arXiv:2301.04821v2 [cond-mat.stat-mech] 3 May 2023

《统计物理学杂志》190, 96 (2023) http://doi.org/10.1007/s10955-023-03104-8

关于应 用非高斯噪声进行随机朗之万模拟

尼尔斯·格伦贝克-詹森*
Department of Mechanical & Aerospace Engineering and Department of Mathematics
University of California, Davis, CA 95616, U.S.A.

鉴于朗之万方程时间步长无关随机积分器的最新进展，我们重新审视了在离散时间恒温器中使用非高斯分布表示热噪声项的问题。我们发现，现代方法所具有的理想时间步长不变性源于高斯噪声，而偏离该分布将扭曲由积分器涨落-耗散平衡产生的玻尔兹曼统计。我们以 GJ 随机 Verlet 方法为研究重点，因为在现有方法中，这些方法包含最准确的热力学量。在这组方法中，我们发现，任何满足涨落-耗散定理所给出的一阶矩和二阶矩的施加噪声分布，都将产生正确且与时间步长无关的结果，而这些结果由系统坐标的一阶矩和二阶矩决定。然而，如果施加非高斯噪声，系统坐标的高阶矩将出现不理想的偏差，从而损害若干重要的热力学量，尤其是依赖四阶矩的热力学量。由非高斯噪声引起的偏差会随着单步速度衰减而变得显著，进而抑制这些方法的优势。因此，我们得出结论：在使用具有较大时间步长的现代随机恒温器时，必须施加高斯噪声，才能获得可靠的热力学结果。

I. 引言

在过去几十年里，计算统计力学中的离散时间朗之万和布朗模拟一直是分子动力学诸多具体模拟应用的核心，例如材料科学、软物质和生物分子建模 [1–5]。模拟固有的一个根本难题是时间离散化：一方面，时间离散化是推进时间演化所必需的；另一方面，它又不可避免地会扭曲人们关注的连续时间系统特征。因此，模拟任务的一个重要组成部分，是探索并优化两个相互冲突的核心目标之间的平衡：通过增大时间步长提高模拟效率，通过减小时间步长提高模拟精度。

由于离散时间演化算法只能给出连续时间解的近似解，且误差必须在时间步长趋于无穷小时消失，因此，对于实用模拟环境的大多数考量，都集中在时间步长较小的极限下；在此情况下，模拟系统的变量在每一步中的变化都很小。在这一极限下，离散时间系统的行为与连续时间系统十分相似，因此算法开发大多先在连续时间框架下制定目标和衡量标准，随后再对时间进行离散化，并采用相对较小的步长执行，以保持精度。

对于随机微分方程，尤其是朗之万方程，一个额外的复杂之处在于，所关注的并不是某一条具体轨迹，而是由轨迹的统计数据决定所需的连续时间行为是否得到了充分表示。因此，处理这类问题的传统方法是从小时间步长下获得的精确轨迹中获取准确的统计数据，此时连续时间行为得到了很好的近似。然而，随着近年来离散时间随机算法的发展，这些算法能够从不准确的大时间步长轨迹中产生正确的统计数据 [6–8], ，因此，随机噪声相关性和分布等细节如今对于在所有有意义的时间步长下实现统计准确性的优势变得至关重要。这些细节的重要性源于时间的离散化，具体如下所述。

我们在此关注朗之万方程 [9,10]

$$m\dot{v} + \alpha\dot{r} = f + \beta \,, \tag{1}$$

其中， m 是物体的质量，其空间（构型）坐标为 r ，速度（动能）坐标为$v = \dot{r}$。物体受到力 $f = -\nabla E_p(r)$ 的作用，其中 $E_p(r)$ 是关于 r 的势能面，线性摩擦则由非负常数 α 表示。涨落耗散关系指出，热涨落 β 可由一个分布表示，其前两个矩如下所示 [11]:

$$\langle \beta(t) \rangle = 0 \tag{2a}$$

$$\langle \beta(t)\beta(t') \rangle = 2\alpha\, k_B T\, \delta(t - t') \,, \tag{2b}$$

其中 $\delta(t)$是狄拉克δ函数，表示涨落在时间上不相关； k_B 是玻尔兹曼常数， T 是系统通过摩擦常数耦合的热浴的热力学温度

* ngjensen@math.ucdavis.edu

通过摩擦常数 α与热浴耦合。在统计力学中，噪声分布 $\beta(t)$的涨落耗散定理只规定了式（2） [10–12]中给出的两个矩。因此，在朗之万方程中选择噪声分布时有很大的自由度，只要能正确描述热力学性质即可。物理（统计）结果不受这种自由度影响的原因在于，连续时间要求每一瞬间都贡献并累积一个新的、不相关的噪声值 $\beta(t)$；由于中心极限定理，在任何非零时间尺度上都会得到高斯结果（例如参见参考文献 [13]）。因此，采样得到的（一维）麦克斯韦-玻尔兹曼速度分布$\rho_{k,\ e}(v)$和构型玻尔兹曼分布$\rho_{c,\ e}(r)$总会变成 [10]

$$\rho_{k,e}(v) = \sqrt{\frac{m}{2\pi k_B T}} \exp\left(-\frac{\frac{1}{2}mv^2}{k_B T}\right) \tag{3a}$$

$$\rho_{c,e}(r) = \mathcal{C}_c \exp\left(-\frac{E_p(r)}{k_B T}\right), \tag{3b}$$

其中常数 $\mathcal{C}_c$ 由概率的归一化确定

$$\int_{\text{all space}} \rho_{c,e}(r)\, dr = 1\,, \tag{3c}$$

无论式 (1) 中采用何种时间上不相关的分布$\beta(t)$，只要满足式 (2) 即可。同样，对于平坦表面上的扩散输运，其扩散常数为 [1]

$$D_E = \frac{1}{d}\lim_{t\to\infty}\frac{\langle[r(t)-r(0)]^2\rangle}{2t} \tag{4}$$

$$= \frac{k_B T}{\alpha}\,, \tag{5}$$

其中 d 是构型空间的维数。

鉴于中心极限定理保证，对瞬时变量 $\beta(t)$进行足够充分的积分后，噪声分布将呈高斯分布，因此自然会选择符合式（2）的高斯分布作为$\beta(t)$。然而，在离散时间模拟中，直接采用满足式（2）的均匀分布也并不少见。其理由在于，均匀分布伪随机数生成器 [14], 的计算效率较高，而将这些变量转换为高斯分布变量后再使用则需要额外的计算开销。1988年发表的文献 [15], 对在离散时间下使用非高斯变量进行了深入研究，得出的结论是：鉴于当时随机离散时间积分器固有的不精确性，采用非高斯噪声分布（特别是均匀分布）时，统计矩随所用时间步长的缩放关系并未出现不利变化。在此基础上，文献 [16] 于1991年将研究扩展到更广泛的噪声分布；当时间步长趋近于零时，这些分布会产生与使用高斯噪声无法区分的统计结果，而当时方法固有的系统性时间步误差也会随之消失。对于趋近于零的时间步长，使用高斯噪声也会得到相同结果；与此同时，当时现行方法固有的系统性时间步误差也会消失。这些研究指出，采用简单的非高斯噪声具有计算效率方面的优势，因为在较小时间步长下，统计差异被认为微不足道；无论如何，为保证随机积分器的准确性，即使使用高斯噪声，也必须采用较小的时间步长。正如一篇近期比较当代随机积分器质量的文章 [17] 所指出的，只要采用较小的时间步长，使用传统算法时这种做法就十分普遍，例如 Br¨unger-Brooks-Karplus[18] 方法（文献 [19]中也有分析）或 Schneider-Stoll [20]方法。（关于传统及近期朗之万积分器的更详尽列表，请参见文献 [8]。）例如，在广泛使用的分子动力学模拟套件中，这两种方法会采用非高斯噪声，如 LAMMPS [21–23] 和 ESPResSo [24]。

本文旨在结合最新的随机积分器，重新探讨非高斯噪声的应用。这些积分器旨在利用高斯噪声，生成准确且与时间步长无关的统计数据和输运结果。我们将证明，即使构型分布和动能分布的低阶矩看似不受噪声分布细节的影响，所采样的相空间仍可能存在显著差异，从而影响系统的统计采样。我们的分析完全基于近期提出的 GJ 方法集 [8], ，因为在线性系统中使用高斯噪声时，这些方法具有经过优化且不随时间步长变化的统计性质。这样，我们还可以研究使用非高斯噪声在大时间步长下的影响，同时避免受到积分器固有的其他时间步误差的干扰。本文安排如下：第二节回顾以标准 Verlet 和分裂运算形式表述的 GJ 方法的基本特征与结构；第 III 节分析线性系统中构型坐标和动能坐标的前四个统计矩；第四节给出与第 III 节分析相应的模拟；第五节总结讨论并给出具体结论。

II. GJ方法综述

我们采用完整的 GJ 方法集 [8] 来求解方程 (1)，其形式为

$$u^{n+\frac{1}{2}} = \sqrt{c_1}\, v^n + \frac{\sqrt{c_3}\,\Delta t}{2m} f^n + \frac{\sqrt{c_3}}{2m}\beta^{n+1} \tag{6a}$$

$$r^{n+1} = r^n + \sqrt{c_3}\,\Delta t\, u^{n+\frac{1}{2}} \tag{6b}$$

$$v^{n+1} = \frac{c_2}{\sqrt{c_1}}\, u^{n+\frac{1}{2}} + \sqrt{\frac{c_3}{c_1}}\frac{\Delta t}{2m} f^{n+1} + \sqrt{\frac{c_3}{c_1}}\frac{1}{2m}\beta^{n+1}\,, \tag{6c}$$

其中，变量 r、 v、 u 和 f 上标 n 表示数值近似对应的离散时间 $t_n = t_0 + n\,\Delta t$，而 Δt 为离散时间步长。因此， $f^n = f\,(t_n,\ r^n)$。表示的两种速度分别为时刻 $t_{n+\frac{1}{2}}$ 的半步速度 $u^{n+\frac{1}{2}}$，以及时刻 t_n 的格点速度 v^n，其定义为

$$u^{n+\frac{1}{2}} = \frac{r^{n+1} - r^n}{\sqrt{c_3}\,\Delta t} \tag{7a}$$

$$v^n = \frac{r^{n+1} - (1-c_2)r^n - c_2 r^{n-1}}{2\sqrt{c_1 c_3}\,\Delta t} + \frac{1}{4m}(\beta^n - \beta^{n+1})\,. \tag{7b}$$

关于半步和位点的统计定义见参考文献 [8]。函数参数 $c_2 = c_2(\frac{\alpha\Delta t}{m})$ 是单个时间步的速度衰减因子，它是 $\alpha\Delta t/m$ 的函数。另两个函数参数 c_1 和 c_3 定义为

$$2c_1 = 1 + c_2 \tag{8a}$$

$$\frac{\alpha\Delta t}{m}\,c_3 = 1 - c_2\,. \tag{8b}$$

函数 c_2 的选择决定了不同 GJ方法之间的区别（参见参考文献 [8]，其中给出了 c_2 的若干种选择）。对 c_2 的基本要求是，对于 $\frac{\alpha\Delta t}{m} \to 0$，$c_2 \to 1 - \frac{\alpha\Delta t}{m}$，从而确保对于 $\frac{\alpha\Delta t}{m} \to 0$， $c_1 \to 1$ 和 $c_3 \to 1$ 也成立，这与无摩擦 Verlet 方法 [25–29] 一致。最后，离散时间随机噪声变量定义为

$$\beta^{n+1} = \int_{t_n}^{t_{n+1}} \beta(t)\,dt\,, \tag{9}$$

因此，由式 (9) 和 (2) 可得离散时间涨落耗散定理:

$$\langle \beta^n \rangle = 0 \tag{10a}$$

$$\langle \beta^n \beta^\ell \rangle = 2\alpha\Delta t\, k_B T\, \delta_{n,\ell}\,, \tag{10b}$$

其中 $\delta_{n,\ \ell}$ 是克罗内克δ函数。

式（6）中表示的 GJ 方法同样可以写成显式速度形式和蛙跳 Verlet 形式 [8]，也可以写成所谓的分裂形式，将演化划分为惯性、相互作用和热力学操作。参考文献 [30]，给出了一种实现方式，其中将关键的一对热力学操作合并为一个操作，因此未显式包含半步速度。这里我们将 GJ 分裂形式写为如下:

$$v^{n+\frac{1}{4}} = v^n + \sqrt{\frac{c_3}{c_1}}\frac{\Delta t}{2m}\,f^n \tag{11a}$$

$$r^{n+\frac{1}{2}} = r^n + \sqrt{\frac{c_3}{c_1}}\frac{\Delta t}{2}\,v^{n+\frac{1}{4}} \tag{11b}$$

$$u^{n+\frac{1}{2}} = \sqrt{c_1}\,v^{n+\frac{1}{4}} + \frac{1}{\sqrt{2}}\sqrt{(1-c_2)\frac{k_B T}{m}}\,\sigma^{n+1}$$
$$= \sqrt{c_1}\,v^{n+\frac{1}{4}} + \frac{\sqrt{c_3}}{2m}\,\beta^{n+1} \tag{11c}$$

$$v^{n+\frac{3}{4}} = \frac{c_2}{\sqrt{c_1}}\,u^{n+\frac{1}{2}} + \frac{1}{\sqrt{2c_1}}\sqrt{(1-c_2)\frac{k_B T}{m}}\,\sigma^{n+1}$$
$$= \frac{c_2}{\sqrt{c_1}}\,u^{n+\frac{1}{2}} + \sqrt{\frac{c_3}{c_1}}\frac{1}{2m}\,\beta^{n+1} \tag{11d}$$

$$r^{n+1} = r^{n+\frac{1}{2}} + \sqrt{\frac{c_3}{c_1}}\frac{\Delta t}{2}\,v^{n+\frac{3}{4}} \tag{11e}$$

$$v^{n+1} = v^{n+\frac{3}{4}} + \sqrt{\frac{c_3}{c_1}}\frac{\Delta t}{2m}\,f^{n+1}\,, \tag{11f}$$

其中 σ^n 是一个随机变量，其 $\langle\sigma^n\rangle = 0$ 和$\langle\sigma^n\sigma^\ell\rangle = \delta_{n,\ \ell}$，且式（7a）中的半步速度 $u^{n+\frac{1}{2}}$ 在式（11c）中显式出现。

从式（6）或式（11）中消去速度，可得到纯构型的 GJ 随机 Verlet 形式

$$r^{n+1} = 2c_1\,r^n - c_2\,r^{n-1} + \frac{c_3\Delta t^2}{m}\,f^n + \frac{c_3\Delta t}{2m}(\beta^n + \beta^{n+1})\,, \tag{12}$$

这是唯一一种随机 Verlet 形式，满足以下条件：1）对于谐势 $E_p(r)$，其胡克力为 $f = -\kappa r(\kappa > 0)$，且高斯涨落为 β^n，将得到正确的玻尔兹曼分布 $\rho_{c,\ e}(r^n)$，如式（3b）所示，使得

$$\rho_{c,e}(r^n) = \sqrt{\frac{\kappa}{2\pi k_B T}}\exp\left(-\frac{\frac{1}{2}\kappa(r^n)^2}{k_B T}\right) \tag{13a}$$

$$\langle r^n r^n\rangle = \frac{k_B T}{\kappa}\,, \tag{13b}$$

对于稳定范围（$\Omega_0\Delta t\ < 2\sqrt{c_1/c_3}$）内的任意时间步长均成立；2）对于线性势，其中 $f\ =$ 为常数，可得到正确的构型漂移速度

$$v_d\ =\ \left\langle\frac{r^{n+1} - r^n}{\Delta t}\right\rangle = \frac{f}{\alpha}\,; \tag{14}$$

且 3）对于 $f\ = 0$，可得到正确的爱因斯坦扩散，如式（4）所示，使得

$$D_E = \lim_{n\to\infty}\frac{\left\langle (r^n - r^0)^2\right\rangle}{2d\,n\Delta t}\ =\ \frac{k_B T}{\alpha}\,. \tag{15}$$

同样，人们发现 [8] 对于每个时间步长仅有一个高斯噪声值 β^n 时，胡克力 $f = -\kappa r$ $(\kappa\ \geq\ 0)$会产生唯一的半步速度 $u^{n+\frac{1}{2}}$

在式（7a）中，具有正确速度分布函数$\rho_{k,\ e}(u^{n+\frac{1}{2}})$来自式（3a），因此

$$\langle u^{n+\frac{1}{2}}u^{n+\frac{1}{2}}\rangle = \frac{k_BT}{m}\,. \tag{16}$$

相反，参考文献 [8] 也指出，无法定义一种具有相同时间步长无关性质的格点速度 v^n 。鉴于使用高斯噪声时，这些方程的统计性质不随时间步长变化，我们现在将研究当 β^n采用高斯分布以外的其他分布时， r^n 和 $u^{n+\frac{1}{2}}$ 的统计性质。

III. 线性系统的矩分析

本文将研究构型坐标 r^n 和半步速度 $u^{n+\frac{1}{2}}$的相关矩如何随 β^n的施加噪声分布 $\rho(\beta)$变化，该分布满足式 (10)。在此分析中，我们主要假设系统为具有谐势的线性系统

$$E_p(r) = \frac{1}{2}\kappa\, r^2 \tag{17}$$

$$\Rightarrow \quad f = -\kappa r\,, \tag{18}$$

因此，式 (12) 中线性化的构型随机 GJ 方法可以写为

$$r^{n+1} = 2c_1X\, r^n - c_2\, r^{n-1} + \frac{c_3\Delta t}{2m}(\beta^n + \beta^{n+1})\,, \tag{19}$$

with

$$X = 1 - \frac{c_3}{c_1}\frac{\Omega_0^2\Delta t^2}{2}\,. \tag{20}$$

振子的固有频率 Ω_0 为 $\Omega_0^2 = \kappa/m$。

A. 一阶矩和二阶矩

对于变量的一阶矩，显然，方程 (10a) 可确保 $\langle r^n\rangle = 0$（由方程 (19) 得出），进而确保 $\langle u^{n+\frac{1}{2}}\rangle = 0$（由方程 (7a) 得出），而无论为 β^n选择何种分布。同理可知，方程 (10a) 还可确保方程 (14) 中 f =const 的漂移速度保持恒定且正确，而不受时间步长和分布 β^n的影响。

对于二阶矩 $\langle r^n r^n\rangle$和 $\langle u^{n+\frac{1}{2}}u^{n+\frac{1}{2}}\rangle$，我们首先求得 $\langle r^n r^n\rangle$ ，使其成为由方程（19）的统计平均值分别乘以 r^{n-1}、 r^n和 r^{n+1}后所得到的方程组的解。这个线性方程组可以方便地写成

$$\begin{pmatrix} 1 & -2c_1X & c_2 \\ 0 & 2c_1 & -2c_1X \\ c_2 & -2c_1X & 1 \end{pmatrix} \begin{pmatrix} \langle r^{n-1}r^{n+1}\rangle \\ \langle r^n r^{n+1}\rangle \\ \langle r^n r^n\rangle \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 2c_1X+2 \end{pmatrix} \left(\frac{c_3\Delta t}{2m}\right)^2 \langle \beta^n\beta^n\rangle\,. \tag{21}$$

我们立即注意到，只要满足方程（10），该方程就不受 β^n 分布细节的影响。因此，无论时间步长 Δt 如何，二阶矩 $\langle r^n r^n\rangle$ 也不受所选 β^n 分布的影响。方程（7a）中半步速度对应的二阶矩为

$$\langle u^{n+\frac{1}{2}}u^{n+\frac{1}{2}}\rangle = \frac{2}{c_3\,\Delta t^2}(\langle r^n r^n\rangle - \langle r^n r^{n+1}\rangle)\,, \tag{22}$$

其中，右侧的两个相关函数均为由式 (21) 得到的解；只要式 (10b) 成立，该方程就保持不变。因此，无论时间步长 Δt 如何，半步速度的二阶矩也不受所选 β^n 分布的影响。势能（式 (17)）和动能的平均值，

$$\langle E_p(r^n)\rangle = \frac{\kappa}{2}\langle r^n r^n\rangle \;=\; \frac{1}{2}k_BT \tag{23}$$

$$\left\langle E_k(u^{n+\frac{1}{2}})\right\rangle = \frac{1}{2}\frac{m}{c_3\,\Delta t^2}(\langle r^n r^n\rangle - \langle r^n r^{n+1}\rangle) = \frac{1}{2}k_BT\,, \tag{24}$$

因此，对于任何满足式 (10) 的随机变量 β^n ，这些平均值都不受影响。

由于 GJ 方法中坐标 r^n 和 $u^{n+\frac{1}{2}}$ 的二阶矩不受满足方程（10）的噪声分布具体形式影响，我们得出结论：无论采用何种 β^n 分布，GJ 方法在势能和动能（因而温度）方面都具有时间步长不变性。这种不变性也适用于上文提到的扩散方程（4）和（15），因为该量仅取决于噪声变量的二阶矩。

图 2ab 和 6ab 展示了对几个非高斯噪声变量的二阶矩不依赖时间步长的数值验证，相关模拟将在第四节讨论。尽管二阶矩不受所选 β^n 分布的影响这一点颇具吸引力，因为它给人以温度已正确实现的印象，但正如我们将在下文看到的，这可能不足以确保正确的热力学采样。

B. 三阶矩

r^n 和 $u^{n+\frac{1}{2}}$ 的三阶矩通常并非计算热力学量所必需，尤其是在系统势能具有对称性的情况下。然——不过，如果出于某种原因对系统施加非对称涨落，它们可能值得研究。按照计算二阶矩时所采用的相同步骤，我们对式（19）分别与 $r^{\ell_1}r^{\ell_2}$、 $\ell_i = n-1$、 n、 $n+1$ 的六种不同组合的乘积求统计平均值。所得结果为线性系统

$$\mathcal{A}_3 \begin{pmatrix} \langle r^{n-1}r^{n-1}r^{n+1}\rangle \\ \langle r^{n-1}r^{n}r^{n+1}\rangle \\ \langle r^{n-1}r^{n+1}r^{n+1}\rangle \\ \langle r^{n}r^{n}r^{n+1}\rangle \\ \langle r^{n}r^{n+1}r^{n+1}\rangle \\ \langle r^{n}r^{n}r^{n}\rangle \end{pmatrix} = \left(\frac{k_BT}{\kappa}\right)^{\frac{3}{2}} \mathcal{R} = \begin{pmatrix} \langle r^{n-1}r^{n-1}\beta^{n}\rangle \\ \langle r^{n-1}r^{n}\beta^{n}\rangle \\ \langle r^{n-1}r^{n+1}(\beta^{n}+\beta^{n+1})\rangle \\ \langle r^{n}r^{n}\beta^{n}\rangle \\ \langle r^{n}r^{n+1}(\beta^{n}+\beta^{n+1})\rangle \\ \langle r^{n+1}r^{n+1}(\beta^{n}+\beta^{n+1})\rangle \end{pmatrix} \frac{c_3\Delta t}{2m}, \tag{25}$$

其中

$$\mathcal{A}_3 = \begin{pmatrix} 1 & 0 & 0 & -2c_1X & 0 & c_2 \\ 0 & 1 & 0 & c_2 & -2c_1X & 0 \\ c_2 & -2c_1X & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & c_2 & -2c_1X \\ 0 & c_2 & 0 & -2c_1X & 1 & 0 \\ 0 & 0 & c_2 & 0 & -2c_1X & 1 \end{pmatrix} \tag{26}$$

经过一些计算，右侧 $\mathcal{R} = \{\mathcal{R}_i\}$ 可表示为

$$\mathcal{R}_1 = 0 \tag{27a}$$
$$\mathcal{R}_2 = 0 \tag{27b}$$
$$\mathcal{R}_3 = 0 \tag{27c}$$
$$\mathcal{R}_4 = (1-c_2)^3\left(\frac{\kappa\Delta t}{\alpha}\right)^{\frac{3}{2}}\Gamma_3 \tag{27d}$$
$$\mathcal{R}_5 = (2c_1X+1)\mathcal{R}_4 \tag{27e}$$
$$\mathcal{R}_6 = (2c_1X(2c_1X+2)+2)\mathcal{R}_4\,, \tag{27f}$$

其中，涨落 β^n 满足式（10），偏度记为 Γ_3,

$$\Gamma_3 = \frac{\langle \beta^n\beta^n\beta^n\rangle}{\langle \beta^n\beta^n\rangle^{\frac{3}{2}}}\,. \tag{28}$$

与采用任何对称分布的情况一样，当将高斯分布的 β^n 施加于对称势能 E_p（r）时，构型坐标 r^n 的所有奇数阶矩均为零。

速度的三阶矩 $u^{n+\frac{1}{2}}$ 为

$$\left\langle u^{n+\frac{1}{2}}u^{n+\frac{1}{2}}u^{n+\frac{1}{2}}\right\rangle = \frac{1}{\left(\Delta t\sqrt{c_3}\right)^3}\left\langle (r^{n+1}-r^n)^3\right\rangle$$
$$= \frac{3}{\left(\Delta t\sqrt{c_3}\right)^3}\left(\left\langle r^nr^nr^{n+1}\right\rangle - \left\langle r^nr^{n+1}r^{n+1}\right\rangle\right), \tag{29}$$

其中，右侧的两个相关函数是方程（25）解的一部分。

对于给定的 GJ 方法（c_2），上述方程组可以针对给定的系统参数和偏度 Γ_3 轻松进行数值求解。第四节将讨论相关示例，并在图 6cd 中展示这些示例，同时给出针对某一特定非对称噪声分布，对方程（6）和（18）进行数值求解的结果。下面将给出几种特殊情形的闭式表达式。

特殊情形，$X=0$: 矩阵 $\mathcal{A}_3$ 和向量 $\mathcal{R}$大幅简化，可轻松求得解

$$\langle r^nr^nr^n\rangle = \frac{2}{1+c_2^3}\left(\frac{k_BT}{\kappa}\right)^{\frac{3}{2}}\mathcal{R}_4 \tag{30a}$$
$$\langle r^nr^nr^{n+1}\rangle = \frac{1-c_2}{1+c_2^3}\left(\frac{k_BT}{\kappa}\right)^{\frac{3}{2}}\mathcal{R}_4 \tag{30b}$$
$$\langle r^nr^{n+1}r^{n+1}\rangle = \frac{1+c_2^2}{1+c_2^3}\left(\frac{k_BT}{\kappa}\right)^{\frac{3}{2}}\mathcal{R}_4\,, \tag{30c}$$

其中，我们利用了 $\Omega_0^2\Delta t^2 = 2\frac{c_3}{c_1}$ 这一特殊情形可提供 $\mathcal{R}_6 = 2\mathcal{R}_5 = 2\mathcal{R}_4$，并且

$$\mathcal{R}_4 = \left(\frac{1-c_2^2}{2}\right)^{\frac{3}{2}}\Gamma_3\,. \tag{31}$$

方程 (29)、(30b) 和 (30c) 给出

$$\left\langle \left(u^{n+\frac{1}{2}}\right)^3\right\rangle = -\frac{3}{\left(\Delta t\sqrt{c_3}\right)^3}\frac{c_2}{c_2^2-c_2+1}\left(\frac{k_BT}{\kappa}\right)^{\frac{3}{2}}\mathcal{R}_4 \tag{32}$$
$$= -\frac{3}{2\sqrt{2}}\left(\frac{1-c_2^2}{1+c_2}\right)^{\frac{3}{2}}\frac{c_2}{c_2^2-c_2+1}\left(\frac{k_BT}{m}\right)^{\frac{3}{2}}\Gamma_3\,. \tag{33}$$

这些表达式的用法示例以空心标记（○）显示在图 6cd 中，适用于某种特定的非对称噪声分布。

特殊情形，$c_2 = 0$：在这种情况下，我们有 $c_1 = \frac{1}{2}$、$c_3 = m/\alpha\Delta t$ 和 $X = 1 - \Delta t\kappa/\alpha$。简化后的矩阵 $\mathcal{A}_3$ 和向量 $\mathcal{R}$ 得到

$$\langle r^n r^n r^n\rangle = \frac{3X^2+3X+2}{1-X^3}\left(\frac{k_BT}{\kappa}\right)^{\frac{3}{2}}\mathcal{R}_4 \tag{34a}$$

$$\langle r^n r^n r^{n+1}\rangle = \frac{2X^3+3X^2+2X+1}{1-X^3}\left(\frac{k_BT}{\kappa}\right)^{\frac{3}{2}}\mathcal{R}_4 \tag{34b}$$

$$\langle r^n r^{n+1} r^{n+1}\rangle = \frac{X^4+2X^3+2X^2+2X+1}{1-X^3}\left(\frac{k_BT}{\kappa}\right)^{\frac{3}{2}}\mathcal{R}_4\,, \tag{34c}$$

其中

$$\mathcal{R}_4 = \left(\frac{\kappa\Delta t}{2\alpha}\right)^{\frac{3}{2}}\Gamma_3\,. \tag{35}$$

方程 (29)、(34b) 和 (34c) 给出

$$\left\langle \left(u^{n+\frac{1}{2}}\right)^3\right\rangle = \frac{3X^2}{2\sqrt{2}}\frac{1-X^2}{1-X^3}\left(\frac{k_BT}{m}\right)^{\frac{3}{2}}\Gamma_3\,. \tag{36}$$

这些表达式的使用示例以实心标记（●）显示在图 6cd 中，对应一种特定的非对称噪声分布。

C. 四阶矩

计算能量和压力涨落需要用到 r^n 和 $u^{n+\frac{1}{2}}$ 的四阶矩，而能量和压力涨落是若干重要热力学量的基础，例如热容、热膨胀、压缩率和热压力系数 [1]。因此，四阶矩对于系统的恰当热力学表征至关重要。按照计算二阶矩和三阶矩时所用的相同步骤，我们对方程（19）与 $r^{\ell_1}r^{\ell_2}r^{\ell_3}$ 的十种不同组合的乘积求统计平均值，其中 $\ell_i = n-1$、n、$n+1$。结果得到以下线性系统

$$\mathcal{A}_4\begin{pmatrix}\langle r^{n-1}r^{n-1}r^{n-1}r^{n+1}\rangle\\ \langle r^{n-1}r^{n-1}r^{n}r^{n+1}\rangle\\ \langle r^{n-1}r^{n-1}r^{n+1}r^{n+1}\rangle\\ \langle r^{n-1}r^{n}r^{n}r^{n+1}\rangle\\ \langle r^{n-1}r^{n}r^{n+1}r^{n+1}\rangle\\ \langle r^{n-1}r^{n+1}r^{n+1}r^{n+1}\rangle\\ \langle r^{n}r^{n}r^{n}r^{n+1}\rangle\\ \langle r^{n}r^{n}r^{n+1}r^{n+1}\rangle\\ \langle r^{n}r^{n+1}r^{n+1}r^{n+1}\rangle\\ \langle r^{n}r^{n}r^{n}r^{n}\rangle\end{pmatrix} = \left(\frac{k_BT}{\kappa}\right)^2\mathcal{R} = \begin{pmatrix}0\\ \langle r^{n-1}r^{n-1}r^{n}\beta^{n}\rangle\\ \langle r^{n-1}r^{n-1}r^{n+1}(\beta^{n}+\beta^{n+1})\rangle\\ \langle r^{n-1}r^{n}r^{n}\beta^{n}\rangle\\ \langle r^{n-1}r^{n}r^{n+1}(\beta^{n}+\beta^{n+1})\rangle\\ \langle r^{n-1}r^{n+1}r^{n+1}(\beta^{n}+\beta^{n+1})\rangle\\ \langle r^{n}r^{n}r^{n}\beta^{n}\rangle\\ \langle r^{n}r^{n}r^{n+1}(\beta^{n}+\beta^{n+1})\rangle\\ \langle r^{n}r^{n+1}r^{n+1}(\beta^{n}+\beta^{n+1})\rangle\\ \langle r^{n+1}r^{n+1}r^{n+1}(\beta^{n}+\beta^{n+1})\rangle\end{pmatrix}\frac{c_3\Delta t}{2m}\,, \tag{37}$$

其中

$$\mathcal{A}_4 = \begin{pmatrix}1 & 0 & 0 & 0 & 0 & 0 & -2c_1X & 0 & 0 & c_2\\ 0 & 1 & 0 & 0 & 0 & 0 & c_2 & -2c_1X & 0 & 0\\ c_2 & -2c_1X & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & c_2 & -2c_1X & 0\\ 0 & c_2 & 0 & -2c_1X & 1 & 0 & 0 & 0 & 0 & 0\\ 0 & 0 & c_2 & 0 & -2c_1X & 1 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & c_2 & -2c_1X\\ 0 & 0 & 0 & c_2 & 0 & 0 & -2c_1X & 1 & 0 & 0\\ 0 & 0 & 0 & 0 & c_2 & 0 & 0 & -2c_1X & 1 & 0\\ 0 & 0 & 0 & 0 & 0 & c_2 & 0 & 0 & -2c_1X & 1\end{pmatrix}. \tag{38}$$

经过一些繁琐的计算，右端 $\mathcal{R}=\{\mathcal{R}_i\}$ 可表示为

$$\mathcal{R}_1 = 0 \tag{39a}$$
$$\mathcal{R}_2 = \frac{1}{2}(1-c_2^2)(1-X) \tag{39b}$$
$$\mathcal{R}_3 = [2+2c_1X]\mathcal{R}_2 \tag{39c}$$
$$\mathcal{R}_4 = [2c_1X+1-c_2]\mathcal{R}_2 \tag{39d}$$
$$\mathcal{R}_5 = [2c_1X(2c_1X+2-c_2) \ +1-2c_2]\mathcal{R}_2 \tag{39e}$$
$$\mathcal{R}_6 = [(2c_1X)^3+(2c_1X)^2(3-c_2)+4c_1X(1-2c_2)-4c_2]\mathcal{R}_2 \tag{39f}$$
$$\mathcal{R}_7 = [3+(\Gamma_4-3)\mathcal{R}_2]\mathcal{R}_2 \tag{39g}$$
$$\mathcal{R}_8 = [4c_1X+3-c_2+(2c_1X+1)(\Gamma_4-3)\mathcal{R}_2]\mathcal{R}_2 \tag{39h}$$
$$\mathcal{R}_9 = [(2c_1X)^2+2c_1X(3-c_2)+3-2c_2+(1+2c_1X)^2(\Gamma_4-3)\mathcal{R}_2]\mathcal{R}_2 \tag{39i}$$
$$\mathcal{R}_{10} = [3(2c_1X+2)+(1+(1+2c_1X)^3)(\Gamma_4-3)\mathcal{R}_2]\mathcal{R}_2\,, \tag{39j}$$

其中我们使用了式（10b）。注意，式（39g）–（39j）中的 $\mathcal{R}_7$-$\mathcal{R}_{10}$ 取决于峰度 Γ_4

$$\Gamma_4 = \frac{\langle \beta^n\beta^n\beta^n\beta^n\rangle}{\langle\beta^n\beta^n\rangle^2}\,, \tag{40}$$

而峰度无法由式（10）中的涨落耗散关系确定。

虽然式（37）–（39）并未直接提供有用信息，但它们为计算式（17）中势能的涨落 σ_p 提供了一条直接途径

$$\begin{aligned}\sigma_p &= \sqrt{\langle E_p^2(r^n)\rangle - \langle E_p(r^n)\rangle^2} \\ &= \frac{\kappa}{2}\sqrt{\langle r^nr^nr^nr^n\rangle - \langle r^nr^n\rangle^2}\,.\end{aligned} \tag{41}$$

方程 (37)-(39) 也可直接得到由方程 (7a) 中速度所对应的动能涨落σ_k

$$\begin{aligned}\sigma_k &= \sqrt{\left\langle E_k^2(u^{n+\frac12})\right\rangle - \left\langle E_k(u^{n+\frac12})\right\rangle^2} \\ &= \frac{m}{2}\sqrt{\langle u^{n+\frac12}u^{n+\frac12}u^{n+\frac12}u^{n+\frac12}\rangle - \langle u^{n+\frac12}u^{n+\frac12}\rangle^2}\,,\end{aligned} \tag{42}$$

其中 $\langle u^{n+\frac12}u^{n+\frac12}\rangle$ 由方程 (16) 给出，并且

$$\begin{aligned}\langle u^{n+\frac12}u^{n+\frac12}u^{n+\frac12}u^{n+\frac12}\rangle \ &= \ \frac{1}{c_3^2\Delta t^4}\langle (r^{n+1}-r^n)^4\rangle \\ &= \ \frac{2}{c_3^2\Delta t^4}(\langle r^nr^nr^nr^n\rangle - 2\langle r^nr^nr^nr^{n+1}\rangle \\ &+ 3\langle r^nr^nr^{n+1}r^{n+1}\rangle - 2\langle r^nr^{n+1}r^{n+1}r^{n+1}\rangle)\,.\end{aligned} \tag{43}$$

右侧的所有相关函数均由方程 (37)-(39) 的解给出。

对于任意给定的 GJ 方法（c_2），上述方程组都可以针对给定的系统参数和峰度 Γ_4 轻松进行数值求解。第四节将讨论相关示例；图 2cd 和 6ef 展示了这些示例，以及针对特定噪声分布求解方程 (6) 和 (18) 所得到的相应数值解。对于 $X=0$ 和 $c_2=0$ 的特殊情形，可以给出闭式表达式。

特殊情形，$X=0$：此特定时间步长为

$$\Omega_0\Delta t = \sqrt{2}\sqrt{\frac{c_1}{c_3}}\,. \tag{44}$$

经过繁琐但直接的代数运算，可得 $\langle r^nr^nr^nr^n\rangle$ 和$\langle u^{n+\frac12}u^{n+\frac12}u^{n+\frac12}u^{n+\frac12}\rangle$ 关于 $X=0$ 的闭式表达式；即，对于 $\Omega_0\Delta t = \sqrt{2}\sqrt{c1/c3}$。相应的势能和动能涨落可写为

$$\begin{aligned}\sigma_p &= \frac{\kappa}{2}\sqrt{\langle r^nr^nr^nr^n\rangle - \langle r^nr^n\rangle^2} \\ &= \frac{k_BT}{\sqrt{2}}\sqrt{\frac14\frac{1+7c_2^2}{1+c_2^2}+\frac14\Gamma_4\frac{1-c_2^2}{1+c_2^2}}\end{aligned} \tag{45}$$

$$\begin{aligned}\sigma_k &= \frac{m}{2}\sqrt{\langle u^{n+\frac12}u^{n+\frac12}u^{n+\frac12}u^{n+\frac12}\rangle - \langle u^{n+\frac12}u^{n+\frac12}\rangle^2} \\ &= \frac{k_BT}{\sqrt{2}}\sqrt{\frac12\frac{c_2^3}{1+c_2^2}[3-\Gamma_4]+\frac14[1+\Gamma_4]}\,.\end{aligned} \tag{46}$$

请注意，若 β^n 是高斯分布随机变量，且其峰度为 $\Gamma_4=3$，则上述能量涨落给出预期值 $\sigma_p=\sigma_k=k_BT/\sqrt{2}$.

这些表达式的用法示例以空心标记（○）的形式展示在图 2cd 和 6ef 中，适用于特定的噪声分布。

特殊情形，$c_2=0$：此特殊情形（GJ-0，参见参考文献 [8]）对应一阶差分方程

$$r^{n+1} = r^n + \frac{1}{\alpha}\left[\Delta t\, f^n + \frac12(\beta^n+\beta^{n+1})\right]\,, \tag{47}$$

它可以近似求解无惯性的朗之万方程，即布朗运动。线性化的胡克定律系统（$f = -\kappa r$）为

$$r^{n+1} = X r^n + \frac{1}{2\alpha}(\beta^n + \beta^{n+1}), \tag{48}$$

与 $X = 1 - \kappa\Delta t/\alpha$ 相对应

方程 (19) 和 (20)。这种简洁性使方程 (37) 中最后四个方程与其余方程解耦，并可计算四阶矩，从而得到势能的平均值和涨落

$$\langle E_p(r^n)\rangle = \frac{\kappa}{2}\,\langle r^n r^n\rangle \;=\; \frac{1}{2}k_B T \tag{49}$$

$$\sigma_p = \frac{\kappa}{2}\sqrt{\langle r^n r^n r^n r^n\rangle - \langle r^n r^n\rangle^2} \;=\; \frac{k_B T}{\sqrt{2}}\sqrt{\frac{\frac{1}{2}X^4[3-\Gamma_4] + \frac{1}{4}X^3[7-\Gamma_4] + \frac{1}{4}(X^2+X+1)[1+\Gamma_4]}{X^3+X^2+X+1}}. \tag{50}$$

我们还注意到，如果 β^n 是一个高斯分布变量，其峰度为 $\Gamma_4 = 3$，则势能涨落会得到预期的

value, $\sigma_p = k_B T/\sqrt{2}$.
类似地，我们可以根据相关函数求出动能涨落 σ_k

$$\left\langle \left(u^{n+\frac{1}{2}}\right)^4 \right\rangle \;=\; (1-X)^2\frac{4\sigma_p^2 + (k_B T)^2}{m^2} + 6\left(\frac{k_B T}{m}\right)^2 X^2(1-X) - \left(\frac{k_B T}{m}\right)^2 (3+\Gamma_4)X\left(1 - \frac{3}{2}X + X^2\right), \tag{51}$$

使得

$$\sigma_k = \frac{1}{2}m\sqrt{\left\langle \left(u^{n+\frac{1}{2}}\right)^4 \right\rangle - \left(\frac{k_B T}{m}\right)^2}, \tag{52}$$

其中我们使用了式 (16)。因此，在这种特殊情形下也可以轻松计算 σ_k。我们再次注意到，对于高斯涨落，其中 $\Gamma_4 = 3$，动能涨落会得到预期值 $\sigma_k = k_B T/\sqrt{2}$。
这些表达式的应用示例见图 2cd 和 6ef，其中实心标记（●）表示特定的噪声分布。

IV. 模拟

虽然上述表达式提供了直接和间接的方法来评估构型变量和动能变量在热力学上重要的前四阶矩，但我们通过数值模拟检验这些结果，使模拟和分析能够相互验证。此外，模拟还提供了关于实际分布 $\rho_c(r^n)$ 和 $\rho_k(u^{n+\frac{1}{2}})$ 的更多细节；在某些情况下，当涨落 β^n 并非高斯分布时，这些细节揭示了它们与玻尔兹曼统计数据之间的显著偏差。

我们展示了涨落 β^n 的三种不同分布的结果，它们都满足离散时间涨落耗散平衡，该平衡约束了式（10）中给出的一阶和二阶矩，因此都使 r^n 和 $u^{n+\frac{1}{2}}$ 的一阶和二阶矩正确，从而使由动能和势能测得的温度正确（见第III A节）：

高斯分布：$\beta^n = \beta_G^n$ 分布为 $\rho_G(\beta)$：

$$\rho_G(\beta) = \frac{1}{\sqrt{2\pi\langle\beta^n\beta^n\rangle}}\exp\left(-\frac{\beta\beta}{2\langle\beta^n\beta^n\rangle}\right), \tag{53a}$$

具有偏度和峰度

$$\Gamma_3 = 0 \tag{53b}$$
$$\Gamma_4 = 3\,. \tag{53c}$$

均匀：$\beta^n = \beta_U^n$ 分布为 $\rho_U(\beta)$：

$$\rho_U(\beta) = \frac{1}{2\sqrt{3\langle\beta^n\beta^n\rangle}} \times \begin{cases} 1\ , & -\sqrt{3} \le \frac{\beta}{\sqrt{\langle\beta^n\beta^n\rangle}} < \sqrt{3} \\ 0\ , & \text{otherwise} \end{cases}, \tag{54a}$$

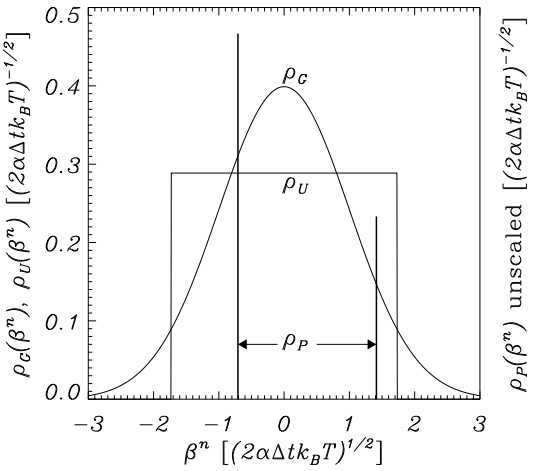


图 1. 式 (53)、(54) 和 (55) 分别给出了三种采用的分布：$\rho_G(\beta^n)$（高斯分布）、$\rho_U(\beta^n)$（均匀）和 $\rho_P(\beta^n)$（峰形）。这三种分布均满足式 (10) 所示的离散时间涨落耗散关系。

具有偏度和峰度

$$\Gamma_3 = 0 \tag{54b}$$

$$\Gamma_4 = \frac{9}{5}\,. \tag{54c}$$

峰形：$\beta^n = \beta_P^n$ 分布为 $\rho_P(\beta)$:

$$\rho_P(\beta) = \frac{2}{3}\,\delta\left(\frac{\beta}{\sqrt{\langle\beta^n\beta^n\rangle}} + \frac{1}{\sqrt{2}}\right) + \frac{1}{3}\,\delta\left(\frac{\beta}{\sqrt{\langle\beta^n\beta^n\rangle}} - \sqrt{2}\right), \tag{55a}$$

偏度和峰度为

$$\Gamma_3 = \frac{1}{\sqrt{2}} \tag{55b}$$

$$\Gamma_4 = \frac{12}{5}\,. \tag{55c}$$

图 1 示意了这三种分布。伪随机数 β^n 由 RANMAR 算法（在 LAMMPS [31] 中使用）生成的均匀分布数 $[0,1)$ 得到。

为了验证上一节的分析，我们采用几种 GJ 方法模拟了受噪声驱动的谐振子方程（6）和（18）。对于高斯噪声，只要根据方法适当调整时间步长，这些方法都会得到相同的统计结果。为了

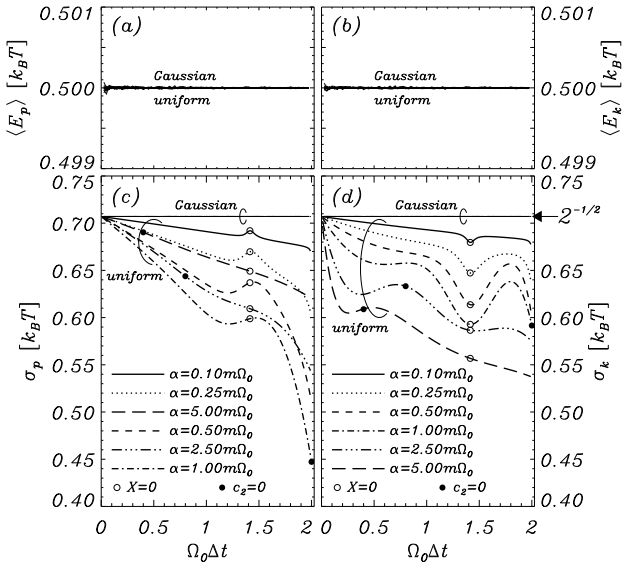


图 2. 对谐振系统（式 (6) 和 (18)）的平均势能和动能进行模拟与分析，分别为式 (23) 中的 $\langle E_p\rangle$（a）和式 (24) 中的 $\langle E_k\rangle$（b）；并分析其涨落，分别为式 (41) 中的 σ_p（c）和式 (42) 中的 σ_k（d）。这些量随约化时间步长变化，系统温度 T 非零。图中展示了高斯分布噪声（式 (53a) ）和均匀噪声分布（式 (54a)）的结果。采用了若干摩擦参数 α，如图所示。模拟结果与解析结果无法区分。标记表示 $c_2 = 0$（式 (50)–(52) 中的 •）和 $X = 0$（式 (45) 和 (46) 中的 ◦）的特殊情形。

简洁起见，我们仅展示 GJ-I（GJF-2GJ）方法的结果 [7, 8], , 该方法

$$c_2 = \frac{1 - \frac{\alpha\Delta t}{2m}}{1 + \frac{\alpha\Delta t}{2m}} \tag{56a}$$

$$c_1 \;=\; c_3 = \frac{1}{1 + \frac{\alpha\Delta t}{2m}}\,. \tag{56b}$$

GJF-2GJ 方法在构型坐标上与 GJF 方法重合 [6]。值得注意的是，GJ-I（GJF-2GJ）方法可在 LAMMPS 中使用 [31]。以特征单位 t_0（即固有频率 Ω 的倒数）归一化时间后$_0$, , 约化阻尼为 $\alpha/m\Omega_0$, , 特征能量为 $E_0 = \kappa r_0^2$, 其中 r_0 是选定的特征长度， r 以此归一化。模拟在 $k_BT = E_0$ 下进行，涵盖稳定范围 $\Omega_0\Delta t < 2$内的多种 $\alpha/m\Omega_0$ 取值。每个统计值均由 1,000 次独立模拟的平均值计算得出，每次模拟包含 10^8 个时间步长。

图2展示了采用均匀噪声分布 β_U^n （式（54a））所得的结果，并以采用高斯分布β_G^n （式（53a））所得的结果作为参考。图2ab分别展示了构型变量和动能变量的结果，验证了第 IIIA 节中的一般结论；也就是说，仅取决于

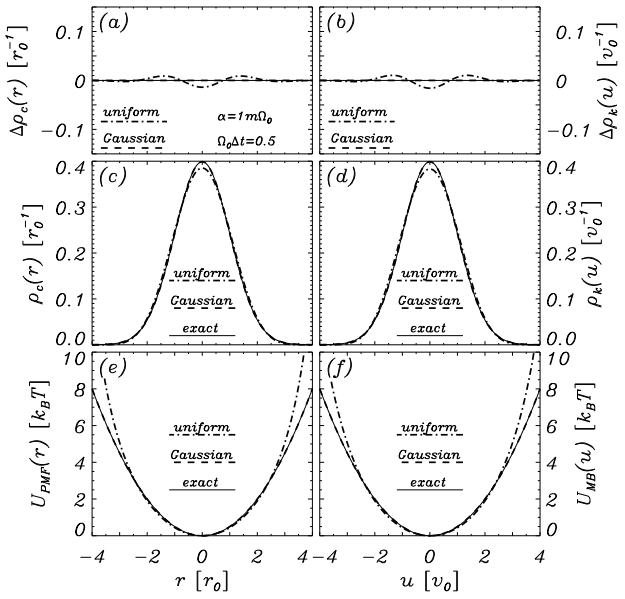


图 3. 对线性系统方程 (6) 和 (18) 施加高斯噪声和均匀噪声进行模拟时的构型分布和动能分布 $\rho_c(r^n)$ 和$\rho_k(u^{n+\frac{1}{2}})$，其中 $\alpha = 1m\Omega_0, \Omega_0\Delta t = 0.5$, $k_BT = E_0$。所示曲线为方程 (3)（实线，标注为精确分布）、使用高斯噪声 β_G^n 的模拟数据（虚线，标注为高斯分布）以及使用均匀噪声 β_U^n 的模拟数据（点划线，标注为均匀分布）。使用高斯噪声进行模拟得到的分布与方程 (3) 的精确参考分布无法区分。(c) 和 (d) 显示分布，(a) 和 (b) 显示与精确分布方程 (3) 的偏差，(e) 和 (f) 显示由这些分布推导出的有效势（平均力势）和动能势（麦克斯韦-玻尔兹曼）（方程 (58)）。

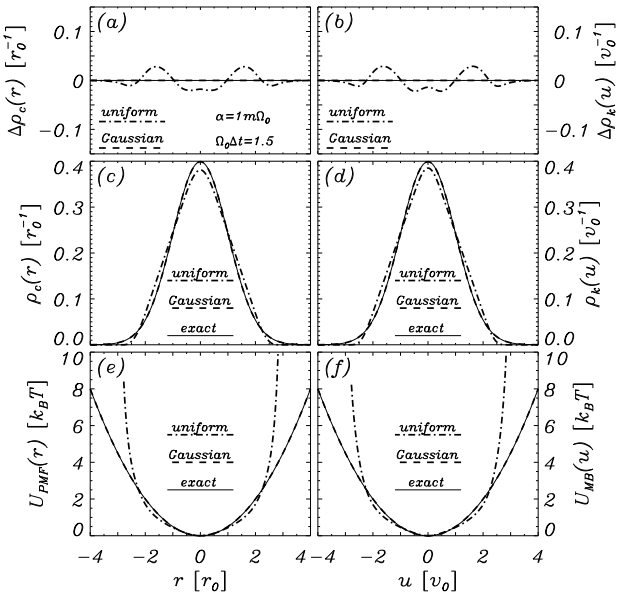


图 4. 对线性系统方程（6）和（18）施加高斯噪声和均匀噪声进行模拟时的构型分布 $\rho_c(r^n)$ 和$\rho_k(u^{n+\frac{1}{2}})$，其中 $\alpha = 1m\Omega_0$, $\Omega_0\Delta t = 1.5$, $k_BT = E_0$。图中曲线为方程（3）（实线，标记为精确分布）、使用高斯噪声 β_G^n 的模拟数据（虚线，标记为高斯分布），以及使用均匀噪声 β_U^n 的模拟数据（点划线，标记为均匀分布）。使用高斯噪声进行模拟得到的分布与方程（3）中的精确参考分布无法区分。（c）和（d）展示分布，（a）和（b）展示相对于精确分布方程（3）的偏差，（e）和（f）展示由这些分布导出的有效势（平均力势，PMF）和动能（麦克斯韦-玻尔兹曼，MB）势（方程（58））。

噪声分布一阶和二阶矩的统计性质，只要噪声满足式（10），就能通过 GJ 方法正确求得。图中每条曲线所示的数据对应的阻尼参数为 $\alpha/m\Omega_0$=0.1、0.25、0.5、1、2.5 和 5；高斯噪声分布和均匀噪声分布的结果都与式（23）和式（24）给出的预期值非常接近。此前已针对高斯噪声 [7, 8]对此进行了广泛验证。

图2cd展示了模拟结果，以及与第III C节分析预期的比较；比较对象是方程（41）和（42）中的构型能量涨落和动能涨落 σ_p andσ_k。由于应用于线性系统的方程（53a）中的高斯噪声β_G^n 会为 r^n 和$u^{n+\frac{1}{2}}$产生高斯分布，且其一阶和二阶矩正确并且不依赖于时间步长（见第III A节），因此四阶矩同样正确且不依赖于时间步长。在能量涨落图中可以清楚地看到这一点：对于 $\Gamma_4 = 3$，模拟结果与求解方程（37）–（39）所得结果无法区分，且均恒定在正确值$\sigma_p = \sigma_k = k_BT/\sqrt{2}$。然而，均匀分布噪声 β_U^n （方程（54a））既不会为 r^n 和$u^{n+\frac{1}{2}}$产生高斯分布，而是会产生依赖于时间步长和阻尼参数的分布，即使满足方程（10）也是如此。这一点可从能量涨落图图2cd中看出，其中展示了模拟结果以及方程（37）–（39）关于 $\Gamma_4 = 3$ 和$\Gamma_4 = \frac{9}{5}$ 的结果；后者是均匀分布的峰度。模拟结果与分析结果无法区分。显然，使用均匀噪声时，能量涨落取决于变量的四阶矩，可能显著偏离高斯噪声所给出的正确值。我们还观察到，当时间步长较小或阻尼参数较小时，涨落会趋近于正确值。这是可以理解的，因为在这些极限情况下，每个时间步内的阻尼都很小，因此 r^n 和$u^{n+\frac{1}{2}}$ 中的涨落由许多时间步的 β_U^n 贡献构成，使得 r^n 和$u^{n+\frac{1}{2}}$ 中的积分噪声根据中心极限定理趋近于高斯分布。反之，如果每个时间步内的阻尼变得不可忽略，那么 r^n 和$u^{n+\frac{1}{2}}$ 中的有效噪声将仅由少数几个 β_U^n 贡献构成，因而无法很好地近似高斯结果。

为了观察模拟分

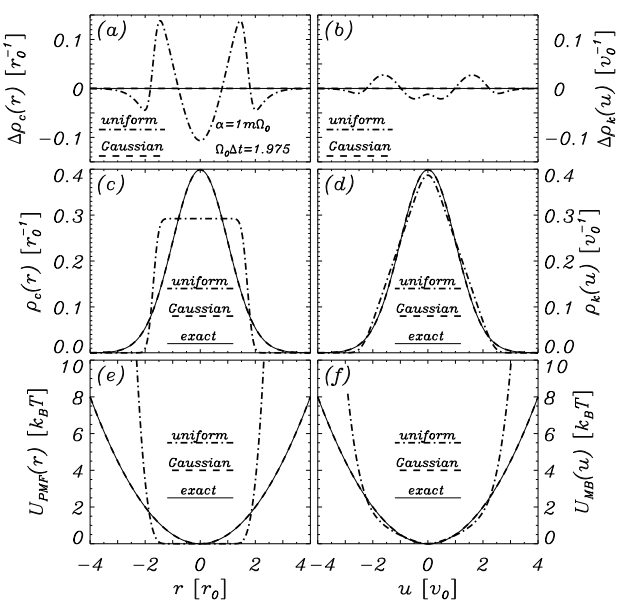


图 5. 对线性系统方程（6）和（18）施加高斯噪声和均匀噪声进行模拟时的构型分布和动能分布 $\rho_c(r^n)$ 和$\rho_k(u^{n+\frac{1}{2}})$，其中 $\alpha = 1m\Omega_0, \Omega_0\Delta t = 1.975$, $k_BT = E_0$。图中曲线为方程（3）（实线，标记为 exact）、使用高斯噪声 β_G^n 的模拟数据（虚线，标记为 Gaussian）以及使用均匀噪声 β_U^n 的模拟数据（点划线，标记为 uniform）。使用高斯噪声进行模拟得到的分布与方程（3）的精确参考分布无法区分。（c）和（d）显示分布，（a）和（b）显示相对于方程（3）精确分布的偏差，（e）和（f）显示由这些分布推导出的有效势（平均力势，PMF）和动能（麦克斯韦-玻尔兹曼，MB）势（方程（58））。

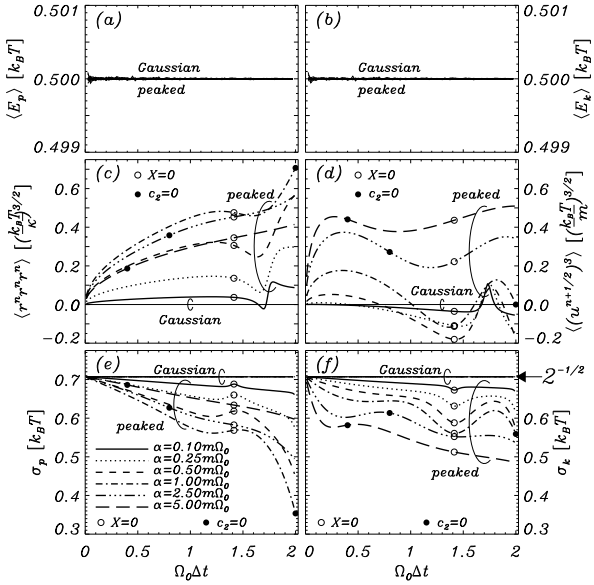


图6. 谐振系统（方程（6）和（18））中，非零温度 T下，平均势能和动能 $\langle E_p\rangle$（由方程（23）给出）（a）和 $\langle E_k\rangle$（由方程（24）给出）（b）、相应的三阶矩 $\langle (r^n)^3\rangle$（由方程（25）–（27）给出）（c）和 $\langle (u^{n+\frac{1}{2}})^3\rangle$（由方程（29）给出）（d），以及相应的能量涨落 σ_p（由方程（41）给出）（e）和 σ_k（由方程（42）给出）（f）随约化时间步长的变化情况及其模拟和分析结果。结果展示了高斯分布 β_G^n（方程（53a））和不对称峰形 β_P^n（方程（55a））噪声分布。采用了若干摩擦参数 α，如图所示。模拟结果与解析结果无法区分。标记表示 $c_2 = 0$（•，由方程（34a）和（36）、方程（50）–（52）给出）以及 $X = 0$（○，由方程（30a）和（33）、方程（45）和（46）给出）的特殊情形。

布的细节，我们选取了几个具有代表性的参数值作为示例。图 3–5 展示了分别在 $\alpha = 1m\Omega_0$ 和 $\Omega_0\Delta t = 0.5$, 1.5,1.975 下，使用高斯分布（虚线）和均匀分布（点划线）噪声分布进行模拟得到的构型分布和动能分布 $\rho_c(r)$ 和 $\rho_k(u)$。图中还展示了连续时间朗之万动力学预期得到的精确高斯分布结果 $\rho_{k,e}(u)$ 和 $\rho_{c,e}(r)$（实线，见公式 (3a) 和 (13a)）。分布 $\rho_c(r^n)$ 和 $\rho_k(u^{n+\frac{1}{2}})$ 显示在图 3–5cd 中，偏差，

$$\Delta\rho_c(r^n) = \rho_c(r^n) - \rho_{c,e}(r^n) \tag{57a}$$

$$\Delta\rho_k(u^{n+\frac{1}{2}}) = \rho_k(u^{n+\frac{1}{2}}) - \rho_{k,e}(u^{n+\frac{1}{2}}), \tag{57b}$$

与预期（正确）分布的偏差如图3–5ab所示；有效的构型势和动能势，

$$U_{PMF}(r^n) = -k_BT\ln\rho_c(r^n) + \tilde{\mathcal{C}}_c \tag{58a}$$

$$U_{MB}(u^{n+\frac{1}{2}}) = -k_BT\ln\rho_k(u^{n+\frac{1}{2}}) + \tilde{\mathcal{C}}_k, \tag{58b}$$

其中，$\tilde{\mathcal{C}}_c$ 和$\tilde{\mathcal{C}}_k$ 的取值满足 $\min[U_{PMF}(r^n)] = \min[U_{MB}(u^{n+\frac{1}{2}})] = 0$，结果如图所示

图 3-5ef。对于谐势，这些势的统计学正确值应为$\frac{1}{2}\kappa\langle r\rangle^2$ 和 $\frac{1}{2}m\langle u\rangle^2$，，如细实线所示。当 $\Omega_0\Delta t = 0.5$时，图 3 显示，即使与纯高斯分布之间看似微小的偏差，也会对四阶热力学量产生相当大的影响；对于图 2cd 中的这些参数，这导致构型能量涨落和动能涨落分别降低约 7% 和 14%。将时间步长增大至 $\Omega_0\Delta t = 1.5$会加剧由 β_U^n生成的分布偏离高斯分布的程度。如图 4ef 所示，愈发明显的差异在于，与高斯分布相比，均匀噪声使相空间的探索范围更受限，这与图 2cd 中四阶矩的降低相符。最后，图 5 展示了时间步长 $\Omega_0\Delta t = 1.975$非常接近稳定性极限时的结果。这个极端情形表明，噪声中的均匀分布也能产生近似均匀的 r^n分布，而 $u^{n+\frac{1}{2}}$ 的分布则近似为三角形，这与两个均匀分布的数值相加对动能涨落的贡献一致

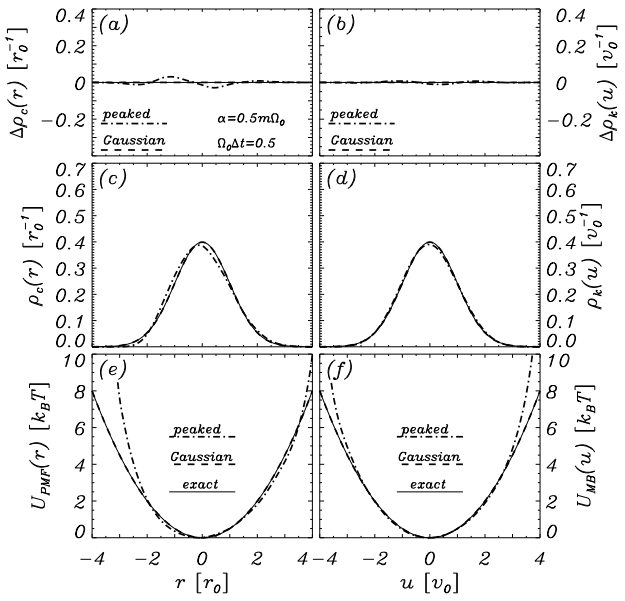


图 7. 对线性系统方程（6）和（18）施加高斯噪声和尖峰噪声进行模拟时的构型分布和动能分布 $\rho_c(r^n)$ 以及 $\rho_k(u^{n+\frac{1}{2}})$，其中 $\alpha = 0.5m\Omega_0, \Omega_0\Delta t = 0.5$，$k_BT = E_0$。图中曲线为方程（3）（实线，标记为精确）、使用高斯噪声 β_G^n 的模拟数据（虚线，标记为高斯分布），以及使用尖峰噪声 β_P^n 的模拟数据（点划线，标记为均匀）。使用高斯噪声进行模拟得到的分布与方程（3）的精确参考分布无法区分。（c）和（d）展示分布，（a）和（b）展示相对于精确分布方程（3）的偏差，（e）和（f）展示由这些分布导出的有效势（平均力势）和动能（麦克斯韦-玻尔兹曼）势（方程（58））。

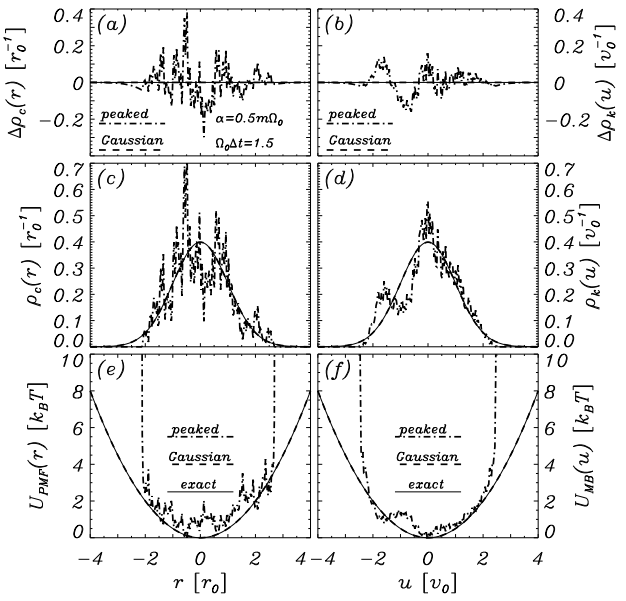


图 8. 对线性系统方程 (6) 和 (18) 施加高斯噪声和均匀噪声时的构型分布和动能分布 $\rho_c(r^n)$ 和$\rho_k(u^{n+\frac{1}{2}})$，其中 $\alpha = 0.5m\Omega_0$，$\Omega_0\Delta t = 1.5$, $k_BT = E_0$。所示曲线为方程 (3)（实线，标注为精确值）、使用高斯噪声 β_G^n 的模拟数据（虚线，标注为高斯分布），以及使用峰形噪声 β_P^n 的模拟数据（点划线，标注为均匀分布）。使用高斯噪声的模拟所得分布与方程 (3) 的精确参考分布无法区分。(c) 和 (d) 显示分布，(a) 和 (b) 显示相对于精确分布方程 (3) 的偏差，(e) 和 (f) 显示由这些分布导出的有效势（平均力势）和动能（麦克斯韦-玻尔兹曼）势（方程 (58)）。

项。再次从图 5ef 可以看出，与高斯噪声相比，使用均匀噪声时相空间的采样范围要受限得多，即使测得的构型温度和动能温度都与正确值相符。

秉承参考文献 [15, 16] 的思路，我们探究一种更具挑战性的噪声分布 β_P^n（即式（55a）中定义的尖峰分布）的应用；这种分布既是离散的，又是不对称的（见图1）。与上文讨论的均匀噪声一样，我们也将结果与高斯噪声的结果进行比较，后者表现正确且不受时间步长影响。按照上文所述，对不同时间步长以及归一化摩擦$\alpha/m\Omega_0 = 0.1$取0.25、0.5、1、2.5、5的情况，对式（6）和式（18）进行数值模拟，并与根据第III节分析计算得到的矩进行比较，得到图6所示的数据。其中，模拟结果与第 III B节和第III C节对三阶矩和四阶矩的分析结果无法区分。正如分析所预期的，图6abr^n 和 $u^{n+\frac{1}{2}}$所示的二阶矩，即势能和动能，与统计力学完全一致，因为 β_P^n 的尖峰分布满足式 (10) 中的两个矩。由于此处施加的噪声是不对称的，非零时间步长下，r^n 和 $u^{n+\frac{1}{2}}$ 的三阶矩也可能变为非零。这一点在图 6cd 中有所体现，我们观察到其随系统参数变化呈现出相当复杂的行为。不过，我们确实看到，当 $\Omega_0\Delta t \to 0$ 时，三阶矩趋近于零，这与中心极限定理一致；该定理保证，当大量噪声值对变量产生贡献时，结果将服从高斯分布。最后，与均匀噪声对应的图 2cd 相比，我们发现，在图 6ef 中应用尖峰分布时，能量涨落（四阶矩）表现出非常相似的行为。尽管均匀分布和尖峰分布的形态截然不同，它们的四阶矩却有相似的结果，这并不令人意外，因为第 III C 节对四阶矩的分析表明，两者的差异仅取决于峰度，其值分别为 $\Gamma_4 = \frac{9}{5}$（均匀分布）和 $\Gamma_4 = \frac{12}{5}$（尖峰分布）。同样，我们发现，当 $\Omega_0\Delta t \to 0$ 时，非高斯噪声的所有特征都会消失；并且我们发现，非零时间步长会导致.....的降低

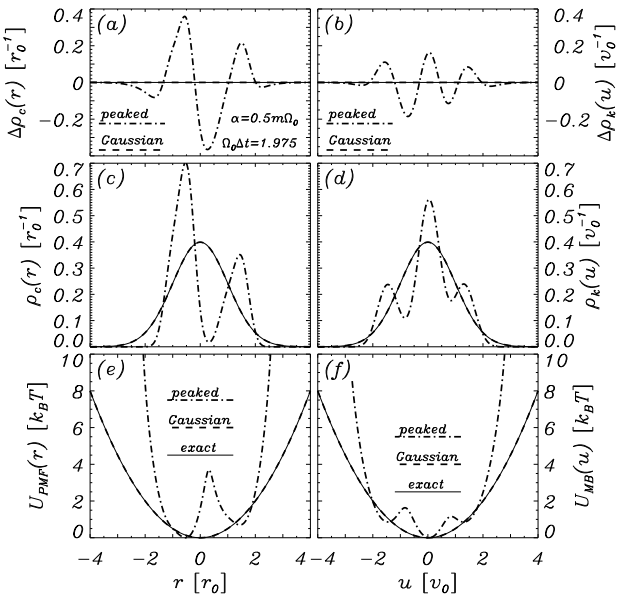


图 9. 对线性系统方程 (6) 和 (18) 施加高斯噪声和均匀噪声进行模拟时的构型分布和动能分布 $\rho_c(r^n)$ 和$\rho_k(u^{n+\frac{1}{2}})$，其中 $\alpha = 0.5m\Omega_0, \Omega_0\Delta t = 1.975$, $k_BT = E_{0}$。图中曲线为方程 (3)（实线，标记为精确值）、使用高斯噪声的模拟数据 β_G^n（虚线，标记为高斯分布）以及使用峰形噪声的模拟数据 β_P^n（点划线，标记为均匀分布）。使用高斯噪声的模拟所得分布与方程 (3) 的精确参考分布无法区分。(c) 和 (d) 显示分布，(a) 和 (b) 显示相对于精确分布（方程 (3)）的偏差，(e) 和 (f) 显示由这些分布导出的有效势（平均力势）和动能（麦克斯韦-玻尔兹曼）势（方程 (58)）。

由构型坐标和动能坐标的四阶矩推导出的、在热力学上具有重要意义的能量涨落。

尖峰噪声产生的模拟坐标分布的详细情况如图7–9所示，分别对应 $\alpha/m\Omega_0 = 0.5$ 和 $\Omega_0\Delta t = 0.5$、1.5、1.975。对于图7所示的较小时间步长，我们看到尖峰噪声导致的坐标分布呈现出看似不明显的偏度和变形。然而，正是这些轻微变形，造成了图6cd和图 6ef 中 $\Omega_0\Delta t = 0.5$对应的三阶矩和四阶矩相对于高斯分布特征的偏差，且偏差有一定幅度。在很大范围的 $\Omega_0\Delta t$内，都能发现相对于高斯分布/玻尔兹曼特征更显著的偏差，包括图8所示的 $\Omega_0\Delta t = 1.5$值。看似不连续的分布并非统计数据不足所致。相反，这是所施加噪声分布的离散性质与离散时间步长相互干涉的结果，恰好使 r^n 和 $u^{n+\frac{1}{2}}$ 中的某些优选值明显地比其他值更容易出现。值得注意的是，正如.....所示，这类特殊分布的出现并未在前四阶矩中留下任何明显或突兀的迹象

图6. 随着时间步长 $\Omega_0\Delta t = 1.975$ 逐渐接近稳定性极限，我们再次发现坐标分布平滑，如图9所示，其中施加的尖峰噪声在各个显示图中均清晰可见。与均匀噪声分布的可视化类似，在这一极限下，每个时间步长内的摩擦相对较大，因此坐标 r^n 和 $u^{n+\frac{1}{2}}$ 的行为在每次仅受少数噪声贡献的影响。

V. 讨论

鉴于近期随机恒温器方面的进展，在使用高斯噪声 [7, 8], 模拟朗之万方程时，即使采用的时间步长处于稳定范围内，也能获得准确的统计结果，因此我们重新审视了离散时间下使用非高斯热噪声的优缺点。传统方法存在系统性的一阶和二阶时间步长误差（例如，参考文献 [18, 20]），为获得准确的模拟结果，必须采用相当小的时间步长。基于此，先前的结论是 [15, 16]，其他分布（包括理想的均匀分布）都可以高效替代，因为中心极限定理可以确保在小时间步长下得到近似高斯结果，从而不会因噪声不完美而进一步显著扭曲模拟结果。多年来，这一结论一直非常有用：无需将随机变量从均匀分布转换为高斯分布，便能提高计算效率。然而，正如本文所展示的，当时间步长增大时——现代 GJ 方法在使用高斯噪声时允许采用较大的时间步长——施加非高斯噪声就无法保留这些方法在涉及高于二阶矩的热力学量方面与时间步长无关的优势。我们发现，当施加的噪声符合方程（10）时，GJ 方法在一阶矩和二阶矩的量度上不受具体噪声分布的影响，例如构型温度和动能温度。因此，仅依据这些矩来判断恒温器的质量会造成误导。从三阶矩和四阶矩，以及 r^n 和 $u^{n+\frac{1}{2}}$实际分布的直观表现中可以明显看出，相空间采样可能会严重失真（偏离玻尔兹曼分布），即使测得的温度值正确也是如此。所有这些结果都印证了离散时间下噪声定义方式的重要性；具体而言，噪声应通过积分方程（9）来定义，这能确保无论 $\beta(t)$ 服从何种底层分布，β^n 都将因中心极限定理而产生高斯结果。因此，在离散时间下使用高斯分布以外的任何分布在形式上都是无效的；除非 $\alpha\Delta t/m$足够小，使单步速度衰减参数为 $c_2 \approx 1$，否则必然会导致显著的采样误差。因此，本文得出的结论是，必须对

现代随机积分器施加高斯噪声，才能利用其大时间步长特性所带来的优势。

VI. 致谢

作者感谢 Charlie Sievers 分享使用 GJF-2GJ 方法获得的 LAMMPS 模拟结果，这些结果支持了本文关于更复杂分子动力学系统的结论；同时感谢 Lorenzo Mambretti 在 RANMAR 随机数生成器方面提供的帮助。作者还感谢最初就……进行的讨论与 Chungho Cheng 就非高斯噪声进行的讨论。

VII. 数据可用性声明

本研究中呈现和讨论的数据可根据合理请求向作者索取。

VIII. 资助和/或利益冲突

本研究未获得任何资助。作者没有需要披露的相关财务或非财务利益。

[1] M.P. Allen，D.J. Tildesley，*Computer Simulation of Liquids*，1989，牛津大学出版社，Inc.，.[2] D. Frenkel 和 B. Smit，*Understanding Molecular Simulations: From Algorithms to Applications*，（学术出版社，圣迭戈，2002年）。[3] D. C. Rapaport，*The Art of Molecular Dynamics Simulations*，2004，（剑桥大学出版社，剑桥，）。[4] W. M. Hoover，*Computational Statistical Mechanics*，（Elsevier Science B.V.，1991年）。[5] A. Leach，*Molecular Modelling: Principles and Applications*，2001，第版，（普伦蒂斯·霍尔：哈洛，英格兰，）。[6] N. Grønbech-Jensen 和 O. Farago，《分子物理学》，*A simple and effective Verlet-type algorithm for simulating Langevin dynamics* **111**，983，2013，，（）。[7] L. F. G. Jensen 和 N. Grønbech-Jensen，*Accurate configurational and kinetic statistics in discrete-time Langevin systems* **117**，2511，2019，《分子物理学》（）。[8] N. Grønbech-Jensen，*Complete set of stochastic Verlet-type thermostats for correct Langevin simulations*，《分子物理学》118，e1662506（2020年）。[9] P. Langevin，*On the Theory of Brownian Motion*，《巴黎科学院报告》146，530（1908年）。[10] W. T. Coffey 和 Y. P. Kalmykov，*The Langevin Equation*，3，2012，第版，世界科学出版社，新加坡，。[11] G. Parisi，*Statistical Field Theory*，（艾迪生-韦斯利，门洛帕克，1988年）。[12] 虽然并非必要，但通常可以合理地预期，$\beta(t)$的分布是对称的，因此所有奇数阶矩均为零。[13] A. Papoulis，*Probability Random Variables and Stochastic Processes*，1965，，，（麦格劳-希尔，伦敦，）。[14] 有关伪随机数生成器的一般介绍，参见例如 W. H. Press、S. A. Teukolsky、W. T. Vetterling 和 B. P. Flannery，*Numerical Recipes* 3第版。（剑桥大学出版社，剑桥，2007年）。[15] A. Greiner、W. Strittmatter 和 J. Honerkamp，*Numerical integration of stochastic differential equations*，，《统计物理学杂志》51，95（1988年）。[16] B. Dünweg 和 W. Paul，*Brownian dynamics simulations without Gaussian random numbers*，，《现代物理学国际期刊 C 辑》2，817（1991年）。[17] J. Finkelstein、G. Fiorin 和 S. Seibold，*Comparison of modern Langevin integrators for simulations of coarse-grained polymer melts*，《分子物理学》118，e1649493 (2020). [18] A. Brünger, C. L. Brooks, and M. Karplus，*Stochastic boundary conditions for molecular dynamics simulations of ST2 water* **105**，495，1984，，化学物理快报（）。[19] R. W. Pastor, B. R. Brooks, and A. Szabo，*An analysis of the accuracy of Langevin and molecular dynamics algorithms* **65**，1409，1988，，《分子物理学》（）。[20] T. Schneider and E. Stoll，*Molecular-dynamics study of a three-dimensional one-component model for distortive phase transitions* **17**，1302，1978，，Phys. Rev. B（）。[21] S. Plimpton，*Fast parallel algorithms for short-range molecular dynamics* **117**，1，1995，，J. Comp. Phys.（）。[22] A. P. Thompson, H. M. Aktulga, R. Berger, D. S. Bolintineanu, W. M. Brown, P. S. Crozier, P. J. in 't Veld, A. Kohlmeyer, S. G. Moore, T. D. Nguyen, R. Shan, M. Stevens, J. Tranchida, C. Trott, S. J. Plimpton，*LAMMPS – a flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales* **271**，108171，2022，，，计算物理通讯（）。[23]有关如何使用*fix_langevin*参考文献 [20]中的 Schneider 和 Stoll 恒温器，请参阅 LAMMPS文档 [31] 。[24] F. Weik, R. Weeber, K. Szuttor, K Breitsprecher, J. de Graaf, M. Kuron, J. Landsgesell, H. Menke, D. Sean, and C. Holm，*ESPResSo 4.0—an extensible software package for simulating soft matter systems* **227**，1789，2019，，欧洲物理学杂志特刊（）.另请参阅 https://espressomd.github.io/ 中关于 ESPResSo-4.2.0 第 6.3.1 节的文档。[25]L. Verlet，*Computer experiments on classical fluids. I. Thermodynamical properties of Lennard-Jones molecules*，Phys. Rev. 159, 98 (1967). [26] W. C. Swope, H. C. Andersen, P. H. Berens, K. R. Wil- son，*A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules: Application to small water clusters*，《化学物理杂志》76, 637 (1982). [27] D. Beeman，*Some multistep methods for use in molecular dynamics calculations* **20**，130，1976，，J. Comp. Phys.（）. [28] O. Buneman，*Time-reversible difference procedures*，J. Comp. Phys. 1, 517 (1967). [29] R. W. Hockney，*The potential calculation and some applications* **9**，135，1970，，Methods Comput. Phys.（）.

[30] J. Finkelstein, C. Cheng, G. Fiorin, B. Seibold, and N. Grønbech-Jensen, *Bringing Langevin splitting methods into agreement with correct discrete-time thermodynamics* **155**, 184104 2021 ,《化学物理杂志》 ( ).

[31] 有关“fix langevin”命令的说明，请参阅 http://lammps.sandia.gov/doc/Manual.pdf。