Journal of Statistical Physics, Vol. 114, Nos. 1/2, January 2004 (© 2004)

Fractional Kinetics in Kac–Zwanzig Heat Bath Models

Raz Kupferman[1]

Received April 17, 2003; accepted June 16, 2003

We study a variant of the Kac–Zwanzig model of a particle in a heat bath. The heat bath consists of n particles which interact with a distinguished particle via springs and have random initial data. As $n \to \infty$ the trajectories of the distinguished particle weakly converge to the solution of a stochastic integro-differential equation—a generalized Langevin equation (GLE) with power-law memory kernel and driven by $1/f^{\alpha}$-noise. The limiting process exhibits fractional sub-diffusive behaviour. We further consider the approximation of non-Markovian processes by higher-dimensional Markovian processes via the introduction of auxiliary variables and use this method to approximate the limiting GLE. In contrast, we show the inadequacy of a so-called fractional Fokker–Planck equation in the present context. All results are supported by direct numerical experiments.

KEY WORDS: Fractional diffusion; Hamiltonian systems; heat bath; stochastic differential equations; Markovian approximation; weak convergence.

1. INTRODUCTION

Anomalous diffusion is a well-studied phenomenon applicable to a broad variety of fields (e.g., particles moving through media with internal degrees of freedom, such as actin networks[3]). A random process $X(t)$ is said to exhibit anomalous diffusion when the variance of its displacement after time t has the asymptotic form

$$\mathbb{E}\,|\Delta X(t)|^2 \sim t^{\gamma}, \qquad t \to \infty,$$

where $\mathbb{E}$ denotes averaging, or expectation, and $\gamma \neq 1$. The process is called sub-diffusive when $\gamma < 1$ and super-diffusive when $\gamma > 1$; the case $\gamma = 1$ corresponds to regular diffusion.

[1] Institute of Mathematics, The Hebrew University, Jerusalem, 91904, Israel; e-mail: raz@math.huji.ac.il

0022-4715/04/0100-0291/0 © 2004 Plenum Publishing Corporation

Early work on anomalous diffusion dates back to the 1960s with the Montroll–Weiss model of continuous-time random walk.[39] Random walks serve as standard models for normal diffusion processes (e.g., a discrete time random walk may weakly converge to Brownian motion[18]). Continuous time random walk are characterized by two parameters: a characteristic waiting time between jump events and a jump distance. Anomalous diffusion arises when the characteristic waiting time diverges.[5] The divergence of relaxation times is the key for the occurrence of anomalous diffusion. Systems that exhibit anomalous diffusion are characterized by either a diverging waiting time, or equivalently, by a strong non-Markovian nature—the evolution of the system at time t depends on its past, and the range of this "memory" is long compared with the characteristic timescale of its motion. In contrast, normal diffusion occurs when the microscopic timescale is small compared to the observation time.

Mechanical models of a particle immersed in a heat bath were introduced by Ford, Kac, and Mazur[15, 14] and Zwanzig[54] as simple models to study kinetics and irreversible statistical mechanics. The "heat bath" is a collection of n particles which interact with a "distinguished" particle through springs; the heat bath particles are assumed to have random initial data distributed according to the laws of statistical mechanics. There exists a huge amount of literature in this subject. Heat bath models have received renewed interest in recent years in the context of variable reduction,[32, 33, 48] coarse time stepping,[25, 48] and transition state models.[29]

In ref. 33 a variant of the Kac–Zwanzig model was considered. For a certain regime of parameters, the trajectory of the distinguished particle tends, as $n \to \infty$ (the "thermodynamic limit"), to a limiting process $Q(t)$, which satisfies a stochastic integro-differential equation (SIDE), known as a generalized Langevin equation (GLE). The convergence is in a weak sense (in distribution) in the space of continuous functions.[8] The parameters of the model, namely, the masses and spring constants, determine the parameters of the limiting GLE, which are the driving noise and the memory kernel.

In this paper we construct a Kac–Zwanzig model within the setting of ref. 33, with the heat bath parameters chosen such that the limiting GLE has a memory kernel that decays as a power-law (Section 2). The limiting GLE for $Q(t)$ has the form

$$\ddot{Q}(t) + k_0 \int_0^t (t-s)^{-\gamma} \dot{Q}(s)\, ds + V'(Q(t)) = z(t), \tag{1.1}$$

where $z(t)$ is a stationary centered Gaussian process with auto-covariance $\mathbb{E} z(t)\, z(s) = \beta^{-1} k_0 |t-s|^{-\gamma}$, with β being an inverse temperature. The

process $z(t)$ is a generalized random process which can be identified with the derivative of fractional Brownian motion;[35] it is often referred to as a $1/f^{\alpha}$-noise. The weak convergence of the trajectories of the distinguished particle is proved in Section 3.

There are two simple, but instructive cases for which (1.1) can be solved analytically: in the case of a free particle, $V(Q) = 0$, and in the case of a quadratic potential well, $V(Q) = \frac{1}{2} Q^2$; in both cases the equation is linear and can be solved by standard methods (Section 4). In the case of a free particle the distribution of $Q(t)$ exhibits anomalous diffusion with exponent γ. Although anomalous diffusion is often associated with non-Gaussian behaviour, here an initial Gaussian distribution on Q and $\dot{Q}$ remains Gaussian for all times. In the case of a quadratic potential well the Boltzmann equilibrium distribution is reached as $t \to \infty$. The rate of equilibration is sub-exponential (i.e., slower than exponential) as expected for a sub-diffusive system.

In order to make quantitative predictions about solutions of (1.1) it is necessary to derive equations for the one- and multi-dimensional probability density functions (PDFs). For Markovian systems driven by white noise the equation that determines the one-dimensional PDF is the Fokker–Planck equation (FPE).[17] There is, however, no simple differential equation for the one- and multi-dimensional PDFs for non-Markovian processes. The derivation of such equations is at the heart of the Mori–Zwanzig approach;[19, 42] see also ref. 28 and references therein for a review of non-Markovian GLEs. An alternative is to approximate the non-Markovian system by a higher-dimensional Markovian system through the introduction of auxiliary variables. This approach dates back to Mori[40] and has been applied extensively by Kłosek–Dygas *et al.*,[11, 12] and recently, in the context of $1/f^{\alpha}$-noise, by Landis *et al.*[34] In Section 5 we describe how to approximate, in general, a Gaussian non-Markovian system by a higher-dimensional Markovian SDE. We follow Ref. 34 to construct a particular example where the memory kernel decays as a power-law. The whole approach may seem circular: a high-dimensional system is proved to converge to the solution of a non-Markovian SIDE, which is then approximated by a higher-dimensional SDE. As we show in this paper, a (deterministic) system of very large size may thus be accurately approximated by a (stochastic) system of, say, 6 variables.

In Section 6 we consider the fractional Fokker–Planck equation (FFPE), which is a partial integro-differential equation often viewed as a natural generalization of the FPE for strongly non-Markovian systems which exhibit anomalous diffusion.[4–6, 37, 45, 52] For continuous-time random walk, the FFPE governs the distribution of the process in a certain limit where the jump distance tends to zero.[5] It is however unclear whether it

has a much wider scope as does the FPE for Markovian system. We solve the FFPE for the case of a free particle, and show that its solution differs substantially from the solution found in Section 4. This simple example shows that the FFPE does not describe systems governed by the GLE (1.1).

In Section 7 we present numerical experiments which test the convergence of the trajectories to the solution of (1.1), and the Markovian approximation considered in Section 5.

2. A KAC-ZWANZIG HEAT BATH MODEL

We introduce a mechanical model of a particle immersed in a heat bath of n particles. The model is defined by the Hamiltonian:

$$H(Q_n, P_n, q, p) = \frac{1}{2} P_n^2 + V(Q_n) + \frac{1}{2} \sum_{j=1}^{n} \frac{p_j^2}{m_j} + \frac{1}{2} \sum_{k=1}^{n} k_j (q_j - Q_n)^2, \quad (2.1)$$

where (Q_n, P_n) are the position and momentum of the distinguished particle, which has unit mass and resides in a potential field $V(Q_n)$; $q = (q_1, q_2, \ldots, q_n)$ and $p = (p_1, p_2, \ldots, p_n)$ are the positions of momenta of the n heat bath particles. The jth particle in the heat bath has mass m_j, and interacts with the distinguished particle through a linear spring with stiffness constant k_j. We also define $\omega_j = (k_j/m_j)^{1/2}$, which is the characteristic frequency of the jth heat bath particle. All motions take place in one dimension. The subscript n in Q_n, P_n labels the size of the heat bath, which will eventually be taken infinitely large as we consider the thermodynamic limit.

Hamilton's equations are

$$\begin{aligned} \dot{Q}_n &= P_n \\ \dot{P}_n &= -V'(Q_n) + \sum_{k=1}^{n} k_j (q_j - Q_n) \\ \dot{q}_j &= p_j / m_j \\ \dot{p}_j &= -k_j (q_j - Q_n), \end{aligned} \quad (2.2)$$

supplemented with initial conditions $Q_n(0) = Q_0$, $P_n(0) = P_0$, $q_j(0) = q_{j0}$, and $p_j(0) = p_{j0}$. The initial data for the heat bath particles are assumed to be randomly drawn from a Gibbsian distribution with inverse temperature β, conditioned by the non-random initial data Q_0, P_0. For fixed Q_n, P_n the

Hamiltonian (2.1) is quadratic in q, p, hence the corresponding measure is Gaussian. It is easily verified that

$$q_{j0} = Q_0 + \beta^{-1/2} k_j^{-1/2} \xi_j$$
$$p_{j0} = \beta^{-1/2} m_j^{1/2} \eta_j,$$

where $\xi_j, \eta_j \sim \mathcal{N}(0, 1)$ are mutually independent sequences of i.i.d. random variables.

The system (2.2) describes a particle interacting with a collection of n oscillators with spectrum ω_j. Conceptually, a heat bath is a mechanical system characterized by a broad and dense spectrum. A natural way to realize such a scenario, without being too restrictive, is to choose the parameters k_j, m_j such that the corresponding frequencies ω_j are random, uniformly distributed in the range $1/n^c + [0, n^a]$, with $0 < a, c < 1$. Specifically,

$$\omega_j = n^{-c} + n^a \nu_j, \qquad \nu_j \text{ i.i.d.}, \quad \nu_1 \sim \mathcal{U}[0, 1].$$

Thus, as $n \to \infty$, the spectrum covers an increasingly large range of frequencies in an increasingly dense manner. Note the use of a low-frequency cutoff, $\omega_j \geqslant 1/n^c$, which was not necessary in ref. 33. This is because of the singularity in the auto-correlation of fractional Brownian motion.

Having chosen the frequencies ω_j, it remains to choose either the masses m_j, or the spring constants k_j. In this paper we take

$$k_j = f^2(\omega_j)\, \Delta\omega, \qquad f^2(\omega) = \frac{2a_0}{\pi} \Gamma(1-\gamma) \sin\left(\frac{\gamma\pi}{2}\right) \frac{1}{\omega^{1-\gamma}}, \tag{2.3}$$

where $\Delta\omega = n^a/n$ is the mean spectral density, $\Gamma(z)$ is the Euler Gamma function,[1] and $\gamma \in (0, 1)$ and $a_0 > 0$ are parameters. The reason for these choices will become apparent in the following.

The probability space is defined by the three mutually independent sequences of random variables ν_j, ξ_j, and η_j; the first is related to the model parameters—the spectrum of the heat bath—and the two other are related to the initial data. Our results for the limiting behaviour of the system as $n \to \infty$ hold almost surely with respect to the choice of frequencies, which we denote by ν-almost-surely, or, ν-a.s. Probabilities, expectations and variances with respect to the ν variables are denoted by $\mathbb{E}_\nu$, $\mathbb{P}_\nu$, and Var_ν. Similarly, $\mathbb{P}_{\xi\eta}$, $\mathbb{E}_{\xi\eta}$, and $\mathrm{Var}_{\xi\eta}$ denote probabilities, expectations and variances with respect to the initial data; when no confusion should arise, we use the shorter notation $\mathbb{P}$, $\mathbb{E}$, and Var.

As usual, the q_j, p_j variables in (2.2) can be integrated, giving rise to an integro-differential equation for the trajectory $Q_n(t)$ of the distinguished particle:

$$\ddot{Q}_n(t)+\int_0^t \kappa_n(t-s)\,\dot{Q}_n(s)\,ds+V'(Q_n(t))=z_n(t), \tag{2.4}$$

where

$$\kappa_n(t)=\sum_{j=1}^{n} f^2(\omega_j)\cos(\omega_j t)\,\Delta\omega \tag{2.5}$$

$$z_n(t)=\beta^{-1/2}\sum_{j=1}^{n} f(\omega_j)[\xi_j\cos(\omega_j t)+\eta_j\sin(\omega_j t)](\Delta\omega)^{1/2}. \tag{2.6}$$

The function $\kappa_n(t)$ is the memory kernel, which encapsulates the dependence of the force that the heat bath exerts on the distinguished particle at time t on the history of its trajectory up to that time; it is random only through the frequencies ω_j. The function $z_n(t)$ is a random forcing which depends both on the frequencies and on the initial data. In the present setting it is a stationary centered Gaussian process; its auto-covariance satisfies the fluctuation-dissipation relation

$$\mathbb{E}_{\xi\eta} z_n(t)\,z_n(s)=\beta^{-1}\kappa_n(t-s),$$

irrespective of the choice of frequencies.

It will be shown below that $z_n(t)$ tends, as $n\to\infty$, to a generalized (distribution valued) random process. This suggests that we should consider an integrated version of (2.4):

$$\dot{Q}_n(t)+\int_0^t K_n(t-s)\,\dot{Q}_n(s)\,ds+\int_0^t V'(Q_n(s))\,ds=\dot{Q}_0+Z_n(t), \tag{2.7}$$

where $K_n(t)=\int_0^t \kappa_n(s)\,ds$ and $Z_n(t)=\int_0^t z_n(s)\,ds$, i.e.,

$$K_n(t)=\sum_{j=1}^{n}\omega_j^{-1}f^2(\omega_j)\sin(\omega_j t)\,\Delta\omega \tag{2.8}$$

$$Z_n(t)=\beta^{-1/2}\sum_{j=1}^{n}\omega_j^{-1}f(\omega_j)\{\xi_j\sin(\omega_j t)+\eta_j[1-\cos(\omega_j t)]\}(\Delta\omega)^{1/2}. \tag{2.9}$$

The integrated forcing function, $Z_n(t)$, is a centered Gaussian process with auto-covariance,

$$\mathbb{E}_{\xi\eta} Z_n(t)\,Z_n(s)=\beta^{-1}[R_n(t)+R_n(s)-R_n(t-s)],$$

where

$$R_n(t) = \int_0^t K_n(s)\, ds = \sum_{j=1}^{n} \omega_j^{-2} f^2(\omega_j)[1-\cos(\omega_j t)]\, \Delta\omega. \tag{2.10}$$

Thus, $Z_n(t)$ has stationary (but not independent) increments,

$$Z_n(t) - Z_n(s) \sim \mathcal{N}(0, \beta^{-1} R_n(t-s)).$$

3. THE THERMODYNAMIC LIMIT

We now turn to analyze the "thermodynamic limit" of the above model, that is, the asymptotic limit as $n \to \infty$. We derive a SIDE for the limiting process $Q(t)$. The basic method is the same as in refs. 32 and 33. Some technical differences arise due to the singularity of $\kappa_n(t)$ as $n \to \infty$. As a result, the second-order SIDE has to be interpreted in the sense of generalized functions, in contrast with the situation in refs. 32 and 33.

Consider first the kernel $\kappa_n(t)$, defined by (2.5). It can be viewed as a Monte-Carlo approximation of the integral

$$\int_{1/n^c}^{n^a + 1/n^c} f^2(\omega) \cos(\omega t)\, d\omega,$$

which, as $n \to \infty$, tends to the Fourier cosine transform of $f^2(\omega)$:

$$\kappa(t) = \int_0^\infty f^2(\omega) \cos(\omega t)\, d\omega = \frac{a_0}{|t|^\gamma}. \tag{3.1}$$

It is precisely in order to obtain this asymptotic limit that $f^2(\omega)$ was chosen as in (2.3).

Similarly, the kernel $K_n(t)$, given by (2.8), can be viewed as a Monte-Carlo approximation of

$$K(t) = \int_0^t \kappa(s)\, ds = \frac{a_0}{1-\gamma} |t|^{1-\gamma}, \tag{3.2}$$

whereas $R_n(t)$ approximates

$$R(t) = \int_0^t K(s)\, ds = \frac{a_0}{(1-\gamma)(2-\gamma)} |t|^{2-\gamma}. \tag{3.3}$$

Finally, $Z_n(t)$ can be viewed as a Monte-Carlo approximation of the stochastic integral:

$$Z(t) = \beta^{-1/2} \int_0^\infty \omega^{-1} f(\omega) \sin(\omega t)\, dB_1(\omega)$$

$$+ \beta^{-1/2} \int_0^\infty \omega^{-1} f(\omega)[1 - \cos(\omega t)]\, dB_2(\omega),$$

where $B_1(\omega)$, $B_2(\omega)$ are independent standard Brownian motions. This stochastic integral represents the (non-stationary) centered Gaussian process whose auto-covariance is $\beta^{-1}[R(t) + R(s) - R(t-s)]$, a process which we identify as

$$Z(t) = \frac{2\beta^{-1/2} a_0}{(1-\gamma)(2-\gamma)} B^H(t), \tag{3.4}$$

where $B^H(t)$ is fractional Brownian motion with Hurst parameter $H = 1 - \frac{1}{2}\gamma$.[35] Unlike standard Brownian motion, the increments of fractional Brownian motion are not independent; for Hurst parameter $H > \frac{1}{2}$, as is the case here, increments are positively correlated, meaning that if $B^H(t)$ is increasing in a certain interval, it is likely to remain increasing in the future. The derivative of fractional Brownian motion is often called a "$1/f^\alpha$-noise," in reference to the power-law behaviour of its auto-covariance. All these arguments are made rigorous in the remaining part of this section.

We first prove that K_n, given by (2.8), converges to K, given by (3.2), in $L^2[0, T]$; the interval $[0, T]$ is bounded, but arbitrary. Convergence occurs for almost every set of frequencies (ν-a.s.). Having proven the convergence of K_n, the convergence of R_n and the (weak) convergence of Z_n are immediate consequences.

The kernel $K_n(t)$ is of the form

$$K_n(t) = \sum_{j=1}^{n} h(\omega_j, t)\, \Delta\omega, \tag{3.5}$$

where $h(\omega, t) = \omega^{-1} f^2(\omega) \sin(\omega t)$ satisfies

$$|h(\omega, t)| \leqslant C \min(\omega^{\gamma-1}, \omega^{\gamma-2}) \equiv h^*(\omega), \tag{3.6}$$

and $C > 0$ is a constant that may depend on T, γ, a, and c, but not on n; throughout this section we use C as a generic notation for a finite positive constant independent of n. We also define

$$\mu_n(t) = \mathbb{E}_\nu h(\omega, t) = \frac{1}{n^a} \int_{1/n^c}^{1/n^c + n^a} h(\omega, t)\, d\omega,$$

and observe that (3.6) implies that

$$|\mu_n(t)| \leqslant \frac{C}{n^a} \equiv \mu_n^*. \tag{3.7}$$

It follows that $\mathbb{E}_v K_n = n^a \mu_n$ converges uniformly to K; indeed,

$$\begin{aligned} |\mathbb{E}_v K_n(t) - K(t)| &\leqslant \left| \int_0^{1/n^c} h(\omega, t)\, d\omega \right| + \left| \int_{1/n^c + n^a}^{\infty} h(\omega, t)\, d\omega \right| \\ &\leqslant C \left| \int_0^{1/n^c} \omega^{\gamma - 1}\, d\omega \right| + C \left| \int_{n^a}^{\infty} \omega^{\gamma - 2}\, d\omega \right| \\ &\leqslant C[n^{-c\gamma} + n^{-a(1-\gamma)}], \end{aligned} \tag{3.8}$$

which, for $0 < a, c < 1$, converges to zero uniformly on $[0, T]$.

Lemma 3.1. Let $K_n(t)$ and $K(t)$ be given by (2.8) and (3.2), respectively. Then v-a.s. $K_n \to K$ in $L^2[0, T]$:

$$\mathbb{P}_v(\lim_{n \to \infty} \|K_n - K\|_{L^2[0,T]} = 0) = 1.$$

Proof. The proof follows the lines of Lemma 3.1 in ref. 33. Since by (3.8) $\mathbb{E}_v K_n(t)$ converges uniformly to $K(t)$, it is sufficient to show that v-a.s. $K_n - \mathbb{E}K_n \to 0$ in $L^2[0, T]$, i.e., that for any $\epsilon > 0$

$$\mathbb{P}_v(\|K_n - \mathbb{E}_v K_n\|_{L^2[0,T]} > \epsilon \text{ i.o.}) = 0$$

(i.o. = infinitely often), which, by the Borel–Cantelli theorem (see, e.g., ref. 7), holds if there exists an integer b for which

$$\sigma_n \equiv \mathbb{E}\, \|K_n - \mathbb{E}_v K_n\|_{L^2[0,T]}^{2b} = \mathbb{E} \left(\int_0^T |K_n(t) - \mathbb{E}K_n(t)|^2\, dt \right)^b$$

is summable. To evaluate σ_n we rewrite it as follows:

$$\sigma_n = \int_0^T \cdots \int_0^T \mathbb{E}\{|K_n(t_1) - \mathbb{E}K_n(t_1)|^2 \cdots |K_n(t_b) - \mathbb{E}K_n(t_b)|^2\}\, dt_1 \cdots dt_b.$$

For any $p = 1, 2, \ldots, b$,

$$K_n(t_p) - \mathbb{E}K_n(t_p) = \Delta\omega \sum_{j=1}^{n} [h(\omega_j, t_p) - \mu_n(t_p)],$$

hence

$$\sigma_n = (\Delta\omega)^{2b} \int_0^T \cdots \int_0^T \sum_{j_1=1}^{n} \cdots \sum_{j_{2b}=1}^{n} V_{j_1,\ldots,j_{2b}}(t_1,\ldots,t_b)\, dt_1 \cdots dt_b, \qquad (3.9)$$

where

$$V_{j_1,\ldots,j_{2b}}(t_1,\ldots,t_b) = \mathbb{E}\{[h(\omega_{j_1}, t_1) - \mu_n(t_1)][h(\omega_{j_2}, t_1) - \mu_n(t_1)] \cdots [h(\omega_{j_{2b-1}}, t_b) - \mu_n(t_b)][h(\omega_{j_{2b}}, t_b) - \mu_n(t_b)]\}, \qquad (3.10)$$

are the centered joint moments of degree $2b$ of $h(\omega_j, t_p)$.

Since ω_j and ω_i are independent for $j \neq i$, then many of these moments vanish; every $V_{j_1,\ldots,j_{2b}}(t_1,\ldots,t_b)$ that contains an index j which appears only once vanishes. We estimate σ_n by regrouping the $2b$-tuple sum (3.9) by the number k of distinct indices in the product (3.10); k assumes values from 1 to b because each index must occur at least twice, otherwise (3.10) is zero. The number of terms corresponding to a given k can be bounded by Cn^k, where C is a constant that depends on b, but neither on k nor n (there are n^k ways to "decode" a k-letter pattern with an n-letter alphabet). Now, each of the $V_{j_1,\ldots,j_{2b}}(t_1,\ldots,t_b)$ which corresponds to a given k is of the form

$$V_{j_1,\ldots,j_{2b}}(t_1,\ldots,t_b) = \prod_{r=1}^{k} \frac{1}{n^a} \int_{1/n^c}^{1/n^c+n^a} \prod_{s=1}^{m_k} [h(\omega, t_{r,s}) - \mu_n(t_{r,s})],$$

where $m_1, m_2, \ldots, m_k \geqslant 2$ and $m_1 + m_2 + \ldots + m_k = 2b$; the times $t_{r,s}$ belong to the set $(t_1,\ldots,t_b)$. Using the bounds (3.6) and (3.7) on $h(\omega, t)$ and $\mu_n(t)$, $V_{j_1,\ldots,j_{2b}}(t_1,\ldots,t_b)$ can be bounded as follows:

$$\begin{aligned}
|V_{j_1,\ldots,j_{2b}}(t_1,\ldots,t_b)| &\leqslant \prod_{r=1}^{k} \left(\frac{1}{n^a} \int_{1/n^c}^{1/n^c+n^a} [h^*(\omega) + \mu_n^*]^{m_r}\, d\omega \right) \\
&\leqslant \frac{1}{n^{ak}} \prod_{r=1}^{k} \sum_{\ell_r=0}^{m_r} \binom{m_r}{\ell_r} (\mu_n^*)^{m_r-\ell_r} \int_{1/n^c}^{n^a} [h^*(\omega)]^{\ell_r}\, d\omega \\
&\leqslant \frac{C}{n^{ak}} \prod_{r=1}^{k} \sum_{\ell_r=0}^{m_r} \binom{m_r}{\ell_r} (\mu_n^*)^{m_r-\ell_r}\, n^{c(1-\gamma)\,\ell_r} \\
&\leqslant C\, \frac{n^{2c(1-\gamma)\,b}}{n^{ak}}.
\end{aligned}$$

Combining all together we have

$$\sigma_n \leqslant C(\Delta\omega)^{2b} \sum_{k=1}^{b} n^k \frac{n^{2c(1-\gamma)\,b}}{n^{ak}}$$
$$\leqslant C\left(\frac{n^a}{n}\right)^{2b} n^b \frac{n^{2c(1-\gamma)\,b}}{n^{ab}} = Cn^{b[2c(1-\gamma)-(1-a)]}.$$

If we choose $c < \frac{1}{2}(1-a)/(1-\gamma) < \frac{1}{2}(1-a)$ then we can always take b large enough so that σ_n is summable. ∎

Corollary 3.1.

1. v-a.s. $R_n(t)$ converges to $R(t)$ pointwise.

2. v-a.s. $R_n(t)$, $R(t)$ are uniformly (both in n and t) Hölder continuous with exponent $\alpha = \frac{1}{2}$.

Proof. The first statement is an immediate consequence of the v-a.s. L^2-convergence of $K_n \to K$,

$$|R_n(t) - R(t)| = \left| \int_0^t [K_n(s) - K(s)]\, ds \right| \leqslant T^{1/2} \|K_n - K\|_{L^2[0.T]} \to 0.$$

The second statement follows from Cauchy–Schwarz,

$$|R_n(t) - R_n(s)| = \left| \int_s^t K_n(s)\, ds \right|$$
$$\leqslant \left(\int_s^t ds' \right)^{1/2} \left(\int_s^t K_n^2(s')\, ds' \right)^{1/2}$$
$$\leqslant |t-s|^{1/2} \|K_n\|_{L^2[0,\,T]},$$

and $\|K_n\|_{L^2[0,\,T]}$ is uniformly bounded since $K_n \to K$. ∎

Theorem 3.2. v-a.s. the random processes $Z_n(t)$, given by (2.9), converge weakly in $C[0, T]$ to the fractional Brownian motion $Z(t)$, given by (3.4).

Proof. The proof relies on the following theorem (Gikhman and Skorokhod,[18] p. 450): Let Z_n be a collection of real-valued almost-surely continuous stochastic processes on $[0, T]$, such that

1. The finite dimensional distributions of Z_n weakly converge to those of an almost-surely continuous process Z.

2. There exist positive constants b, η, M such that for all n

$$\mathbb{E}\,|Z_n(t+u)-Z_n(t)|^b \leqslant M\,|u|^{\eta}$$

(the tightness condition).

Then $Z_n \Rightarrow Z$ in $C[0,T]$.

Kolmogorov's condition (ref. 31, p. 22) applied to Gaussian processes states that Z has a continuous version if its auto-covariance is Hölder continuous, as is the case since $\mathbb{E}_v Z(t)\,Z(s) = \beta^{-1}[R(t)+R(s)-R(t-s)]$. To show that the finite dimensional distributions of Z_n converge weakly to those of Z we need to show that for any collection of times $0 \leqslant t_1 < t_2 < \ldots < t_k \leqslant T$, the joint probability density function of $(Z_n(t_1),\ldots, Z_n(t_k))$ converges pointwise to the joint probability density function of $(Z(t_1),\ldots, Z(t_k))$. For Gaussian processes this task is greatly simplified as it is sufficient to show the pointwise convergence of the auto-covariance, which was established in Corollary 3.1.

It remains to show the tightness property. For integer b,

$$\begin{aligned}\mathbb{E}_v\,|Z_n(t+u)-Z_n(t)|^{2b} &= (2n-1)!!\,(\mathbb{E}_v\,|Z_n(t+u)-Z_n(t)|^2)^b \\ &= 2^b\beta^{-b}(2n-1)!!\,[R_n(u)]^b.\end{aligned}$$

Since the R_n are uniformly Hölder continuous with exponent $\alpha = \frac{1}{2}$ and $R_n(0) = 0$, then $|R_n(u)| \leqslant C\,|u|^{1/2}$, and the tightness condition is satisfied by taking $b > 2$. ∎

Comment. Tauberian theorems relate asymptotic properties of functions to asymptotic properties of their integral transforms (e.g., ref. 23, p. 91). Thus, the power-law decay of $\kappa(t)$, as $t \to \infty$, results only from the power-law divergence of $f^2(\omega)$, as $\omega \to 0$ (in addition to the requirement that the Fourier integral exists), and is insensitive to the precise structure of $f^2(\omega)$ at finite ω.

Having established the (v-a.s.) convergence of $K_n \to K$ and $Z_n \Rightarrow Z$, we can now prove the weak convergence of Q_n to a limiting process:

Theorem 3.3. Suppose that $V'(Q)$ is globally Lipschitz continuous, then v-a.s the random processes $Q_n(t)$, defined by (2.4), weakly converge in $C^1[0,T]$ to the process $Q(t)$ satisfying the SIDE:

$$\dot{Q}(t)+\int_0^t K(t-s)\,\dot{Q}(s)\,ds+\int_0^t V'(Q(s))\,ds = P_0+Z(t). \qquad (3.11)$$

Proof. Since ν-a.s. $K_n \to K$ in $L^2[0,T]$ (in particular, in $L^1[0,T]$) and $Z_n \Rightarrow Z$ in $C[0,T]$, then the required result follows if the mapping $(K,Z) \mapsto Q$ defined by (3.11) is a continuous mapping from $L^1[0,T] \times C[0,T]$ to $C^1[0,T]$ (weak convergence is preserved under continuous mappings). This continuity is a well-known property of the Volterra equation (see refs. 21, 38, and Section 12 in ref. 25). ▌

Comment. Since the original system (2.4) is a second order IDE, we rewrite (3.11) as

$$\ddot{Q}(t) + \int_0^t \kappa(t-s)\, \dot{Q}(s)\, ds + V'(Q(t)) = z(t), \tag{3.12}$$

where $z(t) = \dot{Z}(t)$ is a generalized random process, namely, a Gaussian noise with auto-covariance $\mathbb{E}z(t)\, z(s) = \beta^{-1}\kappa(t-s)$.

4. SOLUTIONS OF THE GENERALIZED LANGEVIN EQUATION

In this section we study the generalized Langevin equation (3.12), which, as just shown, governs the weak limit $Q(t)$ of the trajectories $Q_n(t)$; the weak limit is with respect to the initial data ξ_j, η_j, and is attained almost surely with respect to the random frequencies ν_j. Throughout this section, expectations and variances are with respect to the initial data, but will be denoted for convenience by simply $\mathbb{E}$ and Var.

In two particular cases (3.12) can be solved analytically: for a free particle, $V(Q) = 0$, and for a particle in a quadratic potential, $V(Q) = \frac{1}{2} Q^2$. In the first case the process diffuses anomalously from its initial position. In the second case it approaches, as $t \to \infty$, a stationary distribution—the equilibrium Boltzmann distribution. These two cases are solvable since (3.12) is then linear, and can be solved by standard methods; see, e.g., ref. 44. Since $z(t)$ is a Gaussian (generalized) process, $Q(t)$ obtained by a linear mapping of $z(t)$ is also Gaussian. It follows that the statistics of the trajectories $Q(t)$ are fully determined by the mean $\mathbb{E}Q(t)$ and the auto-covariance $\mathrm{Cov}(Q(t), Q(s))$.

4.1. Free Particle

For a free particle, $V(Q) = 0$, $Q(t)$ solves the SIDE

$$\ddot{Q}(t) + \int_0^t \kappa(t-s)\, \dot{Q}(s)\, ds = z(t). \tag{4.1}$$

Equations of this type are commonly solved using the Laplace transform. We denote the Laplace transform of a function $Y(t)$ by

$$\hat{Y}(p) = \int_0^\infty Y(t)\, e^{-pt}\, dt.$$

Introducing the functions $H(t)$ and $h(t) = \dot{H}(t)$, which we define by their transforms:

$$\hat{H}(p) = p^{-1}[p + \hat{\kappa}(p)]^{-1}, \qquad \hat{h}(p) = [p + \hat{\kappa}(p)]^{-1},$$

the solution to (4.1) is

$$\begin{aligned} Q(t) &= Q_0 + P_0 H(t) + \int_0^t H(t-s)\, z(s)\, ds \\ P(t) &= P_0 h(t) + \int_0^t h(t-s)\, z(s)\, ds. \end{aligned} \tag{4.2}$$

Since $\mathbb{E}z(t) = 0$ it follows that

$$\begin{aligned} \mu_{\mathrm{Q}}(t) &= \mathbb{E}Q(t) = Q_0 + P_0 H(t) \\ \mu_{\mathrm{P}}(t) &= \mathbb{E}P(t) = P_0 h(t). \end{aligned} \tag{4.3}$$

The variance of the displacement and the momentum is further found to be given by

$$\begin{aligned} \beta\sigma_{\mathrm{QQ}}(t) &= 2\int_0^t H(s)\, ds - H^2(t) \\ \beta\sigma_{\mathrm{PP}}(t) &= 1 - h^2(t). \end{aligned} \tag{4.4}$$

If $\kappa(t) = a_0\, t^{-\gamma}$ then $h(t)$ is expressible in terms of a special function

$$h(t) = E_{2-\gamma}(-k_0 t^{2-\gamma}),$$

where $E_\alpha(t)$ are the Mittag–Leffler functions[13] defined by the series expansion,

$$E_\alpha(t) = \sum_{n=0}^{\infty} \frac{t^n}{\Gamma(\alpha n + 1)}.$$

We also introduce the generalized Mittag–Leffler functions,

$$E_{\alpha,\beta}(t) = \sum_{n=0}^{\infty} \frac{t^n}{\Gamma(\alpha n + \beta)},$$

so that $E_\alpha(t) = E_{\alpha,1}(t)$. The Mittag–Leffler functions play an important role in fractional differential calculus; they are a generalization of the exponential function, and reduce to it for $\alpha = 1$. The long-time behaviour of $E_{\alpha,\beta}(t)$ is

$$E_{\alpha,\beta}(t) \sim -\frac{t^{-1}}{\Gamma(\beta-\alpha)} \qquad t \to \infty. \tag{4.5}$$

After some straightforward manipulations we find

$$\begin{aligned} \beta\sigma_{QQ}(t) &= 2t^2\, E_{2-\gamma,3}(-k_0 t^{2-\gamma}) - [t\, E_{2-\gamma,2}(-k_0 t^{2-\gamma})]^2 \\ \beta\sigma_{PP}(t) &= 1 - [E_{2-\gamma}(-k_0 t^{2-\gamma})]^2. \end{aligned} \tag{4.6}$$

The long-time asymptotic behaviour of the variances is obtained by substituting (4.5),

$$\beta\sigma_{QQ}(t) \sim \frac{2}{k_0\Gamma(1+\gamma)}\, t^\gamma, \qquad \beta\sigma_{PP}(t) \sim 1 - \frac{t^{2(\gamma-1)}}{[k_0\Gamma(\gamma-1)]^2}. \tag{4.7}$$

Thus, the limiting behaviour of free particle is sub-diffusive with exponent γ. Graphs of $\beta\sigma_{QQ}(t)$ and $\beta\sigma_{PP}(t)$ for three values of γ are shown in Fig. 1. Normal diffusive behaviour is recovered as $\gamma \to 1$. The particularly remarkable property of sub-diffusion is the slow approach of $\sigma_{PP}(t)$ toward its equilibrium value β^{-1}. This is in contrast with the exponential equilibration rate of the velocity variance in a diffusive regime.

4.2. Quadratic Potential Well

We next consider the case of a particle in a quadratic potential well, $V(Q) = \frac{1}{2}Q^2$. The trajectory $Q(t)$ satisfies the SIDE

$$\ddot{Q}(t) + \int_0^t \kappa(t-s)\, \dot{Q}(s)\, ds + Q(t) = z(t). \tag{4.8}$$

Here again, the equation is linear hence $Q(t)$ is a Gaussian process. Introducing the functions $H_1(t)$ and $h_1(t) = \dot{H}_1(t)$ defined by their Laplace transform,

$$\hat{H}_1(p) = [p^2 + p\hat{\kappa}(p) + 1]^{-1}, \qquad \hat{h}_1(p) = p[p^2 + p\hat{\kappa}(p) + 1]^{-1}.$$

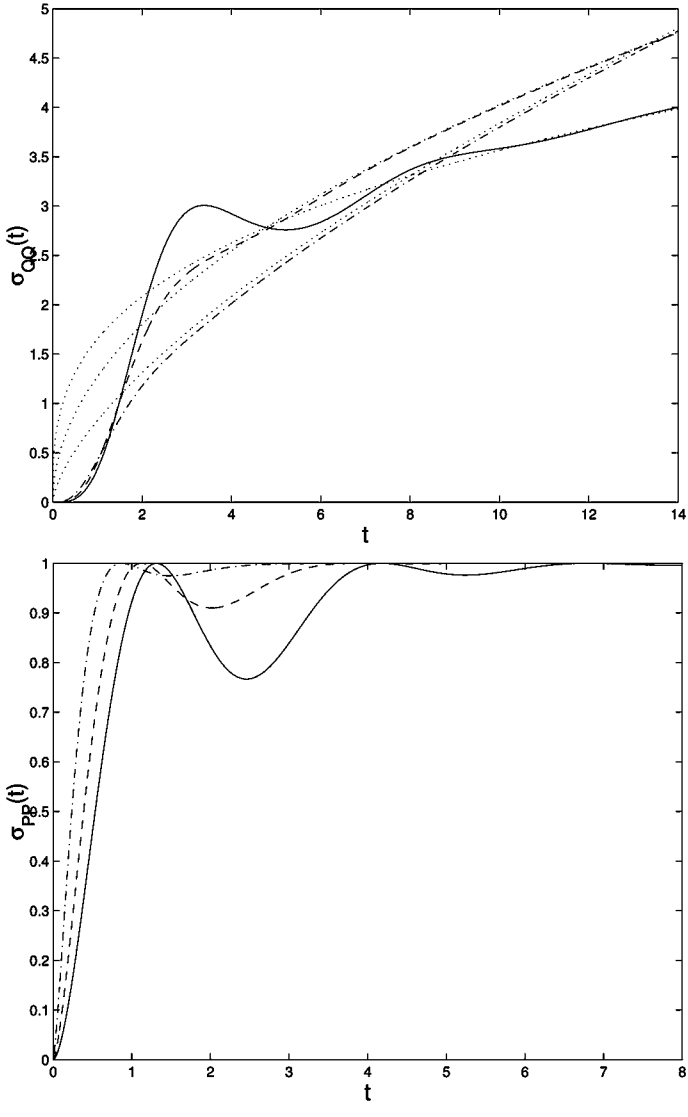

Fig. 1. $\sigma_{QQ}(t)$ (top) and $\sigma_{PP}(t)$ (bottom) for $k_0 = \Gamma(1-\gamma)$, $\beta = 1$, and $\gamma = 1/3$ (solid lines), $\gamma = 1/2$ (dashed lines), and $\gamma = 2/3$ (dash-dotted lines). The dotted lines in the top graph show the asymptotic solutions (4.7).

The solution to (4.8) is

$$\begin{aligned} Q(t) &= Q_0 + P_0 H_1(t) - Q_0 \int_0^t H_1(s)\, ds + \int_0^t H_1(t-s)\, z(s)\, ds \\ P(t) &= P_0 h_1(t) - Q_0 H_1(t) + \int_0^t h_1(t-s)\, z(s)\, ds, \end{aligned} \tag{4.9}$$

so that

$$\begin{aligned} \mu_{\mathrm{Q}}(t) &= Q_0 + P_0 H_1(t) - Q_0 \int_0^t H_1(s)\, ds \\ \mu_{\mathrm{P}}(t) &= P_0 h_1(t) - Q_0 H_1(t), \end{aligned}$$

and the variance of $Q(t)$ and $P(t)$ is given by

$$\begin{aligned} \beta\sigma_{\mathrm{QQ}}(t) &= 2\int_0^t H_1(s)\, ds - H_1^2(t) - \left[\int_0^t H_1(s)\, ds\right]^2 \\ \beta\sigma_{\mathrm{PP}}(t) &= 1 - h_1^2(t) - H_1^2(t). \end{aligned} \tag{4.10}$$

Note the structural similarity between this case and the free particle case. There is however a fundamental difference, which arises from the different asymptotic behaviours of $\hat{H}(p)$ and $\hat{H}_1(p)$ as $p \to 0$. Small p asymptotic behaviour in the Laplace domain determines the large t asymptotic behaviour in the time domain. Thus, trajectories diffuse away for the free particle, and remain bounded for a confining potential (a potential $V(Q)$ is called confining if the Boltzmann factor $e^{-\beta V(Q)}$ is normalizable).

Figure 2 shows graphs of the functions $h_1(t)$, $H_1(t)$, and $\int_0^t H_1(s)\, ds$ for $\gamma = 0.6$; these curves were obtained by numerically inverting the Laplace transform.[10] In Fig. 3 we show $\beta\sigma_{\mathrm{QQ}}(t)$ and $\beta\sigma_{\mathrm{PP}}(t)$, given by (4.10). Both approach, as $t \to \infty$, to the equilibrium value 1. Note the very slow algebraic decay of $\sigma_{\mathrm{QQ}}(t)$.

5. MARKOVIAN APPROXIMATION OF THE GENERALIZED LANGEVIN EQUATION

The generalized Langevin equation (3.12) defines a non-Markovian process. In the previous section we considered particular cases where the equation is linear, and can therefore be solved by analytical methods. Markovian systems can be analyzed through the differential equations that determine the evolution on the one- and multi-dimensional PDFs—the

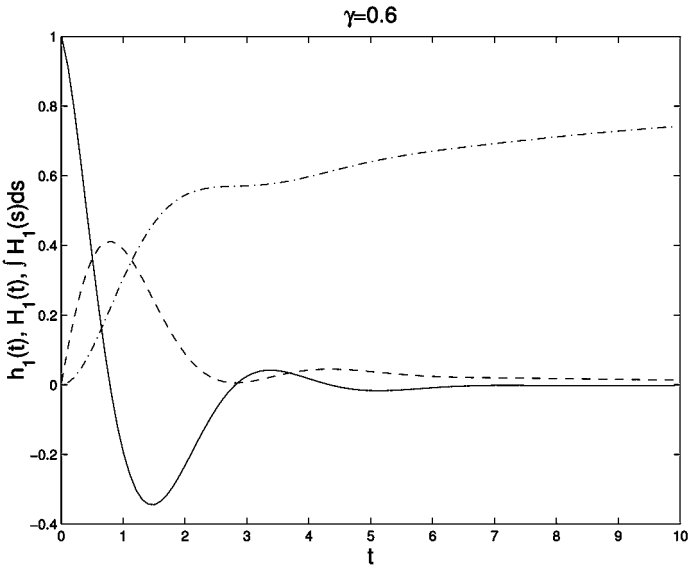

Fig. 2. The functions $h_1(t)$ (solid line), $H_1(t)$ (dashed line), and $\int_0^t H_1(s)\,ds$ (dash-dotted line) for $\gamma = 0.6$.

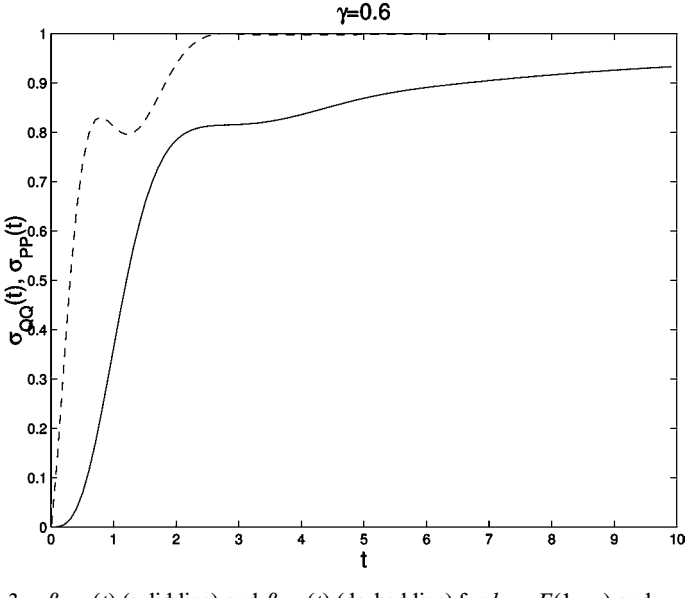

Fig. 3. $\beta\sigma_{\mathrm{QQ}}(t)$ (solid line) and $\beta\sigma_{\mathrm{PP}}(t)$ (dashed line) for $k_0 = \Gamma(1-\gamma)$ and $\gamma = 0.6$.

Fokker–Planck equation for system driven by white noise.[17] It is not known, in general, how to derive such differential equations for non-Markovian system (a general formalism does exist,[19,42] but does not provide closed equations without further approximation). A classical approach is to approximate the non-Markovian system by a Markovian one, through the addition of auxiliary variables. This approach was already proposed by Mori[40] in the context of the Mori–Zwanzig projection formalism;[41,53] see also ref. 20. It was further developed and applied in a series of papers by Kłosek–Dygas *et al.*[11,12] Recently, a Markovian approximation was used for $1/f^{\alpha}$-noise;[34] it is this approximation which we adopt in this section.

We rewrite the generalized Langevin equation as a first-order system:

$$\dot{Q}(t) = P(t)$$

$$\dot{P}(t) = -V'(Q(t)) - \int_0^t \kappa(t-s)\, P(s)\, ds + z(t),$$

recalling that $z(t)$ is a stationary centered Gaussian process with autocovariance

$$\mathbb{E} z(t)\, z(s) = \beta^{-1} \kappa(t-s).$$

The goal is to approximate the trajectories $(Q(t), P(t))$ by trajectories $(\tilde{Q}(t), \tilde{P}(t))$ that solve a Markovian system with m auxiliary variables:

$$\begin{aligned} d\tilde{Q}(t) &= \tilde{P}(t)\, dt & \tilde{Q}(0) &= Q_0 \\ d\tilde{P}(t) &= -[V'(\tilde{Q}(t)) - g^T u(t)]\, dt & \tilde{P}(0) &= P_0 \\ du(t) &= -[\tilde{P}(t)\, g + A u(t)]\, dt + C\, dB(t) & u(0) &\sim \mathcal{N}(0, \Sigma), \end{aligned} \tag{5.1}$$

where $u(t) = (u_1(t), \ldots, u_m(t))^T$, g is a constant m-vector, A, C, Σ are constant $m \times m$ matrices, and $B(t)$ is a vector of m independent standard Brownian motions.

The equation for $u(t)$ is linear, and its solution can be written explicitly; after multiplication by the row vector g^T we obtain

$$g^T u(t) = -\int_0^t \tilde{\kappa}(t-s)\, \tilde{P}(s)\, ds + \tilde{z}(t),$$

where

$$\tilde{\kappa}(t) = g^T e^{-At} g$$

$$\tilde{z}(t) = g^T e^{-At} u(0) + \int_0^t g^T e^{-A(t-s)} C\, dB(s).$$

Since $g^T u(t)$ represents in (5.1) the interaction term between the heat bath and the distinguished particle, the parameter A, C, Σ, g have to be chosen such that $\tilde{z}(t)$ is a stationary centered Gaussian process with auto-covariance $\beta^{-1}\tilde{\kappa}(t)$, the latter being an approximation of $\beta^{-1}\kappa(t)$.

By our choice of $u(0)$, $\tilde{z}(t)$ is a centered Gaussian process with auto-covariance

$$\mathbb{E}\tilde{z}(t)\,\tilde{z}(s) = g^T e^{-At}\left[\Sigma + \int_0^{t\wedge s} e^{A\tau} C C^T e^{A^T\tau}\,d\tau\right] e^{-A^T s} g.$$

Comparing with the expression for $\tilde{\kappa}(t)$ we need

$$\Sigma + \int_0^s e^{A\tau} C C^T e^{A^T\tau}\,d\tau = \beta^{-1} e^{As} e^{A^T s},$$

which is satisfied if

$$\Sigma = \beta^{-1} I, \qquad CC^T = \beta^{-1}(A + A^T). \tag{5.2}$$

Thus, A and g determine Σ and C.

The relation between the parameters A, g, and the kernel is best viewed through the Laplace transform of $\tilde{\kappa}(t)$,

$$\hat{\tilde{\kappa}}(p) = \int_0^\infty e^{-pt} g^T e^{-At} g\,dt = g^T (A + pI)^{-1} g. \tag{5.3}$$

The right hand side is a rational function of p; the numerator is a polynomial of degree $m-1$ and the denominator is a polynomial of degree m. Thus, the Markovian approximation consists of two steps: First one has to approximate the Laplace transform of the memory kernel by a rational function. Then, a matrix A and a vector g have to be constructed such to yield the required $\hat{\tilde{\kappa}}(p)$ via (5.3). The first step invokes approximation theory whereas the second step is made systematic within linear system theory (see, e.g., Zadeh and Desoer[51]).

Exponential Approximation. The simplest case occurs when $\kappa(t)$ can be approximated by a sum of exponentials:

$$\tilde{\kappa}(t) = \sum_{k=1}^{m} \Delta_k^2\, e^{-\alpha_k t},$$

$\alpha_k > 0$, which corresponds in p-space to

$$\hat{\tilde{\kappa}}(p) = \sum_{k=1}^{m} \frac{\Delta_k^2}{p+\alpha_k}.$$

This approximation is realized by taking A diagonal with elements $A_{ii} = \alpha_i > 0$, and $g = (\Delta_1, \Delta_2, \ldots, \Delta_m)^T$.

Jacobi-Fractions. Another class of problems is when $\hat{\kappa}(p)$ has a continued fraction representation in the form of a Jacobi fraction:[49]

$$\hat{\tilde{\kappa}}(p) = \cfrac{\Delta_1^2}{p+\alpha_1 + \cfrac{\Delta_2^2}{p+\alpha_2 + \cfrac{\Delta_3^2}{p+\alpha_3 + \ddots}}}, \tag{5.4}$$

where the α_k are positive. Approximating $\hat{\kappa}(p)$ by its mth convergent, this approximation can be realized by taking A tridiagonal of the form

$$A = \begin{pmatrix} \alpha_1 & -\Delta_2 & & & \\ \Delta_2 & \alpha_2 & -\Delta_3 & & \\ & \Delta_3 & \alpha_3 & -\Delta_4 & \\ & & \ddots & \ddots & \ddots \\ & & & \Delta_m & \alpha_m \end{pmatrix}$$

and $g = (\Delta_1, 0, \ldots, 0)^T$. Since $A + A^T$ is diagonal, so is C with $C_{ii} = \sqrt{2\alpha_i}$. Continued fraction approximations of this type are used by Kłosek–Dygas *et al.*[11, 12] A GLE based on a continued fraction expansion was also constructed by Adelman and Doll for a particle interacting with an harmonic chain.[2]

Comment. If A is diagonalizable, $A = T^{-1} \Lambda T$, then (5.1) can be brought into the equivalent form

$$\begin{aligned} d\tilde{Q} &= \tilde{P}(t)\, dt & \tilde{Q}(0) &= q_0 \\ d\tilde{P} &= -[V'(\tilde{Q}(t)) - g^T T v(t)]\, dt & \tilde{P}(0) &= P_0 \\ dv(t) &= -[\tilde{P}(t)\, T^{-1} g + \Lambda v(t)]\, dt + T^{-1} C\, dB(t) & v(0) &\sim T^{-1} \mathcal{N}(0, \Sigma), \end{aligned}$$

where the equations for $v_1(t), \ldots, v_m(t)$ are now decoupled, i.e., an exponential approximation of the memory kernel can be constructed.

In this paper the memory kernel has Laplace transform $\hat{\kappa}(p) = \Gamma(1-\gamma)\, p^{\gamma-1}$, $0 < \gamma < 1$. A Markovian approximation for the case $\gamma = 1/2$

is constructed by Landis *et al.*,[34] and the same construction can be applied for arbitrary γ. For $\gamma = 1/2$ we have the continued-fraction representation:

$$\hat{\tilde{\kappa}}(p) = \cfrac{\Gamma(1-\gamma)}{1+\cfrac{1}{\cfrac{2}{p-1}+\cfrac{1}{2+\cfrac{1}{\cfrac{2}{p-1}+\cfrac{1}{2+\ddots}}}}}$$

whose first four convergents are

$$\hat{\tilde{\kappa}}_1(p) = \frac{2\Gamma(1-\gamma)}{p+1}$$

$$\hat{\tilde{\kappa}}_2(p) = \frac{2\Gamma(1-\gamma)\,(2p+2)}{p^2+6p+1}$$

$$\hat{\tilde{\kappa}}_3(p) = \frac{2\Gamma(1-\gamma)\,(3p^2+5p+3)}{p^3+15p^2+15p+1}$$

$$\hat{\tilde{\kappa}}_4(p) = \frac{2\Gamma(1-\gamma)\,(4p^3+28p^2+28p+4)}{p^4+28p^3+70p^2+28p+1}.$$

A direct verification shows that this approximation is realized by the choice

$$A = \begin{pmatrix} 1 & 2 & 2 & 2 & \cdots & 2 \\ 2 & 5 & 6 & 6 & \cdots & 6 \\ 2 & 6 & 9 & 10 & \cdots & 10 \\ \vdots & \vdots & \vdots & & \ddots & \\ 2 & 6 & 10 & 14 & \cdots & 4m-3 \end{pmatrix} \qquad g = \sqrt{2\Gamma(1-\gamma)} \begin{pmatrix} 1 \\ 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}. \tag{5.5}$$

The matrix C is subsequently extracted from (5.2).

In Fig. 4 we compare the function $1/p^{1/2}$ with the mth convergent $\hat{\tilde{\kappa}}_m(p)$ for $m = 2, 4, 8, 16$. For the 16th convergent the graphs are indistinguishable over almost five decades.

The Markovian approximation (5.1) with A, g given by (5.5) can be solved for the two problems considered in Section 4. One only has to replace $\hat{\kappa}(p)$ by $\hat{\tilde{\kappa}}(p)$ and numerically invert the Laplace transform. Consider the free particle: in Fig. 5 we compare $\sigma_{QQ}(t)$ given by (4.7) with $\beta = 1$ and $\gamma = 1/2$ (solid line) with the result of the Markovian approximation

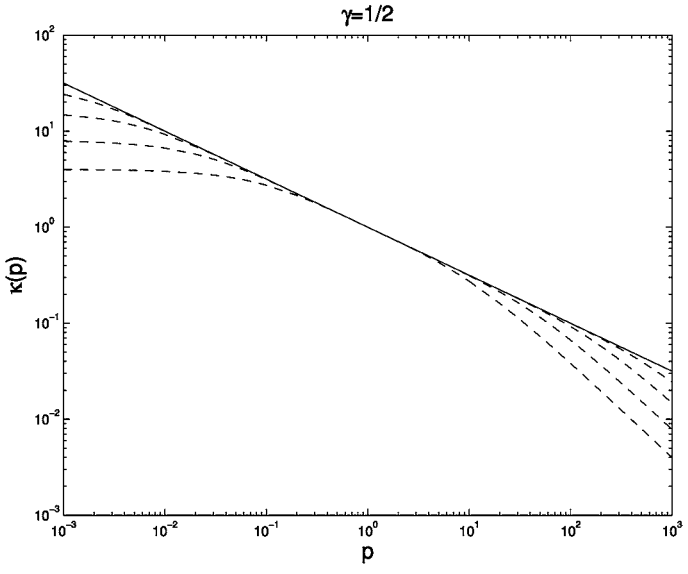

Fig. 4. Solid line: Log-log graph of the function $1/p^{1/2}$. Dashed lines: continued fraction approximation with $m = 2$, 4, 8, and 16 terms.

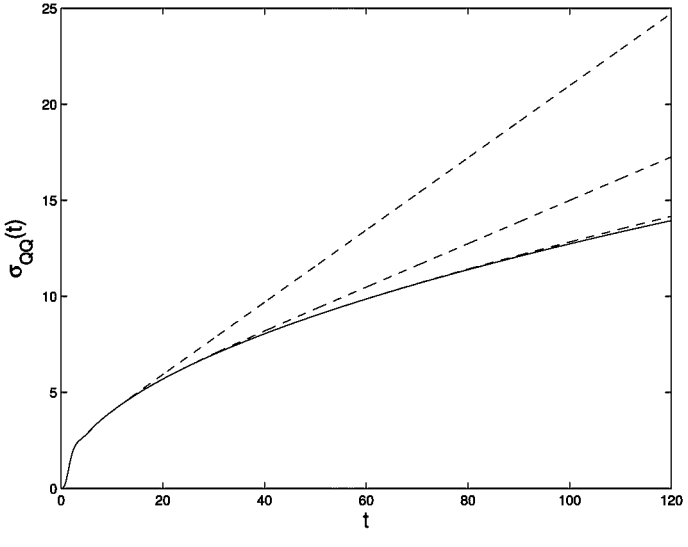

Fig. 5. Solid line: $\sigma_{QQ}(t)$ for a free particle with $\gamma = 1/2$ and $\beta = 1$. Dashed lines: $\sigma_{QQ}(t)$ for the Markovian approximation with $m = 2$, 4, 8.

with $m = 2, 4, 8$ (dashed lines). The larger m the longer is the intermediate asymptotic regime of anomalous diffusion. For $m = 8$ the curves almost coincide up to time $t = 120$.

6. COMPARISON WITH THE FRACTIONAL FOKKER–PLANCK EQUATION

For a particle satisfying a Markovian equation driven by random noise, the phase space density $w(Q, P, t)$ is governed by a Fokker–Planck equation (FPE).[17] In the limit of high friction the FPE may be reduced into a Smulochowsky equation for the marginal density $W(Q, t)$:

$$\frac{\partial}{\partial t} W(Q, t) = \mathscr{L} W(Q, t),$$

where

$$\mathscr{L} g(Q) = K_D \left[-\frac{\partial}{\partial Q} V'(Q) + \frac{\partial^2}{\partial Q^2} \right] g(Q),$$

and K_D is a diffusion constant. (We are assuming here unit temperature and a dimensionless setting in which dimensional constants, such as the Boltzmann constant, can be ignored.) The terminology used in the literature is non-uniform; the equation for $w(Q, P, t)$ is sometimes called a Klein–Kramers equation, whereas the equation for $W(Q, t)$ is sometimes called a Fokker–Planck equation.

Several authors addressed generalizations of the Fokker–Planck equation for non-Markovian systems that exhibit fractional diffusion. One such generalization is the fractional Fokker–Planck equation (FFPE):

$$\frac{\partial}{\partial t} W(Q, t) = {}_0D_t^{1-\gamma} \mathscr{L} W(Q, t), \tag{6.1}$$

where ${}_0D_t^{1-\gamma}$ is the Riemann–Liouville fractional derivative[43]

$${}_0D_t^{1-\gamma} g(t) = \frac{1}{\Gamma(\gamma)} \frac{\partial}{\partial t} \int_0^t \frac{g(s)}{(t-s)^{1-\gamma}} ds.$$

The formal solution of (6.1) can be expressed in terms of the Mittag–Leffler function:

$$W(Q, t) = E_\gamma(\mathscr{L} t^\gamma) W(Q, 0).$$

The attractive features of the FFPE are: (i) $W(Q, t)$ remains nonnegative and normalized, as required by a probability density. (ii) In the presence of a stationary potential, $V(Q)$, the distribution tends as $t \to \infty$ to the Boltzmann distribution $\exp[-V(Q)]$. (iii) The relaxation rate is subexponential; in fact, the solution operator of the FFPE is expressible in terms of Mittag–Leffler functions. (iv) The FFPE tends to a normal Fokker–Planck equation as $\gamma \to 1$. (v) Finally, the FFPE can be shown to be the correct governing equation in certain cases. For example, it is the limiting master equation for continuous time random walk.[5] The literature on the FFPE is too vast to be fully covered. We address to reader to refs. 4, 5, 37, and 45 for recent developments.

Comment. Several authors derived fractional diffusion equations for solutions to linear SIDE.[26, 36, 44, 50] The drawback of these derivations is that they are restricted to the problem at hand; in order to construct the diffusion equation one has to know beforehand the evolution of the probability density. These fractional diffusion equations are not used as a predictive tool; they are merely a reformulation of a result that can be obtained independently.

In the rest of this section we solve the FFPE (6.1) for a free particle and compare the solution with the Gaussian distribution derived in Section 4. A general method for solving the FFPE, based on an integral transform that maps fractional diffusion into normal diffusion, was presented by Barkai.[4] The case of a free particle can be solved by more direct methods; we follow here ref. 5.

The FFPE for a free particle is

$$\frac{\partial}{\partial t} W(Q, t) = {}_0D_t^{1-\gamma} \left[K_D \frac{\partial^2}{\partial Q^2} W(Q, t) \right], \tag{6.2}$$

where we take $W(Q, 0) = \delta(Q)$. A solution for this problem was derived by Schneider and Wyss,[45] expressible in terms of the Fox H function.[16] From a computational point of view it is easier to solve (6.2) in p-space and revert back to t-space using a numerical inversion of the Laplace transform. Using the fact that fractional derivatives satisfy the transformation rule

$${}_0D_t^{1-\gamma} g(t) \mapsto p^{1-\gamma} \hat{g}(p),$$

Eq. (6.2) reads in p-space

$$-\delta(Q) + p\hat{W}(Q, p) = K_D p^{1-\gamma} \frac{\partial^2}{\partial Q^2} \hat{W}(Q, p),$$

and $W(Q, p)$ vanishes as $|Q| \to \infty$. This equation can be solved by standard methods, yielding

$$\hat{W}(Q, p) = \frac{1}{2\, K_D^{1/2} p^{1-\gamma/2}} \exp\left(-\frac{p^{\gamma/2}}{K_D^{1/2}} |Q| \right).$$

In Fig. 6 we plot $W(Q, t)$ for $\gamma = 0.6$ and $t = 0.5, 1, 2, 4, 8$. Note the sharp non-Gaussian shape of the solution, in contrast with the Gaussian distribution of $Q(t)$. In addition to the cusp at $Q = 0$, the large-Q tail of the distribution also exhibits non-Gaussian behaviour; as $Q \to \infty$

$$W(Q, t) \sim Q^{-1} \left(\frac{Q^2}{t^\gamma} \right)^{1-\gamma/2} \exp\left[-c \left(\frac{Q^2}{t^\gamma} \right)^{1/(2-\gamma)} \right],$$

where c is a constant that depends on γ and on K_D.[5]

Thus, we conclude that the FFPE does not apply to the GLE (3.12). As mentioned above, it is possible to derive diffusion equations for $W(Q, t)$ both for the free particle and for the quadratic potential, but these equations do not generalize for other cases.

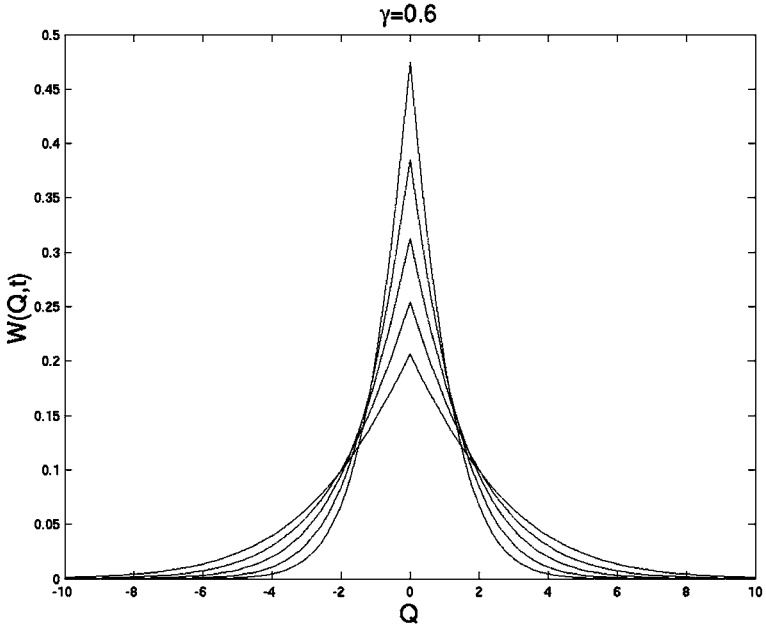

Fig. 6. Solution of (6.2): $W(Q, t)$ versus Q for $t = 0.5, 1, 2, 4, 8$ and $\gamma = 0.6$.

7. NUMERICAL RESULTS

In this section we present numerical results for the Hamiltonian system (2.2) and provide some comparison with the Markovian approximation (5.1) with A, g, C, Σ given by (5.5) and (5.2). To integrate (2.2) we used a symplectic Euler scheme (ref. 24, p. 312), which is known to exhibit good stability properties, and in particular, allows the use of relatively large time steps.[25, 48] The results presented below are for heat baths consisting of $n = 2000$ particles; we used $a = 1/3$ and $c = \frac{1}{2}(1-a)$ in the selection of the random frequencies. Ensemble averages were calculated by averaging over collections of 10^4 realizations. Only small variations were found when the number of particles was increased to $n = 5000$ or when the size of the ensemble was increased.

In Fig. 7 we show a sample path of $Q_n(t)$ for the case of a free particle, $V(Q) = 0$ and $\gamma = 0.6$.

In Fig. 8 we show the evolution of the variance of $Q_n(t)$ and $P_n(t)$ for a free particle with initial conditions $Q_0 = P_0 = 0$. The thick line corresponds to the statistics generated by 10^4 realizations. The dotted lines are

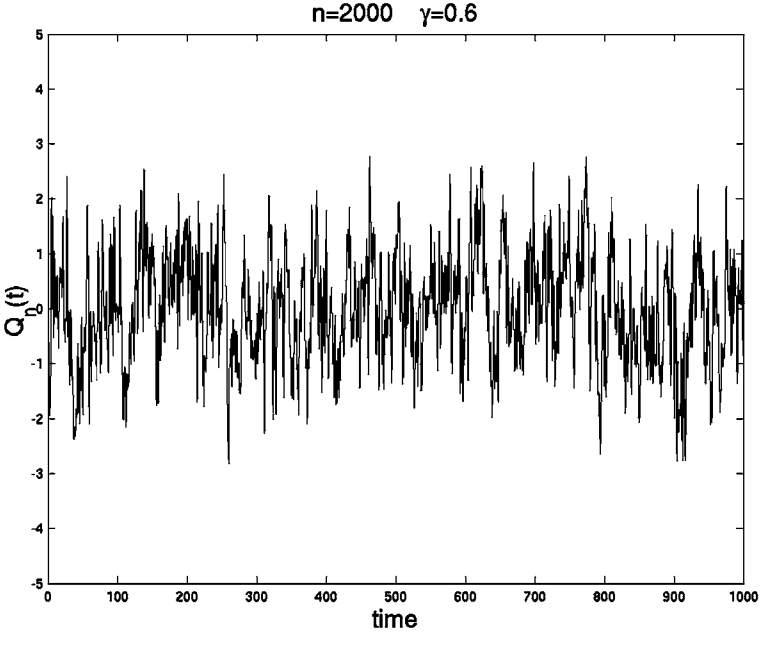

Fig. 7. Sample path of $Q_n(t)$ for a system of $n = 2000$ oscillators and $\gamma = 0.6$.

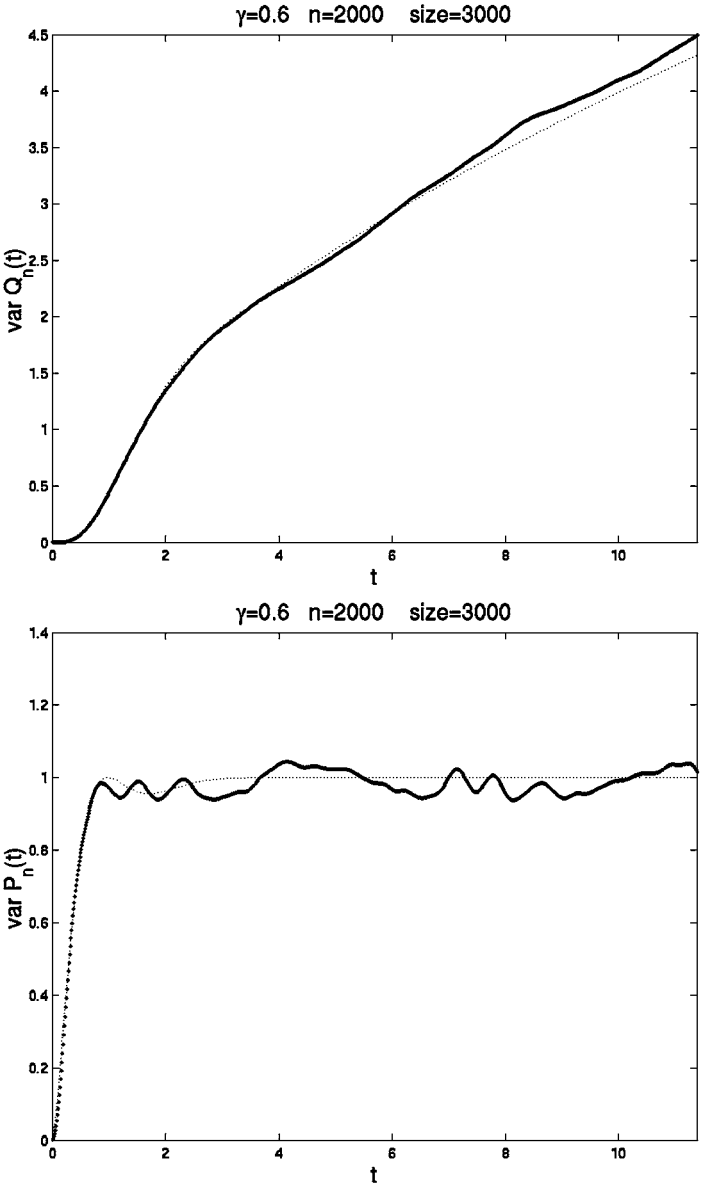

Fig. 8. The mean square displacement $Q_n(t)$ (top) and mean square momentum $P_n(t)$ (bottom) for a free particle in initial state $Q_0 = P_0 = 0$ and $\gamma = 0.6$. The thick lines were obtained by averaging over 3000 realizations. The dotted lines represent the asymptotic predictions (4.6).

the $n \to \infty$ theoretical predictions $\sigma_{QQ}(t)$ and $\sigma_{PP}(t)$, given by (4.6). For Var $Q_n(t)$ the two curves are almost indistinguishable up to $t = 6$. The agreement between Var $P_n(t)$ and $\sigma_{PP}(t)$ is less sharp; in particular, Var $P_n(t)$ oscillates around its asymptotic value 1, which we interpret as a sampling error.

In Fig. 9 we show the evolution of the distribution of $Q_n(t)$ generated by an ensemble of 10^4 realizations. The top figure shows snapshots of the distribution, which, as t increases, approaches the Boltzmann distribution (thick dashed line). The figure on the bottom show the evolution of Var $Q_n(t)$ and Var $P_n(t)$, which we compare to the $n \to \infty$ predictions $\sigma_{QQ}(t)$ and $\sigma_{PP}(t)$, given by (4.10). The agreement is again very good, except for the persistent fluctuations in Var $P_n(t)$.

In Fig. 10 we show the evolution of the distribution of $Q_n(t)$ for a particle in a double-well potential, $V(Q) = Q^4/4 - Q^2/2$. Here too, the distribution approaches, as $t \to \infty$, the equilibrium Boltzmann distribution.

So far, the statistical analysis was generated by evolving ensembles. It is also of interest to compare ensemble averages with long term time averages generated by sample paths. Jakšić and Pillet studied the ergodicity of GLEs,[30] but it is not known whether the finite n system is ergodic and if it is, whether the sequence of invariant measures converges to the invariant measure of the limiting equation (recall that our theorems are restricted to bounded time intervals). In Fig. 11 we show the empirical distribution for a particle in a single-well potential (open circles), and compare it with the Boltzmann distribution (dashed line). In Fig. 12 we repeat the calculation for a double-well potential. Our results indicate that the process is indeed ergodic, however, the convergence of the empirical distribution requires very long averaging intervals (cf. ref. 33). This is due, presumably, to the very long correlation time associated with the algebraic decay of the memory kernel.

Finally, we compare the Hamiltonian system (2.2) with the Markovian approximation (5.1). In Fig. 13 we show the distribution of exit times,[2] τ, from a potential well. We used the double-well potential $V(Q) = Q^4/4 - Q^2/2$ and started the distinguished particle in the left well, assigning initial conditions $Q_0 = -1$, $P_0 = 0$. We ran 10^4 realizations, defining the exit time as the first crossing time of the point $Q = 0$. The solid lines represent the

[2] The problem of exit times for non-Markovian process was studied extensively in the context of activation rate theory. Kramers' theory and its extension to systems with memory kernels is reviewed by Hänggi *et al.*[28] The original work can be found in Grote and Hynes,[22] Hänggi and Mujtabai,[27] and Carmeli and Nitzan;[9] numerical tests for an exponential memory function were conducted by Straub *et al.*[46, 47].

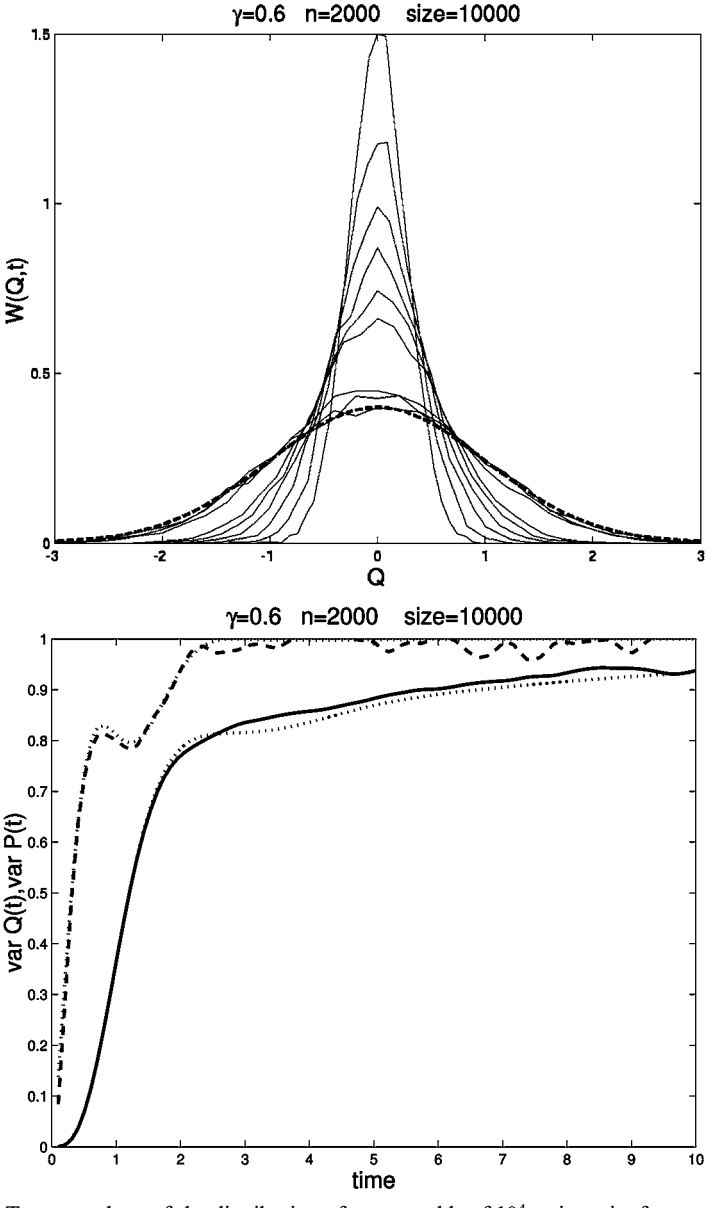

Fig. 9. Top: snapshots of the distribution of an ensemble of 10^4 trajectories for a particle in a single-well potential $V(Q) = Q^2/2$, initial data $Q_0 = P_0 = 0$, and $\gamma = 0.6$. The solid lines show the distribution at times $t = 0.5, 0.6, 0.7, 0.8, 0.9, 1, 2, 4, 8$. The dashed line corresponds to the Boltzmann distribution. Bottom: time evolution of Var $Q_n(t)$ (thick solid line) and Var $P_n(t)$ (thick dashed line). The dotted lines are the $n \to \infty$ predictions (4.10).

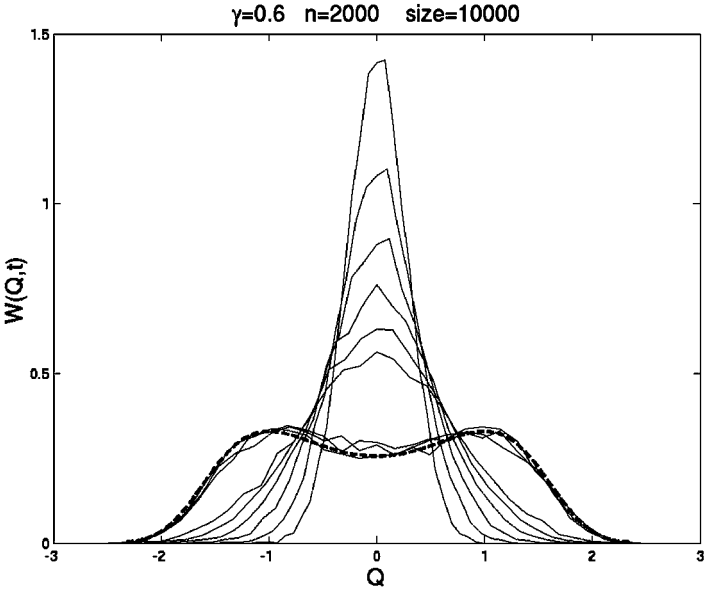

Fig. 10. Snapshots of the distribution of an ensemble of 10^4 trajectories for a particle in a double-well potential $V(Q) = Q^4/4 - Q^2/2$, initial data $Q_0 = P_0 = 0$, and $\gamma = 0.6$. The solid lines show the distribution at times $t = 0.5, 0.6, 0.7, 0.8, 0.9, 1, 2, 4, 8$. The dashed line corresponds to the Boltzmann distribution.

distribution of τ for the Hamiltonian system with $n = 2000$ and $\gamma = 1/2$. The symbols represent the distribution of τ for the SDE (5.1) with $m = 4$ (circles) and $m = 8$ (crosses). The agreement is very good, and the difference between 4 and 8 auxiliary variables is very small, probably dominated by sampling errors.

8. CONCLUDING REMARKS

We presented a simple particle-in-a-heat-bath model, which gives rise to fractional kinetics. While anomalous diffusion strictly occurs only in the limit of an infinitely large heat bath, thousands of heat bath particle suffice to observe anomalous diffusion over tens of time units. Indeed, anomalous diffusion should be viewed as intermediate asymptotics when the relaxation time is larger than the observation time. The limiting process was found to satisfy a SIDE driven by $1/f^{\alpha}$-noise. Since the limiting driving noise was Gaussian, the SIDE could be approximated by a Markovian SDE through the addition of auxiliary variables. Our numerical results show that the

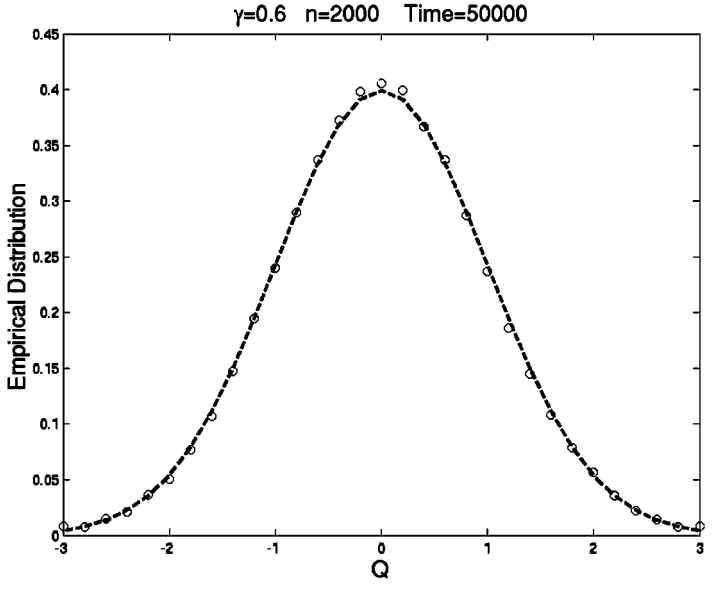

Fig. 11. Open circles: empirical distribution of Q_n for a particle in a single-well potential $V(Q) = Q^2/2$ and $\gamma = 0.6$; the distribution was calculated over a sample path of length $T = 50000$. Dashed line: the Boltzmann distribution.

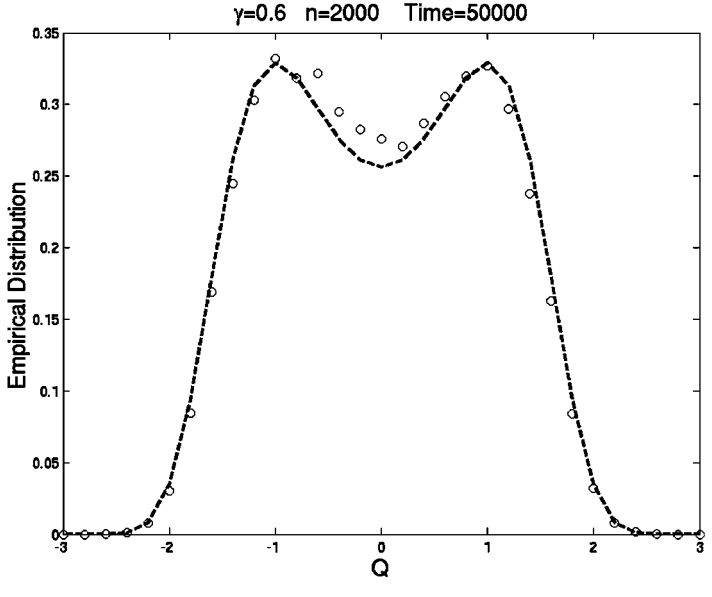

Fig. 12. Same as Fig. 11 for a double-well potential $V(Q) = Q^4/4 - Q^2/2$.

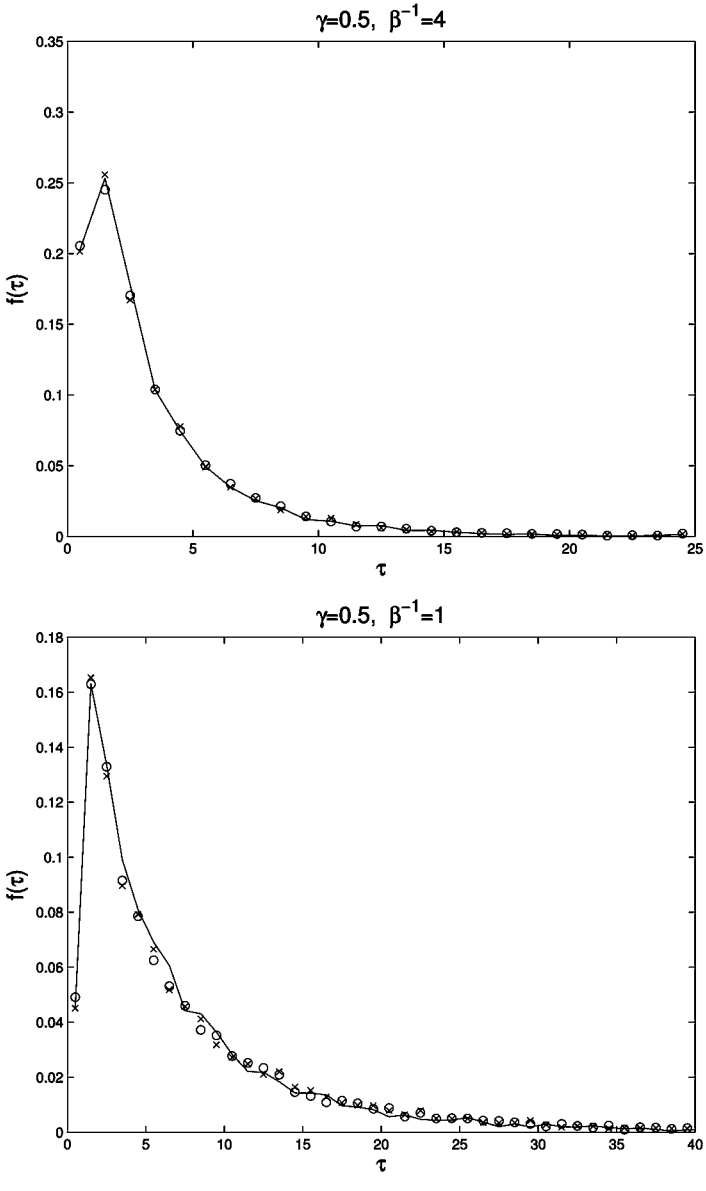

Fig. 13. Distribution of exit times from the potential-well $V(Q) = Q^4/4 - Q^2/2$ for initial data $Q_0 = -1$, $P_0 = 0$; the exit time is defined as the first passage time through the point $Q = 0$. The distributions are based on ensembles of 10^4 realizations. The solid line corresponds to the Hamiltonian system (2.2) with $n = 2000$ particles and $\gamma = 1/2$; the dashed line corresponds to the Markovian approximation (5.1) with $m = 4$ (circles) and $m = 8$ (crosses). The temperature is $\beta^{-1} = 1$ (top) and $\beta^{-1} = 4$ (bottom).

trajectories induced by a heat bath of a large number of particles may be well approximated by an SDE with, say, four extra variables. This work generalizes some of the results in ref. 33 where an exponential memory was eliminated by the addition of one extra variable.

We note that the method of Markovian approximation applies only for the case where the driving noise is Gaussian and the memory term is linear. The treatment of SIDEs driven by non-Gaussian noise is an open problem. While the FFPE was found inadequate in the present problem is may well be appropriate for other classes of non-Markovian systems with diverging relaxation times.

ACKNOWLEDGMENTS

The application of heat bath models to the study of fractional diffusion was suggested independently by A. Majda and W. E. Special thanks to Z. Schuss for introducing me to the method of Markovian approximation and for a critical reading of the manuscript. I am grateful to G. I. Barenblatt, D. Givon, and O. Hald for fruitful discussions and to A. Stuart for useful comments. This research was supported in part by the Israel Science Foundation founded by the Israel Academy of Sciences and Humanities, and by the Applied Mathematical Sciences subprogram of the Office of Energy Research of the US Department of Energy under Contract DE-AC03-76-SF00098.

REFERENCES

1. M. Abramowitz and I. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables* (Dover, New York, 1971).
2. S. Adelman and J. Doll, Generalized Langevin equation approach for atom/solid-surface scattering: Collinear atom/harmonic chain model, *J. Chem. Phys.* **61**:4242–4245 (1974).
3. F. Amblard, A. Maggs, B. Yurke, A. Pargellis, and S. Leibler, Subdiffusion and anomalous local viscoelasticity in actin networks, *Phys. Rev. Lett.* **77**:4470–4473 (1996).
4. E. Barkai, Fractional Fokker–Planck equation, solution and application, *Phys. Rev. E* **63**:046118 (2001).
5. E. Barkai, R. Metzler, and J. Klafter, From continuous time random walks to the fractional Fokker–Planck equation, *Phys. Rev. E* **61**:132–138 (2000).
6. E. Barkai and R. Silbey, Fractional Kramers equation, *J. Phys. Chem. B* **104**:3866–3874 (2000).
7. P. Billingsley, *Probability and Measure*, 3rd ed. (Wiley, New York, 1995).
8. P. Billingsley, *Convergence of Probability Measures*, 2nd ed. (Wiley, New York, 1999).
9. B. Carmeli and A. Nitzan, Non-Markovian theory of activated rate processes. III. Bridging between the Kramers limits, *Phys. Rev. A* **29**:1481–1495 (1984).
10. F. de Hoog, J. Knight, and A. Stokes, An improved method for numerical inversion of Laplace transforms, *SIAM J. Sci. Stat. Comp.* **3**:357–366 (1982).

11. M. Dygas, B. Matkowsky, and Z. Schuss, A singular perturbation approach to non-Markovian escape rate problems, *SIAM J. Appl. Aath.* **46**:265–298 (1986a).
12. M. Dygas, B. Matkowsky, and Z. Schuss, A singular perturbation approach to non-Markovian escape rate problems with state dependent friction, *J. Chem. Phys.* **84**:3731–3738 (1986b).
13. A. Erdélyi, *Tables of Integral Transforms* (McGraw–Hill, New York, 1954).
14. G. Ford and M. Kac, On the quantum Langevin equation, *J. Stat. Phys.* **46**:803–810 (1987).
15. G. Ford, M. Kac, and P. Mazur, Statistical mechanics of assemblies of coupled oscillators, *J. Math. Phys.* **6**:504–515 (1965).
16. C. Fox, The *G* and *H*-functions as symmetrical Fourier fernels, *Trans. Amer. Math. Soc.* **98**:395–429 (1961).
17. I. Gikhman and A. Skorokhod, *Stochastic Differential Equations* (Springer-Verlag, Berlin, 1972).
18. I. Gikhman and A. Skorokhod, *Introduction to the Theory of Random Processes* (Dover, Mineola, NY, 1996).
19. H. Grabert, P. Hänggi, and P. Talkner, Microdynamics and nonlinear stochastic processes of gross variables, *J. Stat. Phys.* **22**:537–552 (1980).
20. P. Grigolini, A generalized Langevin equation for dealing with nonadditive fluctuations, *J. Stat. Phys.* **27**:283–316 (1982).
21. G. Gripenberg, S.-O. Londen, and O. Staffans, *Volterra Integral and Functional Equations* (Cambridge University Press, Cambridge, 1990).
22. R. Grote and J. Hynes, The stable states picture of chemical reactions: II. Rate constants for condensed and gas phase reaction models, *J. Chem. Phys.* **73**:2715–2732 (1980).
23. P. Guest, *Laplace Transform and an Introduction to Distributions* (Ellis Horwood Ltd, Chichester, England, 1991).
24. E. Hairer, S. Norsett, and G. Wanner, *Solving Ordinary Differential Equations: Nonstiff Problems* (Springer-Verlag, Berlin, 1993).
25. O. Hald and R. Kupferman, Asymptotic and numerical analyses for mechanical models of heat baths, *J. Stat. Phys.* **106**:1121–1184 (2002).
26. P. Hanggi, The functional derivative and its use in the description of noisy dynamical systems, in *Stochastic Processes Applied to Physics*, L. Pesquera and M. Rodriguez, eds. (Santander, Spain, 1985), pp. 69–95.
27. P. Hänggi and F. Mojtabai, Thermally activated escape rate in presence of long-time memory, *Phys. Rev. A* **26**:1168–1170 (1982).
28. P. Hänggi, P. Talkner, and M. Bokovec, Reaction rate theory: Fifty years after Kramers, *Rev. Mod. Phys.* **62**:251–341 (1990).
29. W. Huisinga, C. Schütte, and A. Stuart, Extracting macroscopic stochastic dynamics: Model problems, *Comm. Pure Appl. Math.* **56**:234–269 (2003).
30. V. Jakšić and C.-A. Pillet, Ergodic properties of the non-Markovian Langevin equation, *Lett. Math. Phys.* **41**:49–57 (1997).
31. I. Karatzas and S. Shreve, *Brownian Motion and Stochastic Calculus* (Springer, New York, 1991).
32. R. Kupferman and A. Stuart, Fitting SDE models to nonlinear Kac–Zwanzig heat bath models, submitted to *Physica D* (2003).
33. R. Kupferman, A. Stuart, J. Terry, and P. Tupper, Long term behaviour of large mechanical systems with random initial data, *Stoch. Dyn.* **2**:533–562 (2002).
34. S. Landis, B.-Z. Borovsky, and Z. Schuss, The influence of $1/f^{\alpha}$ phase noise on a second-order delay lock loop: Model construction and analysis, preprint, (2003).
35. B. Mandelbrot and J. van Ness, Fractional Brownian motion, fractional Gaussian noise and applications, *SIAM Rev.* **10**:422 (1968).

36. J. Masoliver and K. Wang, Free inertial processes driven by Gaussian noise: Probability distributions, anomalous diffusion, and fractal behavior, *Phys. Rev. E* **51**:2987–2995 (1995).
37. R. Metzler, E. Barkai, and J. Klafter, Anomalous diffusion and relaxation close to thermal equilibrium: A fractional Fokker–Planck equation approach, *Phys. Rev. Lett.* **82**:3563–3567 (1999).
38. R. Miller, *Nonlinear Volterra Integral Equations* (W. A. Benjamin, Philippines, 1971).
39. E. Montroll and G. Weiss, Random walks on lattices. II, *J. Math. Phys.* **6**:167–181 (1965).
40. H. Mori, A continued-fraction representation of the time-correlation function, *Prog. Theor. Phys.* **34**:399–416 (1965a).
41. H. Mori, Transport, collective motion, and Brownian motion, *Prog. Theor. Phys.* **33**:423–450 (1965b).
42. H. Mori, H. Fujisaka, and H. Shigematsu, A new expansion of the master equation, *Prog. Theor. Phys.* **51**:109–122 (1974).
43. K. Oldham and J. Spanier, *The Fractional Calculus* (Academic Press, New York, 1974).
44. J. Porrà, K.-G. Wang, and J. Masoliver, Generalized Langevin equation: Anomalous diffusion and probability distributions, *Phys. Rev. E* **53**:5872–5881 (1996).
45. W. Schneider and W. Wyss, Fractional diffusion and wave equation, *J. Math. Phys.* **30**:134–144 (1989).
46. J. Straub, M. Borkovec, and B. Berne, Shortcomings of current theories of non-Markovian activated rate processes, *J. Chem. Phys.* **83**:3172–3174 (1985).
47. J. Straub, M. Borkovec, and B. Berne, Non-Markovian activated rate processes: Comparison of current theories with numerical simulation data, *J. Chem. Phys.* **84**:1788–1794 (1986).
48. A. Stuart and J. Warren, Analysis and experiments for a computational model of a heat bath, *J. Stat. Phys.* **97**:687–723 (1999).
49. H. Wall, *Analytic Theory of Continued Fractions* (D. Van Nostrand, New York, 1948).
50. K. Wang, Long-time-correlation effects and biased anomalous diffusion, *Phys. Rev. A* **45**:833–837 (1992).
51. L. Zadeh and C. Desoer, *Linear System Theory: The State Space Approach* (McGraw–Hill, New York, 1963).
52. G. Zaslavsky, Chaos, fractional kinetics and anomalous transport, *Physics Reports* **371**:461–580 (2002).
53. R. Zwanzig, Nonlinear generalized Langevin equations, *J. Stat. Phys.* **9**:215–220 (1973).
54. R. Zwanzig, Problems in nonlinear transport theory, in *Systems far from Equilibrium*, L. Garrido, ed. (Springer, New York, 1980), pp. 198–225.