Perturbation method for solving a certain class of partial differential equations

1999 J. Phys. A: Math. Gen. 32 4955

(http://iopscience.iop.org/0305-4470/32/26/313)

J. Phys. A: Math. Gen. **32** (1999) 4955–4962. Printed in the UK PII: S0305-4470(99)00346-7

Perturbation method for solving a certain class of partial differential equations

Richard V Durand and Carl Franck

Laboratory of Atomic and Solid State Physics and Center for Materials Research, Cornell University, Ithaca, NY 14853-2501, USA

Received 23 December 1998, in final form 13 April 1999

Abstract. We show how one can express the solution to a certain type of partial differential equation (PDE) in the form of an infinite operator series acting on the solution to another PDE of the same form. This technique can be thought of as a perturbation method in the sense that the solution to a difficult problem is written in terms of a solution to an analytically tractable problem or 'reference' problem. For diffusion problems, where the solution must have a constant integral with respect to the spatial coordinates (probability conservation), this method is especially useful because the infinite operator series conserves probability term by term. Two example diffusion problems are solved using this method to illustrate the concepts.

1. Introduction

During previous experimental work dealing with colloidal systems [1], we found that we needed an approximate analytical solution of the mutual diffusion equation describing two hydrodynamically interacting spheres suspended in an aqueous medium. One incorporates hydrodynamic interactions into the free diffusion equation by using a tensor diffusivity in place of the usual scalar diffusivity, and the resulting equation is difficult to solve analytically. In [1], we found an analytic solution to this equation in the form of an infinite operator series acting on a solution to the free diffusion equation, but we only calculated the first few terms in the series. In this paper, we give the full operator series, and we show that this technique is not just limited to diffusion equations, but can be applied to an entire class of partial differential equations (PDEs). This technique can be thought of as a perturbation method in the sense that the solution to a difficult problem is written in terms of a solution to an analytically tractable problem or 'reference' problem.

The organization of the paper is as follows. In section 2, we derive the perturbation theory, and show that it has certain desirable properties when used in the context of diffusion problems. In section 3, we illustrate the concepts through two examples, and in section 4, we draw some conclusions and give ideas for future work.

2. Perturbation theory

Consider PDEs of the form

$$\frac{\partial}{\partial t}u(\boldsymbol{r},t) = \mathcal{L}u(\boldsymbol{r},t) \tag{1}$$

0305-4470/99/264955+08$30.00

with initial condition

$$u(\boldsymbol{r}, 0) = f(\boldsymbol{r}) \tag{2}$$

where $\mathcal{L}$ is a linear partial differential operator which depends only on the coordinates $\boldsymbol{r}$ and not on t, and $f(\boldsymbol{r})$ is any function. For example, the mutual diffusion equation for two hydrodynamically interacting spheres with separation vector $\boldsymbol{r}$ is of the form (1) if $u(\boldsymbol{r}, t)$ denotes the probability distribution of the separation coordinate $\boldsymbol{r}$ at time t, and one puts [2]

$$\mathcal{L} = \nabla \cdot \boldsymbol{D}(\boldsymbol{r}) \cdot \nabla \tag{3}$$

where $\boldsymbol{D}(\boldsymbol{r})$, the mutual diffusion tensor, is a function of the separation vector. Of course, in the free diffusion case, this reduces to $\mathcal{L} = 2D_0\nabla^2$, where D_0 is the diffusivity of a single particle.

In the spirit of perturbation theory, we would like to write the solution to equation (1) in terms of a solution to an equation of the same form but with a different operator $\tilde{\mathcal{L}}$ (a 'reference' system). For instance, in the example given above, we could choose our reference system to be the free diffusion system and write the solution to the hydrodynamically interacting diffusion problem in terms of this reference solution. In the next section, we show how this can be done using an operator formalism.

2.1. Operator series

The formal solution to the problem of interest (equation (1)) is given by

$$u(\boldsymbol{r}, t) = \mathrm{e}^{\mathcal{L}t} f(\boldsymbol{r}) \tag{4}$$

and the formal solution to the 'reference problem'

$$\frac{\partial}{\partial t}\tilde{u}(\boldsymbol{r}, t) = \tilde{\mathcal{L}}\tilde{u}(\boldsymbol{r}, t) \tag{5}$$

$$\tilde{u}(\boldsymbol{r}, 0) = u(\boldsymbol{r}, 0) = f(\boldsymbol{r}) \tag{6}$$

is given by

$$\tilde{u}(\boldsymbol{r}, t) = \mathrm{e}^{\tilde{\mathcal{L}}t} f(\boldsymbol{r}). \tag{7}$$

Note that the operators $\mathcal{L}$ and $\tilde{\mathcal{L}}$ do not necessarily commute, and the complexity of the manipulations that follow stem from this fact.

Let us write the exponential on the right-hand side of equation (4) as $\mathrm{e}^{(\mathcal{L}-\tilde{\mathcal{L}})t+\tilde{\mathcal{L}}t}$, and expand it as follows [3]:

$$\mathrm{e}^{(\mathcal{L}-\tilde{\mathcal{L}})t+\tilde{\mathcal{L}}t} = \mathrm{e}^{\tilde{\mathcal{L}}t} + \int_0^t \mathrm{e}^{\mathcal{L}(t-s)}(\mathcal{L} - \tilde{\mathcal{L}})\mathrm{e}^{\tilde{\mathcal{L}}s}\,\mathrm{d}s. \tag{8}$$

If we apply the operator (8) to $f(\boldsymbol{r})$, and use equation (7), we can write the solution to the original problem in terms of the 'reference' solution $\tilde{u}(\boldsymbol{r}, t)$ as follows:

$$u(\boldsymbol{r}, t) = \tilde{u}(\boldsymbol{r}, t) + \int_0^t \mathrm{e}^{\mathcal{L}(t-s)}(\mathcal{L} - \tilde{\mathcal{L}})\tilde{u}(\boldsymbol{r}, s)\,\mathrm{d}s. \tag{9}$$

If one expands $\tilde{u}(\boldsymbol{r}, s)$ in a Taylor series about $s = t$ (using equation (5) to replace each time derivative with $\tilde{\mathcal{L}}$), substitutes this in equation (9), and formally expands the exponential in equation (9) as well, then, after integrating over s and collecting terms of the same power of t, one finds that

$$u(\boldsymbol{r}, t) = \sum_{n=0}^{\infty} A_n \tilde{u}(\boldsymbol{r}, t) \tag{10}$$

where $A_0 \equiv 1$, and the operators A_n for $n \geqslant 1$ are given by

$$A_n = \sum_{k=0}^{n-1} \frac{(-1)^k t^n}{k!(n-1-k)!n} \mathcal{L}^{n-k-1}(\mathcal{L} - \tilde{\mathcal{L}})\tilde{\mathcal{L}}^k. \tag{11}$$

A recursion relation exists among the A_n, given by

$$A_{n+1} = \frac{t}{n+1}(\mathcal{L}A_n - A_n\tilde{\mathcal{L}}) \tag{12}$$

which can be proven directly using equation (11).

Let us define $u^{(n)}(\boldsymbol{r}, t) \equiv A_n\tilde{u}(\boldsymbol{r}, t)$. Using equations (12) and (5), we see that

$$u^{(n+1)}(\boldsymbol{r}, t) = \frac{t}{n+1}\left(\mathcal{L}u^{(n)}(\boldsymbol{r}, t) - A_n\frac{\partial}{\partial t}\tilde{u}(\boldsymbol{r}, t)\right). \tag{13}$$

Using the relation

$$\frac{\partial}{\partial t}(A_n\tilde{u}) = A_n\frac{\partial\tilde{u}}{\partial t} + \frac{n}{t}A_n\tilde{u} \tag{14}$$

we can write equation (13) as

$$u^{(n+1)}(\boldsymbol{r}, t) = \frac{1}{n+1}\left[n + t\left(\mathcal{L} - \frac{\partial}{\partial t}\right)\right]u^{(n)}(\boldsymbol{r}, t). \tag{15}$$

Thus, the solution to equation (1) can be written as a series of functions as follows:

$$u(\boldsymbol{r}, t) = \sum_{n=0}^{\infty} u^{(n)}(\boldsymbol{r}, t) \tag{16}$$

where $u^{(0)}(\boldsymbol{r}, t) \equiv \tilde{u}(\boldsymbol{r}, t)$, and each successive $u^{(n)}$ is given by equation (15). Note that one may also write equation (16), in a purely formal way, as

$$u(\boldsymbol{r}, t) = \sum_{n=0}^{\infty} \binom{n + t(\mathcal{L} - \frac{\partial}{\partial t}) - 1}{n}\tilde{u}(\boldsymbol{r}, t) \tag{17}$$

where the symbol ( ) denotes a binomial coefficient.

Thus, we have written the solution to equation (1) in terms of the solution to the reference problem equation (5). If one can choose an analytically soluble reference problem that is in some sense 'close' enough to the problem of interest, then one hopes that the perturbation solution (equation (16)) will converge rapidly. In the next section, we show that the $u^{(n)}$ have properties advantageous for solving diffusion problems.

2.2. Conservation of probability

If one wants to use this technique to solve unbounded diffusion problems (where $u(\boldsymbol{r}, t)$ is a probability density), it would be desirable if the integral over $\boldsymbol{r}$ of any truncation of the perturbation solution (equation (16)) remained unity. If that were the case, one would not have to worry about probability conservation when adding successive terms $u^{(n)}$. Since $\int u^{(0)} = \int \tilde{u} = 1$ (assuming our reference solution $\tilde{u}(\boldsymbol{r}, t)$ is a probability density), this requires that $\int u^{(n)} = 0$ for all $n \geqslant 1$.

To prove this, integrate both sides of equation (15) over the spatial coordinates $\boldsymbol{r}$ to obtain

$$\int u^{(n+1)} = \frac{n}{n+1}\int u^{(n)} + \frac{t}{n+1}\int \mathcal{L}u^{(n)} - \frac{t}{n+1}\frac{\partial}{\partial t}\int u^{(n)}. \tag{18}$$

Let us make an argument based on induction, and assume that $\int u^{(n)} = 0$ for some $n \geqslant 1$. With this assumption, we see that the first and last terms on the right-hand side of equation (18)

are identically zero for that particular n. For diffusion problems, $\mathcal{L}$ is the divergence of another operator, and thus, after an integration by parts, and assuming that $u(\boldsymbol{r}, t)$ and its derivatives with respect to $\boldsymbol{r}$ go to zero at infinity, we get $\int \mathcal{L}u^{(n)} = 0$, which means that the entire right-hand side of equation (18) vanishes. Thus, $\int u^{(n)} = 0$ implies that $\int u^{(n+1)} = 0$. To complete the induction argument, it remains to be shown that $\int u^{(1)} = 0$. To see this, substitute $n = 0$ in equation (18), and use the facts that $\int \mathcal{L}u^{(0)} = 0$ and $\int u^{(0)} = 1$.

3. Two example problems

In this section, we apply the procedure developed in the previous section to two example diffusion problems in order to demonstrate the technique's utility. The first example is a physically reasonable diffusion problem involving two particles interacting hydrodynamically. The second problem was constructed to have convenient mathematical properties, which will be discussed later.

3.1. Physically reasonable problem

We will solve a one-dimensional diffusion equation describing two particles initially separated by a surface to surface distance $x = 5$, characterized by a mutual diffusivity $D(x)$ that varies with the separation distance in a way that is characteristic of hydrodynamic interactions, namely

$$D(x) = \frac{x}{x+2}. \tag{19}$$

This mutual diffusivity vanishes at contact ($x = 0$), increases linearly for small separations, and algebraically approaches a constant as $x \to \infty$, all consistent with the behaviour of the mutual diffusivity of a pair of hydrodynamically interacting spheres in three dimensions [4]. This expression for the diffusivity gives negative, and hence unphysical, values for $-2 \leqslant x < 0$, but this is outside the domain of the problem, because as discussed below, we do not allow the particles to overlap, and so there is no particle density for $x < 0$. The PDE describing the time evolution of the probability distribution $u(x, t)$ of the separation distance x is given by:

$$\frac{\partial u}{\partial t} = \frac{\partial}{\partial x}\left(\frac{x}{x+2}\right)\frac{\partial u}{\partial x} \tag{20}$$

with the initial condition

$$u(x, 0) = \delta(x - 5) \tag{21}$$

where $\delta(x)$ is the Dirac delta function.

In order to compare the approximate solutions with the exact solution, equation (20) was solved numerically using a Crank–Nicholson method [5]. We imposed a no flux boundary condition at $x = 0$ to prevent the particles from overlapping. Unfortunately, the perturbation method cannot, in general, handle boundary conditions, and so our approximate solution will develop a probability density in the unphysical region $x < 0$. However, for the timescale of interest here, the boundary condition does not affect the overall solution appreciably, because there is a very small probability of the particles diffusing that closely together.

We choose, as our reference system, a pair of freely diffusing particles with mutual diffusivity $D = 1$ and initial separation $x = 5$, which results in the reference equation

$$\frac{\partial \tilde{u}}{\partial t} = \frac{\partial^2 \tilde{u}}{\partial x^2} \tag{22}$$

$$\tilde{u}(x, 0) = \delta(x - 5) \tag{23}$$

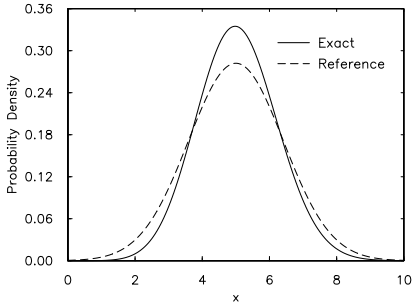

Figure 1. The pair probability distribution at $t = 1$ for an initial separation of $x = 5$ for both the freely diffusing (reference) and hydrodynamically interacting (exact) systems.

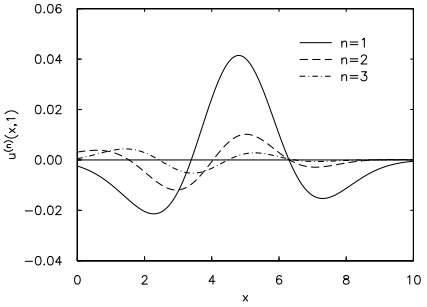

Figure 2. The $n = 1, 2, 3$ terms in the perturbation solution (equation (16)) for the example problem given in the text.

with solution

$$\tilde{u}(x,t) = u^{(0)}(x,t) = \frac{1}{\sqrt{4\pi t}} \exp[-(x-5)^2/(4t)]. \tag{24}$$

The reference solution along with the exact solution at $t = 1$ is shown in figure 1. The presence of the 'hydrodynamic interactions' has caused the particles to not stray as far from their initial positions as in the freely diffusing case, and has led to an asymmetry in the probability distribution.

We used the software package `Mathematica` [6] and equations (15) and (24) to analytically compute the correction terms $u^{(1)}$ through $u^{(10)}$†, where the operator $\mathcal{L}$ in equation (15) is given by

$$\mathcal{L} = \frac{\partial}{\partial x}\left(\frac{x}{x+2}\right)\frac{\partial}{\partial x}. \tag{25}$$

Calculating the analytic form of $u^{(10)}(x,t)$ required approximately 20 min on a computer equipped with a Pentium II microprocessor running at 266 MHz. In contrast, the calculation of $u^{(5)}(x,t)$ required 10 s. The terms $u^{(1)}$, $u^{(2)}$, and $u^{(3)}$ are shown graphically in figure 2 at $t = 1$. The fourth- and higher-order terms continue to diminish with increasing order, with the amplitude of the oscillations decreasing to the point that $u^{(10)}(x,1)$ is barely distinguishable from zero on the scale of figure 2. Figure 3 shows the difference between the exact solution and the zeroth-, first-, second-, and third-order approximate solutions at $t = 1$. Note that the difference between the approximate and exact solutions diminishes as more terms are kept. This trend continues through the higher-order solutions, at least until the tenth order for $t = 1$.

In order to quantitatively compare the exact solution with the approximate solutions, we compute the L^1 norm of the difference between the exact and approximate solutions, given by

$$L^1(N,t) \equiv \int_0^\infty \left| u(x,t) - \sum_{n=0}^{N} u^{(n)}(x,t) \right| \mathrm{d}x. \tag{26}$$

† $u^{(1)}(x,t)$ is given by

$$u^{(1)}(x,t) = \left(\frac{1}{x+2} - \frac{(x-5)^2}{2t(x+2)} - \frac{x-5}{(x+2)^2}\right) u^{(0)}(x,t)$$

where $u^{(0)}(x,t)$ is given by equation (24). The higher-order terms quickly become very algebraically complicated, and will not be given here.

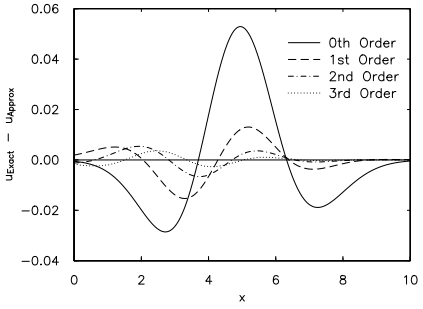

Figure 3. The difference between the exact solution and the approximate analytical solution obtained by keeping successively more terms in equation (16).

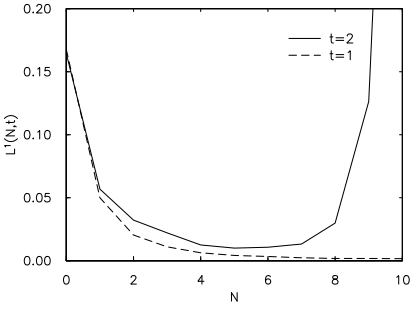

Figure 4. The L^1 norm $L^1(N,t)$ of the difference between the exact and approximate solutions as a function of the number of terms N used in equation (16).

This quantity will approach zero as $\sum u^{(n)}(x,t)$ approaches the exact solution $u(x,t)$. Figure 4 shows $L^1(N,t)$ for $t = 1, 2$ and $N = 0$–10. For $t = 1$, the absolute difference between the approximate and exact solutions decreases with increasing N up to at least tenth order. However, for $t = 2$, the difference initially decreases with increasing N, then increases for $N \geqslant 6$, as is characteristic of an asymptotic series.

3.2. Mathematically convenient example

Let us again consider a one-dimensional diffusion problem. As we have said before, the perturbation method cannot, in general, handle boundary conditions. However, if the diffusivity $D(x)$ has the property that $D(x) = D(-x)$, then we can construct a perturbation solution that satisfies a no flux boundary condition at $x = 0$. For, if $D(x)$ has this symmetry property, then the operator $\mathcal{L} = \partial_x D(x)\partial_x$ is invariant under reflection about $x = 0$. Therefore, if $u(x,t)$ is a solution to the PDE, then $u(-x,t)$ is also a solution, and $u(x,t) + u(-x,t)$ is the solution that satisfies the no flux boundary condition at $x = 0$. If we choose a reference solution of this form, then because of the symmetry of $\mathcal{L}$, the $u^{(n)}(x,t)$ generated by repeated application of equation (15) will also satisfy the no flux boundary condition at $x = 0$, and thus the perturbation solution (equation (16)) will satisfy the no flux boundary condition†.

We choose a $D(x)$ similar to equation (19), but which is symmetric about $x = 0$, namely‡

$$D(x) = \frac{x^2 + 1}{x^2 + 6}. \tag{27}$$

We start with a particle separation of $x = 2$, and thus the PDE describing the time evolution of the probability distribution $u(x,t)$ of the separation distance x is given by:

$$\frac{\partial u}{\partial t} = \frac{\partial}{\partial x}\left(\frac{x^2 + 1}{x^2 + 6}\right)\frac{\partial u}{\partial x} \tag{28}$$

$$u(x,0) = \delta(x - 2) \tag{29}$$

$$\frac{\partial}{\partial x}u(0,t) = 0 \tag{30}$$

where $\delta(x)$ is the Dirac delta function.

We choose, as our reference system, a pair of freely diffusing particles with mutual diffusivity $D = 0.5$, which seems a natural choice, since the initial particle separation is $x = 2$,

† We would like to thank one of the referees of this paper for pointing out this fact.

‡ This form of $D(x)$ was suggested by a referee of this paper.

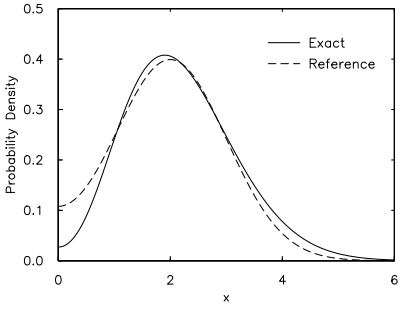

Figure 5. The pair probability distribution at $t = 1$ for an initial separation of $x = 2$ for both the freely diffusing (reference) and interacting (exact) systems.

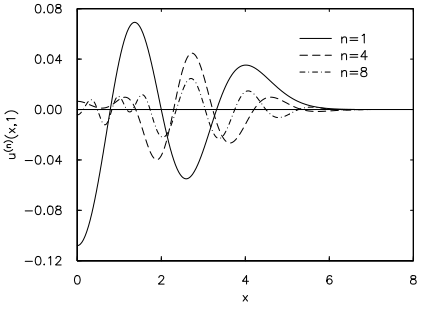

Figure 6. The $n = 1, 4, 8$ terms in the perturbation solution (equation (16)) for the example problem given in the text.

and the diffusivity (equation (27)) at this separation is 0.5. Thus, the reference equation is given by

$$\frac{\partial \tilde{u}}{\partial t} = \frac{1}{2}\frac{\partial^2 \tilde{u}}{\partial x^2} \tag{31}$$

$$\tilde{u}(x, 0) = \delta(x - 2) \tag{32}$$

$$\frac{\partial}{\partial x}\tilde{u}(0, t) = 0 \tag{33}$$

with solution

$$\tilde{u}(x, t) = u^{(0)}(x, t) = \frac{1}{\sqrt{2\pi t}}(\exp[-(x - 2)^2/(2t)] + \exp[-(x + 2)^2/(2t)]). \tag{34}$$

As before, equation (28) was solved numerically using a Crank–Nicholson method [5]. The reference solution along with the exact solution at $t = 1$ is shown in figure 5.

We used the software package `Mathematica` [6] and equations (15) and (34) to analytically compute the correction terms $u^{(1)}$ through $u^{(8)}$, where the operator $\mathcal{L}$ in equation (15) is given by

$$\mathcal{L} = \frac{\partial}{\partial x}\left(\frac{x^2 + 1}{x^2 + 6}\right)\frac{\partial}{\partial x}. \tag{35}$$

The ninth- and higher-order terms were not calculated due to long computation times. The functions $u^{(1)}$, $u^{(4)}$, and $u^{(8)}$ are shown in figure 6 at $t = 1$, and figure 7 shows the difference between the approximate and exact solutions at $t = 1$ for the zeroth-, first-, fourth-, and eighth-order approximations.

Figure 8 shows $L^1(N, t)$ (equation (26)) for $t = 0.2, 0.5, 1, 2$ and $N = 0$–8. For $t = 0.2, 0.5, 1$, the absolute difference between the approximate and exact solutions decreases with increasing N up to at least eighth order. In contrast, for $t = 2$, the series solution rapidly diverges from the exact solution after the second-order term, which is again characteristic of an asymptotic series. These results suggest that, as in the first example, the perturbation method works better for smaller times.

4. Conclusions

In this paper, we have presented a technique for writing the solution to a certain type of PDE in terms of the solution to another PDE of the same type. We see that this method is useful

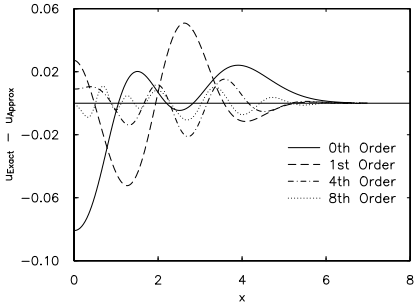

Figure 7. The difference between the exact solution and the approximate analytical solution obtained by keeping successively more terms in equation (16).

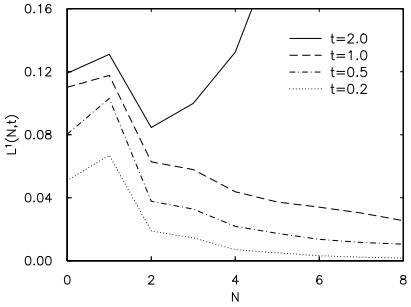

Figure 8. The L^1 norm $L^1(N,t)$ of the difference between the exact and approximate solutions as a function of the number of terms N used in equation (16).

as a perturbation technique, if one can find a PDE 'close' to the PDE of interest, for which an analytic solution can be found. Then, all one must do is successively apply equation (15) to the reference solution in order to generate a successively more accurate analytic solution to the original problem. Due to the complexity of numerical solutions of PDEs in higher numbers of dimensions, the advantages of having an analytic solution as opposed to a numerical one become more apparent as the number of spatial dimensions increases.

Some further work remains to be done, however. First of all, it remains to be determined under which conditions the perturbation solution (equation (16)) actually converges. Figures 4 and 8 seem to indicate that in the two examples given, the series of approximate solutions approaches the exact solution more rapidly for smaller t, and is asymptotic for some, if not all, values of t. This behaviour could stem from the fact that the reference solutions in the two example problems both have essential singularities at $t = 0$, possibly limiting the radius of convergence in the time domain. Perhaps equation (16) can be resummed via Padé summation to yield more useful answers when the series is asymptotic. Also, the current scheme does not allow one to incorporate boundary conditions (except in very special cases, as in the second example problem), and so it only works, in general, with unbounded problems. We present this result in the hopes that other workers will also find it useful and possibly expand upon it.

Acknowledgments

We would like to thank Veit Elser and Frank Hicks for helpful discussions. RD wishes to acknowledge fellowship support from the National Science Foundation. Financial support was provided by the National Science Foundation through grant No DMR-9320910, and we also made use of the Cornell Center for Materials Research facilities supported by the NSF under grant No DMR-9632275.

References

[1] Durand R V and Franck C 1997 *Phys. Rev.* E **56** 1998
[2] Felderhof B U and Jones R B 1983 *Physica* A **119** 591
[3] Hansen J-P and McDonald I R 1991 *Theory of Simple Liquids* (London: Academic) p 245
[4] Batchelor G K 1976 *J. Fluid Mech.* **74** 1
[5] Press W H, Flannery B P, Teukolsky S A and Vetterling W T 1992 *Numerical Recipes in C: The Art Of Scientific Computing* 2nd edn (Cambridge: Cambridge University Press) ch 19
[6] *Mathematica* 1995 version 3.0 Wolfram Research Inc. Champaign, IL