: Let

$$g(\Delta t, N) = e^{T} \| u^0 - \tilde{u}^0 \|_1 + C_1 \Delta t + C_2 \frac{\ln N}{\sqrt[4]{N}}. \tag{8.7}$$

Theorem 8.2 and Lemma 8.3 together imply that

$$E\left[\| F_{\Delta t}^k u^0 - (\widetilde{D}_{\Delta t} \widetilde{R}_{\Delta t})^k \tilde{u}^0 \|_1 \right] \leq g(\Delta t, N)\left(1 + 5\, T\, N^{\frac{5}{4}} \sum_{r=1}^{\infty} N^{-\frac{\varrho}{4} r}\right).$$

Note that $\sum_{r=1}^{\infty} N^{-\frac{\varrho}{4} r}$ is a ratio series with ratio $N^{-\frac{\varrho}{4}} < 1$. Therefore, since $N \geq 10$ (assumption A_2), we have

$$5\, T\, N^{\frac{5}{4}} \sum_{r=1}^{\infty} N^{-\frac{\varrho}{4} r} = 5\, T\, N^{\frac{5}{4}} \frac{N^{-\frac{\varrho}{4}}}{1 - N^{-\frac{\varrho}{4}}} = \frac{5\, T\, N^{\frac{5}{4}}}{N^{\frac{\varrho}{4}} - 1} \leq T.$$

The theorem follows immediately.

The final result in this section is a bound on the variance of the error. The proof, which is very similar to the proof of Theorem 8.4, is based on the following two points:

1) If Z is a random variable, then

$$var\,(Z) = E[Z^2] - E[Z]^2 \leq E[Z^2].$$

2) For all random variables $Z \geq 0$ and all real numbers a,

$$P(Z^2 \geq a^2) = P(Z \geq a).$$

THEOREM 8.5: Assume that hypotheses A_1, A_2, and A_3 hold. Let C_1 be given by (5.16) and C_2 by (8.5). Then

$$var\left(\| F_{\Delta t}^{k} u^0 - (\widetilde{D}_{\Delta t} \widetilde{R}_{\Delta t})^k \tilde{u}^0 \|_1 \right) \leq (1+T)\left[e^T \| u^0 - \tilde{u}^0 \|_1 + C_1 \Delta t + C_2 \frac{\ln N}{\sqrt[4]{N}} \right]^2. \quad (8.8)$$

Proof: Set $Z = \| F_{\Delta t}^{k} u^0 - (\widetilde{D}_{\Delta t} \widetilde{R}_{\Delta t})^k \tilde{u}^0 \|_1$. As noted above, it suffices to bound $E[Z^2]$. Let $g(\Delta t, N)$ be defined by (8.7). From Theorem 8.2 we have

$$P(Z^2 \geq \gamma^2 g(\Delta t, N)^2) = P(Z \geq \gamma g(\Delta t, N)) \leq 5 T N^{\frac{5}{4} - \frac{9}{4}\gamma}.$$

Setting $\gamma = \sqrt{r}$ for each $r \in \mathbb{Z}^+ = \{1, 2, \ldots\}$ and applying Lemma 8.3 we find

$$E[Z^2] \leq g(\Delta t, N)^2 \left(1 + 5 T N^{\frac{5}{4}} \sum_{r=1}^{\infty} N^{-\frac{9}{4}\sqrt{r}}\right).$$

All that remains is to estimate $\sum_{r=1}^{\infty} N^{-\frac{9}{4}\sqrt{r}}$. To accomplish this set $b = N^{-\frac{9}{4}}$, $\beta = -\ln b$ and note that $\beta > 0$. Then

$$\begin{aligned}
\sum_{r=1}^{\infty} b^{\sqrt{r}} &\leq b + \int_1^{\infty} e^{-\beta\sqrt{x}}\, dx \\
&= b - \left[\left(\frac{2}{\beta^2} + \frac{2\sqrt{x}}{\beta}\right) e^{-\beta\sqrt{x}} \right]_1^{\infty} \\
&= b + \left(\frac{2}{\beta^2} + \frac{2}{\beta}\right) e^{-\beta} \\
&\leq 2 N^{-\frac{9}{4}}
\end{aligned}$$

where we have used the hypothesis that $N \geq 10$ and hence, $\ln N > 2$. Thus,

$$E[Z^2] \leq g(\Delta t, N)^2 (1 + 10 T N^{-1}) \leq (1+T) g(\Delta t, N)^2.$$

8.4. Dependence of the Error on Arbitrary ν We will now remove the restriction $\nu = 1$. For arbitrary $\nu \leq 1$ let u_ν be the solution of equation (1.3a,b) with initial data u^0_ν. Define

$$u(x,t) \equiv u_\nu(\sqrt{\nu}x, t).$$

Then u satisfies (1.3a,b) with diffusion coefficient 1 and initial data $u^0(x) \equiv u^0_\nu(\sqrt{\nu}x)$. Note that

$$\| \partial_x u \|_\infty = \sqrt{\nu} \, \| \partial_x u_\nu \|_\infty \tag{8.9}$$

and

$$\| \partial_x u \|_1 = \| \partial_x u_\nu \|_1. \tag{8.10}$$

The random gradient method scales in the same manner. In other words, let $\tilde{u}_\nu$ be the random gradient solution of (1.3a,b) with diffusion coefficient ν at time $T = k\,\Delta t$. Denote the initial particle positions by $X^0_i(\nu)$, $i = 1, \ldots, N$. Then for any $k \geq 0$,

$$\tilde{u}^k(x) \equiv \tilde{u}^k_\nu(\sqrt{\nu}x)$$

is the random gradient solution of (1.3a,b) with diffusion coefficient 1 and initial particle positions

$$X^0_i = \frac{1}{\sqrt{\nu}} X^0_i(\nu).$$

This statement follows immediately from the fact that if η is a Gaussian distributed random variable with variance $2\nu\Delta t$ then $\eta/\sqrt{\nu}$ is a Gaussian random variable with variance $2\Delta t$.

Thus we have

$$\| u_\nu(T) - \tilde{u}^k_\nu \|_1 = \sqrt{\nu} \, \| u(T) - \tilde{u}^k \|_1$$

and using (8.6) we find

$$\mathrm{E}\ \| u_\nu(T) - \tilde{u}_\nu^k \|_1 = \sqrt{\nu}\, \mathrm{E}\ \| u(T) - \tilde{u}^k \|_1$$

$$\leq \sqrt{\nu}\,(1 + T)\,[\ e^T\ \| u^0 - \tilde{u}^0 \|_1 + \mathrm{C}_1\, \Delta t + \mathrm{C}_2 \frac{\ln N}{\sqrt[4]{N}}\ ].$$

It is necessary to investigate the dependence of the constants C_1 and C_2 on ν. From (5.16), (8.9) and (8.10) we have

$$\mathrm{C}_1 = T\ e^{3T}\ \{e^T\ \| \partial_x u^0 \|_\infty + \frac{4\sqrt{2\Delta t}}{\sqrt{\pi}}\}\ \| \partial_x u^0 \|_1$$

$$= T\ e^{3T}\ \{\sqrt{\nu}\, e^T\ \| \partial_x u_\nu^0 \|_\infty + \frac{4\sqrt{2\Delta t}}{\sqrt{\pi}}\}\ \| \partial_x u_\nu^0 \|_1 \tag{8.11}$$

Thus, the splitting error is $O\,(\sqrt{\nu})$. Note that if one is modeling a wave front of the form

$$g_\nu(x) \equiv g\,(x\,/\sqrt{\nu}) \tag{8.12}$$

for some arbitrary C^1 function g, then even though $\partial_x g_\nu = O\,(\nu^{-1/2})$ the constant C_1 remains $O\,(1)$ due to the factor $\sqrt{\nu}$ multiplying $\| \partial_x u_\nu^0 \|_\infty$ in (8.11) above.

In order to examine the dependence of C_2 on ν let K^ν be chosen so that $-\mathrm{K}^\nu \leq \mathrm{X}_i^0(\nu) \leq \mathrm{K}^\nu$ for all i. Then $\mathrm{K} = \mathrm{K}^\nu/\sqrt{\nu}$ satisfies (8.4) and we find

$$\mathrm{C}_2 = \frac{\sqrt{3}}{9}\,(\mathrm{K} + 3\sqrt{T}\,) + 2\,[\,(\mathrm{K} + 6\sqrt{T}\,)\,(1 + 2e^T) + \frac{\sqrt{\Delta t}}{\sqrt{\pi}}\,]$$

$$= \frac{\sqrt{3}}{9}\,(\frac{\mathrm{K}^\nu}{\sqrt{\nu}} + 3\sqrt{T}\,) + 2\,[\,(\frac{\mathrm{K}^\nu}{\sqrt{\nu}} + 6\sqrt{T}\,)\,(1 + 2e^T) + \frac{\sqrt{\Delta t}}{\sqrt{\pi}}\,].$$

Hence, $\mathrm{C}_2' = \sqrt{\nu}\,\mathrm{C}_2$ is bounded uniformly in ν for $\nu \leq 1$ and we have

$$\mathrm{E}\,\| u_\nu(T) - \tilde{u}_\nu^k \|_1 \leq (1+T)\sqrt{\nu}\,[\; e^T \;\| u^0 - \tilde{u}^0 \|_1 + \mathrm{C}_1\,\Delta t \;+ \mathrm{C}_2\,\frac{\ln N}{\sqrt[4]{N}}\,].$$

$$= (1+T)\,[\; e^T \;\| u_\nu^0 - \tilde{u}_\nu^0 \|_1 + \mathrm{C}_1\sqrt{\nu}\,\Delta t \;+ \mathrm{C}_2'\,\frac{\ln N}{\sqrt[4]{N}}\,]$$

as claimed. The conclusion that C_2 is $O\,(\nu^{-1/2})$ may be misleading however. For example, with waves of the form (8.12) one generally chooses the approximate initial data so that $\mathrm{K}^\nu = O\,(\sqrt{\nu})$ in which case $\mathrm{C}_2 = O\,(1)$.

Finally we remark that a similar argument applied to (8.8) can be used to establish the validity of the bound on the variance of the error in (1.7).

9. Numerical Results In order to compare the theoretical bounds with the actual performance of the method we used the random gradient method to compute a known exact solution. In addition, we present the results of experiments with a second order solution of the reaction equation (3.1a,b) and a second order operator splitting (Strang splitting). These experiments allow us to test the 'sharpness' of our theoretical estimates and our understanding of the way various sources of error behave.

9.1. The Test Problem Let $\nu = 1$. Then the Kolmogorov equation (1.3a,b) has a traveling wave solution of the form

$$u(x,y) = g(x - \alpha t) \tag{9.1}$$

with speed

$$\alpha = \frac{5}{\sqrt{6}}$$

and wave form

$$g(x) = (1 - (\sqrt{2} - 1)\, e^{x/\sqrt{6}})^{-2}.$$

Our approximation to u^0 was determined by placing N particles, each with weight $w_i^0 = N^{-1}$, at

$$X_i^0 = \begin{cases} g^{-1}(1 - \frac{i}{N}) & i = 1, \ldots, N-1, \\ g^{-1}(\frac{1}{2N}) & i = N. \end{cases}$$

For this choice of $\tilde{u}^0$ we have

$$\| u^0 - \tilde{u}^0 \|_1 = O(N^{-1}). \tag{9.2}$$

For u given by (9.1) we define the center of the wave at time t to be the point $x_c = x_c(t)$ such that

$$u(x_c, t) = \frac{1}{2}.$$

We measured the error at time t on a grid of 1001 equally spaced points, centered at x_c and spaced a distance $\Delta x = 0.02$ apart. We measured the error in each of the discrete L^1, L^2, and L^∞ norms. Errors in all three norms decreased at roughly the same rate and therefore we present only the results in the L^1 norm.

9.2. Numerical Results Table 1 contains the results of a computation with the basic random gradient method described in Chapter 2. Each entry in the table is the L^1 norm of the error at time $T = 1.0$ after one run. Note that N increases by 4 as one moves to the right along a row while Δt decreases by 2 as one moves down a column. Further note that, on the average, the error decreases by 2 as one moves diagonally down one row and right one column. We conclude that for the present problem, the proper relationship between Δt and N is

$$\Delta t = O\left(\frac{1}{\sqrt{N}}\right). \tag{9.3}$$

The relation $\Delta t = O(\sqrt[4]{N}^{-1})$ arrived at by theoretical considerations appears to be an underestimate of the dependence of the error on N. In other words, for $\Delta t = O(\sqrt[4]{N}^{-1})$ the errors that depend on N will decrease twice as fast as the errors that depend on Δt, until eventually these latter sources of error dominate all others. The method still converges if we take $\Delta t = O(\sqrt[4]{N}^{-1})$ but we will be doing four times as much work to get the same results. In methods such as the vortex sheet method, in which the work is proportional to $O(N^2)$, we would be doing *sixteen times as much work.* One should be careful when basing their choice of parameters on theoretical estimates alone.

Note that the errors in Table 1 decrease at a rate remarkably close to $O(\Delta t) = O(\sqrt{N}^{-1})$. We interpret this to mean that when Δt and N are chosen so that (9.3) holds, the expected value of the error is $O(\Delta t) = O(\sqrt{N}^{-1})$ and that, for $N \geq 1000$, the

variance of this error is relatively small. Thus, for $N \geq 1000$ one obtains reasonable results with one trial. This corroborates the estimates we made in the introduction to Chapter 8.

First Order Solution of the ODE $u_t = u(1-u)$.

	- Number of Particles -					
Δt	1000	4000	16000	64000	256000	1024000
1	0.5504	0.3863	0.4548	0.4394	0.4447	0.4439
2^{-1}	0.2287	0.2110	0.2168	0.2192	0.2202	0.2210
4^{-1}	0.0995	0.1148	0.0989	0.1056	0.1125	0.1135
8^{-1}	0.1116	0.0531	0.0575	0.0579	0.0551	0.0576
16^{-1}	0.0976	0.0419	0.0276	0.0359	0.0300	0.0288
32^{-1}	0.1034	0.0453	0.0192	0.0256	0.0116	-

Table 1 (L^1 norm)

Fix Δt and consider the error as a function of N. Note that at first it decays like $\sqrt{N}^{-1}$ but then eventually levels out. Further note that this 'plateau' occurs further and further to the right as Δt decreases. The plateau is due to those sources of error, such as the splitting error and the error due to approximating $R_{\Delta t}$ by $\widetilde{R}_{\Delta t}$, that depend on Δt alone. Similarly, one can isolate plateaus that depend only N. Since the $O(\sqrt{N}^{-1})$ errors dominate the $O(N^{-1})$ error in (9.2), these errors are primarily if not exclusively due to the random walk.

While we know of no way to improve the accuracy of the random walk there are several ways to obtain a method which is higher order in time. We begin by considering a second order ODE solver. Define

$$\widetilde{R}_{\Delta t}^{2nd}\,\tilde{u}^0(x) = \tilde{u}^0(x) + \frac{\Delta t}{2}\left[\, f\,(\tilde{u}^0(x)) + f\,(\tilde{u}^0(x) + \Delta t\ f\,(\tilde{u}^0(x))\,\right]. \tag{9.4}$$

This is simply a second order solution of the ODE (3.1a,b) ([18], p. 364). Table 2 contains the results of a numerical experiment with $\widetilde{R}_{\Delta t}$ replaced by $\widetilde{R}_{\Delta t}^{2nd}$. It is immediately apparent that there has been an overall decrease in the error. Note, however, that the *rate* of convergence has not changed, even as a function of Δt alone. On the average the errors still decay like $O(\Delta t)$.

Second Order Solution of the ODE.

	- Number of Particles -				
Δt	1000	4000	16000	64000	256000
1	0.1034	0.0663	0.0682	0.0634	0.0660
2^{-1}	0.0439	0.0275	0.0223	0.0215	0.0230
4^{-1}	0.0297	0.0189	0.0104	0.0088	0.0095
8^{-1}	0.0250	0.0139	0.0063	0.0040	0.0044
16^{-1}	0.0273	0.0165	0.0076	0.0036	0.0031
32^{-1}	0.0309	0.0137	0.0078	0.0045	0.0022

Table 2 (L^1 norm)

We interpret this data in the following way. When $\widetilde{R}_{\Delta t}$ is replaced by $\widetilde{R}_{\Delta t}^{2nd}$ the $(\Delta t)^2$ in (3.10) is replaced by $(\Delta t)^3$.† Thus, the dependence of the last term on the right in (3.6) on Δt is now $O((\Delta t)^2)$ rather than $O(\Delta t)$. However, its dependence on N is still $O(\sqrt{N}^{-1})$. From (9.2) we see that the middle term is $O(N^{-1})$ and hence is presumably negligible compared to the last term. What remains is the first term, the error due to operator splitting.

Following Strang [54] we now employ the following operator splitting algorithm,

$$\tilde{u}_s^{j+1} = \widetilde{R}_{\Delta t/2}^{2nd}\, \widetilde{D}_{\Delta t}\ \widetilde{R}_{\Delta t/2}^{2nd}\, \tilde{u}_s^{j}\,. \tag{9.5}$$

We will refer to this as *Strang splitting.* We believe that when Strang splitting is used the error due to exact operator splitting becomes $O(\sqrt{\nu}\,(\Delta t)^2)$,

$$\| F_{\Delta t}^k u^0 - (R_{\Delta t/2} D_{\Delta t}\, R_{\Delta t/2})^k\, u^0 \|_{L^1} \ \leq\ \mathrm{C}\,\sqrt{\nu}\,(\Delta t)^2.$$

In fact, we think that this can be shown by applying the argument in the proof of Theorem 5.8 to the function

$$w_s(x,t) = F_t\, u^0(x) - R_{t/2}\, D_t\ R_{t/2} u^0(x).$$

† This statement is easily proved. We simply use the well known fact that the local truncation error for a second order ODE solver is $O((\Delta t)^3)$ to replace the right hand side of (6.7) by $(\Delta t)^3$ times the appropriate constant.

In this regard we note Beale and Majda have shown that Strang splitting for the Navier-Stokes equations is second order accurate [7].

Second Order Solution of the ODE with Strang Splitting.

	- Number of Particles -				
Δt	1000	4000	16000	64000	256000
1	0.0545	0.0243	0.0139	0.0083	0.0109
2^{-1}	0.0293	0.0144	0.0090	0.0057	0.0062
4^{-1}	0.0300	0.0140	0.0097	0.0034	0.0028
8^{-1}	0.0242	0.0146	0.0064	0.0028	0.0018
16^{-1}	0.0266	0.0175	0.0071	0.0026	0.0020
32^{-1}	0.0311	0.0133	0.0083	0.0045	0.0022

Table 3 (L^1 norm)

In Table 3 we present the results of using this algorithm on the test problem. We note a further decrease in the error as compared to Tables 1 and 2. This can be explained by the fact that the first term on the right hand side in (3.6) is now $O((\Delta t)^2)$ rather than $O(\Delta t)$. In fact, for $\Delta t \leq 1/8$ the errors that depend on Δt appear to be so small there is little further decrease in the error if one fixes N and lets Δt go to 0. It is important to note that the error still depends on N like $O(\sqrt{N}^{-1})$. In other words, the overall dependence of the error on N has not changed. However, now the choice of parameters which results in the first and last terms on the right in (3.6) decreasing at the same rate is

$$\Delta t = O(\frac{1}{\sqrt[4]{N}}).$$

It is important to note that Strang splitting costs no more than first order splitting. One simply, takes half a time step at the beginning and another half time step at the end,

$$\tilde{u}_\epsilon^j = \tilde{R}_{\Delta t/2}^{2nd} \tilde{D}_{\Delta t} (\tilde{R}_{\Delta t}^{2nd} \tilde{D}_{\Delta t})^{k-1} \tilde{R}_{\Delta t/2}^{2nd} \tilde{u}^0.$$

Table 4 contains the data from columns 1, 3, and 5 of Table 3 organized so that the most efficient way to decrease the errors now lies on the diagonal. Thus, we see that for a given fixed

N one can now achieve the same level of accuracy as the original method with fewer time steps. This results in a small savings in computational effort. To decrease the error by four the original version requires $N \rightarrow 16N$ and $\Delta t \rightarrow \Delta t/4$ resulting in 64 times as much work.† On the other hand, the higher order method requires 32 times as much work to achieve one fourth the error. We remark that for methods in which the work required at each time step is $O(N^2)$ the savings is proportionally smaller.

Second Order Solution of the ODE with Strang Splitting.

	- Number of Particles -		
Δt	1000	16000	256000
1	0.0545	0.0139	0.0109
2^{-1}	0.0293	0.0090	0.0062
4^{-1}	0.0300	0.0097	0.0028
8^{-1}	0.0242	0.0064	0.0018
16^{-1}	0.0266	0.0071	0.0020
32^{-1}	0.0311	0.0083	0.0022

Table 4 (The optimal choice of parameters lies on the diagonals.)

Finally, we replaced $\widetilde{R}_{\Delta t}$ by $R_{\Delta t}$ (it is easy to compute the exact solution of equation (3.1a,b)) and tried this version on the test problem. In this experment we did not use Strang splitting. The results are presented in Table 5. For those choices of N and Δt for which one expects the errors due to Δt to be noticeable we note a moderate improvement over the use of the second order ODE solver without Strang splitting. On the other hand, when the $O(\sqrt{N}^{-1})$ errors dominate, the errors in Table 5 are quite close to those in Table 2. Comparing Table 5 with Table 3 we note the marked difference that the use of Strang splitting makes. We conclude that if one is going to go to the trouble of using a higher order solution of the ODE (3.1a), then one should also use Strang splitting, especially because it results in no appreciable increase in computational effort.

† Here we have neglected the work required to sort the particles at the end of every time step and assumed the work at every time step is $O(N)$.

Exact Solution of the ODE.

	- Number of Particles -					
Δt	1000	4000	16000	64000	256000	1024000
1	0.0756	0.0307	0.0329	0.0267	0.0293	0.0288
2^{-1}	0.0372	0.0199	0.0155	0.0128	0.0144	0.0149
4^{-1}	0.0297	0.0175	0.0102	0.0069	0.0074	0.0075
8^{-1}	0.0248	0.0141	0.0063	0.0037	0.0039	0.0037
16^{-1}	0.0272	0.0166	0.0076	0.0035	0.0031	0.0016
32^{-1}	0.0309	0.0137	0.0079	0.0045	0.0022	0.0007

Table 5 (L^1 norm)

9.3. Conclusions The most obvious conclusion is that the theoretical estimates underestimate the rate of convergence. One can argue that this is a special test problem and more general problems may converge at a slower rate. However, most solutions of (1.3a,b) converge to traveling wave solutions in time (Bramson [8]) and it seems likely that the method's behavior with this particular traveling wave solution is representative of its general behavior when approximating a traveling wave solution.

We believe that the failure of our analysis to accurately predict the true rate of convergence is largely due to the fact that we divided the overall error into the sum of the errors made at each time step and bounded each of these errors separately. In this regard we note that in [33] Hald was able to establish the correct rate of convergence for the method considered there because he could write down the exact solution at any time t. Thus, he simply subtracted this from the computed solution. This example serves to indicate the perils of using the triangle inequality too liberally.

The numerical results above clearly demonstrate that for the original version of the random gradient method the dependence of the overall error on the fractional step is $O(\Delta t)$. Based on this evidence we believe that the bound in (3.6) is sharp and that our analysis simply underestimates the dependence of the last term on N. Furthermore, we believe that the conver-

gence proof provides a sharp estimate of the dependence of the error on Δt .

References

1. B. J. Alder, D. M. Ceperly, W. A. Lester, and P. J. Reynolds, "Fixed-node Quantum Monte Carlo for Molecules," *J. Chem. Phys.*, vol. 77 , no. 11, pp. 5593-5603, 1982.

2. B. J. Alder and D. M. Ceperley, "Quantum Monte Carlo," *Science*, vol. 231, pp. 555-560, 1986.

3. C. R. Anderson and C. A. Greengard, "On Vortex Methods," *Siam J. Numer. Anal.*, vol. 22, 1985.

4. D. G. Aronson and H. F. Weinberger, "Nonlinear Diffusion in Population Genetics, Combustion and Nerve Pulse Propagation," in *Lecture Notes in Mathematics*, ed. J. A. Goldstein, vol. 446, pp. 5-49, Springer, New York, 1975.

5. J. T. Beale and A. Majda, "Vortex Methods I: Convergence in Three Dimensions," *Math Comp*, vol. 39, pp. 1-27, 1982.

6. J. T. Beale and A. Majda, "Vortex Methods II: Higher Order Accuracy in Two and Three Dimensions," *Math Comp*, vol. 39, pp. 29-52, 1982.

7. T. J. Beale and A. Majda, "Rates of Convergence for Viscous Splitting of the Navier-Stokes Equations," *Math Comp*, vol. 37, pp. 243-259, 1981.

8. M. Bramson, "Convergence of Solutions of the Kolmogorov Equation to Travelling Waves," *Memoirs of the American Mathematical Society*, vol. 44 , no. 285, 1983.

9. Y. Brenier, "A Particle Method for One Dimensional Non-Linear Reaction Advection Diffusion Equations," *Technical Report No. 351*, Instituto de Investigaciones en Mathematicas Aplicadas y en Sistemas, Universidad Nacional Autonoma de Mexico, 1983.

10. C. C. Chang, "Numerical Solution of Stochastic Differential Equations," *Ph.D. Thesis*, University of California, Berkeley, 1985.

11. C. C. Chang, "Random Vortex Methods for the Navier-Stokes Equations," *manuscript*, 1986.

12. A. J. Chorin, "Hermite Expansions in Monte-Carlo Computation," *J. Computational Physics*, vol. 27, pp. 472-482, 1971.

13. A. J. Chorin, "Numerical Study of Slightly Viscous Flow," *J. Fluid Mech.*, vol. 57, pp. 785-796, 1973.

14. A. J. Chorin, "Vortex Sheet Approximation of Boundary Layers," *J. Comp. Phys.*, vol. 27, pp. 428-442, 1978.

15. A. J. Chorin, "Numerical Methods For Use in Combustion Modeling," *Proc. Int. Conf. Num. Meth. in Science and Engineering*, Versailles, 1979.

16. K. L. Chung, *A Course in Probability Theory*, Harcourt, Brace & World Inc., 1968.

17. P. Colella, "Glimm's Method for Gas Dynamics," *Siam J. Sci. Stat. Comput.*, vol. 3, no. 1, 1982.

18. S. D. Conte and C. de Boor, *Elementary Numerical Analysis*, McGraw Hill, New York, 1980. third edition

19. G. H. Cottet, "Methodes Particulaires pour l'Equation d'Euler dans le Plan," *These de 3eme cycle*, l'Universite Pierre et Marie Curie, Paris, 1982.

20. G. H. Cottet and S. Gallic, *A Particle Method to Solve Transport-Diffusion Equations, Part I: The Linear Case*, Rapport Interne No: 115, Centre de Mathematiques Appliquees, Ecole Polytechique, Paris, 1984.

21. P. Davis and P. Rabinowitz, *Methods of Numerical Integration*, Academic Press, 1975.

22. W. Feller, "An Introduction to Probability Theory and Its Applications," *2nd ed.*, vol. II, John Wiley & Sons, New York, 1971.

23. P. C. Fife and J. B. McLeod, "The Approach of Solutions of Nonlinear Diffusion Equations to Travelling Wave Front Solutions," *Arch. Rational Mech. Anal.*, vol. 65, pp. 335-362, 1977.

24. G. B. Folland, *Introduction to Partial Differential Equations,* Princeton University Press, 1976.

25. A. Friedman, *Partial Differential Equations of the Parabolic Type,* Prentice-Hall, Englewood Cliffs, N.J., 1964.

26. A. F. Ghoniem and F. S. Sherman, "Grid Free Simulation of Diffusion Using Random Walk Methods," *J. Comp. Phys.*, vol. 61, pp. 1-37, 1985.

27. J. Goodman, "Convergence of the Random Vortex Method in Two Dimensions," *Comm Pure and Applied Math*, to appear.

28. C. A. Greengard, "Three Dimensional Vortex Methods," *Ph.D. Thesis*, U. C. Berkeley, 1984.

29. O. H. Hald, *private communication.*

30. O. H. Hald and V. M. Del Prete, "Convergence of Vortex Methods for Euler's Equations," *Math Comp*, vol. 32, pp. 791-809, 1978.

31. O. H. Hald, "The Convergence of Vortex Methods, II," *Siam J. Num. Anal.*, vol. 16, pp. 726-755, 1979.

32. O. H. Hald, "Convergence of Random Methods for a Reaction-Diffusion Equation," *Siam J. Sci. Stat. Comput.*, vol. 2 , 1981.

33. O. H. Hald, "Convergence of a Random Method With Creation of Vorticity," *Siam J. Sci. Stat. Comput.*, vol. 7 , pp. 1373-1386, 1986.

34. O. H. Hald, "Convergence of Vortex Methods for Euler's Equations, III," *Siam J. Num. Analysis*, to appear.

35. J. H. Halton, "On the Efficiency of Certain Quasi-Random Sequences of Points in Evaluating Multi-Dimensional Integrals," *Numerische Mathematik*, vol. 2, pp. 84-91, 1960.

36. J. M. Hammersley and D. C. Handscomb, *Monte Carlo Methods*, Methuen & Co. Ltd, London, 1964.

37. W. Hoeffding, "Probability Inequalities for Sums of Bounded Random Variables," *American Statistical Association Journal*, pp. 13-30, March, 1963.

38. B. Jansen, *Random Number Generators*, Petersson, Stockholm, 1976.

39. F. John, *Partial Differential Equations*, Springer Verlag, 1982.

40. D. E. Knuth, *The Art of Computer Programming*, 2, Addison and Wesley, 1969.

41. S. N. Kruzkov, "First Order Quasilinear Equations with Several Space Variables," *Math USSR - sb*, vol. 10, pp. 217-243, 1970.

42. M. Loève, *Probability Theory*, Springer Verlag, 1977.

43. F. H. Maltz and D. L. Hitzl, "Variance Reduction in Monte-Carlo Computations Using Multi-Dimensional Hermite Polynomials," *J. Computational Physics*, vol. 32, pp. 345-376, 1979.

44. P. A. Raviart, "An Analysis of Particle Methods," *CIME Course on Numerical Methods in Fluid Dynamics*, Publications du Laboratoire d'Analyse numerique, Universite Pierre et Marie Curie, Paris, 1983.

45. P. A. Raviart, *Particle Approximation of Linear Hyperbolic Equations of the First Order*, Publications du Laboratoire d'Analyse numerique, Universite Pierre et Marie Curie, Paris, 1983.

46. S. G. Roberts, "Accuracy of the Random Vortex Method for a Problem with Non-Smooth Initial Conditions," *J. Comp. Phys.* , vol. 58, pp. 29-43, 1985.

47. S. G. Roberts, "Convergence of a Random Walk Method for Burgers' Equation," *Ph.D. Thesis*, University of California, Berkeley, 1985.

48. G. Rosen, "Brownian-Motion Correspondence Method for Obtaining Approximate Solutions to Nonlinear Reaction-Diffusion Equations," *Physical Review Letters*, vol. 53 , no. 4, pp. 307-310, 1984.

49. R. Y. Rubinstein, *Simulation and the Monte Carlo Method,* John Wiley & Sons, New York, 1981.

50. A. S. Sherman and C. S. Peskin, "A Monte Carlo Method for Scalar Reaction-Diffusion Equations," *SIAM J. Sci Stat Comp*, 1986.

51. A. S. Sherman and C. S. Peskin, "Solving the Hodgkin-Huxley Equations by a Random Walk Method," *manuscript*, 1986.

52. J. Smoller, *Shock Waves and Reaction-Diffusion Equations,* Springer Verlag, New York, 1983.

53. G. Sod, *Numerical Methods in Fluid Dynamics,* Cambridge University Press, New York, 1985.

54. G. Strang, "On the Construction and Comparison of Difference Schemes," *SIAM J. Numer. Anal.*, vol. 5, pp. 506-517, 1968.

This report was done with support from the Department of Energy. Any conclusions or opinions expressed in this report represent solely those of the author(s) and not necessarily those of The Regents of the University of California, the Lawrence Berkeley Laboratory or the Department of Energy.

Reference to a company or product name does not imply approval or recommendation of the product by the University of California or the U.S. Department of Energy to the exclusion of others that may be suitable.

LAWRENCE BERKELEY LABORATORY
TECHNICAL INFORMATION DEPARTMENT
UNIVERSITY OF CALIFORNIA
BERKELEY, CALIFORNIA 94720