Global Optimization, Generalized Canonical Ensembles, and Universal Ensemble Equivalence

Richard S. Ellis
Department of Mathematics and Statistics
University of Massachusetts Amherst

Collaborators:

Bruce Turkington, Marius Costeniuc
Department of Mathematics and Statistics
University of Massachusetts Amherst

Kyle Haven
Courant Institute of Mathematical Sciences
New York University

Hugo Touchette
School of Mathematical Sciences
Queen Mary University of London

June 23, 2004

Presentation at the Berlin Stochastics Colloquium
Research supported by a grant from the National Science Foundation (NSF-DMS-0202309)
Email: rsellis@math.umass.edu

Outline of the Talk

- Three related minimization problems
  - Constrained minimization: microcanonical ensemble
  - Unconstrained minimization with a Lagrange multiplier: canonical ensemble
  - Unconstrained minimization with a Lagrange multiplier and a penalty function: generalized canonical ensemble
- Relationships among the solutions of the three problems
  - Determined by concavity properties of the microcanonical entropy
  - From nonequivalence to universal equivalence
  - Exact comparisons of equilibrium macrostates
- Statistical mechanical ensembles
  - Large deviation methodology
- Two theorems on ensemble equivalence and nonequivalence
- Results for Curie-Weiss-Potts model
- Conclusion and applications
- Bibliography

Talk is based on the 2004 paper by M. Costeniuc, R. S. Ellis, H. Touchette, and B. Turkington: "The generalized canonical ensemble and its universal equivalence with the microcanonical ensemble."

Three Related Minimization Problems

- $\mathcal{X}$ a space
- I a nonnegative function on $\mathcal{X}$
- f a real-valued function on $\mathcal{X}$

Investigate the relationships among solutions of three minimization problems.

1. Constrained minimization for given u:

$$\mathcal{E}^u = \{\nu \in \mathcal{X} : I(\nu) \text{ is minimized for } f(\nu) = u\}$$

2. Unconstrained minimization for given β:

$$\mathcal{E}_\beta = \{\nu \in \mathcal{X} : I(\nu) + \beta f(\nu) \text{ is minimized}\}$$

   β a Lagrange multiplier

3. Unconstrained minimization for given β, γ, u:

$$\mathcal{E}(\gamma)^u_\beta = \{\nu \in \mathcal{X} : I(\nu) + \beta f(\nu) + \gamma[f(\nu) - u]^2 \text{ is minimized}\}$$

   $\gamma[f(\nu) - u]^2$ a penalty function

Rewrite $\mathcal{E}(\gamma)_\beta^u$

$$\mathcal{E}(\gamma)_\beta^u = \{\nu \in \mathcal{X} : I(\nu) + \beta f(\nu) + \gamma[f(\nu) - u]^2 \text{ is minimized}\}$$

- As $\gamma \to \infty$, $\gamma[f(\nu) - u]^2 \to \delta_u(f(\nu))$.
- This gives constraint in

$$\mathcal{E}^u = \{\nu \in \mathcal{X} : I(\nu) \text{ is minimized for } f(\nu) = u\}.$$

- Work with large $\gamma > 0$.

$$\mathcal{E}(\gamma)_\beta = \{\nu \in \mathcal{X} : I(\nu) + \beta f(\nu) + \gamma[f(\nu)]^2 \text{ is minimized}\}$$

$$\mathcal{E}^u = \{\nu \in \mathcal{X} : I(\nu) \text{ is minimized for } f(\nu) = u\}$$

$$\mathcal{E}_\beta = \mathcal{E}(0)_\beta = \{\nu \in \mathcal{X} : I(\nu) + \beta f(\nu) \text{ is minimized}\}$$

$\mathcal{E}^u$, $\mathcal{E}_\beta$, and $\mathcal{E}(\gamma)_\beta$ express the asymptotic behavior of the microcanonical ensemble, the canonical ensemble, and the generalized canonical ensemble. Derive via large deviations.

Theorem (RSE, KH, BT)

$$\mathcal{E}^u = \{\nu \in \mathcal{X} : I(\nu) \text{ is minimized for } f(\nu) = u\}$$

$$\mathcal{E}_\beta = \{\nu \in \mathcal{X} : I(\nu) + \beta f(\nu) \text{ is minimized}\}$$

$$\mathcal{E}(\gamma)_\beta = \{\nu \in \mathcal{X} : I(\nu) + \beta f(\nu) + \gamma [f(\nu)]^2 \text{ is minimized}\}$$

Only 4 relationships between $\mathcal{E}^u$ and $\mathcal{E}_\beta$.

Theorem (*JSP*, 2000)

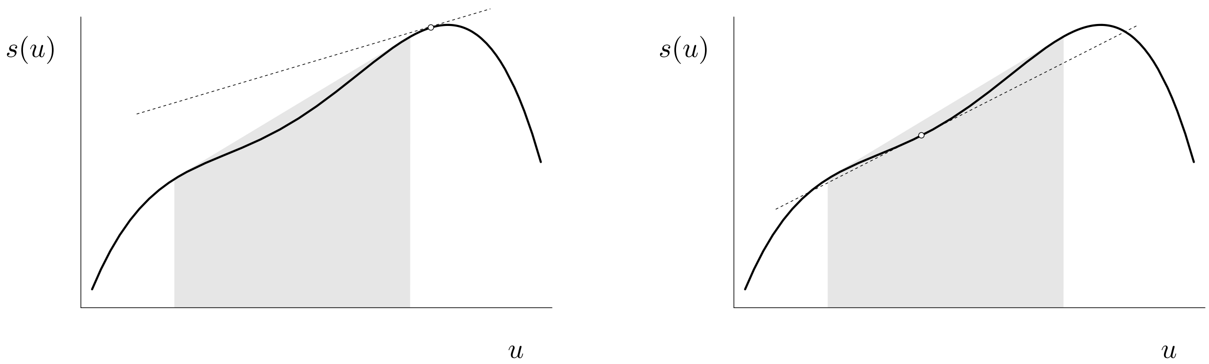
1. Fix β. Then $\exists u$ such that $\nu \in \mathcal{E}_\beta \Rightarrow \nu \in \mathcal{E}^u$.
2. Fix u. Can we always find β such that $\nu \in \mathcal{E}^u \Rightarrow \nu \in \mathcal{E}_\beta$?
   (a) **Full equivalence.** $\exists \beta$ such that $\mathcal{E}^u = \mathcal{E}_\beta$.
   (b) **Partial equivalence.** $\exists \beta$ such that $\mathcal{E}^u \subset \mathcal{E}_\beta$ but $\mathcal{E}^u \neq \mathcal{E}_\beta$.
   (c) **Nonequivalence.** $\forall \beta \;\; \mathcal{E}^u \cap \mathcal{E}_\beta = \emptyset$.

- Give criteria for 2(a), 2(b), and 2(c) in terms of concavity properties of the microcanonical entropy

$$s(u) = -\inf\{I(\nu) : \nu \in \mathcal{X}, f(\nu) = u\}.$$

  - s strictly concave at (all) $u \Rightarrow$ full equivalence for (all) u
  - s not strictly concave at $u \Rightarrow$ partial equivalence for u
  - s not concave on subset $A \Rightarrow$ nonequivalence $\forall u \in A$.
- Is there a similar theorem relating $\mathcal{E}^u$ and $\mathcal{E}(\gamma)_\beta$? Give criteria for full equivalence, partial equivalence, and nonequivalence in terms of concavity properties of what function?

Example of Microcanonical Entropy



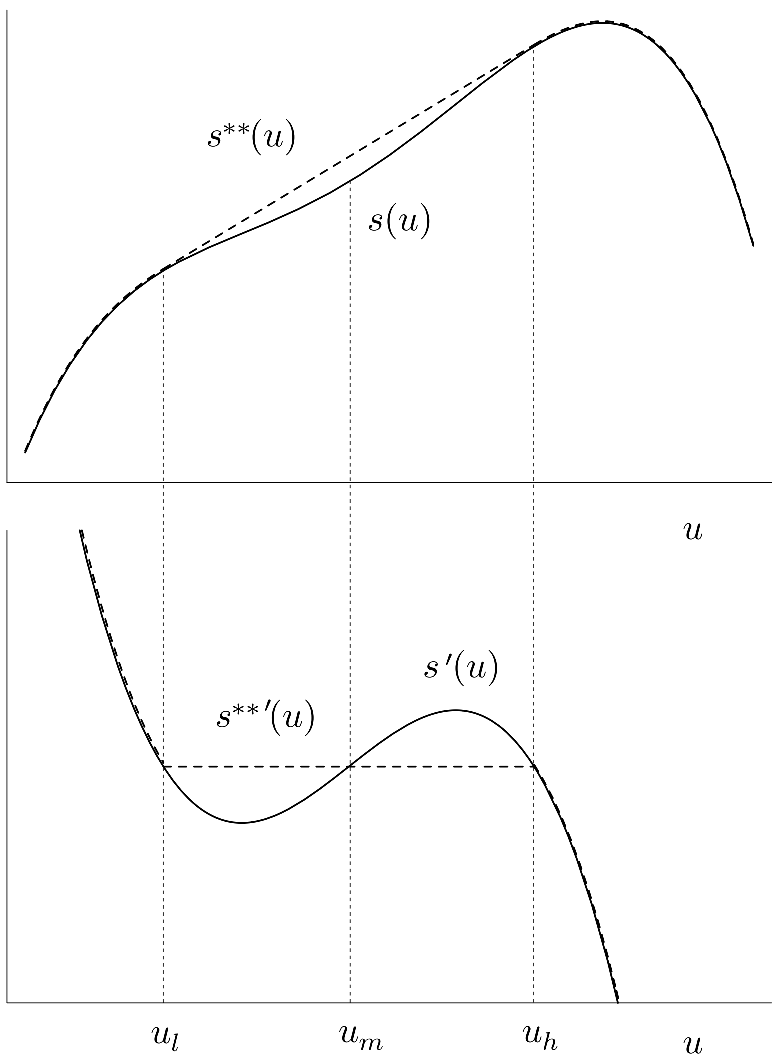
Denote by u_ℓ and u_h the projection of the shaded region onto the u axis.

- For $u < u_\ell$ and $u > u_h$, s is strictly concave (strictly supporting line), and we have full equivalence.
- For $u = u_\ell$ and $u = u_h$, s is not strictly concave (nonstrictly supporting line), and we have partial equivalence.
- For $u_\ell < u < u_h$, s is not concave (no supporting line), and we have nonequivalence.

□ Concave hull s^{**} of s

Define $s^{**} = (s^*)^*$, double-Legendre-Fenchel transform of s.
s^{**} equals the concave, u.s.c. hull of s.

- Define s concave at u if $s(u) = s^{**}(u)$.
- Define s strictly concave at u if $s(u) = s^{**}(u)$ and s^{**} strictly concave at u.
- Define s nonconcave at u if $s(u) \neq s^{**}(u)$.



From Nonequivalence to Universal Equivalence (MC, RSE, HT, BT)

$$\mathcal{E}^u = \{\nu \in \mathcal{X} : I(\nu) \text{ is minimized for } f(\nu) = u\}$$
$$\mathcal{E}_\beta = \{\nu \in \mathcal{X} : I(\nu) + \beta f(\nu) \text{ is minimized}\}$$
$$\mathcal{E}(\gamma)_\beta = \{\nu \in \mathcal{X} : I(\nu) + \beta f(\nu) + \gamma [f(\nu)]^2 \text{ is minimized}\}$$

Problem. Suppose that for all β and a subset A of u

$$\text{nonequivalence: } \mathcal{E}^u \cap \mathcal{E}_\beta = \emptyset.$$

Find $\gamma > 0$ and β such that for all u

$$\text{universal equivalence: } \mathcal{E}^u = \mathcal{E}(\gamma)_\beta.$$

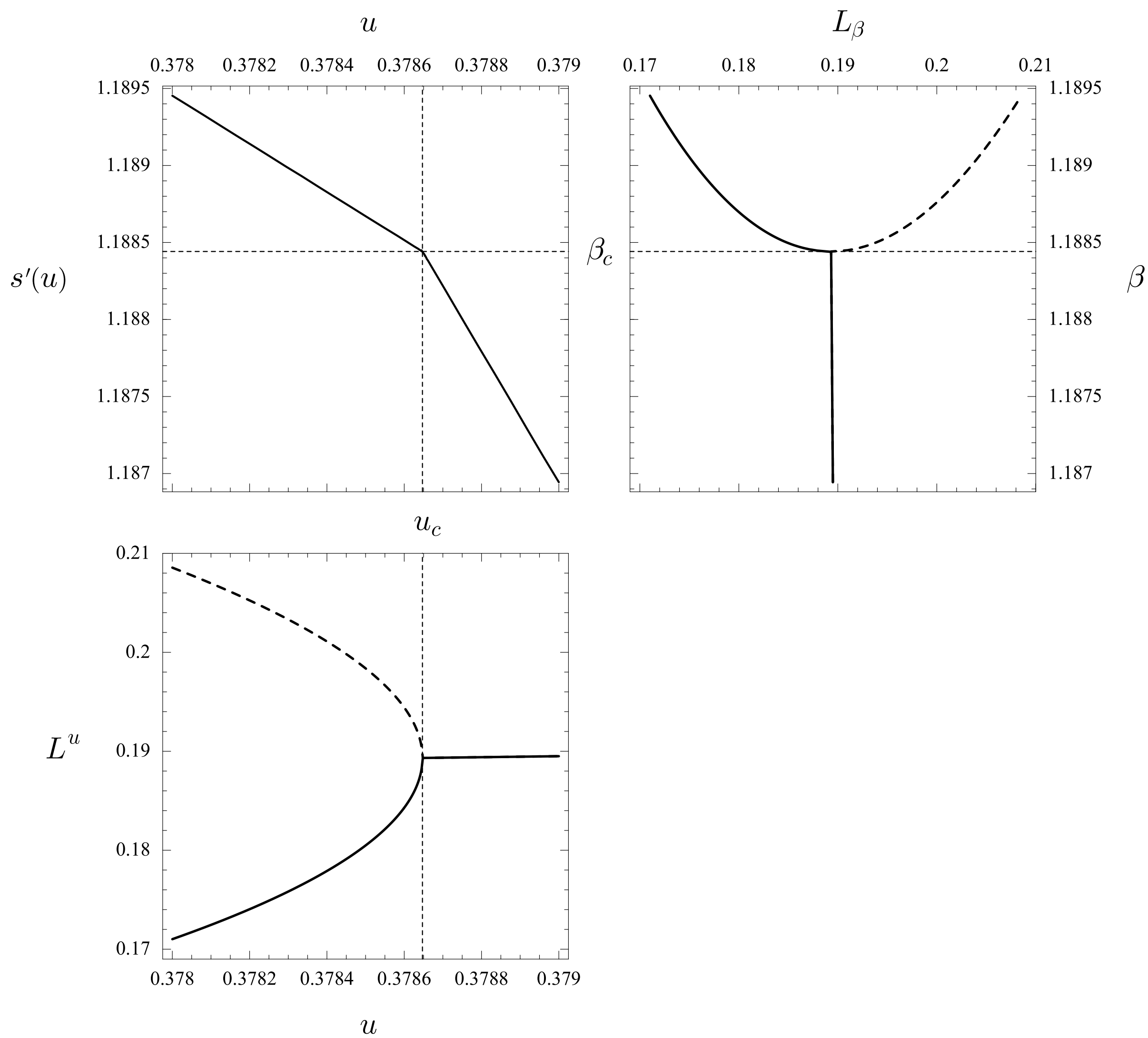
Surprise. The simplicity with which γu^2 enters the formulation.

Theorem. Define $s(u) = -\inf\{I(\nu) : \nu \in \mathcal{X}, f(\nu) = u\}$.

1. $s(u)$ strictly concave $\Rightarrow \exists \beta$ such that $\mathcal{E}^u = \mathcal{E}_\beta$.
2. $s(u) - \gamma u^2$ strictly concave $\Rightarrow \exists \beta$ such that $\mathcal{E}^u = \mathcal{E}(\gamma)_\beta$.
3. Assume: s is C^2, not strictly concave, and s'' is bounded above. Choose $\gamma > \frac{1}{2} s''(u)$ for all u. Then $s(u) - \gamma u^2$ is strictly concave for all u and $\exists \beta$ such that for all u
$$\mathcal{E}^u = \mathcal{E}(\gamma)_\beta.$$
4. Assume: s is C^2, not strictly concave, and s'' is not bounded above. Then for each u $\exists \gamma_0 \geq 0$ and $\exists \beta$ such that $\forall \gamma > \gamma_0$
$$\mathcal{E}^u = \mathcal{E}(\gamma)_\beta.$$
5. As $\gamma \uparrow$, $\mathcal{E}(\gamma)_\beta$ picks up more $\nu \in \mathcal{E}^u$.

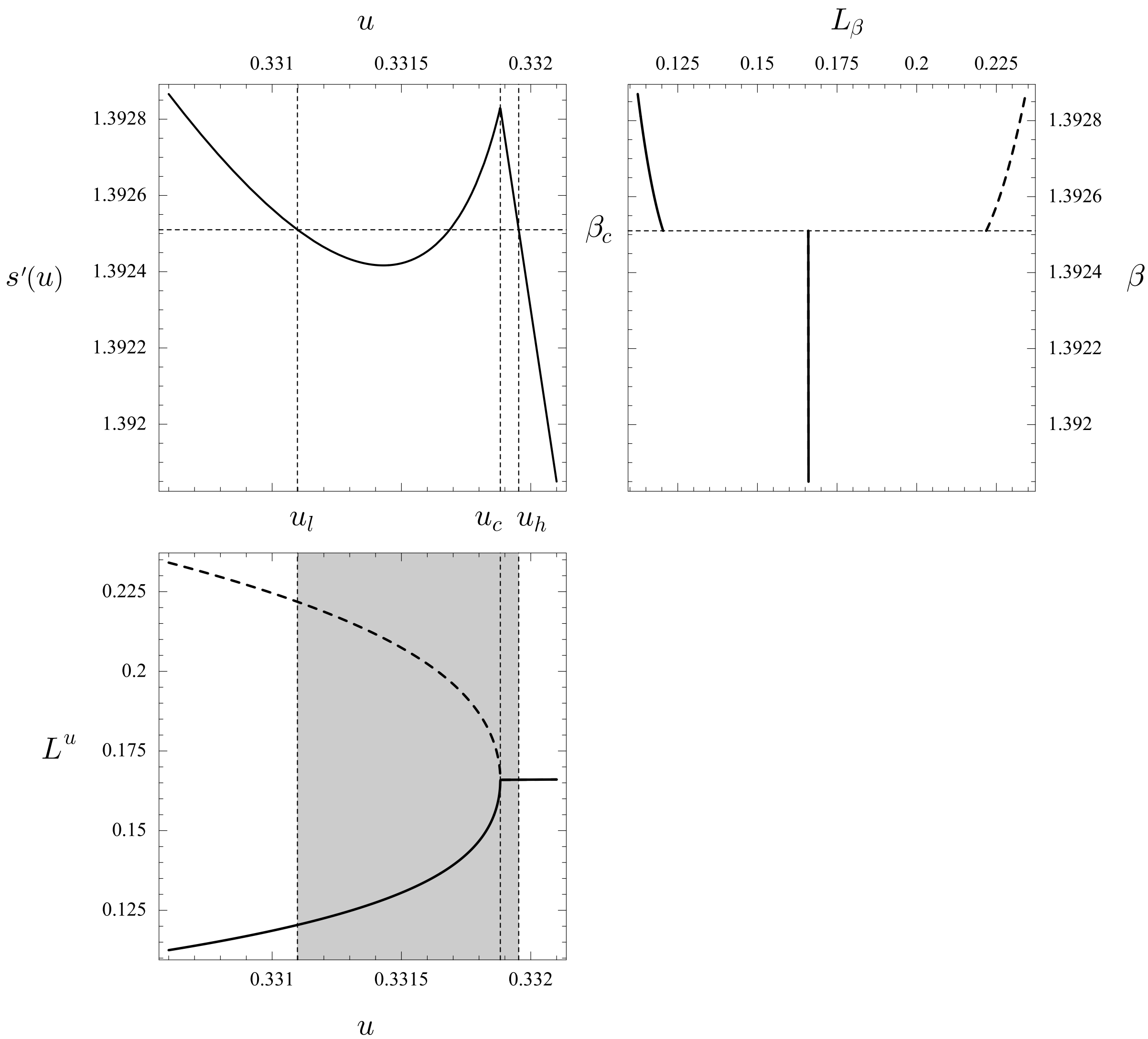
$\mathcal{E}_\beta$, $s'(u)$, and $\mathcal{E}^u$ for the BEG Model

$K = 1.111111111$ in the mean-field Blume-Emery-Griffiths (BEG) model



- s' monotonically decreasing $\Rightarrow s$ strictly concave
- Full equivalence of ensembles
- Continuous phase transitions in β and u

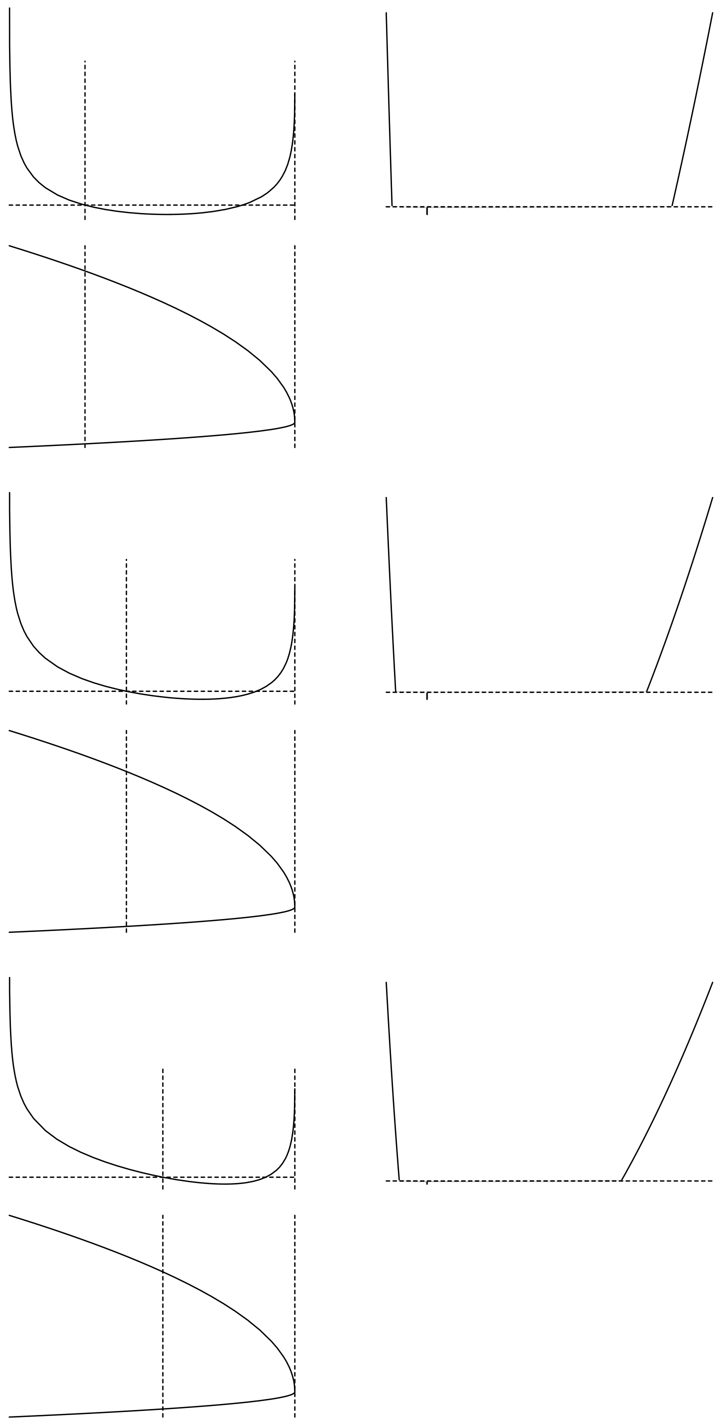
$K = 1.081651726$ in the mean-field BEG model



- s' not decreasing $\Rightarrow s$ not concave
- $s(u)$ not concave for $u_l = 0.3311 < u < u_h = 0.33195$
- Canonical ph. tr. at β_c defined by Maxwell-equal-area line
- Nonequivalence of ensembles: for $u_l < u < u_h$ L^u is not realized by L_β for any β: $\mathcal{E}^u \cap \mathcal{E}_\beta = \emptyset$ for all β.
- First-order phase transition in β versus second-order in u

□ $\mathcal{E}(\gamma)_\beta$, $(s(u) - \gamma u^2)'$, and $\mathcal{E}^u$ for the CWP Model

- $\mathcal{E}(\gamma)_\beta$ for $\gamma = 0, 1, 2$; $(s(u) - \gamma u^2)'$; and $\mathcal{E}^u$
- Full equiv. left of vertical line in $(s(u) - \gamma u^2)'$ figure; as $\gamma \uparrow$, full equiv. region increases.



Physical Background

Two classical choices of probability distributions in equilibrium statistical mechanics:

Microcanonical ensemble
$u = \text{const}$

$\uparrow \; ? \; \downarrow$

Canonical ensemble
β or $T = \text{const}$

Also generalized canonical ensemble = canonical ensemble with a penalty function

- Are the probability distributions equivalent?
- Can microcanonical equilibrium macrostates always be realized canonically?
  - Classical answer: yes.
  - Modern theory: in general no.
- Can microcanonical equilibrium macrostates always be realized generalized-canonically? In general yes.
- Equivalence of ensembles:
  - Example: perfect gas
  - General conditions: short-range interactions

Examples of Systems Having Nonequivalent Ensembles

- Gravitational systems: Lynden-Bell (1968), Thirring (1970), Gross (1997, 2001)
- Lennard-Jones gas: Borges and Tsallis (2002)
- Plasma models: Smith and O'Neil (1990)
- Spin models
  - Curie-Weiss-Potts model: Costeniuc, Ellis, and Touchette (2004)
  - Half-blocked spin model: Touchette (2003)
  - Hamiltonian mean-field model: Latora, Rapisarda, and Tsallis (2001)
  - Mean-field Blume-Emery-Griffiths model
    - Thermo level: Barré, Mukamel and Ruffo (2002)
    - Macro level: Ellis, Touchette, and Turkington (2004)
  - Mean-field XY model: Dauxois, Holdsworth and Ruffo (2000)
- Turbulence models: Robert and Sommeria (1991); Caglioti, Lions, Marchioro, and Pubvirenti (1992); Kiessling and Lebowitz (1997); Ellis, Haven, and Turkington (2002)

Statistical Mechanical Ensembles

Boltzmann (1872), Gibbs (1876, 1902)

1. ω_i, $i = 1, 2, \ldots, n$, each $\omega_i \in \Lambda$ (spins or vorticities or . . .)

2. Microstates: $\omega = (\omega_1, \omega_2, \ldots, \omega_n) \in \Lambda^n$

3. Hamiltonian or energy function: $H_n(\omega)$

4. Energy per particle: $h_n(\omega) = \frac{1}{n} H_n(\omega)$

5. Prior measure P_n; e.g., if Λ is a finite set,

$$P_n(\omega) = \frac{1}{|\Lambda|^n} \text{ for each } \omega$$

6. Macroscopic variable $L_n(\omega)$ bridging microscopic and macroscopic descriptions: $L_n(\omega)$ maps Λ^n into a space $\mathcal{X}$ ($[-1, 1]$ or $\mathcal{P}(\Lambda)$ or $L^2(\Lambda)$ or . . .).

   (a) $\mathcal{X}$ is space of macrostates.

   (b) Require bounded, continuous energy representation function f mapping $\mathcal{X}$ into $\mathbb{R}$: as $n \to \infty$

   $$h_n(\omega) = f(L_n(\omega)) + \text{o}(1) \text{ uniformly over } \omega.$$

   (c) Require basic LDP with respect to P_n:

   $$P_n\{\omega : L_n(\omega) \approx \nu\} \asymp e^{-nI(\nu)},$$

   $I(\nu)$ rate function for macrostates $\nu \in \mathcal{X}$.

□ Example: Curie-Weiss-Potts (CWP) Spin Model

Approximation to the Potts model (Wu (1982))

1. n spins $\omega_i \in \Lambda = \{1, 2, \ldots, q\}$, $q \geq 3$
2. Microstates: $\omega = (\omega_1, \omega_2, \ldots, \omega_n) \in \Lambda^n$
3. Hamiltonian or energy function:
$$H_n(\omega) = -\frac{1}{2n} \sum_{j,k=1}^n \delta(\omega_j, \omega_k)$$
4. Energy per particle:
$$h_n(\omega) = \frac{1}{n} H_n(\omega)$$
5. Prior measure:
$$P_n(\omega) = \frac{1}{q^n} \text{ for each } \omega \in \Lambda^n$$
6. Macroscopic variable (empirical vector):
$$L_n = (L_{n,1}, L_{n,2}, \ldots, L_{n,q}),$$
$$L_{n,i}(\omega) = \frac{1}{n} \sum_{j=1}^n 1_i(\omega_j) = \frac{1}{n} \cdot \#\{j : \omega_j = i\},$$
$$L_{n,i} \geq 0, \sum_{i=1}^q L_{n,i} = 1 \implies L_n(\omega) \in \mathcal{P}(\mathbb{R}^q)$$
   (a) $\mathcal{P}(\mathbb{R}^q)$ is space of macrostates.
   (b) Energy representation function:
$$h_n(\omega) = -\tfrac{1}{2}\langle L_n(\omega), L_n(\omega)\rangle = f(L_n(\omega)),$$
$$f(\nu) = -\tfrac{1}{2}\langle \nu, \nu \rangle \text{ for } \nu \in \mathcal{P}(\mathbb{R}^q)$$
   (c) Basic LDP:
$$P_n\{\omega : L_n(\omega) \approx \nu\} \asymp e^{-nR(\nu)}$$
   Sanov's Theorem gives rate function
$$R(\nu) = \sum_{i=1}^q \nu_i \log(q\nu_i),$$
   relative entropy of $\sum_{i=1}^q \nu_i \delta_i$ with respect to $\sum_{i=1}^q \frac{1}{q}\delta_i$

□ Models to which formalism has been applied

- Miller-Robert model of fluid turbulence based on the 2D Euler equations (CB, RSE, BT)
- Model of geophysical flows based on equations describing barotropic, quasi-geostrophic turbulence (RSE, KH, BT)
- Model of soliton turbulence based on a class of generalized nonlinear Schrödinger equations (RSE, RJ, PO, BT)
- Mean-field Blume-Emery-Griffiths spin model (RSE, HT, BT)
- Curie-Weiss-Potts spin model (MC, RSE, HT)
  - $\mathcal{E}^u$ and $\mathcal{E}_\beta$ known explicitly
  - Check ensemble equivalence and nonequivalence by hand

- Prior measure: $P_n(\{\omega\}) = \frac{1}{|\Lambda|^n}$ for each $\omega \in \Lambda^n$
- Assumption: $L_n(\omega)$ maps Λ^n into $\mathcal{X}$ such that
  - ◦ $h_n(\omega) = f(L_n(\omega)) + \mathrm{o}(1)$ for bdd. cont. f: $\mathcal{X} \to \mathbb{R}$
  - ◦ $\exists$ rate function $I(\nu)$ for macrostates $\nu \in \mathcal{X}$:

$$P_n\{\omega : L_n(\omega) \approx \nu\} \asymp e^{-nI(\nu)}$$

□ Microcanonical ensemble P_n^u

$$P_n^u(d\omega) = P_n(d\omega \mid h_n(\omega) \approx u)$$

- Postulate of equiprobability. If Λ is a finite set and $P_n(\{\omega\}) = \frac{1}{|\Lambda|^n}$ for each ω, then the conditional probability P_n^u is constant on energy shell $\{\omega : h_n(\omega) \approx u\}$.

- Microcanonical entropy $s(u)$:

$$\boxed{P_n\{\omega : h_n(\omega) \approx u\} \asymp e^{ns(u)}, \ s(u) = -\inf\{I(\nu) : f(\nu) = u\}}$$

$$\begin{aligned} P_n\{\omega : h_n(\omega) \approx u\} &\approx P_n\{\omega : f(L_n(\omega)) \approx u\} \\ &\approx P_n\{\omega : L_n(\omega) \in f^{-1}(u)\} \\ &\asymp \sup\left\{\exp[-nI(\nu) : \nu \in f^{-1}(u)]\right\} \\ &= \exp[-n \cdot \inf\{I(\nu) : \nu \in f^{-1}(u)\}] \\ &= \exp[-n \cdot \underbrace{\inf\{I(\nu) : f(\nu) = u\}}_{-s(u)}] \end{aligned}$$

- Asymptotic P_n^u-distribution for $L_n(\omega)$:

If $\nu \in \mathcal{X}$ satisfies $f(\nu) = u$, then

$$\begin{aligned}
&P_n^u\{\omega : L_n(\omega) \approx \nu\} \\
&= P_n\{\omega : L_n(\omega) \approx \nu, h_n(\omega) \approx u\} \cdot \frac{1}{P_n\{\omega : h_n(\omega) \approx u\}} \\
&\approx P_n\{\omega : L_n(\omega) \approx \nu, f(L_n(\omega)) \approx u\} \cdot \frac{1}{P_n\{\omega : h_n(\omega) \approx u\}} \\
&= P_n\{\omega : L_n(\omega) \approx \nu\} \cdot \frac{1}{P_n\{\omega : h_n(\omega) \approx u\}} \\
&\asymp \exp[-n(I(\nu) + s(u))].
\end{aligned}$$

If $f(\nu) \neq u$, then $P_n^u\{\omega : L_n(\omega) \approx \nu\} \asymp 0$.

- LDP for P_n^u-distribution of $L_n(\omega)$:

$$P_n^u\{\omega : L_n(\omega) \approx \nu\} \asymp e^{-nI^u(\nu)}$$

$$I^u(\nu) = \begin{cases} I(\nu) + s(u) & \text{if } f(\nu) = u \\ \infty & \text{otherwise} \end{cases}$$

- Microcanonical equilibrium macrostates defined by $I^u(\nu) = 0$:

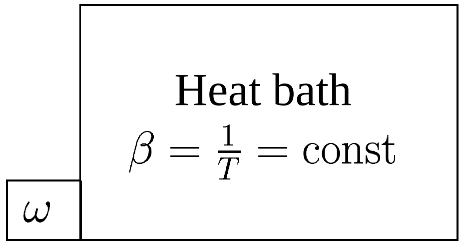
  $I^u(\nu) \geq 0$ for all ν

  $I^u(\nu) > 0 \Longrightarrow P_n^u\{\omega : L_n(\omega) \approx \nu\} \to 0$ exponentially fast

  $I^u(\nu) = 0 \Longleftrightarrow I(\nu) = -s(u) = \inf\{I(\mu) : f(\mu) = u\}$.

$$\mathcal{E}^u = \{\nu \in \mathcal{X} : I(\nu) \text{ is minimized for } f(\nu) = u\}$$

□ Canonical ensemble $P_{n,\beta}$



- Gibbs probability distribution:

$$P_{n,\beta}(d\omega) = \frac{1}{Z_n(\beta)}\, e^{-\beta n h_n(\omega)}\, P_n(d\omega),$$

$$Z_n(\beta) = \int_{\Lambda^n} e^{-\beta n h_n}\, dP_n \asymp e^{-n\varphi(\beta)}$$

  $\varphi(\beta)$ is the canonical free energy per particle.

- LDP for $P_{n,\beta}$-distribution of $L_n(\omega)$:

$$\boxed{\begin{array}{c} P_{n,\beta}\{\omega : L_n(\omega) \approx \nu\} \asymp e^{-nI_\beta(\nu)} \\ I_\beta(\nu) = I(\nu) + \beta f(\nu) - \varphi(\beta) \end{array}}$$

- Canonical equilibrium macrostates defined by $I_\beta(\nu) = 0$:

  $I_\beta(\nu) \geq 0$ for all ν

  $I_\beta(\nu) > 0 \Rightarrow P_{n,\beta}\{\omega : L_n(\omega) \approx \nu\} \to 0$ exponentially fast

$$\boxed{\mathcal{E}_\beta = \{\nu \in \mathcal{X} : I(\nu) + \beta f(\nu) \text{ is minimized}\}}$$

- Microcanonical equilibrium macrostates defined by $I^u(\nu) = 0$:

$$\boxed{\mathcal{E}^u = \{\nu \in \mathcal{X} : I(\nu) \text{ is minimized for } f(\nu) = u\}}$$

□ Generalized Canonical Ensemble $P_{n,\beta,\gamma}$

- Challa and Hetherington (1988), Johal et. al. (2003), Kiessling and Lebowitz (1997)

- Generalized Gibbs probability distribution:

$$P_{n,\beta,\gamma}(d\omega) = \frac{1}{Z_n(\beta,\gamma)} e^{-n\beta h_n(\omega) - n\gamma[h_n(\omega)]^2} P_n(d\omega),$$

$$Z_n(\beta) = \int_{\Lambda^n} e^{-n\beta h_n - n\gamma[h_n]^2} dP_n \asymp e^{-n\varphi(\beta,\gamma)}$$

  $\varphi(\beta,\gamma)$ is the generalized canonical free energy per particle.

- LDP for $P_{n,\beta,\gamma}$-distribution of $L_n(\omega)$:

$$\boxed{\begin{array}{c} P_{n,\beta,\gamma}\{\omega : L_n(\omega) \approx \nu\} \asymp e^{-nI_{\beta,\gamma}(\nu)} \\ I_{\beta,\gamma}(\nu) = I(\nu) + \beta f(\nu) + \gamma[f(\nu)]^2 - \varphi(\beta,\gamma) \end{array}}$$

- Generalized canonical equilibrium macrostates defined by $I_{\beta,\gamma}(\nu) = 0$:
  $I_{\beta,\gamma}(\nu) \geq 0$ for all ν
  $I_{\beta,\gamma}(\nu) > 0 \Rightarrow P_{n,\beta,\gamma}\{\omega : L_n(\omega) \approx \nu\} \to 0$ exponentially fast

$$\boxed{\mathcal{E}(\gamma)_\beta = \{\nu \in \mathcal{X} : I(\nu) + \beta f(\nu) + \gamma[f(\nu)]^2 \text{ is minimized}\}}$$

- Canonical equilibrium macrostates $\mathcal{E}_\beta = \mathcal{E}(0)_\beta$

- Microcanonical equilibrium macrostates:

$$\boxed{\mathcal{E}^u = \{\nu \in \mathcal{X} : I(\nu) \text{ is minimized for } f(\nu) = u\}}$$

Theorem 1: Microcanonical Ensemble More Basic Than Canonical Ensemble

RSE, Kyle Haven, Bruce Turkington (*JSP*, 2000)

$$\mathcal{E}^u = \{\nu \in \mathcal{X} : I(\nu) \text{ is minimized for } f(\nu) = u\}$$

$$\mathcal{E}_\beta = \{\nu \in \mathcal{X} : I(\nu) + \beta f(\nu) \text{ is minimized}\}$$

- **Canonical is always realized microcanonically:**

$$\mathcal{E}_\beta = \bigcup_{u \in f(\mathcal{E}_\beta)} \mathcal{E}^u$$

- **Full equivalence of ensembles:**

  $s(u) = -\inf\{I(\nu) : f(\nu) = u\}$ strictly concave at u

  $\Rightarrow \mathcal{E}^u = \mathcal{E}_\beta$ for unique β

  $\Rightarrow$ canonical $\equiv$ microcanonical

- **Partial equivalence of ensembles:**

  s not strictly concave at $u \;\Rightarrow\; \mathcal{E}^u \subset \mathcal{E}_\beta$ for unique β but $\mathcal{E}^u \neq \mathcal{E}_\beta$

- **Nonequivalence of ensembles:**

  s not concave at u

  $\Rightarrow \mathcal{E}^u \cap \mathcal{E}_\beta = \emptyset$ for all β

  $\Rightarrow$ microcanonical not realized canonically

Theorem 2: Universal Equivalence of Ensembles Is Possible

Marius Costeniuc, RSE, Hugo Touchette, Bruce Turkington (2004)

$$\mathcal{E}^u = \{\nu \in \mathcal{X} : I(\nu) \text{ is minimized for } f(\nu) = u\}$$

$$\mathcal{E}(\gamma)_\beta = \{\nu \in \mathcal{X} : I(\nu) + \beta f(\nu) + \gamma [f(\nu)]^2 \text{ is minimized}\}$$

- **Generalized canonical always realized microcanonically:**

$$\mathcal{E}(\gamma)_\beta = \bigcup_{u \in f(\mathcal{E}(\gamma)_\beta)} \mathcal{E}^u$$

- **Full equivalence of ensembles:**
  $s(u) - \gamma u^2$ strictly concave at u
  $\Rightarrow \mathcal{E}^u = \mathcal{E}(\gamma)_\beta$ for unique β
  $\Rightarrow$ generalized canonical $\equiv$ microcanonical
  - **Universal equivalence of ensembles**: choose γ so that $s(u) - \gamma u^2$ is strictly concave for all u

- **Partial equivalence of ensembles:**
  $s(u) - \gamma u^2$ not strictly concave at u $\Rightarrow$ $\mathcal{E}^u \subset \mathcal{E}(\gamma)_\beta$ for unique β but $\mathcal{E}^u \neq \mathcal{E}(\gamma)_\beta$

- **Nonequivalence of ensembles:**
  $s(u) - \gamma u^2$ not concave at u
  $\Rightarrow \mathcal{E}^u \cap \mathcal{E}(\gamma)_\beta = \emptyset$ for all β
  $\Rightarrow$ micro not realized generalized-canonically

Proof of Theorem 2 from Theorem 1

- Define $P_n^u(d\omega) = P_n(d\omega \mid h_n(\omega) \approx u)$.
- LDP for P_n: $P_n\{\omega : L_n(\omega) \approx \nu\} \asymp e^{-nI(\nu)}$
- Define $s(u) = -\inf\{I(\nu) : \nu \in \mathcal{X}, f(\nu) = u\}$.

Theorem 1. Relate $\mathcal{E}^u$ and $\mathcal{E}_\beta$ via s.

- Introduce new prior measures

$$P_{n,\gamma}(d\omega) = \text{const} \cdot e^{-n\gamma[h_n(\omega)]^2}\, P_n(d\omega).$$

- Rewrite the generalized canonical ensemble:

$$\begin{aligned} P_{n,\beta,\gamma}(d\omega) &= \text{const} \cdot e^{-n\beta h_n(\omega) - n\gamma[h_n(\omega)]^2}\, P_n(d\omega) \\ &= \text{const} \cdot e^{-n\beta h_n(\omega)}\, P_{n,\gamma}(d\omega) \\ &= P_{n,\beta}(d\omega) \text{ with } P_n \text{ replaced by } P_{n,\gamma}. \end{aligned}$$

- Verify $P_n^u(d\omega) \asymp P_{n,\gamma}^u(d\omega)$ since $h_n(\omega) \approx u$.
- Recall $h_n(\omega) = f(L_n(\omega)) + \text{o}(1)$ uniformly over ω.
- LDP for $P_{n,\gamma}$: $P_{n,\gamma}\{\omega : L_n(\omega) \approx \nu\} \asymp e^{-n\left(I(\nu) + \gamma[f(\nu)]^2 - \text{const}\right)}$

Theorem 2. Relate $\mathcal{E}^u$ and $\mathcal{E}(\gamma)_\beta$ via s_γ, where

$$\begin{aligned} s_\gamma(u) &= -\inf\{I(\nu) + \gamma[f(\nu)]^2 : f(\nu) = u\} \\ &= -\inf\{I(\nu) : f(\nu) = u\} - \gamma u^2 \\ &= s(u) - \gamma u^2 \end{aligned}$$

Results for the CWP model

- Prior measure:

$$P_n(\omega) = \frac{1}{q^n} \text{ for each } \omega \in \{1, 2, \ldots, q\}^n$$

- Energy per particle:

$$h_n(\omega) = -\frac{1}{2n^2} \sum_{j,k=1}^n \delta(\omega_j, \omega_k)$$

- Macroscopic variable (empirical vector):

$$L_n = (L_{n,1}, L_{n,2}, \ldots, L_{n,q}),$$

$$L_{n,i}(\omega) = \frac{1}{n} \sum_{j=1}^n 1_i(\omega_j) = \frac{1}{n} \cdot \#\{j : \omega_j = i\}$$

  - Energy representation function:

$$h_n(\omega) = f(L_n(\omega)),\ f(\nu) = -\tfrac{1}{2}\langle \nu, \nu \rangle \text{ for } \nu \in \mathcal{P}(\mathbb{R}^q)$$

  - Basic LDP:

$$P_n\{\omega : L_n(\omega) \approx \nu\} \asymp e^{-nR(\nu)}$$

    Sanov's Theorem gives rate function

$$R(\nu) = \sum_{i=1}^q \nu_i \log(q\nu_i),$$

    relative entropy of $\sum_{i=1}^q \nu_i \delta_i$ w.r.t. $\sum_{i=1}^q \frac{1}{q}\delta_i$

- Equilibrium macrostates:

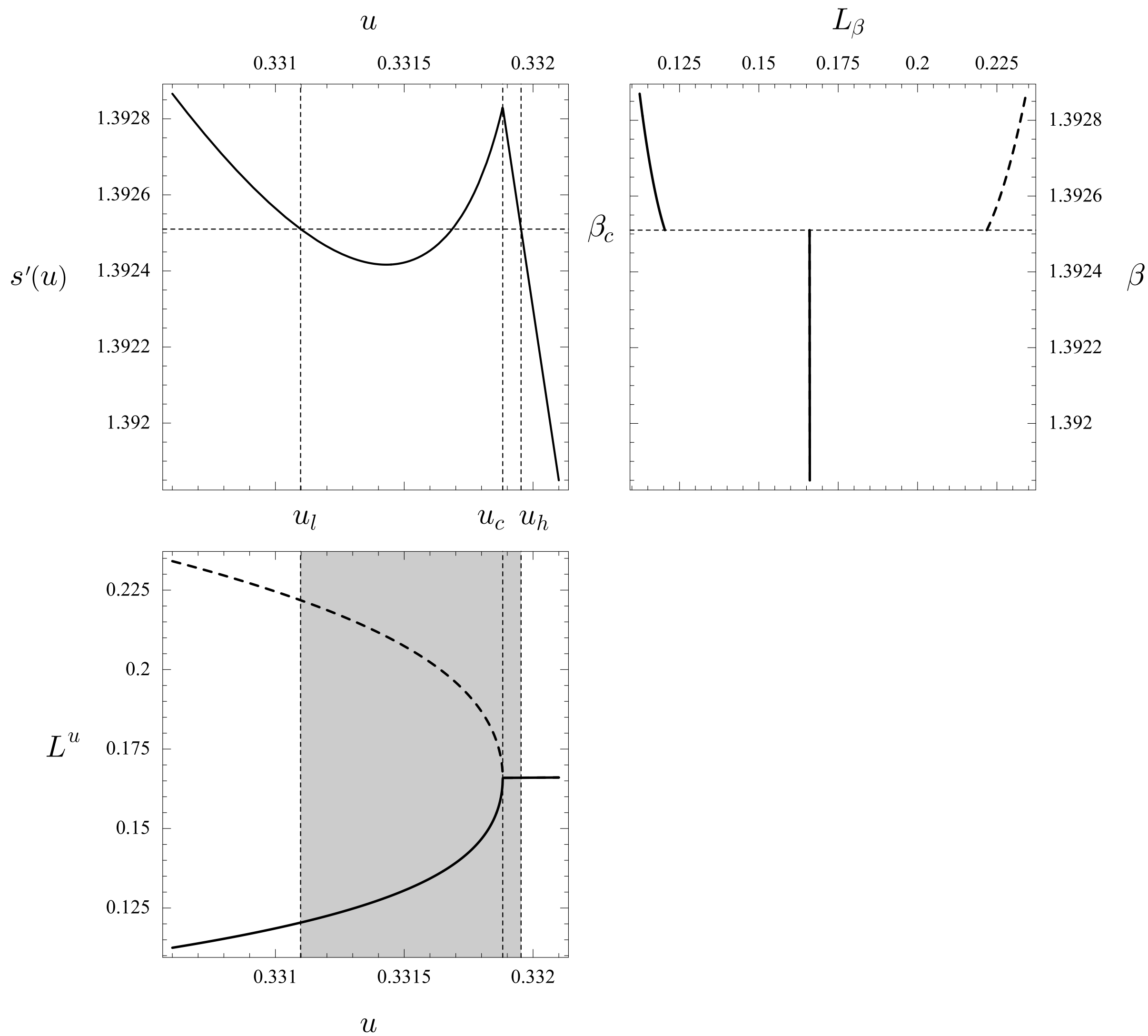
$$L^u = \arg\inf_\nu \{R(\nu) : f(\nu) = u\}$$

$$L_\beta = \arg\inf_\nu \{R(\nu) + \beta f(\nu)\}$$

$$L_{\beta,\gamma} = \arg\inf_\nu \{R(\nu) + \beta f(\nu) + \gamma[f(\nu)]^2\}$$

□ $\mathcal{E}_\beta$, s' , and $\mathcal{E}^u$ for the BEG Model

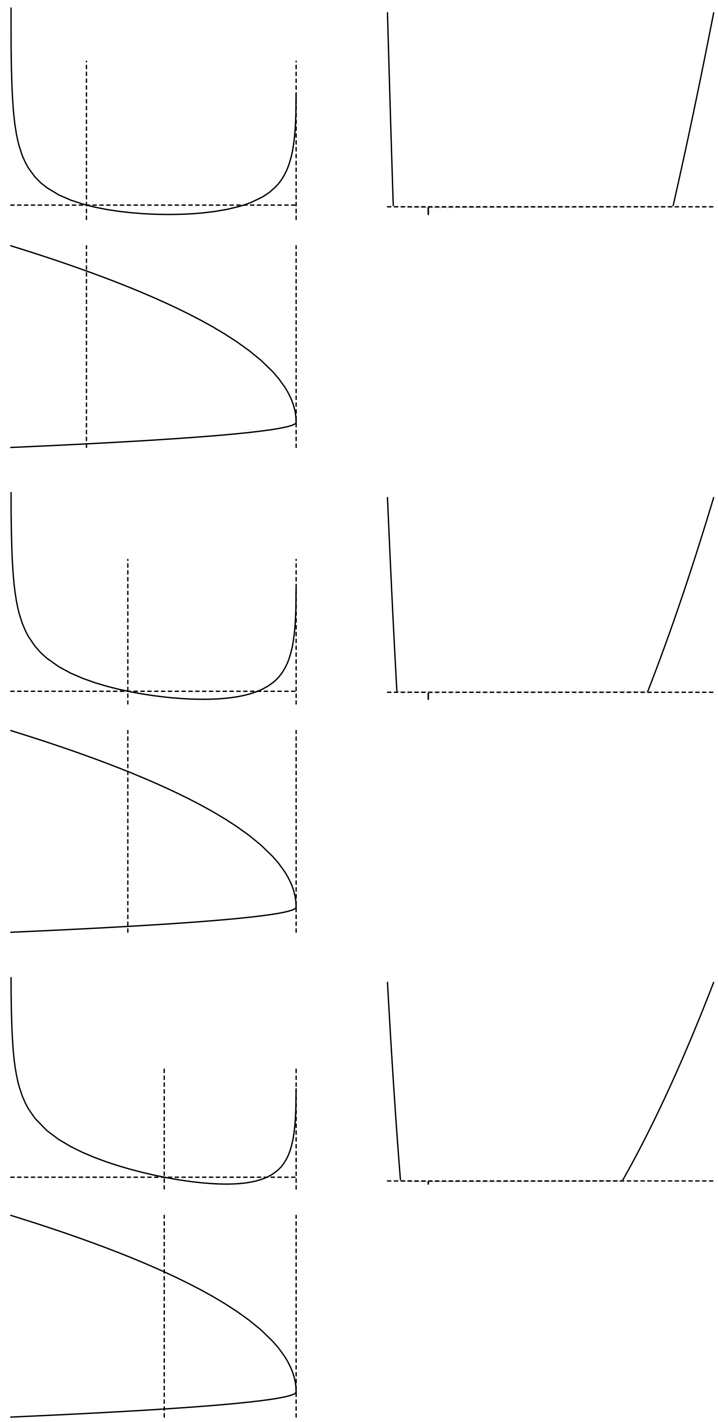
$K = 1.081651726$ in the mean-field BEG model



- s' not decreasing $\Rightarrow s$ not concave
- $s(u)$ not concave for $u_l = 0.3311 < u < u_h = 0.33195$
- Canonical ph. tr. at β_c defined by Maxwell-equal-area line
- Nonequivalence of ensembles: for $u_l < u < u_h$ L^u is not realized by L_β for any β: $\mathcal{E}^u \cap \mathcal{E}_\beta = \emptyset$ for all β.
- First-order phase transition in β versus second-order in u

□ $\mathcal{E}(\gamma)_\beta$, $(s(u) - \gamma u^2)'$, and $\mathcal{E}^u$ for the CWP Model

- $\mathcal{E}(\gamma)_\beta$ for $\gamma = 0, 1, 2$; $(s(u) - \gamma u^2)'$; and $\mathcal{E}^u$ for the CWP model with $q = 8$
- Full equiv. left of vertical line in $s'(u)$ figure; as $\gamma \uparrow$, full equiv. region increases.



Conclusions and Applications

- Canonical equilibrium macrostates are always realized microcanonically. If s is not strictly concave at u, but $s(u) - \gamma u^2$ is, then the microcanonical equilibrium macrostates not realized canonically are realized generalized-canonically.

- Universal equivalence of ensembles can be achieved via the generalized canonical ensemble.

- In classical models such as the Ising spin model, I is affine or convex and f is affine. Thus
$$s(u) = -\inf\{I(\nu) : \nu \in \mathcal{X}, f(\nu) = u\}$$
is concave. Full or partial equivalence of ensembles holds.

- Models of turbulence show additional features.

  - All our results generalize to multidimensional cases in which s is a function of energy u, enstrophy, circulation, and other quantities conserved by the underlying p.d.e.

  - The most spectacular application of statistical theories of turbulence is to the prediction of large scale, coherent structures of the atmosphere of Jupiter including the Great Red Spot.

  - The microcanonical equilibrium macrostates not realized canonically often include macrostates of physical interest; e.g., the Great Red Spot of Jupiter.

Bibliography

Reference for This Talk

- M. Costeniuc, R. S. Ellis, H. Touchette, and B. Turkington, “The generalized canonical ensemble and its universal equivalence with the microcanonical ensemble,” submitted for publication (2004).

Theory of Large Deviations

- A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications*, New York: Springer-Verlag, 1998.
- R. S. Ellis, *Entropy, Large Deviations, and Statistical Mechanics*, New York: Springer-Verlag, 1985.
- R. S. Ellis, “The theory of large deviations: from Boltzmann’s 1877 calculation to equilibrium macrostates in 2D turbulence,” *Physica D* **133**: 106–136 (1999).
- O. E. Lanford, “Entropy and equilibrium states in classical statistical mechanics,” *Statistical Mechanics and Mathematical Problems* 1–113, ed. A. Lenard, Lecture Notes in Physics, Volume 20, Berlin: Springer-Verlag, 1973.

□ Blume-Emery-Griffiths Model

- M. Blume, V. J. Emery, R. B. Griffiths, “Ising model for the λ transition and phase separation in He^3-He^4 mixtures,” *Phys. Rev. A* **4**: 1071–1077 (1971).

- R. S. Ellis, P. Otto, and H. Touchette. ”Analysis of phase transtions in the mean-field Blume-Emery-Griffiths model,” submitted for publication (2004).

□ Curie-Weiss-Potts Model

- M. Costeniuc, R. S. Ellis, and H. Touchette, “Complete analysis of equivalence and nonequivalence of ensembles for the Curie-Weiss-Potts model,” in preparation (2004).

- R. S. Ellis and K. Wang, ”Limit theorems for the empirical cector of the Curie-Weiss-Potts model,” *Stoch. Proc. Appl.* **35**: 59-79 (1990).

□ Generalized Canonical Ensembles

- M. S. S. Challa and J. H. Hetherington, “Gaussian ensemble: an alternate Monte-Carlo scheme,” *Phys. Rev. A* **38**: 6324–6337 (1988).

- M. S. S. Challa and J. H. Hetherington, "Gaussian ensemble as an interpolating ensemble," *Phys. Rev. Lett.* **60**: 77–80 (1988).

  R. S. Johal, A. Planes, and E. Vives, "Statistical mechanics in the extended Gaussian ensemble," *Phys. Rev. E* **68**: 056113 (2003).

□ Nonequivalence of Ensembles

- J. Barré, D. Mukamel, S. Ruffo, "Inequivalence of ensembles in a system with long-range interactions," *Phys. Rev. Lett.* **87**: 030601/1–4 (2001).

- R. S. Ellis, K. Haven, and B. Turkington, "Analysis of statistical equilibrium models of geostrophic turbulence," *J. Appl. Math. Stoch. Anal.* **15**: 341–361 (2002).

- R. S. Ellis, K. Haven, and B. Turkington, "Large deviation principles and complete equivalence and nonequivalence results for pure and mixed ensembles, *J. Stat. Phys.* **101**: 999–1064, 2000.

- R. S. Ellis, K. Haven, and B. Turkington, "Nonequivalent statistical equilibrium ensembles and refined stability theorems for most probable flows," *Nonlinearity* **15**: 239–255 (2002).

- R. S. Ellis, H. Touchette, and B. Turkington, “Thermodynamic versus statistical nonequivalence of ensembles for the mean-field Blume-Emery-Griffiths model, *Physica A* **335**: 518-538 (2004).

- G. L. Eyink and H. Spohn, “Negative-temperature states and large-scale, long-lived vortices in two-dimensional turbulence,” *J. Stat. Phys.* **70**: 833–886 (1993).

- M. K.-H. Kiessling and J. L. Lebowitz, “The micro-canonical point vortex ensemble: beyond equivalence,” *Lett. Math. Phys.* **42**: 43–56 (1997).

- F. Leyvraz and S. Ruffo, “Ensemble inequivalence in systems with long-range interactions,” *J. Math. Phys. A: Math. Gen.* **35**: 285–294 (2002).

- J. T. Lewis, C.-E. Pfister, and W. G. Sullivan, “Entropy, concentration of probability and conditional limit theorems, *Markov Proc. Related Fields* **1**: 319–386 (1995).

- J. T. Lewis, C.-E. Pfister, and W. G. Sullivan, “The equivalence of ensembles for lattice systems: some examples and a counterexample,” *J. Stat. Phys.* **77**: 397–419 (1994).

- W. Thirring, *A Course in Mathematical Physics 4: Quantum Mechanics of Large Systems*, trans. E. M. Harrell, New York: Springer-Verlag, 1983.