Physics 7240: Advanced Statistical Mechanics

Lecture 1: Introduction and Overview

Leo Radzihovsky[*]

Department of Physics, University of Colorado, Boulder, CO 80309

(Dated: August 27, 2024)

---

[*] Electronic address: `radzihov@colorado.edu`

I. INTRODUCTION

A. Background

Statistical mechanics is an extremely powerful set of ideas and tools with a broad range of applications throughout physics[1–7]. The subject was founded in mid 1800s by scientific giants like Gibbs, Boltzmann, Maxwell,. . . Formulated in terms of system's microscopic degrees of freedom, that are governed by basic principles of Newtonian and quantum mechanics, electricity and magnetism, relativity, etc., it allows one to derive macroscopic behavior of extensive systems. These include gases, liquids, crystals, magnets, superfluids and superconductors, and many other states of matter. It thus provides a microscopic theory of thermodynamics for natural phenomena.

Recall that much older phenomenological theory of thermodynamics postulates four laws[1, 2]:

- 0th law: temperature exists, $T_A = T_B$, $T_B = T_C \quad \rightarrow \; T_A = T_C$
- 1st law: energy conservation, $\Delta U = W + Q$
- 2nd law: entropy of a closed system increases, $TdS \geq 0$
- 3rd law: vanishing entropy at zero temperature, $S(T = 0) = 0$,

that led to much of early understanding of gases, liquids, and solids.

A much more powerful microscopic probabilistic theory is statistical mechanics that posits a single basic underlying postulate of *ergodicity*, i.e., that in equilibrium time averages of physical observables can be replaced by an ensemble average, with all microstates of a closed, isolated system are accessible with equal probability, from which all of thermodynamics emerges via detailed calculations.

B. Fundamentals

There are a number of statistical ensembles that can be used to calculate system's properties: microcanonical (for closed system with fixed energy E, number of constituents N, volume of the system V,. . . , and all other extensive variables) and canonical (for open systems, with fixed temperature T, etc) determined by which macroscopic observables are kept fixed in the calculation. In the thermodynamic limit of $N \gg 1$, satisfied by (most) macroscopic systems, central limit theorem dictates that fluctuations around the average are small (scaling as $\sqrt{N} \ll N$), thus the choice of an ensemble is a matter of convenience, leading to equivalent average predictions for a bulk system.

For a closed classical system, described by a microcanonical ensemble, the central quantity is "multiplicity" (density of states)[1, 2],

$$\Omega(E) = \sum_{\{q_n\}} \delta(E - H(\{q_n\})), \tag{1}$$

a total number of microstates at energy E, whose logarithm gives the entropy of the system

$$S(E) = k_B \ln \Omega(E), \tag{2}$$

thereby connecting it to thermodynamics, with $1/T \equiv \beta = \partial S/\partial E$; we will typically work in units in which the Boltzmann constant $k_B = 1$, i.e., temperature is measured in units of energy. Above, $\{q_n\}$ denotes a generic set of *microstates* of the system characterized by a classical Hamiltonian, H. An equilibrium long-time value of a physical observable $O(\{q_n\})$ is given by ensemble average over $\{q_n\}$ with a microcanonical probability distribution $P_E = 1/\Omega(E)$. The thermodynamic entropy can be written in a more general form, when expressed in therms of P_E, namely,

$$\begin{aligned} S(E) &= -k_B \sum_{\{q_n\}} P_E \ln P_E, & (3)\\ &= -k_B \mathrm{Tr}\left[P_E \ln P_E\right] \equiv -k_B \langle \ln P_E \rangle_{P_E}. & (4) \end{aligned}$$

Amazingly, such microcanonical description is believed to extend to a generic closed quantum many-body system in a *single* eigenstate, a postulate known as the Eigenstate Thermalization Hypothesis (ETH), put forth in early 90s by Mark Srednicki[8] and independently by Josh Deutsch[9]. As consequence ETH posits that a quantum mechanical average of an observable $\hat{O}$ for a system in a generic quantum many-body state $|\Psi(t)\rangle$, under unitary evolution, in a long time limit "relaxes" to equilibrium, described by the microcanonical ensemble at energy $E = \langle \Psi | \hat{H} | \Psi \rangle$,

$$\frac{1}{t} \int_0^t dt' \langle \Psi(t') | \hat{O} | \Psi(t') \rangle_{t \to \infty} = \langle \hat{O} \rangle_{P_E}. \tag{5}$$

To see this we can express $|\Psi(t)\rangle$ in terms of many-body eigenstates $|n\rangle$ and note that for a generic state, contributions off-diagonal in n, n' will dephase after some time (set by minimum gaps) for a generic spectrum with no large gaps or degeneracies. This is the "relaxation" to a stationary time-independent state. The diagonal terms appear to depend on the initial state $|\Psi(0)\rangle$ through the expansion coefficients, c_n, according to $\sum_n |c_n|^2 O_{nn}$. ETH proposes that for generic state with a finite excitation energy density (not a ground state) $\mathcal{E} = E/V$, the diagonal matrix element of a local operator is independent of n, depending only on $\mathcal{E}$. Thus O_{nn} can be taken out of the sum, which then reduces to 1, with the result reducing to the microcanonical ensemble. Namely, in a long time limit the expectation value relaxes to an ensemble average of O_{nn} over eigenstates around energy E in a range of fixed energy density $\mathcal{E}$.

It appears somewhat paradoxical that a closed quantum system evolving under pure unitary, norm-preserving evolution, nevertheless exhibits "relaxation", with accompanying growth of sub-system's entanglement entropy (see below). The resolution is that local observables gets "dressed" and thereby spread into operators that are nonlocal in coordinate and in operator space. Thus, while there is no loss of information or many-body norm, under growth of entanglement the information delocalizes and thereby becomes unobservable through local low-order operators measured in typical experiments. Thus from the perspective of accessible local observables, the system relaxes to equilibrium value of these observables.

There are notable exceptional closed quantum systems (typically in the presence of local random heterogeneities, i.e., "quenched disorder"), that are not ergotic. They do not equilibrate and thus violate ETH, exhibiting the so-called "many-body localization" (MBL)[10], that is currently vigorously studied[11]. More generally, "integrable" systems have infinite number of conserved quantities (that thereby cannot relax) violate ETH, with MBL being only one example where such integrability is emergent.

We will assume ETH throughout and will typically use the canonical ensemble (in which calculations are usually simpler), with a Boltzmann probability weight for a state $\{q_n\}$, given by

$$P_T(\{q_n\}) = \frac{e^{-\beta H(\{q_n\})}}{Z}, \tag{6}$$

where all of thermodynamics is defined by a classical partition function,

$$Z = \sum_{\{q_n\}} e^{-\beta H(\{q_n\})} = e^{-\beta F}, \tag{7}$$

and the corresponding free energy F. For quantum system, $\hat{H}$ and $\{\hat{q}_n\}$ are generically non-commuting operators and the partition function is a Hilbert-space trace of the corresponding (unnormalized) Boltzmann density matrix operator,

$$\begin{aligned} Z &= \mathrm{Tr}\left[e^{-\beta H(\{\hat{q}_n\})}\right], & (8)\\ &= \sum_{\psi_n} e^{-\beta E_n}, & (9) \end{aligned}$$

where the $\hat{H}$ eigenstates, $|\psi_n\rangle$ play the role of the classical microstates $\{q_n\}$. The quantum-thermodynamic average of observables is computed via

$$\begin{aligned} \langle\hat{O}\rangle &= Z^{-1}\mathrm{Tr}\left[\hat{O}e^{-\beta H(\{\hat{q}_n\})}\right], & (10)\\ &= \mathrm{Tr}\left[\hat{O}\hat{\rho}\right], & (11)\\ &= Z^{-1}\sum_{\psi_n}\langle\psi_n|\hat{O}|\psi_n\rangle e^{-\beta E_n}, & (12) \end{aligned}$$

with the first line valid for a classical version as well, as the trace “Tr” then reduces to a sum over commuting degrees of freedom (diagonal operators), $\{\hat{q}_n\}$ labeling the microstates.

Finally, we note that a closed quantum mechanical system in a (normalized) state $|\Psi\rangle$ is characterized by a pure density matrix $\hat{\rho} = |\Psi\rangle\langle\Psi|$. By splitting the full closed system A+B into subregions A and B and tracing over degrees of freedom in region B (spanned by basis $|\psi_B\rangle$), we can then define a “reduced density matrix” for subsystem A,

$$\hat{\rho}_A = \mathrm{Tr}_B\left[\hat{\rho}\right] \equiv \sum_{\psi_B}\langle\psi_B|\hat{\rho}|\psi_B\rangle, \tag{13}$$

which generically will be a density matrix for a mixed state, i.e., $\mathrm{Tr}\left[\rho_A^2\right] < 1$. Correspondingly, we can define the entanglement entropy, S_E

$$S_E = -k_B \mathrm{Tr}\left[\hat{\rho}_A \ln \hat{\rho}_A\right], \tag{14}$$

that is a quantum-mechanical analog of the thermodynamic entropy, that measures the extent of entanglement in a state $|\Psi\rangle$ between regions A and B. For example, for a fully entangled two-spin singlet state $|\Psi\rangle = \frac{1}{2^{1/2}}\left[|\uparrow\rangle_A|\downarrow\rangle_B - |\downarrow\rangle_A|\uparrow\rangle_B\right]$, the entanglement entropy is maximum, $S_E = k_B \ln 2$, quantifying the fact that in such a state there is no local information about a single spin. In contrast, a fully unentangled product state $|\Psi\rangle = |\uparrow\rangle_A |\downarrow\rangle_B$ has a vanishing entanglement entropy, $S_E = 0$. For a more general case of an N spin-chain, starting with an initial state that is unentangled, the initial entropy $S_E(t = 0, L_A) = 0$, but will evolve in time in an ergotic ETH system to $S_E(t >> t_{equil}, L_A) = L_A s$, that is extensive and given by the thermodynamic entropy. We note that since $S_E(L) = 0$, $S_E(L_A) = S_E(L - L_A)$, maximized at $L_A = L/2$.

As is clear from (9),(12), a canonical Hamiltonian approach to computation of thermodynamic averages for quantum systems requires an additional step of a diagonalization of the Hamiltonian. As we will see in more detail in subsequenty lectures, an alternative approach allows one to bypass the diagonalization step, but at the cost of working with a path-integral over (imaginary) time-dependent microscopic degrees of freedom $\{q_n(\tau)\}$, with an exponential of the corresponding Euclidean classical action, $\mathcal{S}_E[\{q_n(\tau)\}] = \int_0^{\beta\hbar} d\tau L_E[\{q_n(\tau)\}]$ as the counter-part of the Boltzmann weight, with $L_E[\{q_n(\tau)\}] = \frac{1}{2}m(\partial_\tau q_n)^2 + V[\{q_n(\tau)\}]$ the classical Euclidean Lagrangian, with $it \equiv \tau$. Namely, for a quantum system the partition function can be written as a path-integral,

$$Z = \sum_{\{q_n(\tau)\}} e^{-\mathcal{S}_E[\{q_n(\tau)\}]/\hbar}, \tag{15}$$

where the imaginary time ranges over a finite interval, $0 \leq \tau \leq \beta\hbar$. This thereby reduces quantum statistical mechanics of a d-dimensional system to a $d+1$-dimension path-integral formulation over commuting degrees of freedom. We also note that in the limit of high temperature (much higher than any natural frequencies of the system), $\{q_n(\tau)\}]$ are τ independent and the Euclidean action reduces to $\int_0^{\beta\hbar} d\tau L[\{q_n(\tau)\}] \rightarrow \beta H[\{q_n\}]$, and the partition function reduces to the classical one in Eq. (7).

C. Relation to information theory

Amazingly, as was first discussed by "the father" of information theory, Claude Shannon in his landmark 1948 paper, "A Mathematical Theory of Communication", thermodynamic entropy as defined by the microcanonical ensemble is intimately related to information contained in e.g., a transmitted signal or text of a book. The profound Shannon's assertion is that "information" (Shannon entropy) is the von Neumann entropy of the signal or data, viewed as stochastic data characterized by probability distribution.

To see this more explicitly, we consider a binary signal of N bits of 0s and 1s

$$\ldots 0001101111000110000 \ldots \tag{16}$$

with probability p_0 of a 0 bit and $p_1 = 1 - p_0$ of a 1 bit appearing in the signal. We then would like to define "Information" to be consistent with the colloquial intuition, namely, the most predictible signal of, e.g., all 1s, $p_1 = 1$ carrying no, i.e., zero information (since each bit is completely predictable to be a 1) and the least predictable signal of random 0s and 1s, $p_0 = p_1 = 1/2$ carrying maximum Information. Following Shannon we then define Information S (Shannon entropy of the signal) in terms of the number of possible length N bit signals that can be created subject to the statistics of the signal characterized by $p_{0,1}$. The latter is nothing but the multiplicity,

$$\begin{aligned} \Omega &= \frac{N!}{M_0! M_1!} = \frac{N!}{(p_0 N)!(p_1 N)!}, & (17) \\ &= \frac{N^N}{(p_0 N)^{p_0 N} (p_1 N)^{p_1 N}} \equiv 2^{NS}, & (18) \\ &= p_0^{-p_0 N} (1 - p_0)^{-(1-p_0)N}, & (19) \end{aligned}$$

where in the second line we used Stirling approximation. This then gives Shannon's information per bit, $S = N^{-1} log_2 \Omega$,

$$S = -p_0 \log_2(p_0) - (1 - p_0) \log_2(1 - p_0), \tag{20}$$

just like the entropy of a Ising magnet at fixed magnetization. In the more general case of m (rather than just 2) letters a_i, with $i = 1, \ldots, m$ and corresponding probability for each letter p_i, Shannon's information is given by,

$$S = -\sum_{i=1}^{m} p_i \log_2 p_i. \tag{21}$$

This and many other aspects of classical and quantum informaitons are well discussed by J. Preskill in his Caltech notes https://www.lorentz.leidenuniv.nl/quantumcomputers/literature/preskill_1to_6.pdf as well as by Ed Witten in his "mini-review of information theory", that appeared on the arXiv.

II. OVERVIEW

A. Technicalities

With the background in place, we now overview the systems that we will cover in this course. These will include particle systems, such as gases, liquids and solids, whose microscopic variables are a set of continuous positions vectors, $\{\vec{r}_i\}$, of N particles labelled by a discrete index i. The classical partition function will then involve dN-dimensional integrals over these particle coordinates. (For a classical system the corresponding momenta variables, $\{\vec{p}_i\}$ decouple and can be integrated out of the full phase-space integral, leaving only the coordinate averages). For the quantum particle systems, the averages are $(d+1)N$-dimensional traces over $\{\vec{r}_i(\tau)\}$ coordinates.

For classical magnetism, the microscopic degrees of freedom are fixed-length spins, $\{\vec{S}_i\}$, integrated over surface of N, d-dimensional spheres. For quantum magnetism the trace can be performed in the S_z basis, corresponding to a discrete $2S+1$-dimensional Hilbert space of $-S, -S+1, \ldots, S-1, S$ for each spin i.

In some applications, it will be convenient to transition to a continuum limit of (effectively) infinite number of degrees, with $n \to \mathbf{x}$, for example labelling phonon displacements of a continuum description of an elastic medium, $\vec{u}(\mathbf{x})$. With this, the trace over the macroscopic degrees of freedom will reduce to a d- and $d+1$-dimensional path-integral in the classical and quantum limits, respectively. In these cases, we will have to pause to review the path-integral calculus.

B. Topics

In this course we will discuss a number of systems, as outlined in the syllabus. We will begin with magnetism, reviewing paramagnetism of decoupled spins[1] and demonstrating how ferromagnetism (and other magnetic orders) emerges from interacting lattice of spins. We will first treat such magnetic system using a microscopic Weiss mean-field (MF) approximation[4, 7], that will predict a continuous, finite temperature paramagnet (PM) to ferromagnet (FM) phase transition.

Building on this microscopic analysis, we will derive Landau MF theory (MFT) of continuous phase transitions, utilizing path-integral formulation and, following Landau by simply postulating it based on symmetry. Motivated by this and guided by symmetries, we will discuss a number of other Landau MFTs, such as the Ising, XY, Heisenberg and N-component $O(N)$ PM-FM[7, 14], liquid-gas, liquid-crystal, isotropic-nematic liquid, normal-superfluid, and other phase transitions. We will leave more detailed analysis of these for later lectures.

We will then study the regime of validity of Landau MFT, uncovering the so-called Ginzburg regime, where fluctuations, neglected by MFT are large and thus cannot be neglected[4, 7, 14].

While for harmonic (noninteracting, Gaussian,. . .) systems taking fluctuations into account is not a problem, the challenge comes when the Hamiltonian is nonlinear (interacting), as partition function cannot then generically be computed. Because of strong fluctuations, the perturbation theory in nonlinearities (interactions) diverges and nonperturbative approach is necessary. We will learn how to deal with this challenge, by following its founding fathers, Migdal, Kadanoff, Wilson, Fisher,. . . [12, 13], and develop the renormalization group (RG) to treat these nonlinearities in the presence of strong fluctuations, necessary near continuous phase transitions[4, 7, 14] and other exotic regimes.

To this end, we will first digress to develop field-theoretic path integral methods[7, 14], that are necessary to calculate partition function and thermodynamic averages over continuum fields $\phi(\mathbf{x})$.

We them move on to a variety of other systems, such as superfluids, bulk and films, studying Kosterlitz-Thouless transition and its dual roughenning transition described by the sine-Gordon model. To this end, we will develop particle-vortex (topological defects) and other type of duality methods. We will also discuss fluctuations in two-dimensional crystals and their melting into a hexatic and isotropic fluids by unbinding of dislocations and disclinations, respectively. We will then study a charged superconductor and its associated vortex states of matter[15]. We will explore the physics of basic liquid crystal phases, nematic, cholesteric, smectic-A, smectic-C and their associated phases transitions.[16].

We will then turn to a study of phases and phase transitions, utilizing RG field-theoretic methods in the presence of local random heterogeneities, "quenched disorder", that is an unavoidable ingredient in most condensed matter systems[17]. This will allow us to treat quenched disorder in a variety of systems such as random-field XY ferromagnets, randomly-pinned charge-density waves (CDW) and vortex lattices, liquid crystals in aerogel and on a dirty substrate, and disordered polymerized membranes.

Time permitting, we will conclude this course with studies of equilibrium and non-equilibrium dynamics, utilizing stochastic quantization of a stochastic Langevin equation, derivation and study of the Fokker-Planck equation and development of the dynamic renormalization group[4, 7, 18, 19].

[1] Pathria: *Statistical Mechanics*, Butterworth-Heinemann (1996).

[2] L. D. Landau and E. M. Lifshitz: *Statistical Physics*, Third Edition, Part 1: Volume 5 (Course of Theoretical Physics, Volume 5).

[3] Mehran Kardar: *Statistical Physics of Particles*, Cambridge University Press (2007).

[4] Mehran Kardar: *Statistical Physics of Fields*, Cambridge University Press (2007).

[5] J. J. Binney, N. J. Dowrick, A. J. Fisher, and M. E. J. Newman : *The Theory of Critical Phenomena*, Oxford (1995).

[6] John Cardy: *Scaling and Renormalization in Statistical Physics*, Cambridge Lecture Notes in Physics.

[7] P. M. Chaikin and T. C. Lubensky: *Principles of Condensed Matter Physics*, Cambridge (1995).

[8] "Chaos and Quantum Thermalization", Mark Srednicki, *Phys. Rev. E* **50** (1994); arXiv:cond-mat/9403051v2; "The approach to thermal equilibrium in quantized chaotic systems", *Journal of Physics A* **32**, 1163 (1998).

[9] "Quantum statistical mechanics in a closed system", J. M. Deutsch, *Phys. Rev. A* **43**, 2046.

[10] D. M. Basko, I. L. Aleiner and B. L. Altshuler, *Annals of Physics* **321**, 1126 (2006).

[11] "Many body localization and thermalization in quantum statistical mechanics", *Annual Review of Condensed Matter Physics* **6**, 15-38 (2015).

[12] M. E. Fisher, *Rev. Mod. Phys.* **42**, 597 (1974).

[13] K. G. Wilson and J. Kogut, *Phys. Rep.* **12** C, 77 (1974).

[14] J. Zinn-Justin: *Quantum Field Theory and Critical Phenomena*, Oxford (1989).

[15] P. G. de Gennes: *Superconductivity of Metals and Alloys*, Addison-Wesley (1989).

[16] P. G. de Gennes and J. Prost: *The Physics of Liquid Crystals*, Oxford (1993).

[17] For a review of heterogeneous systems, see for example an article by D. S. Fisher in *Physics Today* (1989).

[18] S. K. Ma, “Dynamic Critical Phenomena”.

[19] P. Hohenberg, B. I. Halperin, “Critical dynamics”, *Rev. Mod. Phys.*.