Discrete singular convolution for the solution of the Fokker–Planck equation

G. W. Wei
Department of Computational Science, National University of Singapore, Singapore 119260

(Received 22 December 1998; accepted 11 February 1999)

This paper introduces a discrete singular convolution algorithm for solving the Fokker–Planck equation. Singular kernels of the Hilbert-type and the delta type are presented for numerical computations. Various sequences of approximations to the singular kernels are discussed. A numerical algorithm is proposed to incorporate the approximation kernels for physical applications. Three standard problems, the Lorentz Fokker–Planck equation, the bistable model and the Henon–Heiles system, are utilized to test the accuracy, reliability, and speed of convergency of the present approach. All results are in excellent agreement with those of previous methods in the field. © *1999 American Institute of Physics.* [S0021-9606(99)50518-7]

I. INTRODUCTION

One of the formal approaches to the real world phenomena is to start from the microscopic theory such as the classical Liouville equation or the quantum Liouville equation. These equations describe many-body systems and their general solutions are unfeasible to obtain at present. Reduced descriptions in terms of one- or few-particle distribution functions or density operators lead to the Bogoliubov–Born–Green–Kirkwood–Yvon (BBGKY) hierarchy, Zwanzig's equation, or equivalently, Mori's generalized Langevin equation. These equations formally provide better physical insight for macroscopic phenomena. However, they are still exact and are not soluble in general. Appropriate truncation of these equations results in nonlinear mesoscopic kinetic equations, such as the Boltzmann equation. The Boltzmann equation describes a system in terms of a typical one-particle reduced density operator which is coupled to the rest of the system via binary or triple collisions. Hydrodynamic (or transport) equations, such as the equation of continuity, the equation of motion, and the equation of energy conservation, can be derived from kinetic equations. Hydrodynamic equations are at the macroscopic level and can be used to simulate the real world phenomena as what have usually been done with the Navier–Stokes equation in computational fluid dynamics. However, the involved transport coefficients are formally still determined by the kinetic equations. The Boltzmann equation is nonlinear and is still hard to solve in general. A further simplified model is the celebrated Fokker–Planck equation (FPE).

The theory of the Fokker–Planck equation has a relatively long history starting from the early work by Einstein,[1] Langevin,[2] Fokker,[3] and Planck,[4] as well as many others.[5–7] The modern theory of the Fokker–Planck equation is very rich in its content due to rapid development in analytical and computational analyses and a broad spectrum of applications. It is capable of describing a broad range of scientific phenomena, such as the relation of fluctuations and random force, nonlinearity in pattern formation and various mode coupling effects. Many physical, chemical, biological, and economical systems can be described by the Fokker–Planck equation with various coefficients. There has been a great deal of theoretical attention[8–24] on the FPE in the recent years. Numerous algorithms have been explored for the numerical solution of the Fokker–Planck equation. As one may expect, each of proposed methods has its advantages and limitations. Path integral methods have been utilized by a number of authors.[25–27] Wehner and Wolfer[16] have presented an elegant formalism where the path integrals involving the Onsager–Machlup functionals[11] are numerically evaluated. Monte Carlo techniques[28] are very useful for providing information about certain properties of a system, usually in terms of the moments of the underlying stochastic process. Swendsen and Wang have recently developed a cluster Monte Carlo algorithm[29] which has potential for handling critical slowing down phenomena in nonlinear Fokker–Planck systems. Since Monte Carlo approaches do not directly refer to the probability density distribution, more direct approaches, such as the finite difference method and spectral method are frequently employed[30–32] when the entire distribution function is required. It is well-known that the finite difference method often suffers from stiff systems. Chang and Cooper[33] were the first to discuss a practical finite difference procedure in which the number density of the system is preserved in each step, and the distribution function evolves in a quasiequilibrium manner. The Chang–Cooper method has been generalized by Larsen *et al.*[34] to allow a larger time increment and to achieve greater numerical stability for a wide class of systems, which include the nonlinear Compton scattering problem. A drawback of their approach is that analytic expressions for the collision parameters are required at each time step, which may not be available for a practical application. Recently, Epperlein[19] further generalized the Chang–Cooper method by taking into account the energy conservation. His fully conservative scheme has been applied to a spatially homogeneous plasma system involving the Coulomb collision. Most recently, Drozdov and Morillo have utilized a *K*-point Stirling interpolation formula for accurately calculating the distribution

function.[23] These authors demonstrated that their results are much more reliable than those of a previous cumulant expansion. For a wide class of Fokker–Planck operators, the eigenvalue expansion approach is applicable.[11,36,37] In such an approach, various spectral methods and pseudospectral methods can be used to provide extremely accurate results for the eigenvalues of the Fokker–Planck equation. In cases like a Lorentz gas system, the full set of eigenfunctions and eigenvalues determines completely the dynamics of the Fokker–Planck equation. In particular, Shizgal's method,[35] using nonclassical weight functions, is often optimized in terms of accuracy and speed of convergence for certain problems under study.[36,37] A distributed approximating functional approach[38,39] was used for solving the Fokker–Planck equation with comparable results. Park and Petrosian[40] have recently provided a detailed comparison of several different methods for the solution of the Fokker–Planck equation.

In the eigenfunction expansion approach, use is made of a mathematical transformation to arrive at a particular form of the Fokker-Planck operator which is formally similar to the Schrödinger operator. Then the solution of the Fokker–Planck equation can be treated in a manner analogous to a bound state expansion treatment of the Schrödinger equation. Conceptually, there are some essential differences between the Schrödinger equation and the Fokker–Planck equation. The former is a first principle approach to a closed system at the microscopic level and is an exact treatment. The latter is a statistical approach to an open system at the mesoscopic level and is an approximated description.

Analytical approximation approaches, such as the scaling theory,[10] WKB analysis,[12] and normal mode analysis[41] are extremely useful for obtaining approximate solutions under certain conditions. For example, the interesting aspect associated with the long time behavior of a system, determined by the occurrence of exponentially small eigenvalues, can be very accurately analyzed by the WKB method.[12] In contrast, most numerical methods encounter difficulties in treating exponentially small eigenvalues. Ryskin[24] has given an analytical procedure for correcting equations of evolution, including the Markov processes.

The purpose of the present paper is to introduce a discrete singular convolution (DSC) algorithm for solving the Fokker–Planck equation. The underlying mathematical structure of such an algorithm is the theory of distributions which was informally used by physicists and engineers, and was later presented in rigorous mathematical form by Schwartz,[42] Korevaar,[43,44] and others. There are three parallel descriptions for the theory of distributions. One description is to characterize them as an equivalence class, or generalized limit of various Cauchy sequences (fundamental sequences) and fundamental families as rigorously defined by Korevaar.[43] This approach is particularly convenient for the delta distribution. Another description is to formulate them as continuous linear functionals on the space of test functions as introduced by Schwartz.[42] The vector space of test functions is obtained from a class of test functions with compatible convergence or topology. The third description is based on generalized derivatives of integrable functions. Generalized derivatives are distributions rather than well-behaved functions. The first description is intuitive and convenient for various applications. The functional description is particularly elegant and concise. It is also very convenient for higher dimensional applications. The third description is useful for certain practical applications involving derivatives and antiderivatives. These three methods are formally equivalent and are commonly used for theoretical descriptions in general. The present work explores computational aspects of the distribution theory, namely a discrete singular convolution algorithm. Numerical tests indicate that present approach is extremely efficient and robust for solving the Fokker–Planck equation describing various physical phenomena.

This paper is organized as the follow. The formalism of discrete singular convolution is introduced in Sec. II. Two types of singular kernels, namely, singular kernels of the Hilbert type and the delta type, are discussed for the point of view of distributions. Sequences of approximations to these distributions are presented. A computational algorithm is proposed to realize the concept of the discrete singular convolution. Numerical techniques regarding discretization, truncation, boundary, and matrix representation of operators are discussed in detail. The application of the present approach is illustrated by numerically solving the Fokker–Planck equation in Sec. III. Important models such as the Lorentz gas, and the quartic potential system are utilized to test the present algorithm. Another important benchmark problem, the Henon–Heiles system, is also included to illustrate the present approach. Conclusions are given in Sec. IV.

II. DISCRETE SINGULAR CONVOLUTION

Singular convolutions appear in many science and engineering problems. Discrete singular convolution is a general approach for numerically solving singular convolution problems. By appropriate realizations of a singular convolution kernel, the discrete singular convolution can be an extremely efficient, accurate, and reliable algorithm for scientific computations. Computational philosophy is presented in the first subsection. A number of examples of approximation sequences are given in the second subsection. These are followed by a numerical algorithm describing the detailed implementation of the present approach.

A. Singular convolution

The simplest way to introduce theory of *singular convolution* (SC) is to work in the context of distributions. Let T be a distribution and $\eta(t)$ be an element of the space of test functions [e.g., $\eta(t) \in \mathcal{D}$]. A singular convolution is defined as

$$F(t) = (T * \eta)(t) = \int_{-\infty}^{\infty} T(t-x)\,\eta(x)\,dx. \tag{1}$$

Here $T(t-x)$ is a singular kernel. Depending on the form of the kernel T, the singular convolution is the central issue for a wide range of science and engineering problems. For example, singular kernels of the Hilbert-type have a general form of

$$T(x)=\frac{1}{x^n}\ (n=1,2,\ldots). \tag{2}$$

Here, kernel $T(x)=(1/x)$ is commonly occurred in electrodynamics, molecular spectroscopy, theory of analytic functions and the Hilbert transform; $T(x)=(1/x^2)$, is the kernel used in tomography. Other interesting examples are singular kernels of the delta type

$$T(x)=\delta^{(n)}(x)\ (n=0,1,2,\ldots). \tag{3}$$

Here, kernel $T(x)=\delta(x)$ is important for interpolation of surfaces and curves (including atomic and molecular potential energy surface); and $T(x)=\delta^{(n)}(x)$, $(n=1,2,\ldots)$ are essential for numerically solving partial differential equations. Singular kernels of Abel-type

$$T(x)=\frac{1}{x^\beta}\ (0<\beta<1), \tag{4}$$

was introduced in the tautochrone problem. It has applications in the area of holography and interferometry with phase objects and is of practical importance in aerodynamics, heat and mass transfer, and plasma diagnostics. Since these three types of kernels are singular, they cannot be directly digitalized in computer. Hence, the singular convolution, Eq. (1), is of little numerical merit. To avoid the difficulty of using singular expressions directly in computer, sequences of approximations (T_α) of the distribution T can be constructed

$$\lim_{\alpha\to\alpha_0} T_\alpha(x)\to T(x), \tag{5}$$

where α_0 is a generalized limit. Obviously, in the case of $T(x)=\delta(x)$, the sequence, $T_\alpha(x)$, is a delta sequence.

It is interesting to note that the concept of approximation sequences is intimately related to wavelet idea of dilation. For example, a family of Shannon's (wavelet) scaling functions $\{T_\alpha(x)=(\sin\alpha x/\pi x)\}$ is a delta sequence

$$\lim_{\alpha\to\infty}<\frac{\sin\alpha x}{\pi x},\ \ \eta(x)>=\eta(0), \tag{6}$$

where η is a test function and $\langle\cdot,\ \cdot\rangle$ denotes the standard inner product. Therefore, Shannon's scaling functions provide approximations of the delta distribution. The advantage is that, unlike the original delta distribution, a delta sequence is well behaved. As a consequence, $\eta(x)$ is actually no longer required to be an element of the space of test functions. Most importantly, with a sufficiently smooth approximation, it makes sense to consider a *discrete singular convolution* (DSC),

$$F_\alpha(t)=\sum_k T_\alpha(t-x_k)f(x_k), \tag{7}$$

where $F_\alpha(t)$ is an approximation to $F(t)$ and $\{x_k\}$ is an appropriate set of discrete points on which the DSC (7) is well defined. Note that, the original test function $\eta(x)$ has been replaced by $f(x)$. The mathematical property or requirement of $f(x)$ is determined by the approximate kernel T_α. In particular, if T_α is Shannon's sampling kernel $[T_\alpha(x)=(\sin\alpha x/\pi x)$ and $\alpha<\infty]$, then the requirement for $f(x)$ is relaxed to a band-limited L^2 function. In general, the convolution defined by the pair of functions is required being Lebesque integrable.

A sequence of approximation can be improved by a regularizer

$$\lim_{\sigma\to\infty} R_\sigma(x)=1. \tag{8}$$

The role of a regularizer is to increase the regularity of convolution kernels. For the delta sequence, it follows from Eq. (5) that

$$\int \lim_{\alpha\to\alpha_0} T_\alpha(x)R_\sigma(x)dx=R_\sigma(0)=1, \tag{9}$$

where $R_\sigma(0)=1$ is the special requirement for a *delta regularizer*.

B. Examples

1. Sequences for singular kernels of the Hilbert-type

The Hilbert transform links between the imaginary part and the real part of an analytic function. It has important applications in a wide variety of science and engineering fields mostly associated with causality, such as in the linear response theory and the concept of analytic signals. The Hilbert transform is defined by a singular convolution

$$\mathcal{H}[f](x)\equiv\frac{1}{\pi}P\int_{-\infty}^{\infty}\frac{f(x-y)}{y}dy,\quad f\in L^2(R)\cap C(R), \tag{10}$$

where P denotes taking the Cauchy principal value and $L^2(R)$ denotes the space of square integrable functions on the real axis R with the standard $\|\cdot\|_2$ norm. Finally, $C(R)$ denotes the set of all continuous and bounded functions on R. Note that the resulting function of the Hilbert transform again belongs to $L^2(R)$. The Hilbert kernel, H, can be uniformly approximated by a Hilbert sequence H_α,

$$\begin{aligned}H(x)&=\frac{1}{\pi}\frac{1}{x}=\lim_{\alpha\to\infty}H_\alpha(x)\\&=\lim_{\alpha\to\infty}\frac{1}{\pi}\frac{1-\cos(\alpha x)}{x},\end{aligned} \tag{11}$$

where $H_\alpha(x)$ is a well-behaved except for the case of $\alpha=\alpha_0$. Asymptotically, the Hilbert sequence decays slowly and the approximation can be improved by using a regularizer

$$H_{\alpha,\sigma}(x)=H_\alpha(x)R_\sigma(x)=\frac{1}{\pi}\frac{1-\cos(\alpha x)}{x}R_\sigma(x), \tag{12}$$

where $R_\sigma(x)$ satisfies Eq. (8). Another possible restriction on a Hilbert regularizer is $R_\sigma(\pi/\alpha)=1$. A good choice of the Hilbert regularizer is $R_\sigma(x)=e^{-[(|x|-\pi/\alpha)^2/2\sigma^2]}$. The derivatives of the Hilbert transform, $(d^q/dx^q)\mathcal{H}$ can be expressed as derivatives of the Hilbert kernel, $(d^q/dx^q)H_{\alpha,\sigma}(x)$, in the convolution. In fact, all other Hilbert-type singular kernels, $H_n\,(n=2,3,\ldots)$, can be expressed as the derivatives of $(1/\pi)(1/x)$,

$$H_n(x)=\frac{1}{\pi}\frac{1}{x^n}$$
$$=\frac{(-1)^{n-1}}{(n-1)!}\frac{d^{n-1}}{dx^{n-1}}H(x)$$
$$=\lim_{\alpha\to\infty}\lim_{\sigma\to\infty}\frac{(-1)^{n-1}}{(n-1)!}\frac{d^{n-1}}{dx^{n-1}}H_{\alpha,\sigma}(x)$$
$$=\lim_{\alpha\to\infty}\lim_{\sigma\to\infty}\frac{(-1)^{n-1}}{(n-1)!}\frac{d^{n-1}}{dx^{n-1}}\frac{1}{\pi}\frac{1-\cos(\alpha x)}{x}R_\sigma(x). \tag{13}$$

2. Sequences for singular kernels of the delta type

The delta distribution or so-called Dirac delta function δ is a generalized function which follows from the fact that it is an integrable function inside a particular interval but itself needs not to have a value. Heaviside introduced both the unit step Heaviside function and the Dirac delta function as its derivative and referred to the latter as the *unit impulse*. Dirac, for the first time, explicitly discussed the properties of δ in his classic text on quantum mechanics; for this reason δ is often called Dirac delta function. However, delta distribution has a history which antedates both Heaviside and Dirac. It appeared in explicit form as early as 1822, in Fourier's *Théorie Analytique de la Chaleur*. The work of Heaviside, and subsequently of Dirac, in the systematic but informal exploitation of the step function and delta function, has made delta distribution familiar to physicists and engineers before Sobolev, Schwartz,[42] Korevaar,[43] and others put it into a rigorous mathematical form. The Dirac delta function is the most important special case of distributions or generalized functions.

In particular, the Hermite function expansion of Dirac delta function was proposed by Schwartz[42] and Korevaar[43] over 40 years ago and was introduced by Hoffman *et al.* for numerical simulations.[38] General analysis of the delta distribution by means of orthogonal series has been studied by Walter[45] and others.[46–48] The use of many delta sequences as probability density estimators was discussed by Walter and Blum[48] and others.[47,49,50]

Example 1: Dirichlet's delta sequence.

The most important example of a delta sequence of the Dirichlet-type is Dirichlet's delta sequence

$$\delta_\alpha(x)=\begin{cases}D_\alpha(x) & \text{for } |x|\leqslant\pi \quad \text{for } \alpha=0,1,2,\ldots\\ 0 & \text{otherwise}\end{cases}, \tag{14}$$

where D_α is the Dirichlet kernel

$$D_k(x)=\frac{1}{\pi}\left[\frac{1}{2}+\cos(x)+\cos(2x)+\cdots+\cos(kx)\right]$$
$$=\frac{\sin[(k+\frac{1}{2})x]}{2\pi\sin\frac{1}{2}x}, \quad k=0,1,2,\ldots. \tag{15}$$

Dirichlet's delta sequence plays an important role in approximation theory and is the key element in trigonometric polynomial approximations. Physically, it describes the diffraction of light passing a regular series of pinholes in which the kth pinhole's contribution is proportional to e^{ik}.

Example 2: Modified Dirichlet's delta sequence.

Sometimes there is a slight advantage in taking the last term in D_α with a factor of $\frac{1}{2}$,

$$D^*_\alpha(x)=D_\alpha-\tfrac{1}{2}\cos(\alpha x)$$
$$=\frac{\sin(\alpha x)}{2\pi\tan(\frac{1}{2}x)}, \quad \alpha=0,1,2,\ldots. \tag{16}$$

This is the so-called modified Dirichlet kernel. The difference $D_\alpha-D^*_\alpha$ tends uniformly to zero on $(-\pi,\ \pi)$ as $\alpha\to\infty$. They are equivalent with respect to convergence.

The expression given by

$$\delta_\alpha(x)=\begin{cases}D^*_\alpha(x) & \text{for } |x|\leqslant\pi a \quad \text{for } \alpha=0,1,2,\ldots\\ 0 & \text{otherwise}\end{cases} \tag{17}$$

is a delta sequence of the Dirichlet type as $\alpha\to\infty$.

Example 3: Shannon's delta sequence.

Shannon's delta sequence or Dirichlet's continuous delta sequence is given by the following Fourier transform of the characteristic function, $\chi_{[-\alpha,\alpha]}$,

$$\delta_\alpha(x)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\chi_{[-\alpha,\alpha]}e^{-i\xi x}d\xi$$
$$=\frac{\sin(\alpha x)}{\pi x}. \tag{18}$$

This converges to the delta distribution as $\alpha\to\infty$. Equation (18) is related to Shannon's sampling theory in the information theory and theory of sampling. It actually forms an orthogonal basis for a reproducing kernel Hilbert space. Shannon's sampling kernel can be derived from the generalized Lagrange interpolating formula. Alternatively, Shannon's delta sequence can be given as an integration

$$\delta_\alpha(x)=\frac{1}{\pi}\int_0^\alpha\cos(xy)dy, \tag{19}$$

or as the limit of a continuous product

$$\delta_\alpha(x)=\lim_{N\to\infty}\frac{\alpha}{\pi}\prod_{k=1}^{N}\cos\left(\frac{\alpha}{2^k}x\right)=\lim_{N\to\infty}\frac{1}{2^N\pi}\frac{\sin(\alpha x)}{\sin\left(\frac{\alpha}{2^N}x\right)}. \tag{20}$$

To improve the asymptotic behavior of Shannon's sampling kernel, regularized Lagrange interpolating kernel, and regularized Shannon's sampling kernel were proposed recently.[51,52]

Example 4: The de la Vallée Poussin delta sequence.

The de la Vallée Poussin kernel is given by

$$P_{n,p}(x)=\frac{1}{p+1}\sum_{k=n-p}^{n}D_k(x)$$
$$=\frac{1}{2\pi}+\frac{1}{\pi}\sum_{k=1}^{n-p}\cos kx$$
$$+\frac{1}{\pi}\sum_{k=1}^{p}\left[1-\frac{k}{p+1}\right]\cos[(n-p+k)x]$$
$$=\frac{\sin\left[(2n+1-p)\frac{x}{2}\right]\sin\left[(p+1)\frac{x}{2}\right]}{2\pi(p+1)\sin^2\left(\frac{x}{2}\right)}, \tag{21}$$

$$p=0,\ldots,n;\quad n=0,1,\ldots, \tag{22}$$

where $D_k(x)$ are Dirichlet kernels given by Eq. (15). It is interesting to note that de la Vallée Poussin kernels reduce to Ferér's positively definite kernel when $p=n$. The de la Vallée Poussin delta sequence is given by

$$\delta_{n,p}(x)=\begin{cases}P_{n,p}(x) & \text{for } |x|\leqslant x \text{ for } p=0,\ldots,n;\ n=0,1,\ldots,\\ 0 & \text{otherwise}\end{cases} \tag{23}$$

as $n,\ p\to\infty$. The de la Vallée Poussin delta sequence is of the Dirichlet type when $p<n$.

A simplified de la Vallée Poussin kernel given by

$$\delta_\alpha(x)=\frac{1}{\pi\alpha}\frac{\cos(\alpha x)-\cos(2\alpha x)}{x^2} \tag{24}$$

is very useful numerically. Expression (24) is used in our further discussion.

C. Computational aspects

To use approximation sequences for certain computations, a numerical algorithm is required. In fact, such an algorithm is not unique. Different algorithms can be proposed and they can be improved. The purpose of this subsection is to provide a simple and efficient numerical approach. Various computational aspects are discussed in this subsection.

1. Discretization

For a given kernel, there are many possible ways of discretizations. Important examples are interpolating, quasi-interpolating, and noninterpolating discretizations. Moreover, the grid in each discretization can be either uniform or nonuniform. Since uniform discretization can lead to a single integration kernel on a grid, it is simple, robust and efficient. In a nonuniform discretization, the kernel must vary on a grid. This produces no problem for a global method in which the grid and the kernel are systematically determined by zeros of the highest degree polynomial. However, it can be very inconvenient for a local method to be discretized nonuniformly because the lack of a simple procedure to prescribe the kernel over the whole domain with high computational accuracy.

It seems to us that for sequences of the delta type and the Hilbert-type, an interpolating (or quasi-interpolating) algorithm sampling at *Nyquist frequency*, $\alpha=(\pi/\Delta)$, has great advantage over a noninterpolating discretization. Not only the interpolating (or quasi-interpolating) nature guarantees the highest accuracy on the set of grid points, but it also provides the highest possible computational efficiency off a grid. This is because the *Nyquist interval* given by $[-(\pi/\Delta),(\pi/\Delta)]$ is the largest possible sampling interval that is free of alias whenever the L^2 function $f(x)$ under study satisfies the *Nyquist condition*,

$$\operatorname{supp}\hat{f}(k)\subset\left\{-\frac{\pi}{\Delta},\frac{\pi}{\Delta}\right\}. \tag{25}$$

This fact can actually be phrased in the following Shannon's sampling theorem[53]

$$f(x)=\sum_{k=-\infty}^{\infty}f(x_k)\frac{\sin\frac{\pi}{\Delta}(x-x_k)}{\frac{\pi}{\Delta}(x-x_k)}. \tag{26}$$

The significance of Shannon's sampling theorem is that by a discrete, but infinite set of sampling data $\{f(x_k)\}$ one can actually recover a band-limited L^2 function on a real line. This is particularly significant for the information theory. Shannon's sampling theorem also has great impact on signal and image processing because the Fourier transform of Shannon's delta sequence kernel is an ideal low-pass filter for signals band-limited to $[-(\pi/\Delta),(\pi/\Delta)]$.

Another important aspect of Shannon's delta sequence kernel is that it is a reproducing kernel

$$\frac{\sin\frac{\pi}{\Delta}(x-y)}{\frac{\pi}{\Delta}(x-y)}=\sum_{n=-\infty}^{\infty}\frac{\sin\frac{\pi}{\Delta}(x-x_n)}{\frac{\pi}{\Delta}(x-x_n)}\frac{\sin\frac{\pi}{\Delta}(y-x_n)}{\frac{\pi}{\Delta}(y-x_n)}. \tag{27}$$

It can be used to generate an orthonormal basis for the Paley–Wiener reproducing kernel Hilbert space $\mathbf{B}_\alpha^2$. A form which is more useful for a quadrature filter representation is

$$\frac{\sin\pi x}{\pi x}=\sum_{n=-\infty}^{\infty}\frac{\sin\frac{\pi}{2}x_n}{\frac{\pi}{2}x_n}\frac{\sin\pi(2x-x_n)}{\pi(2x-x_n)}. \tag{28}$$

Shannon's delta sequence kernel is also related to the earlier Whittaker's cardinal series[54,55]

$$\frac{(-1)^n\sin\frac{\pi}{\Delta}x}{\frac{\pi}{\Delta}x-n\pi}=\frac{\sin\frac{\pi}{\Delta}(x-x_n)}{\frac{\pi}{\Delta}(x-x_n)}. \tag{29}$$

Whittaker's cardinal series has played an extremely important role in the development of mathematical theory of sampling. Moreover, the Nth power of Shannon's kernel $(\sin\pi x/\pi x)^N$ is a B-spline of order N in its Fourier representation. In this sense, various B-spline approximations, which

are widely used in applied mathematics and engineering, are related to various reproducing kernels and theory of delta distribution.

We shall also use the *Nyquist rate*, uniform interpolation for the discretization of Dirichlet delta sequence kernel,

$$\frac{\sin\left[\left(l+\frac{1}{2}\right)(x-x')\right]}{2\pi\sin\left[\frac{1}{2}(x-x')\right]}\to\frac{\sin\left(\frac{\pi}{\Delta}(x-x_k)\right)}{(2M+1)\sin\left(\frac{\pi}{\Delta}\frac{x-x_k}{2M+1}\right)}, \quad (30)$$

In a comparison to the Shannon's delta sequence kernel, Dirichlet's delta sequence kernel has one more parameter M which can be optimized to achieve better results in computations. Usually, we set a sufficiently large M for various numerical applications. Obviously, Dirichlet's delta sequence kernel converts to Shannon's delta sequence kernel at the limit of $M\to\infty$. This uniform interpolating discretization will also be used for discreting modified Dirichlet's delta sequence kernels

$$\frac{\sin\left[\left(l+\frac{1}{2}\right)(x-x')\right]}{2\pi\tan\left[\frac{1}{2}(x-x')\right]}\to\frac{\sin\left(\frac{\pi}{\Delta}(x-x_k)\right)}{(2M+1)\tan\left(\frac{\pi}{\Delta}\frac{x-x_k}{2M+1}\right)}, \quad (31)$$

and for the de la Vallée Poussin delta sequence kernels

$$\frac{1}{\pi\alpha}\frac{\cos[\alpha(x-x')]-\cos[2\alpha(x-x')]}{(x-x')^2}\to\frac{2}{3}\frac{\cos\frac{\pi}{\bar{\Delta}}(x-x_k)-\cos\frac{2\pi}{\bar{\Delta}}(x-x_k)}{\left[\frac{\pi}{\bar{\Delta}}(x-x_k)\right]^2}, \quad (32)$$

where $\bar{\Delta}=\frac{3}{2}\Delta$ is required by interpolation. The efficiency and accuracy of these discretizations will be numerically tested in the next section.

Finally the sequences of the Hilbert type is similarly discretized as

$$\frac{(-1)^{n-1}}{(n-1)!}\frac{d^{n-1}}{dx^{n-1}}\frac{1}{\pi}\frac{1-\cos(\alpha x)}{x}R_\sigma(x)\to\frac{(-1)^{n-1}}{(n-1)!}\frac{d^{n-1}}{dx^{n-1}}\frac{1}{\pi}\frac{\sin^2\frac{\pi}{2\Delta}(x-x_k)}{\frac{\pi}{2\Delta}(x-x_k)}R_\sigma(x-x_k). \quad (33)$$

Hence transforms of the Hilbert-type for a continuous L^2 function band-limited to (π/Δ) are given by the following generalized Hilbert sampling theorem:

$$\mathcal{H}_n[f](x)\equiv\frac{1}{\pi}P\int_{-\infty}^{\infty}\frac{f(x-y)}{y^n}dy = \frac{(-1)^{n-1}}{(n-1)!}\frac{1}{\pi}\sum_{k=-\infty}^{\infty}f(x_k)\frac{d^{n-1}}{dx^{n-1}}\frac{\sin^2\frac{\pi}{2\Delta}(x-x_k)}{\frac{\pi}{2\Delta}(x-x_k)}. \quad (34)$$

The case of $n=1$ can be found in Ref. 56. A proof of this theorem is beyond of the scope of the present paper and is presented elsewhere. This generalized Hilbert sampling theorem has potential for using in many science and engineering problems. Numerical applications of Eq. (34) and its regularized forms are described elsewhere.

Since π/Δ is proportional to the highest frequency can be reached in the Fourier representation, Δ should be very small for a given problem involving very oscillatory functions or very high frequency components.

2. Truncation

It is obvious that the aforementioned approximation sequences are formally defined in the unbounded domain $(-\infty,\infty)$. Hence, the corresponding convolution kernels have a noncompact support. Practically, it is impossible to implement a unbounded domain in a computer. Therefore, it is necessary for numerical purpose to restrict the support of the kernels to a finite region. In practical, this can be achieved either by a conformal mapping or by a spatial truncation of the convolution kernel. Conformal mapping requires a change in the governing equation under study. In comparison, truncation is a very simple approach. The manner of kernel truncation can dramatically affect the numerical performance of the kernel (from the point of view of mathematics, a kernel is not well defined until its domain, range and truncation are all specified). There are two types of truncations. One is *boundary adaptive* truncations. The other is *translationally invariant* truncations. The former can in principle provide a scheme for handling complex boundary. It has an advantage that only information (or grid values) on the set of grid points in the computational domain is ever required for the convolution. No additional boundary extension is needed. A problem is that in order to satisfy boundary condition or geometry requirement, it is necessary for the resulting kernel to vary near the computational boundary which implies more computations and may reduce numerical accuracy. In contrast, a translationally invariant truncation leads to just one kernel in the whole computational domain, which is numerically cheap and efficient. Better numerical efficiency can by achieved by making the kernel *symmetric* (or *antisymmetric*). This leads to a translationally invariant algorithm

$$f^{(n)}(x)\approx\sum_{k=-W}^{W}\delta_\alpha^{(n)}(x-x_k)f(x_k)\quad(n=0,1,2,...), \quad (35)$$

where $2W+1$ is the computational bandwidth, or effective kernel support, which is usually smaller than the whole computational domain, $[a,b]$. Here $\delta_\alpha^{(n)}(x-x_k)$ denotes for the

nth derivative of any of Eqs. (29), (30), (31), and (32). Furthermore, in order for the boundary condition (or complex geometry) to be met, it requires appropriate handling of the function, $f(x_k)$, *outside* the computational domain.

3. Boundary

A complete numerical algorithm has to provide a scheme for handling boundaries. In a global method, its computational domain is the same as its kernel support. In order to satisfy boundary requirements, the kernel must be constructed in adaptive manner. Hence, the kernel varies at different grid points. The major drawback of such an algorithm is its difficulty in the construction of adaptive kernels for problems involving complex geometries and boundary conditions. Thus, global methods have been relatively unsuccessful for dealing with these problems comparing to their achievements in solving problems of simple geometries and boundary conditions. Whereas most local methods use an entirely different philosophy for their kernel construction. For instance, a differentiation kernel in a finite difference method can be the same everywhere and it is translationally invariant on the grid. Local methods are very flexible for dealing with boundary and geometry.

In Eq. (35), if the kernel, $\delta_\alpha^{(n)}(x-x_k)$, is fixed to be symmetric (or antisymmetric) and translationally invariant, there must be cases where $f(x_k)$ are located outside of the computational domain, $[a,b]$, and their values are undefined there. In the present algorithm, such $f(x_k)$ are to be obtained by boundary conditions. For examples, in the Fokker–Planck equation, the natural boundary condition is used and $f(x_k)$ outside the domain $[a,b]$ are set identically to zero. In the Dirichlet boundary condition, such $f(x_k)$ are taken to be $f(a)$ (or $f(b)$). In periodic boundary condition, such $f(x_k)$ are replaced by their corresponding values inside the domain $[a,b]$. In the Neumann boundary condition, such $f(x_k)$ are determined by $f(a)$ and $f'(a)$ [or $f(b)$ and $f'(b)$]. There are special cases, such as in Burgers' equation, where $f(x)$ is antisymmetric around the boundary point and such $f(x_k)$ are replaced by their corresponding $f(x_k)$ inside the domain $[a,b]$ with a negative sign. Symmetric extensions of $f(x_k)$ are used if boundaries are symmetric planes.

4. Discrete representation of operators

In numerically solving the Fokker–Planck equation, it is necessary to give a matrix representation to an operator so that the action of the operator can be numerically realized. Property of an operator matrix is crucial for numerical accuracy and efficiency. Choosing a numerically appropriate matrix representation for an operator is quite complicated. It involves not only the property of the operator, the initial value and boundary conditions, but also the numerical scheme to be utilized. From the computational point of view, the most important property of an operator matrix is its effective boundness, diagonality, and symmetry. The effective boundness of an operator often depends on the initial value and/or boundary conditions. To a desired level of accuracy, the truncated operator matrix should cover the effective energy range of the system. Operator symmetry is useful for simplifying a matrix representation and reducing computer memory required. Numerically, it is simple and convenient to have a diagonal matrix representation of an operator. This, however, often cannot be achieved. For example, the Hamiltonian of the Schrödinger equation often has its potential part being diagonal in the position representation and its kinetic energy part being diagonal in the momentum representation. In principle, there is a basis representation in which a given Hamiltonian is fully diagonal. However, finding this representation is equivalent to solving eigenstates of the Hamiltonian system. In the DSC approach we choose a grid representation for the coordinate so that the potential part, $V(x)$, of the Hamiltonian is diagonal. Hence, we choose a direct interpolation on the grid for its discretization

$$V(x)\rightarrow V(x_k)\,\delta_{m,k}. \tag{36}$$

The differentiation matrix of the Hamiltonian on the coordinate grid is then given in terms of the distribution derivative

$$-\frac{\hbar^2}{2\mu}\frac{d^2}{dx^2}\rightarrow-\frac{\hbar^2}{2\mu}\,\delta_\alpha^{(2)}(x_m-x_k), \tag{37}$$

where μ is the mass of the Hamiltonian system and $\delta_\alpha^{(2)}\times(x_m-x_k)$ are *analytically* given by

$$\delta_\alpha^{(2)}(x_m-x_k)=\left[\left(\frac{d}{dx}\right)^2\delta_\alpha(x-x_k)\right]_{x=x_m}. \tag{38}$$

Thus, the full DSC grid representation for the Hamiltonian operator, H, is given by

$$H(x_m,x_k)=-\frac{\hbar^2}{2\mu}\,\delta_\alpha^{(2)}(x_m-x_k)+V(x_m)\,\delta_{m,k}. \tag{39}$$

This expression is referred as a Hamiltonian-DSC matrix element. Although the Hamiltonian is used as an example, any other operator which consists of a nondiagonal differentiation part and/or a diagonal part can be treated similarly. Note that unlike its global method counterpart, the present matrix representation is banded. The bandedness is particularly important to large scale computations. Obviously, it is unnecessary to restrict the grid to the coordinate grid; it can be the momentum grid or any other appropriate grid (even for time discretization). This symmetric (or antisymmetric), translation invariant algorithm will be used in the next section for solving the Fokker–Planck equation.

III. APPLICATIONS

In the present study we limit our attention to the delta sequence kernels of Shannon (Shannon), the de la Vallée Poussin (DLVP), Dirichlet (Dirichlet), and the modified Dirichlet (MD). Nevertheless, various other delta sequence kernels can be similarly employed. It is noted that the delta sequence kernels of Shannon (Shannon) and de la Vallée Poussin (DLVP) are parameter-free, which is an important and convenient feature for applications, in particular for unexperienced users. The $2M+1$ parameter used for the Dirichlet and modified Dirichlet kernels is chosen as 71 for all calculations. As long as the $2M+1$ value is chosen suffi-

ciently large ($2M+1>W$, where $2W+1$ is the matrix bandwidth), the numerical results are not sensitive to the specific values used.

The goal of the present section is to test the DSC algorithm for the solutions of the Fokker–Planck equation. The DSC method can been applied to the Fokker–Planck equation in many different ways. Two possible approaches are direct time-dependent treatment and eigenfunction expansion (see Appendix). The latter is used in this paper. The Fokker–Planck-DSC matrix, whose elements are of the form of expression (39), is directly diagonalized to generate eigenfunctions and eigenvalues. These can be used to determine the dynamics of the Fokker–Planck system as described in the Appendix. The Fokker–Planck equation treated in this work is of the common form[11]

$$\frac{\partial f(x,t)}{\partial t}=\frac{\partial}{\partial x}\left[A(x)+\frac{\partial}{\partial x}B(x)\right]f(x,t), \tag{40}$$

where $A(x)$ is a drift coefficient and $B(x)$ a diffusion coefficient. Note that both A and B can be nonlinear in the variable x. The distribution function $f(x,t)$ is normalized to one and its dynamics is governed by Eq. (40). The present results, which are in terms of an eigenfunction expansion, are compared with those of other established methods in the field. It is found that the DSC algorithm can provide results which are better than, or at least equivalent to those of previous global spectral methods in terms of accuracy and speed of convergence. However, the DSC algorithm is considerably simpler than global spectral approaches.

For the convenience of the description and completeness, the formalisms of the Fokker–Planck equation, particularly the formal relation between the FPE and the Schödinger equation, is reviewed in the Appendix. The theory of the Fokker–Planck equation has been extensively studied in the last a few decades. The formalism of the eigenfunction expansion of the Fokker–Planck equation also has been discussed by a number of authors.[11,36] We select only those details that are most relevant for our purposes. This section is divided into three subsections; Sec. III A is devoted to the treatment of the Lorentz gas. The benchmark bistable model is numerically studied in Sec. III B. Finally, the Henon–Heiles system is treated in Sec. III C.

A. Lorentz gas

The Lorentz gas model is an interesting problem representing the kinetic motion of a typical light particle in a heat bath of heavy particles. The interaction of two types of particles is modeled by the binary hard sphere collision processing. The Lorentz gas model is a standard example of the Fokker–Planck equation derived from more general theory of the Boltzmann equation by using the hard sphere collision assumption. This model has been applied to thermalization of electrons in a gaseous matrix.[57] The eigenvalue problem of the Lorentz gas has been studied by many authors recently using various methods, such as the Rayleigh–Ritz variational procedure,[58,59] the Wigner–Wilkins kernel,[60] and the Sonine polynomial expansion.[61,62] Particularly, Shizgal and Chen[36] have reported the quadrature discretization method (QDM) calculation employing a nonclassic weight function. They reported very rapid convergence with the first 40 eigenvalues accurately calculated by using 70 or fewer QDM points.

TABLE I. Eigenvalues of the Lorentz gas Fokker–Planck operator.

	Shizgal *et al.* ($N=70$)	Shannon ($N=60$)	Dirichlet ($N=60$)	MD ($N=60$)	DLVP ($N=90$)
λ_1	4.683 40	4.683 395	4.683 395	4.683 395	4.683 394
λ_2	10.112 52	10.112 52	10.112 52	10.112 52	10.112 52
λ_3	16.429 68	16.429 68	16.429 68	16.429 68	16.429 67
λ_6	40.052 38	40.052 38	40.052 38	40.052 38	40.052 30
λ_{10}	80.447 94	80.447 94	80.447 94	80.447 94	80.447 55
λ_{15}	142.446 1	142.446 1	142.446 1	142.446 1	142.444 6
λ_{20}	215.163 1	215.162 7	215.162 7	215.162 8	215.159 2
λ_{30}	387.623	387.616 6	387.616 6	387.616 7	387.607 0
λ_{40}	590.867	590.822 9	590.821 6	590.824 3	590.839 1

The corresponding Fokker–Planck equation of Lorentz gas is given by

$$\frac{\partial f(x,t)}{\partial t}=\frac{\partial}{\partial x}[(2x^2-3)f(x,t)]+\frac{\partial^2}{\partial x^2}[xf(x,t)]. \tag{41}$$

The effective potential $V(z)$ for this system is

$$V(z)=\frac{z^6}{64}-z^2+\frac{15}{4z^2}. \tag{42}$$

[Note the functional relation between z and x is given by Eq. (A17) in the Appendix.] Mathematically, this potential is bounded below but it has singularities both at zero and at infinity. The operator domain for the problem is $[0,\infty)$. Shizgal and Chen[36] noted that the computational interval has to be sufficiently large in order for the large eigenvalues to be convergent.

In the present work we examine the performance of the DSC algorithm by using delta sequence kernels of Shannon (Shannon), Dirichlet (Dirichlet), modified Dirichlet (MD) and the de la Vallée Poussin (DLVP). The computations utilizing the first three delta sequence kernels are conducted using 60 grid points in z and are found to be in excellent agreement with those given by Shizgal and Chen[36] obtained by using up to 70 QDM points. These results for the first 40 eigenvalues are listed in Table I together with those of Shizgal and Chen.[36] We found that the de la Vallée Poussin (DLVP) delta sequence kernel requires 1.5 times of grid points ($N=90$) to achieve similar accuracy.

B. Bistable system

To demonstrate further the reliability and robustness of the DSC method for the Fokker–Planck equations, we consider the bistable system, a model that has significant theoretical applications and is numerically difficult for a certain parameter region. It has received a lot of attention in the literature recently. Van Kampen, Dekker, and van Kampen[8] have performed very detailed theoretical analysis and a finite difference scheme calculation of this system. Suzuki studied this system by using his scaling theory[10] approach. Risken examined the system by using a matrix continued fraction technique.[11] Caroli *et al.* consider this system as a test model for the WKB treatment of the Fokker–Planck equation.[12] In

TABLE II. Eigenvalues of the Fokker–Planck operator for a bistable system ($\epsilon=0.1$).

	Shizgal *et al.* ($N=60$)	Shannon ($N=42$)	Dirichlet ($N=42$)	MD ($N=42$)	DLVP ($N=63$)
λ_1	3.354 530 0(−2)	3.354 528 7(−2)	3.354 528 7(−2)	3.354 528 7(−2)	3.354 528 7(−2)
λ_2	0.927 372	0.927 372	0.927 372	0.927 372	0.927 372
λ_3	1.680 264	1.680 264	1.680 264	1.680 264	1.680 264
λ_5	3.733 985	3.733 985	3.733 985	3.733 985	3.733 985
λ_{10}	11.687 442	11.687 441	11.687 441	11.687 441	11.687 441
λ_{15}	22.639 908	22.639 909	22.639 909	22.639 909	22.639 909
λ_{20}	36.031 815	36.031 785	36.031 785	36.031 785	36.031 787
λ_{25}	51.541 9	51.535 92	51.535 92	51.535 92	51.536 13

their systematic study, they concluded that the final approach to equilibrium is governed by the Kramers high-viscosity rate. Larson and Kostin[14] conducted a formal analysis of this system from the point of view of chemical kinetics. Indira *et al.*[15] obtained a numerical solution for the system using both finite-element and Monte Carlo methods. Blackmore and Shizgal[18] have applied their QDM approach to the model and given a detailed numerical analysis for various physical behaviors of the system. Recently Shizgal and Chen[37] have designed a number of new nonclassic QDM weight functions to attain a superior rate of convergence for this problem. The purpose of the present study is to demonstrate that the DSC algorithm with certain interpolating delta sequence kernels can achieve an extremely high speed of convergence.

The Fokker–Planck equation describing a typical bistable system is given by

$$\frac{\partial f(x,t)}{\partial t}=-\frac{\partial}{\partial x}(\gamma x-gx^3)f(x,t)+\epsilon\frac{\partial^2}{\partial x^2}f(x,t). \tag{43}$$

Here the parameters γ, g, and ϵ are positive and are related to one another through the fluctuation-dissipation theorem at equilibrium. However, when the system is far from equilibrium, γ, g and ϵ become independent of each other. For simplicity and a comparison with previous work, we set $\gamma=g=1$ in the present work. The corresponding effective potential for Eq. (43) (see Appendix) is

$$V(z)=\frac{(z^3-z)^2}{4\epsilon^2}-\frac{1}{2\epsilon}(3z^2-1), \tag{44}$$

where the size of ϵ characterizes the physical regime described in the problem. For ϵ sufficiently small, one has a triple-well potential which supports three isolated, approximately harmonic systems at low energy. These triple wells lead to nearly threefold degenerated states which are very sensitive to computational intervals. For intermediate ϵ, the potential has three shallow wells coupled to each other with the three minima at the points[18]

$$z=0,\ \ \pm\left[\tfrac{2}{3}+(\tfrac{1}{9}+2\epsilon)^{1/2}\right]^{1/2}. \tag{45}$$

This potential supports nearly twofold degenerated states. For sufficiently large ϵ, the potential transforms into a double-well type with a maximum at the origin and two minima at $z=\pm\left[\frac{2}{3}+(\frac{1}{9}+2\epsilon)^{1/2}\right]^{1/2}$. The potential asymptotically behaves as z^6 and can obviously support infinitely many discrete states with positive energies. Formally, the problem is defined in the domain of $(-\infty,\infty)$. Computationally, a relatively large numerical interval is required when the size of ϵ is large.

In order to compare the present results with those obtained previously, we restrict the present calculations for the cases of $\epsilon=0.1$, 0.01. The delta sequence kernels of Shannon (Shannon), Dirichlet (Dirichlet), and modified Dirichlet (MD) are first employed in the present calculations. As usual, it takes about 1.5 times of grid points for the de la Vallée Poussin (DLVP) delta sequence kernel to reach the same level of accuracy. For the convenience of comparison, all eigenvalues in this subsection are measured in units of ϵ (so, $\lambda_n=\epsilon\varepsilon_n$).

In the case of $\epsilon=0.1$, the DSC calculations of Shannon, Dirichlet, and modified Dirichlet using only 42 grid points, yield the results of the first 25 eigenvalues that are in very good agreement with the converged results obtained by QDM using up to 60 grid points.[18] A comparison of these results are listed in Table II. The grid points required for the de la Vallée Poussin (DLVP) delta sequence kernel is 63.

The case of $\epsilon=0.01$ is significantly more difficult to calculate. The DSC results obtained by using four delta sequence kernels for the first 25 eigenvalues is listed in Table III. Similar calculations by Blackmore and Shizgal[18] using 100 QDM grid points are also included in Table III for reference. Except for the first excited state, all eigenvalues are converged up to 6 significant figures using only 52 grid points when the delta sequence kernels of Shannon, Dirichlet, and modified Dirichlet are employed. When ϵ decreases, the shape of the potential bottom become increasely unflat and irregular. As a result, the first few excited states are very sensitive to the number of grid points used and converge very slowly as noted by Blackmore and Shizgal. In general, a dramatical, nonmonotonic change in potential shape requires a very large number of grid points to describe accurately. Therefore, computationally, certain related eigenstates become very sensitive to the number of grid points used in the calculation. The first few eigenvalues approach zero as $\epsilon\to 0$. Shizgal and Chen[37] have recently obtained $\lambda_1=6.077(-8)$ with 60 QDM grid points associated with a Gaussian weight of $e^{-x^2/(2\epsilon)}$. This is compared to the present result of $\lambda_1=1.28(-8)$ with 52 DSC points of Shannon or Dirichlet or modified Dirichlet. A test calculation for $\epsilon=0.001$ indicates that significantly more grid points are needed to converge the first 50 eigenvalues. Since the

TABLE III. Eigenvalues of the Fokker–Planck operator for a bistable system ($\epsilon=0.01$).

	Shizgal *et al.* ($N=100$)	Shannon ($N=52$)	Dirichlet ($N=52$)	MD ($N=52$)	DLVP ($N=78$)
λ_1	6.154 650(−12)	1.278 247(−10)	1.276 095(−10)	1.282 279(−10)	1.271 301(−10)
λ_2	0.967 865	0.967 864	0.967 864	0.967 864	0.967 864
λ_3	1.864 542	1.864 542	1.864 542	1.864 542	1.864 542
λ_5	1.866 975	1.866 975	1.866 975	1.866 975	1.866 975
λ_{10}	3.943 531	3.943 531	3.943 531	3.943 531	3.943 531
λ_{15}	5.960 839	5.960 839	5.960 839	5.960 839	5.960 839
λ_{20}	8.793 163	8.793 146	8.793 147	8.793 147	8.793 146
λ_{25}	12.269 3	12.268 697	12.268 696	12.268 697	12.268 704

Fokker–Planck operator for the bistable potential is unbounded (but bounded below), it has infinity many discrete eigenvalues. The higher energy eigenstates exhibit the typical global behavior and the corresponding eigenvalues become more and more nearly degenerated. Computationally, these large eigenvalues generally converge more slowly and require more grid points for smaller ϵ values. In other words, to maintain certain level of accuracy for higher energy eigenvalues, the grid mesh size Δ should inversely proportional to the energies of the states $[(1/\Delta)\rightarrow 0 \text{ as } (\lambda\rightarrow\infty)]$.

C. The Henon–Heiles system

As the last testing problem for this section, we consider the benchmark Henon–Heiles system[63] to demonstrate the reliability and robustness of the DSC algorithm for calculating the eigenfunctions. Physically, the 2D anharmonic Henon–Heiles resonating system is one of the most important systems for chaotic analysis and it provides a simple and convenient example for the understanding and description of Poincare surfaces and trajectories. Numerically, it is a widely-used benchmark problem. A variety of numerical approaches have been tested on the Henon–Heiles potential system. Earlier work by Marcus *et al.*[64] used basis sets with 990 and 1225 basis states. Feit *et al.*[65] have reported an accurate calculation using their split operator-FFT method; that required 16 384 time steps. A recent computation by Shizgal and Chen[36] achieved six significant figure accuracy for eigenvalues as high as close to the disassociation limit, using 50 or fewer grid points in each dimension. In general, Shizgal's approach converges extremely fast since it can be optimized according to the problem under study.[37] Discrete variable representation[66] and its improved version[67] achieve a similar level of accuracy as Shizgal's quadrature discretization method (QDM). Similar level of accuracy was obtained by Zhang *et al.* by means of distributed approximating functionals.[39] The purpose of the present work is to demonstrate that an extremely high speed of convergence for this 2D system can be achieved by using the DSC algorithm with various interpolating kernels.

The govern equation for the Henon–Heiles system is given by

$$\left[-\frac{1}{2}\frac{\partial^2}{\partial x^2}-\frac{1}{2}\frac{\partial^2}{\partial y^2}+\frac{1}{2}x^2+\frac{1}{2}y^2-\lambda x\left(\frac{1}{3}x^2-y^2\right)\right]\psi_{nm}(x,y) = \epsilon_{nm}\psi_{nm}(x,y), \tag{46}$$

where the parameter λ is a measure of the anharmonicity of the two-dimensional system and is here taken to be $\sqrt{0.0125}$ to enable the comparison of the present results with those of other authors. Noid and Marcus[68] discussed the quasibound states of this system in terms of the principal quantum number n and angular momentum quantum number m. Their classification of states is used in the present description. There is a C_{3v} symmetry in the Henon–Heiles potential which can be utilized in computations to reduce the number of grid points required. However, for the purpose of testing the DSC algorithm, we do not take this advantage. In the cases that the delta sequence kernels of Shannon (Shannon), Dirichlet (Dirichlet), and modified Dirichlet (MD) are used, the present results are calculated using only 40 grid points in each dimension ($N_x=N_y=40, \Delta_x=\Delta_y=0.444\,878\,05$). These are listed in Table IV together with those by Shizgal and Chen[36] obtained by using up to 50 QDM grid points in each dimension. It is seen that all the present results obtained using different delta sequence kernels are in excellent agreement. The same level of accuracy is attained by using the de la Vallée Poussin (DLVP) delta sequence kernel with 1.5 times of grid points in each dimension ($N_x=N_y=60, \Delta_x=\Delta_y=0.299\,016\,39$).

IV. CONCLUSION

This paper introduces a discrete singular convolution (DSC) algorithm for the numerical solution of the Fokker–Planck equation. Singular kernels of the Hilbert-type and the delta type are considered. Various sequences of approximations are constructed for numerical analysis. Computational techniques are discussed for singular kernels of both the Hilbert-type and the delta type. A generalized Hilbert sampling theorem is given. By focusing on the delta sequence kernels of Shannon, Dirichlet, modified Dirichlet, and the de la Vallée Poussin, the utility of the DSC algorithm is explored for solving the Fokker–Planck equation. Important numerical issues examined in this paper are the accuracy of approximation, the speed of convergence, the simplicity of implementation. The DSC algorithm performs extremely well for all issues. The present results are in excellent agreement with those of previous approaches.

Three typical examples, which cover a variety of physical situations, are chosen to demonstrate the usefulness and to test the accuracy of the present algorithm. The first example is the Lorentz gas problem with a quadratic drift co-

TABLE IV. Eigenvalues of the Henon–Heiles system.

n	*m*	Feit *et al.*	Shizgal *et al.*	Shannon	Dirichlet	MD	DLVP
3	3	3.9825	3.982 417	3.982 417	3.982 417	3.982 417	3.982 417
3	−3	3.9859	3.985 761	3.985 761	3.985 761	3.985 761	3.985 761
5	3	5.8672	5.867 015	5.867 015	5.867 015	5.867 015	5.867 015
5	−3	5.8816	5.881 446	5.881 446	5.881 446	5.881 446	5.881 446
6	6	6.9991	6.998 932	6.998 932	6.998 932	6.998 932	6.998 932
6	−6	6.9996	6.999 387	6.999 387	6.999 387	6.999 387	6.999 387
7	3	7.6979	7.697 721	7.697 721	7.697 721	7.697 721	7.697 721
7	−3	7.7371	7.736 885	7.736 885	7.736 885	7.736 885	7.736 885
8	6	8.8116	8.811 327	8.811 327	8.811 327	8.811 327	8.811 327
8	−6	8.8154	8.815 188	8.815 188	8.815 188	8.815 188	8.815 188
9	3	9.4670	9.466 773	9.466 773	9.466 773	9.466 773	9.466 773
9	−3	9.5526	9.552 382	9.552 382	9.552 382	9.552 382	9.552 382
9	9	10.0356	10.035 413	10.035 413	10.035 413	10.035 413	10.035 413
9	−9	10.0359	10.035 592	10.035 592	10.035 592	10.035 592	10.035 592
10	6	10.5727	10.572 480	10.572 480	10.572 480	10.572 480	10.572 480
10	−6	10.5907	10.590 470	10.590 470	10.590 470	10.590 470	10.590 470
11	3	11.1603	11.160 258	11.160 258	11.160 259	11.160 259	11.160 259
11	−3	11.3253	11.325 231	11.325 231	11.325 231	11.325 231	11.325 231
11	9	11.7497	11.749 519	11.749 519	11.749 519	11.749 519	11.749 519
11	−9	11.7525	11.752 297	11.752 297	11.752 297	11.752 297	11.752 297
12	6	12.3335	12.333 785	12.333 785	12.333 785	12.333 785	12.333 785
12	−6	12.2771	12.277 192	12.277 192	12.277 192	12.277 192	12.277 192
12	12	12.7474	12.748 445	12.748 431	12.748 431	12.748 431	12.748 421
12	−12	13.0310	13.032 062	13.032 062	13.032 062	13.032 062	13.032 062
13	3	13.0868	13.086 873	13.086 873	13.086 873	13.086 873	13.086 873
13	−3	13.0800	13.081 196	13.081 196	13.081 196	13.081 196	13.081 196

efficient and a nonclassical diffusion term. The transformed Fokker–Planck equation has a singular potential describing hard sphere collisions of a light particle with a heavy particle bath. The DSC algorithm performs extremely well for this system. In the calculation using the interpolating delta sequence kernels of Shannon, Dirichlet, and modified Dirichlet, the first 40 eigenvalues converge to six significant figures, using only 60 grid points. These results are in excellent agreement with those of Shizgal and Chen[36] obtained by using up to 70 QDM grid points.

In the second example, we utilized a bistable model for two different ϵ values ($\epsilon=0.1$, 0.01). Since the ϵ value measures the ratio of dissipation and convection in the Fokker–Planck equation, a small ϵ means a fast convective motion of the system. The corresponding small Fokker–Planck eigenvalues are difficult to evaluate by numerical approaches due to irregular shape at the bottom of the effective potential. As in the previous example, the DSC algorithm converges faster than that of the QDM approach reported by Shizgal *et al.*[18,36] for this model when the interpolating delta sequence kernels of Shannon, Dirichlet, and modified Dirichlet are utilized. We used only 42 and 52 grid points for $\epsilon=0.1$ and $\epsilon=0.01$, respectively. Similar results reported by Shizgal and Chen[36] were calculated by using up to 60 and 100 QDM grid points, respectively. We noted that the QDM approach with some specifically designed weight function can attain an extremely fast speed of convergence.[37]

The last example considered is the Henon–Heiles anharmonic oscillator potential, which is another standard problem for various numerical methods. The DSC algorithm converges very rapidly for this problem too. For the delta sequence kernels of Shannon, Dirichlet, and modified Dirichlet, the first 97 eigenvalues converge to eight significant figures, when only 42 grid points are used in each dimension. These results are in excellent agreement with those of other methods, such as those of Feit *et al.*[65] obtained using the symmetric split operator-FFT and, in particular, of Shizgal and Chen.[36] The latter were calculated using up to 50 QDM grid points in each dimension. The present results indicate that the DSC algorithm is an efficient, reliable, and robust method for numerically solving the Fokker–Planck equation and eigenvalue problems in general.

ACKNOWLEDGMENT

This work was supported in part by the National University of Singapore.

APPENDIX: THE FORMAL RELATION BETWEEN THE FOKKER–PLANCK EQUATION AND THE SCHRÖDINGER EQUATION

The Fokker–Planck equation (40) considered in this work is a second order linear partial differential equation of the parabolic type. It is convenient to rewrite Eq. (40) as

$$\frac{\partial f(x,t)}{\partial t}=L_{\mathrm{FP}}f(x,t), \tag{A1}$$

where the Fokker–Planck operator L_{FP} is given by

$$L_{\mathrm{FP}}\equiv\frac{\partial A(x)}{\partial x}+\frac{\partial^2 B(x)}{\partial x^2}. \tag{A2}$$

The quantity $f(x,t_0)$ is an initial distribution function at time t_0. Important solutions for the Fokker–Planck equation are the formal one

$$f(x,t)=e^{L_{\mathrm{FP}}(t-t_0)}f(x,t_0), \tag{A3}$$

and the stationary one

$$f_{\mathrm{st}}(x)=\frac{e^{-U(x)}}{\int e^{-U(x)}dx}, \tag{A4}$$

where the function U is given in terms of the drift and diffusion coefficients

$$U(x)=\int_0^x \frac{A(y)}{B(y)}dy+\ln[B(x)]. \tag{A5}$$

These expressions are often discussed in the theoretical analysis of the Fokker–Planck equation. For practical purposes, the Fokker–Planck equation is often numerically solved by certain computational algorithms. One particularly useful way of obtaining the numerical solution is the eigenfunction expansion approach. The crucial assumption in such an approach is that the Fokker–Planck operator has a complete set of discrete spectrum

$$L_{\mathrm{LP}}f_n(x)=-\lambda_n f_n(x), \tag{A6}$$

where the eigenfunctions $f_n(x)$ form a biorthogonal basis such that the initial probability distribution function $f(x,t_0)$ can be expressed as

$$f(x,t_0)=\sum_n a_n f_n(x,t_0), \tag{A7}$$

where the coefficients a_n are determined by

$$a_n=\int_{-\infty}^{\infty} f_n(x,t_0)f(x,t_0)[f_0(x,t_0)]^{-1}dx, \tag{A8}$$

with $f_0(x,t_0)\equiv f_0(x)$ being the equilibrium distribution. In its eigenfunction expansion, the full (time-dependent) solution, Eq. (A3) of the Fokker–Planck equation is written as

$$f(x,t)=\sum_n a_n \exp[-\lambda_n(t-t_0)]f_n(x,t_0). \tag{A9}$$

More detailed discussion of Eq. (A9) can be found in standard references.[11,18] If the Fokker–Planck system behaves normally and has an equilibrium distribution when $t\to\infty$, then eigenvalues λ_n are positive semidefinite, $\lambda_0=0$ and $\lambda_n>0\ \forall n>0$.

Since the drift and diffusion coefficients can be a complex function of x, the Fokker–Planck operator L_{FP} is in general not a Hermitian operator. However, in the eigenvalue problem, it is more convenient to work with a self-adjoint operator. In order to construct a self-adjoint Fokker–Planck operator, we consider a particular case of Eq. (A4),

$$f_0(x)=\frac{1}{B(x)}\exp\left(-\int_0^x \frac{A(y)}{B(y)}dy\right). \tag{A10}$$

By separating $f_0(x)$ from Eq. (A1), one can obtain a new equation

$$\frac{\partial\Phi(x,t)}{\partial t}=-A(x)\frac{\partial\Phi(x,t)}{\partial x}+B(x)\frac{\partial^2\Phi(x,t)}{\partial x^2}$$
$$=-L\Phi(x,t), \tag{A11}$$

where $\Phi(x,t)$ is defined by

$$f(x,t)=f_0(x)\Phi(x,t). \tag{A12}$$

The quantity $\Phi(x,t)$ can be expanded by a complete set of eigenfunctions for the operator L in Eq. (A11) according to

$$\Phi(x,t)=e^{-L(t-t_0)}\Phi(x,t_0)$$
$$=\sum_n b_n e^{-\epsilon_n(t-t_0)}\phi_n(x), \tag{A13}$$

where expansion coefficient b_n is determined by the initial condition $f(x,t_0)$. Note that the new Fokker–Planck operator, L, is a self-adjoint operator on the space spanned by basic functions ϕ_n. The inner product in such a space is defined with respect to the weight function $f_0(x)$. Operator L satisfies a desired eigenvalue equation

$$L\phi_n(x)=\epsilon_n\phi_n(x). \tag{A14}$$

In principle, the Fokker–Planck eigenvalue problem (A14) can be solved directly by numerical methods. However, computationally it is more convenient to work with a Schrödinger equation-like Fokker–Planck equation which is given by

$$-\frac{d^2\psi_n(z)}{dz^2}+V(z)\psi_n(z)=\epsilon_n\psi_n(z), \tag{A15}$$

where $\psi_n(z)$ is given by

$$\psi_n(z)=(f_0[x(z)]\sqrt{B[x(z)]})^{1/2}\phi_n[x(z)], \tag{A16}$$

and the functional relation between x and z is

$$z(x)=\int^x [B(y)]^{-(1/2)}dy. \tag{A17}$$

The potential in the Schrödinger equation-like Fokker–Planck equation, Eq. (A15) is given in terms of the drift and diffusion coefficients in the Fokker–Planck equation,

$$V(z)=\tfrac{1}{4}[W^2(z)-2W_z'(z)], \tag{A18}$$

where the function $W(z)$ can be obtained from the general drift and diffusion functions as

$$W(z)=\frac{1}{\sqrt{B}}\left(A+\frac{1}{2}B'\right), \tag{A19}$$

and

$$B'=\frac{dB[x(z)]}{dz}. \tag{A20}$$

The effective potential V derived in this manner belongs to the class of potentials that occur in supersymmetric quantum mechanics. It is the final form, Eq. (A15), that has been used in this work.

[1] A. Einstein, Ann. Phys. (Leipzig) **17**, 549 (1905); **19**, 371 (1906).
[2] P. Langevin, Comptes Rendus **146**, 530 (1908).
[3] A. D. Fokker, Ann. Phys. (Leipzig) **43**, 810 (1914).
[4] M. Planck, Sitzungsber. K. Preuss. Akad. Wiss. **1917**, 324.
[5] H. A. Kramers, Physica (Amsterdam) **7**, 284 (1940).
[6] S. Chandrasekhar, Rev. Mod. Phys. **15**, 1 (1943).
[7] K. Itô, Proc. Imp. Acad. Japan **20**, 519 (1944).

[8] N. G. van Kampen, J. Stat. Phys. **17**, 71 (1977); H. Dekker and N. G. van Kampen, Phys. Lett. **73A**, 374 (1979).
[9] R. Kubo, K. Matsuo, and K. Kitahara, J. Stat. Phys. **9**, 51 (1973).
[10] M. Suzuki, Adv. Chem. Phys. **46**, 195 (1981).
[11] H. Risken, *The Fokker–Planck Equation: Methods of Solution and Application* (Springer, Berlin, 1984).
[12] B. Caroli, C. Caroli, and B. Roulet, J. Stat. Phys. **21**, 26 (1979).
[13] B. Caroli, C. Caroli, and B. Roulet, J. Stat. Phys. **26**, 83 (1981).
[14] R. S. Larson and M. Kostin, J. Chem. Phys. **69**, 4821 (1978).
[15] R. Indira, M. C. Valsakumar, K. P. N. Murthy, and G. Ananthakrishna, J. Stat. Phys. **33**, 181 (1981).
[16] M. F. Wehner and W. G. Wolfer, Phys. Rev. A **27**, 2663 (1983).
[17] H. Brand, A. Schenzle, and G. Schröder, Phys. Rev. A **25**, 2324 (1982).
[18] R. Blackmore and B. Shizgal, Phys. Rev. A **31**, 1855 (1985).
[19] E. M. Epperlein, J. Comput. Phys. **112**, 291 (1994).
[20] H. L. Frisch and B. Nowakowski, J. Chem. Phys. **98**, 8963 (1993).
[21] M. A. Muñoz and P. L. Garrido, Phys. Rev. E **50**, 2458 (1994).
[22] T. H. Kho, Phys. Rev. A **32**, 666 (1985).
[23] A. N. Drozdov and M. Morillo, Phys. Rev. E **54**, 931 (1996).
[24] G. Ryskin, Phys. Rev. E **56**, 5123 (1997).
[25] H. Haken, Rev. Mod. Phys. **47**, 67 (1975).
[26] H. Grabert and M. S. Green, Phys. Rev. A **19**, 1747 (1979).
[27] H. Dekker, Phys. Rev. A **19**, 2102 (1979).
[28] D. L. Ermak and H. Buckholtz, J. Comput. Phys. **35**, 169 (1980).
[29] R. H. Swendsen and J. S. Wang, Phys. Rev. Lett. **58**, 86 (1987).
[30] P. D. Lax, Commun. Pure Appl. Math. **6**, 231 (1953).
[31] G. E. Forsythe and W. R. Wasow, *Finite Difference Methods for Partial Differential Equations* (Wiley, New York, 1967).
[32] V. Palleschi, F. Sarri, G. Marcozzi, and M. R. Torquati, Phys. Lett. A **146**, 378 (1990).
[33] J. S. Chang and G. Cooper, J. Comput. Phys. **6**, 1 (1970).
[34] E. W. Larson, C. D. Levermore, G. C. Pomraning, and J. G. Sanderson, J. Comput. Phys. **61**, 359 (1985).
[35] B. Shizgal, J. Comput. Phys. **41**, 309 (1981).
[36] B. Shizgal and H. Chen, J. Chem. Phys. **104**, 4137 (1996).
[37] B. Shizgal and H. Chen, J. Chem. Phys. **107**, 8051 (1997).
[38] D. K. Hoffman, N. Nayar, O. A. Sharafeddin, and D. J. Kouri, J. Phys. Chem. **95**, 8299 (1991).
[39] D. S. Zhang, G. W. Wei, D. J. Kouri, and D. K. Hoffman, J. Chem. Phys. **106**, 5216 (1997).
[40] B. T. Park and V. Petrosian, Astrophys. J., Suppl. Ser. **103**, 255 (1996).
[41] B. Caroli, C. Caroli, and B. Roulet, J. Stat. Phys. **26**, 83 (1981).
[42] L. Schwartz, *Théore des Distributions* (Hermann, Paris, 1951).
[43] J. Korevaar, Nederl. Akad. Westensh. Proc. Ser. A **58**, 368 (1955); **58**, 483 (1955); **58**, 663 (1955).
[44] J. Korevaar, Am. Math. Soc. Trans. **91**, 53 (1959).
[45] G. G. Walter, Trans. Am. Math. Soc. **116**, 492 (1965).
[46] G. W. Walson and M. R. Leadbetter, Sankhya **26**, 101 (1965).
[47] B. B. Winter, Ann. Statist. **3**, 759 (1975).
[48] G. G. Walter and J. Blum, Ann. Statist. **7**, 328 (1977).
[49] G. Wahba, Ann. Statist. **3**, 15 (1975).
[50] R. Kronmal and M. Tarter, J. Am. Stat. Assoc. **63**, 925 (1968).
[51] G. W. Wei, D. S. Zhang, D. J. Kouri, and D. K. Hoffman, Phys. Rev. Lett. **79**, 775 (1997).
[52] G. W. Wei, Chem. Phys. Lett. **296**, 215 (1998).
[53] C. E. Shannon, Proc. IRE **137**, 10 (1949).
[54] E. T. Whittaker, Proc. R. Soc. Edinburgh **35**, 181 (1915).
[55] J. M. Whittaker, Proc. Edinburgh Math. Soc. **1**, 41 (1929).
[56] S. L. Hahn, in *The Transforms and Applications Handbook*, edited by A. D. Poularikas (Chemical Rubber, Cleveland, 1996).
[57] K. Andersen and K. E. Shuler, J. Chem. Phys. **40**, 633 (1964).
[58] C. S. Shapiro and N. Corngold, Phys. Rev. A **137**, 1686 (1970).
[59] H. Oser, K. E. Shuler, and G. Weiss, J. Chem. Phys. **41**, 1661 (1964).
[60] M. R. Hoare, Adv. Chem. Phys. **20**, 135 (1971).
[61] B. Shizgal, J. Chem. Phys. **70**, 1948 (1979).
[62] M. R. Hoare and C. H. Kaplinsky, J. Chem. Phys. **52**, 3336 (1970).
[63] M. Henon and C. Heiles, Astron. J. **69**, 73 (1964).
[64] D. W. Noid, M. L. Koszykowski, M. Tabor, and R. A. Marcus, J. Chem. Phys. **72**, 6169 (1980).
[65] M. D. Feit, J. A. Fleck, Jr., and A. Steiger, J. Comput. Phys. **47**, 412 (1982).
[66] I. P. Hamilton and J. C. Light, J. Chem. Phys. **84**, 306 (1988).
[67] J. Echave and D. C. Clary, Chem. Phys. Lett. **190**, 225 (1992).
[68] D. W. Noid and R. A. Marcus, J. Chem. Phys. **62**, 2119 (1975).