Coarsening of Particle Systems

C. David Levermore
Department of Mathematics *and*
Institute for Physical Science and Technology
University of Maryland, College Park, MD
lvrmr@math.umd.edu

presented 9 November 2012 during the KI-Net Workshop
Kinetic Description of Social Dynamics: From Concensus to Flocking
held 5-9 November 2012 at the
Center for Scientific Computation and Mathematical Modeling
University of Maryland, College Park, MD

Classical Particle Systems

A system of N classical particles is governed by an equation of the form

$$\frac{\mathrm{d}X}{\mathrm{d}t} = V(X)\,,$$

where X denotes a point in the N-particle phase space and $V(X)$ is the vector-field that governs the particle dynamics. For example, if the particles are governed by a Hamiltonian system

$$\frac{\mathrm{d}p_i}{\mathrm{d}t} = -\frac{\partial H}{\partial q_i}\,, \qquad \frac{\mathrm{d}q_i}{\mathrm{d}t} = \frac{\partial H}{\partial p_i}\,, \qquad \text{for } i = 1, \cdots, N\,,$$

where $H(p_1, q_1, \cdots, p_N, q_N)$, then $X = (p_1, q_1, \cdots, p_N, q_N)$ and

$$V(X) = \left(-\frac{\partial H}{\partial q_1}, \frac{\partial H}{\partial p_1}, \cdots, -\frac{\partial H}{\partial q_N}, \frac{\partial H}{\partial p_N}\right).$$

Liouville Equation

Most particle systems have more particles than we can afford to simulate. Moreover, most particle systems exhibit chaotic dynamics. This means that it would be impossible to track its orbits accurately for all but short times even if we could afford to simulate the system. However, often we do not want to track orbits accurately. Rather, we want to track robust phenomena that arise from the collective behavior of many particles. By robust phenomena, we mean phenomena that will arise in most members of an ensemble of solutions. Therefore it is natural to consider probability densities $F(X,t)$ over the N-particle phase space. These are governed by the Liouville equation

$$\partial_t F + \nabla_X \cdot \Big(V(X)\, F \Big) = 0\,.$$

The idea is to simulate a smaller particle system that exhibits the same robust phenomena as captured by an approximation to $F(X,t)$. We call the smaller system a coarsening of the original.

Coarsening

We coarsen the N-particle system into an n-particle system through an aggregation mapping $a(X)$. We think of N as the number of particles in the physical system we are modeling and n as the number of particles in a simulation of that system. If we set $x = a(X)$ and let $\partial a(X)$ denote the Jacobian of $a(X)$ then by the chain rule

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \partial a(X)\, V(X)\,.$$

This is not a closed system for x. We will use the Liouville equation to derive a closed system for x. The first step is to build an approximate evolution equation for the reduced pobability density $f(x,t)$ defined by

$$f(x,t) = \int \delta\Big(x - a(X)\Big)\, F(X,t)\, \mathrm{d}X\,.$$

The second step is to identify an evolution for x that corresponds to this evolution equation for $f(x,t)$.

Evolution of the Reduced Density

We see from the Liouville equation and the definition of $f(x,t)$ that the evolution of $f(x,t)$ is governed by

$$\partial_t f(x,t) + \int \delta\Big(x - a(X)\Big)\, \nabla_X \cdot \Big(V(X)\, F(X,t)\Big)\, \mathrm{d}X = 0\,.$$

Because for any vector-valued $U(X)$ we have the divergence identity

$$\int \delta\Big(x - a(X)\Big)\, \nabla_X \cdot U(X)\, \mathrm{d}X = \nabla_x \cdot \int \delta\Big(x - a(X)\Big)\, \partial a(X)\, U(X)\, \mathrm{d}X\,,$$

we can express the evolution equation for $f(x,t)$ in the divergence form

$$\partial_t f(x,t) + \nabla_x \cdot \int \delta\Big(x - a(X)\Big)\, \partial a(X)\, V(X)\, F(X,t)\, \mathrm{d}X = 0\,.$$

Notice that this form does not require $\partial a(X)$ to exist everywhere. In order to close this equation we must express $F(X,t)$ in terms of $f(x,t)$.

Reference Density

We will construct an expression for $F(X,t)$ in terms of $f(x,t)$ with the aid of a reference density. Let $G(X)$ be a positive probability density that satisfies the stationary Liouville equation

$$\nabla_X \cdot \Big(V(X)\, G(X) \Big) = 0\,.$$

For example, if $V(X)$ is Hamiltonian then $G(X)$ might be a Gibbs density

$$G(X) = \frac{1}{Z(\beta)} \exp\Big(-\beta H(X) \Big)\,, \quad Z(\beta) = \int \exp\Big(-\beta H(X) \Big)\, \mathrm{d}X\,,$$

where β can be obtained from the initial data by setting

$$\frac{1}{Z(\beta)} \int H(X)\, \exp\Big(-\beta H(X) \Big)\, \mathrm{d}X = \int H(X)\, F(X,0)\, \mathrm{d}X\,.$$

Here we are assuming that all other conserved quantities are initially zero. In that case the reference density is the thermal equilibrium associated with the initial probability density $F(X,0)$. There are other choices of β.

An Entropy-Based Closure

If we minimize the relative entropy

$$\int F \log\left(\frac{F}{G}\right) \, \mathrm{d}X \, ,$$

subject to the constraint

$$\int \delta\Big(x - a(X)\Big) \, F \, \mathrm{d}X = f(x,t) \, ,$$

then we obtain (steps are skipped) the so-called entropy-based closure

$$F(X,t) = G(X) \, \frac{f(a(X),t)}{g(a(X))} \, ,$$

where $g(x)$ is the positive probability density given by

$$g(x) = \int \delta\Big(x - a(X)\Big) \, G(X) \, \mathrm{d}X \, .$$

This implies that

$$\int \delta\Big(x - a(X)\Big)\, \partial a(X)\, V(X)\, F(X,t)\, \mathrm{d}X$$
$$= \int \delta\Big(x - a(X)\Big)\, \partial a(X)\, V(X)\, G(X)\, \mathrm{d}X\, \frac{f(x,t)}{g(x)}\,,$$

which implies the evolution equation for $f(x,t)$ has the form

$$\partial_t f + \nabla_x \cdot \Big(v(x)\, f\Big) = 0\,,$$

where $v(x)$ is the reduced vector-field given by

$$v(x) = \frac{1}{g(x)} \int \delta\Big(x - a(X)\Big)\, \partial a(X)\, V(X)\, G(X)\, \mathrm{d}X\,.$$

The coarsened n-particle system is therefore governed by

$$\frac{\mathrm{d}x}{\mathrm{d}t} = v(x)\,.$$

Moreover, it can be checked that $g(x)$ satisfies $\nabla_x \cdot \Big(v(x)\, g(x)\Big) = 0$.

Remark. This closure depends largely upon the reference density $G(X)$. The same closure is obtained by minimizing any functional of the form

$$\int G(X)\,\eta\left(\frac{F(X,t)}{G(X)}\right)\,\mathrm{d}X\,,$$

subject to the constraint

$$\int \delta\Big(x-a(X)\Big)\,F(X,t)\,\mathrm{d}X = f(x,t)\,,$$

where $\eta(z)$ is continuously differentiable and strictly convex over $\mathbb{R}_+$ with $\eta(1)=\eta'(1)=0$. For example, for any $p\in\mathbb{R}$ we could have chosen

$$\eta(z)=\begin{cases}\dfrac{z^p-1-p(z-1)}{p\,(p-1)} & \text{for } p\neq 0,1,\\ z-1-\log(z) & \text{for } p=0,\\ z\log(z)-z+1 & \text{for } p=1.\end{cases}$$

The case $p=1$ is equivalent to the entropy functional we used originally.

Hamiltonian Case

Suppose the N-particle dynamics has the Hamiltonian form

$$\frac{\mathrm{d}X}{\mathrm{d}t} = J(X)\,\nabla_X H(X)\,,$$

where the skew-symmetric matrix $J(X)$ gives the symplectic structure and $H(X)$ is the Hamiltonian. Then the coarsened system associated with the Gibbs density will have the Hamiltonian form

$$\frac{\mathrm{d}x}{\mathrm{d}t} = j(x)\,\nabla_x h(x)\,,$$

where the skew-symmetric matrix $j(x)$ gives the symplectic structure and the Hamiltonian $h(x)$ is determined from the relation

$$\exp\Big(-\beta h(x)\Big) = \int \delta\Big(x - a(X)\Big)\exp\Big(-\beta H(X)\Big)\,\mathrm{d}X\,,$$

provided the mapping $a(X)$ satisfies $j(a(X))\partial a(X) = \partial a(X) J(X)$.

Remark. This Hamiltonian example can be extended easily to the more general Lie-Poisson setting. Because most classical N-particle systems fall into the Hamiltonian setting, we do not give this extension here.

Remark. This example illustrates how statistical information can enter the coarsened system. In this case it enters through β, which appears in the formula for the coarsened Hamiltonian $h(x)$ at the bottom of the last page. Because β is essentially reciprocal temperature, this might be related to the quantum statistical potential (QSP) approach in the quantum setting.

Remark. A quantum analog of the foregoing classical formalism can be carried out in the Heisenberg density matrix setting. It is less clear how an analogous formalism might be carried out in the Wigner equation setting.

Mori-Zwanzig Framework

We can improve the foregoing closure within the so-called Mori-Zwanzig framework. We begin by decomposing $F(X,t)$ as

$$F(X,t) = G(X)\,\frac{f(a(X),t)}{g(a(X))} + \widetilde{F}(X,t)\,.$$

Here $\widetilde{F}(X,t)$ is the *deviation* of $F(X,t)$ from the entropy-based closure. The evolution equation for $f(x,t)$ then has the form

$$\partial_t f + \nabla_x \cdot \Big(v(x) f\Big) + \nabla_x \cdot \int \delta\Big(x - a(X)\Big)\,\partial a(X)\,V(X)\,\widetilde{F}(X,t)\,\mathrm{d}X = 0\,,$$

where $v(x)$ is obtained from the entropy-based closure, which is recovered by setting $\widetilde{F}(X,t) = 0$. The idea is to solve the evolution equation for $\widetilde{F}(X,t)$ in terms of $f(x,t)$ in order to eliminate $\widetilde{F}(X,t)$ from this equation.

The evolution equation for the deviation $\widetilde{F}(X,t)$ is

$$\partial_t \widetilde{F} + \widetilde{\mathcal{P}}\nabla_X \cdot \Big(V(X)\widetilde{F}\Big) = -\widetilde{\mathcal{P}}G(X)V(X)\cdot\nabla_X\left(\frac{f(a(X),t)}{g(a(X))}\right),$$

where $\widetilde{\mathcal{P}} = \mathcal{I} - \mathcal{P}$ and $\mathcal{P}$ is the projection operator that acts formally on any function $W(X)$ as

$$\mathcal{P}W(X) = \frac{G(X)}{g(a(X))}\int \delta\Big(a(X) - a(X')\Big)\, W(X')\,\mathrm{d}X'.$$

The right-hand side of the evolution equation for $\widetilde{F}(X,t)$ can be written as

$$\widetilde{R}(X,t) = -G(X)\Big(\partial a(X)V(X) - v(a(X))\Big)\cdot\nabla_x\left(\frac{f(a(X),t)}{g(a(X))}\right).$$

This shows that the evolution of $\widetilde{F}(X,t)$ is driven both by the gradient of the ratio of $f(x,t)$ to its equilibrium $g(x)$ and by the difference between the coarsened vector-field $v(a(X))$ and $\partial a(X)V(X)$.

We can formally solve the evolution equation for $\widetilde{F}(X,t)$ as

$$\widetilde{F}(t) = \exp\Big(-t\widetilde{\mathcal{A}}\Big)\,\widetilde{F}^{\mathsf{in}} + \int_0^t \exp\Big(-(t-t')\widetilde{\mathcal{A}}\Big)\,\widetilde{R}(t')\,\mathrm{d}t'\,,$$

where we have suppressed the dependence on X in all terms, $\widetilde{F}^{\mathsf{in}}$ is the initial-data for $\widetilde{F}(t)$, and the operator $\widetilde{\mathcal{A}}$ is defined by

$$\widetilde{\mathcal{A}}W = \widetilde{\mathcal{P}}\nabla_X \cdot \Big(V(X)\widetilde{\mathcal{P}}W\Big)\,.$$

If the operator $\exp\Big(-t\widetilde{\mathcal{A}}\Big)$ is represented by a kernel $S(X,X',t)$ then $\widetilde{F}(X,t)$ can be expressed as

$$\begin{aligned}\widetilde{F}(X,t) = &\int S(X,X',t)\,\widetilde{F}^{\mathsf{in}}(X')\,\mathrm{d}X' \\ &+ \int_0^t \int S(X,X',t-t')\,\widetilde{R}(X',t')\,\mathrm{d}X'\,\mathrm{d}t'\,.\end{aligned}$$

The Mori-Zwanzig evolution equation for $f(x,t)$ is obtained by placing this expression for $\widetilde{F}(X,t)$ into the equation for $\partial_t f(x,t)$ given two slides ago.

When the foregoing expression for $\widetilde{F}(X,t)$ is placed into the equation for $\partial_t f(x,t)$ given three slides ago while using the fact that

$$\begin{aligned}
&\int \delta\Big(x - a(X)\Big)\, \partial a(X)\, V(X)\, \widetilde{F}(X,t)\, \mathrm{d}X \\
&= \int \delta\Big(x - a(X)\Big) \Big(\partial a(X)\, V(X) - v(x)\Big)\, \widetilde{F}(X,t)\, \mathrm{d}X\,,
\end{aligned}$$

the equation for $\partial_t f(x,t)$ becomes

$$\begin{aligned}
&\partial_t f + \nabla_x \cdot \Big(v(x) f\Big) \\
&\quad + \nabla_x \cdot \iint \delta\Big(x - a(X)\Big) \Big(\partial a(X)\, V(X) - v(x)\Big) \\
&\qquad\qquad S(X, X', t)\, \widetilde{F}^{\mathsf{in}}(X')\, \mathrm{d}X'\, \mathrm{d}X \\
&\quad + \nabla_x \cdot \int \int_0^t \int \delta\Big(x - a(X)\Big) \Big(\partial a(X)\, V(X) - v(x)\Big) \\
&\qquad\qquad S(X, X', t - t')\, \widetilde{R}(X', t')\, \mathrm{d}X'\, \mathrm{d}t'\, \mathrm{d}X\,,
\end{aligned}$$

where $\widetilde{R}(X,t)$ was given in terms of $f(x,t)$ two slides ago.

Remark. This Mori-Zwanzig evolution equation is different than the one introduced by them because we use a reference probability density $G(X)$. They used a uniform reference measure, which corresponds to taking the infinite temperature limit with the Gibbs density. We recover their equation by setting $G(X) = 1$ and requiring $a(X)$ to satisfy

$$g(x) = \int \delta\Big(x - a(X)\Big)\, \mathrm{d}X < \infty \qquad \text{for every } x.$$

Remark. The Mori-Zwanzig evolution equation makes no approximations. It is formally equivalent to the original problem through the appearance of the kernel $S(X, X', t)$ in the two new terms. These are memory terms. The first has memory of the initial-data $\widetilde{F}^{\mathsf{in}}(X)$, while the second has memory of $f(x, t')$ over times t' in $(0, t)$. The Mori-Zwanzig framework suggests we should approximate the actions of $S(X, X', t)$ in these terms.

Markov Approximation

We now assume that the dynamics represented by $S(X, X', t)$ forgets its past fast enough that we can use the so-called *Markov approximation*

$$\widetilde{F}(X,t) = \int K(X,X')\,\widetilde{R}(X',t)\,\mathrm{d}X'\,,$$

where the kernel $K(X, X')$ acts on a sufficiently large class of $\widetilde{R}(X)$ as

$$\begin{aligned}\int K(X,X')\,\widetilde{R}(X')\,\mathrm{d}X' &= \lim_{t\to\infty}\int_0^t\int S(X,X',t-t')\,\widetilde{R}(X')\,\mathrm{d}X'\,\mathrm{d}t' \\ &= \lim_{t\to\infty}\int_0^t\int S(X,X',t')\,\widetilde{R}(X')\,\mathrm{d}X'\,\mathrm{d}t'\,.\end{aligned}$$

The class of $\widetilde{R}(X)$ for which this limit has to be established must include the form of $\widetilde{R}(X,t)$ that appears on the right-hand side of the evolution equation for $\widetilde{F}(X,t)$. This limit cannot be computed generally.

If the Markov approximation for $\widetilde{F}(X,t)$ is placed into the equation for $\partial_t f(x,t)$ then that equation becomes

$$\partial_t f + \nabla_x \cdot \Big(v(x) f \Big) = \nabla_x \cdot \left[\int k(x,x') g(x') \nabla_{x'} \left(\frac{f(x',t)}{g(x')} \right) \, \mathrm{d}x' \right] ,$$

where the matrix-valued kernel $k(x,x')$ is given by

$$\begin{aligned} k(x,x')g(x') = \iint & \delta\Big(x - a(X)\Big) \Big(\partial a(X) V(X) - v(x) \Big) \\ & K(X,X') \, G(X') \Big(\partial a(X') V(X') - v(x') \Big)^{\mathsf{T}} \\ & \delta\Big(x' - a(X')\Big) \, \mathrm{d}X \, \mathrm{d}X' . \end{aligned}$$

To be the generator of a Markov process, this equation should preserve the nonnegativity of its solutions and dissipate all functionals of the form

$$\int g(x) \, \eta\left(\frac{f(x,t)}{g(x)} \right) \, \mathrm{d}x \, , \quad \text{where } \eta(z) \text{ is strictly convex over } \mathbb{R}_+ .$$

Whenever these properties hold we can try to identify a Markov process for which the equation for $f(x,t)$ is the forward Kolmogorov equation. For example, these properties hold when $s(x,x') = s(x',x) \geq 0$, where

$$s(x,x') = -\nabla_{x'} \cdot \Big(\nabla_x \cdot k(x,x') g(x') \Big)^{\mathsf{T}} .$$

In that case the evolution equation for $f(x,t)$ takes the form

$$\partial_t f + \nabla_x \cdot \Big(v(x) f \Big) = \int s(x,x') \left(\frac{f(x',t)}{g(x')} - \frac{f(x,t)}{g(x)} \right) \mathrm{d}x' .$$

This is the forward Kolmogorov equation for the Markov process

$$\frac{\mathrm{d}x}{\mathrm{d}t} = v(x) + \frac{\mathrm{d}z(t)}{\mathrm{d}t} ,$$

where $z(t)$ is the jump process associated with scattering kernel $s(x,x')$. This coarsened dynamics has characteristics of a Monte Carlo simulation.

Whenever these properties do not hold we can try to replace $k(x, x')$ with an appoximation for which they do hold. For example, if $k(x, x')$ is peaked along the diagonal $x' = x$ then we can make the approximation

$$k(x, x') \approx d(x)\, \delta(x - x') \,.$$

where $d(x)$ is the (nonnegative definite) matrix-valued function given by

$$d(x)g(x) = \int k(x, x')g(x')\, \mathrm{d}x' \,.$$

The resulting evolution equation then becomes

$$\partial_t f + \nabla_x \cdot \Big(v(x)f\Big) = \nabla_x \cdot \left[d(x)g(x)\nabla_x \left(\frac{f}{g(x)}\right)\right] \,.$$

This is the forward Kolmogorov equation for the Markov process

$$\frac{\mathrm{d}x}{\mathrm{d}t} = v(x) + \nabla_x \cdot d(x) + d(x)\nabla_x \log(g(x)) + \sqrt{2d(x)}\, \frac{\mathrm{d}w(t)}{\mathrm{d}t} \,,$$

where $w(t)$ is a vector of Wiener processes.

In the Hamiltonian case this coarsened dynamics becomes

$$\frac{\mathrm{d}x}{\mathrm{d}t} = j(x)\nabla_x h(x) + \nabla_x \cdot d(x) - \beta d(x)\nabla_x h(x) + \sqrt{2d(x)}\,\frac{\mathrm{d}w(t)}{\mathrm{d}t}\,.$$

More generally, we might be able to split the kernel $k(x, x')$ into two parts, one that gives rise to a jump process and one that gives rise to a diffusion process. For example, in the Hamiltonian case p might be governed by a jump process while q might be governed by a diffusion process.

Remark. The evolution equation for $f(x, t)$ should be supplemented by an initial condition. The simplest such condition is

$$f(x, 0) = \int \delta\Big(x - a(X)\Big)\, F^{\mathsf{in}}(X)\, \mathrm{d}X\,,$$

where $F^{\mathsf{in}}(X)$ is the initial-data for $F(X, t)$. This initial condition specifies nonnegative initial-data for every nonnegative $F^{\mathsf{in}}(X)$.

Remark. When $\widetilde{F}^{\mathsf{in}}$ is small then another possible initial condition is

$$f(x,0) = \int \delta\Big(x - a(X)\Big)\, F^{\mathsf{in}}(X)\, \mathrm{d}X + \nabla_x \cdot \iint \delta\Big(x - a(X)\Big)\, \Big(\partial a(X)\, V(X) - v(x)\Big) K(X,X')\, \widetilde{F}^{\mathsf{in}}(X')\, \mathrm{d}X'\, \mathrm{d}X\,.$$

This initial condition has the virtue of being more accurate formally, but may not specify nonnegative initial-data for every nonnegative $F^{\mathsf{in}}(X)$.

Remark. Justifying any Markov approximation can be extremely difficult. It requires showing the dynamics generated by $V(X)$ is either sufficiently dissipative or sufficiently chaotic so that for $t = O(1)$ we have

$$\int S(X,X',t)\, \widetilde{F}^{\mathsf{in}}(X')\, \mathrm{d}X' \sim 0\,,$$

$$\int_0^t \int S(X,X',t-t')\, \widetilde{W}(X',t')\, \mathrm{d}X'\, \mathrm{d}t' \sim \int K(X,X')\, \widetilde{W}(X',t)\, \mathrm{d}X'\,.$$

Closing Remarks

Remark. The reference density $G(X)$ was taken to be stationary in the above discussion. Much of what was done goes through if the reference density is slowly varying. For example, it could be a slowly modulated Gibbs density as in the Green-Kubo formalism.

Remark. A similar formalism can be carried out to coarsen a system of N damped stochastic particles to a system of n damped stochastic particles. In that case the Markov approximation is easier to justify. This idea was explored in a one-dimensional setting in the dissertation of J.T. Halbert.

Remark. The fast multipole method provides a framework within which coarsening mappings can be viewed. It shows that close interactions need to be modeled by more than Colomb interactions of aggregate particles.