# Derivation of the GENERIC form of nonequilibrium thermodynamics from a statistical optimization principle

Bruce Turkington
Univ. of Massachusetts Amherst

"An optimization principle for deriving nonequilibrium statistical models of Hamiltonian dynamics." *arXiv 1207.2692*

"A nonequilibrium statistical model of spectrally truncated Burgers-Hopf dynamics." (with Richard Kleeman) *arXiv 1206.6545*

"Best-fit quasi-equilibrium ensembles: a general approach to statistical closure of underresolved Hamiltonian dynamics." (with Petr Plechac) *arXiv 1010.4362*

Nonequilibrium thermodynamics is conceived in a general sense as a *statistical closure* of a complex dynamical system.

- A set of *resolved* (relevant, macroscopic, slow) variables is identified.
- All *unresolved* (irrelevant, microscopic, fast) variables are described by a statistical (probabilistic, stochastic) model.
- A *closed reduced* dynamics approximates (estimates, predicts) the expected (mean, average) evolution of the resolved variables.

The approach discussed in this lecture uses

1. An underlying dynamics that is Hamiltonian with $n$ degrees of freedom.
2. Any independent set of $m$ resolved variables; typically, $m \ll n$.
3. Canonical statistical models of the unresolved fluctuations.
4. An optimization principle to construct the closed reduced equations.

**Key idea:** The reduced model finds the *best fit* of all feasible evolutions of the resolved variables to the underlying dynamics.

## Dynamical system and statistical model

The microstate is a point $z = (q_1, p_1, \ldots, q_n, p_n) \in \Gamma = R^{2n}$.

The microscopic dynamics is *canonical Hamiltonian*:

$$\frac{dq_j}{dt} = \frac{\partial H}{\partial p_j}, \qquad \frac{dp_j}{dt} = -\frac{\partial H}{\partial q_j}.$$

Every *dynamical variable* $F : \Gamma \to R$ evolves under

$$\frac{dF}{dt} = \{F, H\} = \sum_j \frac{\partial F}{\partial q_j}\frac{\partial H}{\partial p_j} - \frac{\partial F}{\partial p_j}\frac{\partial H}{\partial q_j}$$

**Examples of interest for reduction:**

1. A few heavy particles coupled to a "bath" of many light particles.
2. A set of wave modes in a nonlinear wave system interacting with a "fluctuating" wave field.
3. A large-scale coherent structure immersed in a "background" of small-scale turbulent fluctuations.

Sometimes a *non-canonical* Hamiltonian structure is convenient.

An *ensemble* of microscopic trajectories, $z(t)$, is described by a probability density, $\rho(z,t)$, governed by *Liouville's equation*:

$$\frac{\partial \rho}{\partial t} + L\rho = 0\,, \qquad \text{with } L\cdot = \{\,\cdot\,, H\}$$

For every dynamical variable $F$,

$$\frac{d}{dt}\langle F \,|\, \rho(\cdot,t)\rangle = \langle LF \,|\, \rho(\cdot,t)\rangle\,; \qquad \langle F \,|\, \rho\rangle \doteq \int_{\Gamma} F(z)\rho(z)\, dz\,.$$

But the exact solution, $\rho(t) = e^{-tL}\rho(0)$, is not really useful, being too intricate to analyze and too expensive to compute.

Instead impose an approximation to $\rho(z,t)$ that is parametrized by the statistical averages of some selected *resolved variables*

$$A = (A_1, \ldots, A_m) \qquad (\text{ each } A_i : \Gamma \to R\,)$$

## Canonical statistical model

A natural choice of (nonequilibrium) *trial probability densities* is the parametric family

$$\tilde{\rho}(z,\lambda) \,=\, \mathrm{exp}[\lambda^* A(z) \,-\, \phi(\lambda)]\, \rho_{eq}(z)$$

where

$$\phi(\lambda) \,=\, \log \int_\Gamma \mathrm{exp}(\lambda^* A(z))\, \rho_{eq}(z)\, dz \qquad \text{for} \quad \lambda \in R^m$$

$\rho_{eq}$ is a fixed equilibrium density on $\Gamma$; that is, $L\rho_{eq} = 0$.
For instance,

$\rho_{eq}(z) = Z(\beta)^{-1} \mathrm{exp}(-\beta H(z))$ [Gibbs canonical ensemble].

$\lambda = (\lambda^1, \ldots, \lambda^m)$ is the parameter vector for model.

Here and throughout, $\lambda^* A = \lambda^1 A_1 + \ldots + \lambda^m A_m$ .

Two interpretations of $\tilde{\rho}(z,\lambda) = \exp[\lambda^* A(z) - \phi(\lambda)]\, \rho_{eq}(z)$

**Statistics:** $\tilde{\rho}$ defines a parametric statistical model; specifically, an *exponential family with natural parameter* $\lambda$. The vector $A$ of resolved variables is a minimal sufficient statistic for this parametric model.

**Physics:** Each $\tilde{\rho}$ is a *quasi-equilibrium/canonical ensemble*.
It maximizes (relative) entropy

$$S(\rho) = -\langle \log \frac{\rho}{\rho_{eq}} \,|\, \rho \rangle$$

subject to the expected value of the vector of resolved variables

$$\langle A \,|\, \rho \rangle = a\,, \qquad \langle 1 \,|\, \rho \rangle = 1\,.$$

The *entropy function* for this model is $s(a) = S(\tilde{\rho}(\cdot,\lambda))$ and there is a one-to-one correspondence between $a$ and $\lambda$:

$$a = \frac{\partial \phi}{\partial \lambda}\,, \qquad \lambda = -\frac{\partial s}{\partial a} \qquad \text{[Legendre transform].}$$

## Adiabatic closure

A naive way to evolve the parameter $\lambda(t) \in R^m$ is to impose $A_1, \ldots, A_m$ moments of the Liouville equation exactly:

$$\frac{d}{dt}\langle A|\tilde{\rho}\,\rangle \;-\; \langle LA|\tilde{\rho}\,\rangle \;=\; \int_\Gamma A\,(\frac{\partial}{\partial t} + L)\,\tilde{\rho}\,dz \;=\; 0\,.$$

But this closed reduced dynamics is *reversible*:

$$\frac{d}{dt}S(\tilde{\rho}) \;=\; 0 \qquad\qquad [\text{ entropy is conserved }]$$

This memoryless closure neglects the influence of the unresolved fluctuations on the evolution of the resolved variables.

Statistical physicists (Zubarev 1970s and others) advocate using "nonequilibrium statistical operators" with decaying memory

$$\rho(t) \;=\; \exp\left(\int_0^{+\infty} \lambda(t-\tau)\,e^{\tau(L-\alpha)}A\,\alpha d\tau \;-\; \psi[\lambda](t)\right)\,\rho_{eq}$$

## Best-fit closure – a different approach

Retain the quasi-equilibrium trial densities, and evaluate their *Liouville residual* along any feasible parameter path $\lambda(t)$:

$$R \doteq (\frac{\partial}{\partial t} + L) \log \tilde{\rho}(\lambda(t)) = \dot{\lambda}(t)^*(A - a(t)) + \lambda(t)^* LA$$

The ensemble-averaged evolution of any dynamical variable $F$ along such a path of trial densities:

$$\frac{d}{dt}\langle F \,|\, \tilde{\rho}(\lambda(t))\,\rangle - \langle LF \,|\, \tilde{\rho}(\lambda(t))\,\rangle = \langle FR \,|\, \tilde{\rho}(\lambda(t))\,\rangle$$

The statistic $R = R(z; \lambda, \dot{\lambda})$ represents the rate of information loss locally at state $\lambda$ and velocity $\dot{\lambda}$ :

At any sample microstate $z \in \Gamma$, the information for discriminating between the exact density and the trial density after a time increment $\Delta t$ is:

$$\log \frac{e^{-(\Delta t)L}\tilde{\rho}(z, \lambda(t))}{\tilde{\rho}(z, \lambda(t + \Delta t))} = -(\Delta t)\, R(z; \lambda(t), \dot{\lambda}(t)) + O(\,(\Delta t)^2\,) \qquad \text{as } \Delta t \to 0.$$

**Closure concept:** Find paths $\lambda(t)$ in the statistical parameter space that are "best-fit" to the Liouville equation in the sense that they minimize $R$ in some time-integrated norm.

Separate $R$ into its orthogonal components along the resolved and unresolved subspaces of $L^2(\Gamma, \tilde{\rho}(\lambda))$:

$$R(\lambda, \dot{\lambda}) \;=\; P_\lambda R + Q_\lambda R \,, \qquad \text{where}$$

$$\begin{aligned} P_\lambda R \;&=\; [\, \dot{\lambda} - C(\lambda)^{-1} f(\lambda)\,]^*(A - a)\,, \\ Q_\lambda R \;&=\; \lambda^*(Q_\lambda L A) \end{aligned}$$

$C_{ij}(\lambda) = \langle\, (A_i - a_i)(A_j - a_j) \,|\, \tilde{\rho}(\lambda) \,\rangle$ [Fisher information matrix].
$f_i(\lambda) = \langle\, L A_i \,|\, \tilde{\rho}(\lambda) \,\rangle$ [Reversible term in flux].

The *lack-of-fit cost function*:

$$2\,\mathcal{L}(\lambda, \dot{\lambda}) \;=\; \langle\, (P_\lambda R)^2 |\, \tilde{\rho}(\lambda) \,\rangle + \langle\, (W_\lambda Q_\lambda R)^2 |\, \tilde{\rho}(\lambda) \,\rangle\,.$$

The bounded linear operator $W_\lambda$ on $L^2(\Gamma)$ weights the unresolved component $Q_\lambda R$ relative to the resolved component $P_\lambda R$.
$W_\lambda$ contains all the *adjustable parameters* in the closure.

The **best-fit closure** minimizes the lack-of-fit cost functional over all feasible parameter paths $\lambda(t)$, $0 \leq t < +\infty$:

$$\text{Minimize } \int_0^{+\infty} \mathcal{L}(\lambda, \dot{\lambda})\, dt \qquad \text{over paths } \lambda(t) \text{ with } \lambda(0) = \lambda_0 \,.$$

The initial state $\lambda_0 \neq 0$ is specified at time $t = 0$.

The predicted (estimated) macrostate at each time $t > 0$ corresponds to the *extremal* parameter vector $\widehat{\lambda}(t)$.

The predicted evolution $\widehat{\lambda}(t)$ represents a *relaxation* from the nonequilibrium state $\lambda_0 \neq 0$ towards equilibrium $\lambda_{eq} = 0$.

The evolving trial probability density $\tilde{\rho}(\cdot, \widehat{\lambda}(t))$ is the best-fit estimate of the true density $\rho(t) = e^{-tL}\tilde{\rho}(\lambda_0)$ .

## The closed reduced equations

Introduce the *value function* (principal function or "action") of Hamilton-Jacobi theory:

$$v(\lambda_0) = \min_{\lambda(0)=\lambda_0} \int_0^{+\infty} \mathcal{L}(\lambda, \dot{\lambda})\, dt\,,$$

which assigns an optimal lack-of-fit to every state $\lambda_0 \in R^m$.

$v(\lambda)$ solves the (stationary) *Hamilton-Jacobi equation*

$$\mathcal{H}\left(\lambda, -\frac{\partial v}{\partial \lambda}\right) = 0\,,$$

where $\mathcal{H}(\lambda, \mu)$ is the Legendre transform of $\mathcal{L}(\lambda, \dot{\lambda})$:

$$\mu = \frac{\partial \mathcal{L}}{\partial \dot{\lambda}} = C(\lambda)\dot{\lambda} - f(\lambda)$$

$$\mathcal{H} = \dot{\lambda}^* \mu - \mathcal{L} = \frac{1}{2}\mu^* C(\lambda)^{-1}\mu + f(\lambda)^* C(\lambda)^{-1}\mu - \frac{1}{2}\lambda^* D(\lambda)\lambda$$

and

$$D_{ij}(\lambda) = \langle\, (W_\lambda Q_\lambda L A_i)(W_\lambda Q_\lambda L A_j) \,|\, \tilde{\rho}(\lambda)\, \rangle\,.$$

Along an extremal $\widehat{\lambda}(t)$ the corresponding $\widehat{\mu}(t)$ satisfies

$$\widehat{\mu}(t) = -\frac{\partial v}{\partial \lambda}(\widehat{\lambda}(t)) \,.$$

Replacing $\mu$ by its expression in terms of $\dot{\lambda}$ and $\lambda$ produces the **closed reduced equations** (first-order DEs in $t$):

$$\frac{d\widehat{a}}{dt} = C(\widehat{\lambda})\frac{d\widehat{\lambda}}{dt} = f(\widehat{\lambda}) - \frac{\partial v}{\partial \lambda}(\widehat{\lambda}) \,, \qquad \text{with} \qquad \widehat{\lambda} = -\frac{\partial s}{\partial a}(\widehat{a}) \,.$$

Recall that

$$a = \langle\, A \,|\, \tilde{\rho}(\lambda) \,\rangle \,, \qquad C(\lambda) = \langle\, (A-a)(A-a)^* \,|\, \tilde{\rho}(\lambda) \,\rangle \,, \qquad f(\lambda) = \langle\, LA \,|\, \tilde{\rho}(\lambda) \,\rangle.$$

**Physical interpretation** of canonically conjugate $\lambda$ and $\mu$:

$-\lambda^i$ is the *thermodynamic force (affinity)* associated with $A_i$.

$\mu_i = -\frac{\partial v}{\partial \lambda_i}$ is the corresponding *thermodynamic flux*.

$-\sum_i \lambda^i \mu_i \;=\; \sum_i \lambda^i \frac{\partial v}{\partial \lambda^i}$ is the *entropy production*.

## Relation to GENERIC Nonequilibrium Thermodynamics

Various authors (Morrison, Beris and Edwards, Grmela and Öttinger, and others) have proposed a generic format for nonequilibrium macrodynamics:

$$\frac{da}{dt} = L(a)\frac{\partial E}{\partial a} + M(a)\frac{\partial S}{\partial a}$$

where

$$E = \text{ energy}, \;\; S = \text{ entropy}, \;\; L = -L^*, \;\; M = M^* \geq 0\,.$$

The first term is reversible, a generalized Hamiltonian vector field.
The second term is irreversible, a generalized gradient vector field.

Our statistical optimization principle produces expressions for these reversible and irreversible terms:

$$L(a) = \langle\, \{A, A^*\}\, \rangle\,, \quad \frac{\partial E}{\partial a} = \frac{\partial h}{\partial a}(s,a) \;\text{ with }\; h(s,a) = \langle H \rangle\,.$$

$$M(a)\frac{\partial S}{\partial a} = -\frac{\partial v}{\partial \lambda}\,.$$

**Near equilibrium (linear-response) regime**

For small $|\lambda|$, and $a = \langle A|\tilde{\rho}\,\rangle \approx \langle A\rangle_{eq} = 0$, the governing equations of the closed reduced model simplify:

$$\frac{d\hat{a}}{dt} = (\,J \,-\, M\,)\,C^{-1}\,\hat{a}\,,$$

$$JC^{-1}M \,-\, MC^{-1}J + MC^{-1}M \,=\, D\,,$$

where $\hat{a} = C\hat{\lambda}$ and

$$C = \langle AA^{*}\rangle_{eq}\,, \quad J = \langle (LA)A^{*}\rangle_{eq}, \quad D = \langle\,(WQLA)\,(WQLA)^{*}\,\rangle_{eq}\,.$$

*Properties* in the near-equilibrium case:

$$v(\lambda) = \frac{1}{2}\lambda^{*}M\lambda\,, \quad \text{hence} \quad M\frac{\partial s}{\partial a} \,=\, -\frac{\partial v}{\partial \lambda} \qquad \text{[standard GENERIC].}$$

$$\frac{d\hat{s}}{dt} \,=\, 2\,v(\hat{\lambda}(t)) \,=\, \int_{t}^{+\infty} \|R(\hat{\lambda}, d\hat{\lambda}/dt')\|_{W}^{2}\,dt' \qquad \text{[entropy production].}$$

## Nonlinear closed reduced equations

The best-fit closure applies beyond the near-equilibrium regime.

The governing equations are

$$\frac{d\hat{a}}{dt} = f(\hat{\lambda}) - \frac{\partial v}{\partial \lambda}(\hat{\lambda}) \, ,$$

$$\frac{1}{2}\left(\frac{\partial v}{\partial \lambda}\right)^{*} C(\lambda)^{-1}\left(\frac{\partial v}{\partial \lambda}\right) + f(\lambda)^{*} C(\lambda)^{-1}\left(\frac{\partial v}{\partial \lambda}\right) = \frac{1}{2}\lambda^{*} D(\lambda)\lambda \, ,$$

along with the equilibrium conditions

$$v(0) = 0 \, , \qquad \frac{\partial v}{\partial \lambda}(0) = 0 \, .$$

*Entropy production inequality* valid wherever $v(\lambda)$ is convex:

$$\frac{d\hat{s}}{dt} = \sum_{i=1}^{m} \hat{\lambda}^{i} \frac{\partial v}{\partial \lambda^{i}}(\hat{\lambda}) \geq v(\hat{\lambda}) \geq 0 \, .$$

## Application to truncated Burgers-Hopf "turbulence"

Model devised and used by Majda etal. to test stochastic mode reductions.

Project Burgers-Hopf equation onto $n$ Fourier modes, $u(x,t) = \sum_k z_k(t)\, e^{ikx}$ :

$$\frac{\partial u}{\partial t} + \frac{\partial}{\partial x}\left(\frac{1}{2}u^2\right) = 0, \qquad \text{projects to} \qquad \frac{dz_k}{dt} + ik \sum_{k_1+k_2=k} z_{k_1} z_{k_2} = 0 \quad (-n \le k \le n).$$

This Hamiltonian dynamics is ergodic and mixing for $n > 20$, with a Gaussian canonical density $\rho_{eq} \propto \exp(-\beta \sum_k |z_k|^2/2)$.

Reduce to $m \ll n$ resolved modes, using $m$ lowest modes, $A_k = z_k$.

Trial densities are also Gaussian, and consequently nonlinear closed reduced equations are accessible analytically. *Explicit closure:*

$$\frac{d\hat{a}_k}{dt} + ik \sum_{k_1+k_2=k} [1 + \omega(k_1,k_2)]\ \hat{a}_{k_1}\hat{a}_{k_2} = -\sqrt{\frac{\gamma}{\beta}}\,|k|\,\hat{a}_k\,,$$

with $\omega(k_1,k_2) = \dfrac{k_1|k_1| + k_2|k_2| - (k_1+k_2)}{(k_1+k_2)(|k_1|+|k_2|+k_1+k_2)}$ [ modified nonlinearity ],

and $\gamma > 0$ [ single closure parameter scales fractional diffusion].

Closure predictions validated by large ensemble ($10^6$ samples) simulations. Results in *arXiv 1206.6545*.

## More to say

1. Nonstationary version of best-fit closure – "plateau effect."
2. Relate weight operator $W$ to Mori-Zwanzig or Green-Kubo.

## Much more to do

1. Need good applied problems. Suggestions, please!
2. Use data or simulations to estimate parameters in $W$.
3. Criterion for choice of resolved variables $A_1, \ldots, A_m$ and $m$.
4. Develop similar theory for forced systems and nonequilibrium steady states.