

A Reading Course on Statistical Physics: Summer 2018

William Porteous

Updated: January 12, 2021

The purpose of these notes is to first, review the framework of classical and quantum mechanics which is needed to begin studying statistical physics. Then, to introduce kinetic theory as a classical theory, and the hydrodynamic equations therefrom. Then, these notes demonstrate the probabilistic approach to this theory taken in the introduction to classical statistical mechanics. Finally, a quantum theory of statistical mechanics is introduced.

Contents

1	The Variational Framework of Classical Mechanics	2
1.1	Principle of Virtual Work	2
1.2	D'Alembert's Principle and Lagrange's Equations	2
1.3	Hamilton's Principle	3
2	Review of Quantum Mechanics	6
2.1	Inner Product Spaces, Adjoints, and Operator Subgroups	6
2.2	The Postulates and Immediate Consequences	15
2.3	Measurements: Evaluating a Dynamical Variable	16
2.4	The Postulates: Multiple Degrees of Freedom	20
2.5	The Schrodinger Equation: Generic Solutions and Simple 1-D Systems	21
2.6	The Uncertainty Principle and the (Semi)Classical Limit	26
3	Kinetic Theory of Gases	31
3.1	Consequences of Liouville's Theorem	32
3.2	BBKGY Hierarchy - Many Particle Density Evolution	33
3.3	The Boltzmann Equation	34
3.4	Conservation Laws	35
3.5	The Hydrodynamic Equations	36
4	Classical Statistical Mechanics	39
4.1	The Microcanonical Ensemble	39
4.2	Laws of Thermodynamics	40
4.3	The Ideal Gas	42
4.4	The Canonical Ensemble	44
4.5	Introduction to Quantum Statistical Mechanics	49

1 The Variational Framework of Classical Mechanics

In this context, we consider the motion of a classical system in "configuration space", a linear space which can contain the generalized coordinates we find from lagrange's equations.

Definition (Configuration Space and Motion of the System). *If there are n generalized coordinates $\{q_i\}_{i \in [n]}$ that express the instantaneous state of the system, $(\mathbb{R}^n, \{q_i\})$ (\mathbb{R}^n with this basis) is called configuration space. Over a time, the point in configuration space instantaneously giving the state of the system changes (defined by the cartesian path), defining the path of the motion of the system. Time can be formally considered a parameter of this curve.*

1.1 Principle of Virtual Work

A **virtual displacement** of a system is a change in the configuration due to an *arbitrary* and *infinitesimal* change of the coordinates $\delta \mathbf{r}_i$ consistent with the *forces* and *constraints* on the system at an instant in time t . This does not take place in an interval of time dt , and rather than time derivatives being computable from $\delta \mathbf{r}_i$, the displacement samples the physical environment of the system at that time.

If the system is in equilibrium at that time,

$$\mathbf{F}_i \cdot \delta \mathbf{r}_i = 0$$

and

$$\sum_i \mathbf{F}_i \cdot \delta \mathbf{r}_i = 0$$

Take \mathbf{f}_i to be the force of constraint, and \mathbf{F}_i^a the applied force, on the i th element of the system. Restricting to systems for which the forces of constraint do no work gives the principle of virtual work; note that this excludes, for example, a friction. The principle is then:

$$\sum_i \mathbf{f}_i \cdot \delta \mathbf{r}_i = 0 \implies \sum_i \mathbf{F}_i^a \cdot \delta \mathbf{r}_i = 0 \tag{1}$$

In general, the $\delta \mathbf{r}_i$ are no longer independent displacements, related, for example, by the above equation, and the constraints.

A similar condition, but accounting for general motion of the system (not only in equilibrium) is desirable.

Consider then that the equations of motion may be rearranged to read as a force plus a force opposing the change in momentum will bring the system to equilibrium. Consider a virtual displacement under these conditions:

$$\mathbf{F}_i - \dot{\mathbf{p}}_i = 0 \implies \sum_i (\mathbf{F}_i - \dot{\mathbf{p}}_i) \cdot \delta \mathbf{r}_i = 0$$

1.2 D'Alembert's Principle and Lagrange's Equations

$$\sum_i (\mathbf{F}_i^a - \dot{\mathbf{p}}_i) \cdot \delta \mathbf{r}_i = 0$$

Now, using the equations of constraint, the most general form of Lagrange's equations can be obtained. Transform, using the equations of constraint to n independent coordinates:

$$\mathbf{r}_i = \mathbf{r}_i(q_1, \dots, q_n, t)$$

$$\sum_i \mathbf{F}_i \cdot \delta \mathbf{r}_i = \sum_{i,j} \mathbf{F}_i \cdot \sum_j (\partial_{\mathbf{q}_j} \mathbf{r}_i) \delta \mathbf{q}_j = \sum_j \sum_i (\mathbf{F}_i \cdot \partial_{\mathbf{q}_j} \mathbf{r}_i) \cdot \delta \mathbf{q}_j$$

where Q_j is called force corresponding to q_j with

$$Q_j = \sum_i \mathbf{F}_i \cdot \partial_{q_j} \mathbf{r}_i$$

For the work calculated by the time derivative of momentum, this says:

$$\begin{aligned} - \sum_i \mathbf{p}_i \cdot \delta \mathbf{r}_i &= - \sum_i m_i \ddot{\mathbf{r}}_i \cdot \delta \mathbf{r}_i = - \sum_i d_t(p_i \cdot \delta \mathbf{r}_i) - \mathbf{p}_i \cdot \dot{\delta \mathbf{r}}_i \\ &= - \sum_{i,j} [d_t(\mathbf{p}_i \cdot \partial_{\mathbf{q}_j} \mathbf{r}_i) - \mathbf{p}_i \cdot \partial_{\mathbf{q}_j} \dot{\mathbf{r}}_i] \delta \mathbf{q}_j \end{aligned}$$

Since

$$\partial_{q_\alpha} \mathbf{v}_i = \partial_{q_\alpha} \left(\sum_j \partial_{q_j} \mathbf{r}_i \cdot \dot{q}_j + \partial_t \mathbf{r}_i \right) = \partial_{q_\alpha} \mathbf{r}_i$$

and

$$\partial_{q_j} \dot{\mathbf{r}}_i = \partial_{q_j} \mathbf{v}_i$$

by unfolding the expression in terms of $\{\dot{q}_k\}$ and then changing coordinate dependence, the above may be expressed:

$$\begin{aligned} - \sum_i \mathbf{p}_i \cdot \delta \mathbf{r}_i &= - \sum_{i,j} [d_t(\mathbf{p}_i \cdot \partial_{\mathbf{q}_i} \mathbf{v}_i) - \mathbf{p}_i \cdot \partial_{q_i} \mathbf{v}_i] \delta \mathbf{q}_j = - \sum_j d_t \left(\sum_i \mathbf{p}_i \partial_{q_i} \mathbf{q}_i \right) - \sum_i \mathbf{p}_i \partial_{q_i} \mathbf{v}_i \\ &= - \sum_j [d_t(\partial_{q_j} \mathcal{T}) - \partial_{q_j} \mathcal{T}] \delta \mathbf{q}_j \end{aligned}$$

Since the $\delta \mathbf{q}_j$ are linearly independent displacements

$$Q_j = d_t(\partial_{q_j} \mathcal{T}) - \partial_{q_j} \mathcal{T} \quad (2)$$

1.3 Hamilton's Principle

Theorem. *Hamilton's Principle: If a system is described by a generalized potential, the motion of the system (path in configuration space) from t_1 to t_2 is such that*

$$I = \int_{t_1}^{t_2} L \, dt$$

is an extremum over paths of motion (paths in configuration space) or equivalently

$$\delta I = 0$$

To show that Hamilton's principle implies Lagrange's equations, a technique is found in the following discussion of the basic variational problem.

Consider fixed $x_1, x_2 \in \mathbb{R}$, and $y \in C^1(\mathbb{R})$ with $y(x_1) = p_1$ and $y(x_2) = p_2$. Then given an $f: \mathbb{R}^3 \rightarrow \mathbb{R}$, form its associated function $J(y) = \int_{x_1}^{x_2} f(y(x), y'(x), x)$. To obtain paths with $y(x_1) = p_1$ and $y(x_2) = p_2$ which are extrema for J , first, suppose that at least some extrema exist. If $y_0(x)$ is such an extremum of the integral, and y is any other path, consider the homotopy of paths $H(x, \alpha) = y_0(x) + \alpha\eta(x)$ where $\eta(x) = y(x) - y_0(x)$. Note that since $y_0(x)$ is given, the variant η is associated to each $y(x)$. Then, form

$$\tilde{J}(\alpha, y) = J(H(x, \alpha))$$

and observe that the condition that $y_0(x)$ be an extremum is that

$$(\partial_\alpha \tilde{J})(0) = 0$$

for arbitrary η or y . Unfolding this definition reads:

$$\partial_\alpha \tilde{J}(y, 0) = 0 \implies \int_{x_1}^{x_2} \frac{\partial f}{\partial H} \cdot \frac{\partial H}{\partial \alpha} + \frac{\partial f}{\partial H'} \frac{\partial H'}{\partial \alpha} \Big|_{\alpha=0} = 0$$

But $\frac{\partial f}{\partial H'} \frac{\partial H'}{\partial \alpha} = \frac{\partial f}{\partial y'} \cdot \partial_x(\partial_\alpha H)$ integration by parts gives

$$\frac{\partial f}{\partial H'} \cdot \frac{\partial H}{\partial \alpha} \Big|_{x_1}^{x_2} - \int_{x_1}^{x_2} (\partial_x \frac{\partial f}{\partial H'}) \frac{\partial H'}{\partial \alpha} dx = - \int_{x_1}^{x_2} (\partial_x \frac{\partial f}{\partial H'}) \frac{\partial H}{\partial \alpha} dx$$

since $\frac{\partial H}{\partial \alpha} \equiv \eta(x)$ which is identically zero at x_1 and x_2 . Therefore,

$$\partial_\alpha \tilde{J} = 0 \implies \int_{x_1}^{x_2} (\partial_y(f(y_0, y'_0, x) - \partial_x \partial_{y'} f(y_0, y'_0, x)) \eta(x) dx = 0$$

Since y_0 is the fixed extremum and η is arbitrary, this is only possible if

$$\partial_y f(y_0, y'_0, x) - \partial_x \partial_{y'} f(y_0, y'_0, x) = 0$$

which defines a differential equation for which y_0 is a solution. The only candidates for extrema, therefore, are curves satisfying this equation.

This is readily generalized to the case where f is a function of many independent variables and their derivatives. For if $f \in C^1(\mathbb{R}^{2n+1}, \mathbb{R})$, $f(y_i, y'_i, x)$ then one may define H_i and variations η_i , all with constant α (the same form of H_i as previously) so that

$$(\partial_\alpha \tilde{J}(y_i, \alpha))|_{\alpha=0} = 0 \implies \int_{x_1}^{x_2} \sum_{i=1}^n (\partial_{y_i} f - \partial_x \partial_{y'_i} f) \eta_i(x) = 0$$

Since this must hold for any η_i , and each of these is independently arbitrary,

$$\partial_{y_i} f - \partial_x \partial_{y'_i} f = 0 \quad \forall i$$

More generality is possible, with f a function of higher order derivatives of the curves, and in fact, multiple parameters x_j and the integral a multiple integral. Furthermore, it is possible to consider variations with endpoints not fixed.

Now, however, Lagrange's equations are a clear consequence of Hamilton's Principle:

Corollary. *Hamilton's Principle Implies Lagrange's Equations in Conservative Systems*

Proof. Take

$$f(y_i, y'_i, t) = L(q_i, q'_i, t)$$

and

$$I = \int_{t_1}^{t_2} L(q_i, q'_i, t) dt$$

then the Euler-Lagrange equations read the same as Lagrange's equations:

$$\partial_{q_i} L(q_i, q'_i, t) - D_t(\partial_{q'_i} L)(q_i, q'_i, t) = 0 \quad i \in [n]$$

□

Now, Hamilton's principle can be generalized to non-conservative systems, say Hamilton's General Principle, to obtain the more general form of lagrange equations:

$$(D_t \partial_{q'_j} - \partial_{q_j}) T = Q_j$$

The extended principle can be written as follows:

Theorem. *Hamilton's General Principle in Holonomic Systems*

$$\delta J = \delta \int_{t_1}^{t_2} (T + W) dt = 0 \iff \delta \int_{t_1}^{t_2} T + \int_{t_1}^{t_2} \left(\sum_{j=1}^n Q_j \delta q_j \right) dt = 0$$

where $W = \sum_{i=1}^m F_i \cdot r_i$.

Proof. Intuitively, δW represents the work done by the forces of the system on the system over the virtual displacement from the path to the varied path. Precisely,

$$\delta W = \sum_{i=1}^m F_i \cdot \delta r_i$$

Furthermore, the variations δq_i and δr_i are the same as the virtual displacements of the coordinates, so that $\delta \int_{t_1}^{t_2} W = \int_{t_1}^{t_2} \delta W$.

□

The generalized form of lagrange's equations for holonomic systems is recovered from this principle:

Since the δ are the same, on the r_i and q_j , the previously proven relationship

$$\sum_{i=1}^m F_i \cdot \delta r_i = \sum_{j=1}^n Q_j \delta q_j$$

Therefore,

$$\delta I = \delta \int_{t_1}^{t_2} T + \int_{t_1}^{t_2} \left(\sum_{j=1}^n Q_j \delta q_j \right) dt$$

And if T is a function of q_j and q'_j , and the variational principle applies to say:

$$\delta \int_{t_1}^{t_2} T = \int_{t_1}^{t_2} \sum_{j=1}^n (\partial_{q_j} - D_t \partial_{q'_j}) T \delta q_j dt$$

and combining the two expressions

$$\delta I = \int_{t_1}^{t_2} \sum_{j=1}^n [(\partial_{q_j} - D_t \partial_{q'_j}) T + Q_j] \delta q_j dt$$

and since the constraints are holonomic, the δq_j are independent, and

$$[\partial_{q_j} - \frac{d}{dt} \partial_{q'_j}] T + Q_j = 0$$

since the constraints are holonomic by assumption.

The analysis of Hamilton's principle thus applied only to holonomic systems. Further still, Lagrange's general equations were derived from *D'Alembert's Principle* for holonomic systems.

Theorem. *The Cauchy-Schwarz Inequality (For Complex Inner-Product Spaces)*

$$\|\langle w | v \rangle\| \leq \|v\| \|w\|$$

Proof. Consider that $\langle z | z \rangle \geq 0$ for any $z \in \mathbb{V}$. Choose

$$|z\rangle = |v\rangle - \left(\frac{\langle v | w \rangle}{\|w\|^2}\right) |w\rangle$$

and applying this condition

$$\begin{aligned} \langle z | z \rangle &= \langle v | v \rangle - \left(\frac{\langle w | v \rangle}{\|w\|^2}\right) \langle v | w \rangle - \frac{\langle w | v \rangle^*}{\|w\|^2} \langle w | v \rangle + \frac{\langle w | v \rangle^* \langle w | v \rangle}{\|w\|^2} \langle w | w \rangle \\ &= \|v\|^2 - \frac{\|\langle w | v \rangle\|^2}{\|w\|^2} \geq 0 \implies \\ &\|v\|^2 \|w\|^2 \geq \|\langle w | v \rangle\|^2 \end{aligned}$$

Alternatively, consider

$$|z\rangle = \|w\| |v\rangle - |w\rangle \|v\|$$

which is zero if v and w are linearly dependent. □

2 Review of Quantum Mechanics

2.1 Inner Product Spaces, Adjoints, and Operator Subgroups

The inner product is written for any two vectors \mathbf{v}, \mathbf{w} $\langle v | w \rangle$ and satisfies

1. Skew-Symmetry: $\langle v | w \rangle = \langle w | v \rangle^*$

2. Positive-Definiteness: $\langle v|v \rangle \geq 0 \wedge 0 \iff \mathbf{v} = 0$

3. Linearity in 1st Argument: $\langle w|\alpha v + \beta u \rangle = \alpha \langle w|v \rangle + \beta \langle w|u \rangle$

In a finite dimensional inner product space, the **Gram-Schmidt Process** is always possible, which admits a basis w.r.t which the inner product is diagonal, and unit on the basis. That is, given $\langle | \rangle$ on any basis, one may choose $\{|j\rangle\}_{j \in S_n}$ so that

$$\langle v|w \rangle = \langle \sum_i v_i |i\rangle | \sum_j w_j |j\rangle \rangle = \sum_{i,j} v_i^* w_j \langle i|j \rangle = \sum_{i,j} v_i^* w_j \delta_{i,j} = \sum_j v_j^* w_j$$

Further, the positive definiteness of the inner product ensures that the inner product gives coordinates in any basis, that is, the inner-product axioms quickly show

$$|v\rangle = \sum_j v_j |j\rangle \iff v_j = \langle v|j\rangle$$

From now on, a basis can be assumed orthonormal, and $v_j = \langle v|j\rangle$ for any basis $\{|j\rangle\}_{j \in S_n}$. Now, in any basis one may associate the complex inner product space with \mathbb{C}^n via some isomorphism $\Phi: \mathbb{V} \longrightarrow \mathbb{C}^n$ and obtain

$$\Phi(|v\rangle) = (v_j)_j$$

where v_j are the coordinates well defined in \mathbb{V} . The inner product may have a complicated representation under this isomorphism, since $\langle \cdot | \cdot \rangle$ is not diagonal in a basis which is not orthonormal. In an orthonormal basis, however, the inner product $\langle w|v \rangle$ is representable under Φ as

$$(w_j^*)^j (v_j)_j = \sum_j w_j^* v_j$$

This allows Φ^{-1} to take $(w_j^*)^j$ back under its representation, to an abstract object, called $\langle w|$, which transforms according to the basis $\{|j\rangle\}_j$, and is hence, independent of the coordinates w_j^* due to the basis $\{|j\rangle\}$. This space is called \mathbb{V}^* , the dual space, composed of all linear maps $f: \mathbb{V} \longrightarrow \mathbb{C}$. This way, $\langle w|$ has an abstract representation, possible in any basis. It is easily shown that, in V^* , $\langle w| = \langle w|\cdot \rangle$.

Clearly, the space that $\langle w|$ lives in is well defined by this Φ correspondence, and observing its basis transformation properties. The function

$$\dagger: |w\rangle \rightarrow \langle w|$$

is well defined, called the **adjoint**. This is commonly written $|w\rangle^\dagger = \langle w|$. The following may be checked using coordinates and an isomorphism:

$$|\alpha v + w\rangle^\dagger = \alpha^* \langle v| + \langle w|$$

i.e. the adjoint is conjugate linear. As defined, this may be used to show \mathbb{V}^* to be a complex vector space, with bases corresponding to each basis in \mathbb{V} . Thus, the inner product may be equivalently expressed as a map $\langle \cdot | \cdot \rangle: \mathbb{V}^* \times \mathbb{V} \longrightarrow \mathbb{C}$. Inner-product coordinate correspondence and Gram Schmidt hold in V^* so that

$$\langle v| = \sum_j \alpha_j \langle j| \iff \alpha_j = \langle v|j\rangle \iff |v\rangle = \sum_i \beta_i |i\rangle \iff \beta_i = \langle i|v\rangle = \alpha_i^*$$

2.1.1 Operator Adjoint

Turning to linear operators, the coordinate definiteness of the inner product also shows that if $\Omega = (\Omega_{i,j})$ in an orthonormal basis $\{|k\rangle\}_k$, then the coordinates Ω_{ij} are given by:

$$\Omega_{ij} = \langle j|\Omega|i\rangle$$

The adjoint is extended to operators via the definition

$$(\Omega|v\rangle)^\dagger = \langle v|\Omega^\dagger$$

In finite dimensional coordinates this is:

$$\begin{aligned}\langle j|\Omega^\dagger|i\rangle &= \langle \Omega j|i\rangle = \langle i|\Omega j\rangle^* = \Omega_{ji}^* \implies \\ \Omega_{i,j}^\dagger &= \Omega_{j,i}^*\end{aligned}$$

It is easily verified by definition that on operators the adjoint satisfies:

$$(\Omega\Lambda)^\dagger = \Lambda^\dagger\Omega^\dagger$$

Operator Classes

Taking a finite dimensional space, if Ω 's coordinates were symmetric about its main diagonal, then, for purely complex, and purely real Ω respectively, the operator satisfies

$$(a). \Omega^\dagger = -\Omega$$

$$(b). \Omega^\dagger = \Omega$$

This formulation generalizes real-ness and imaginary-ness. Case (a) is called **Hermitian** and (b) **anti-Hermitian**.

For an arbitrary operator Ω , observe:

$$\Omega = \frac{1}{2}[(\Omega + \Omega^\dagger) + (\Omega - \Omega^\dagger)]$$

where clearly the first term forms a hermitian operator, say, $\mathcal{H}(\Omega)$, and the second two terms, $\mathcal{T}(\Omega)$ is an anti-hermitian operator.

The eigenvalue problem for hermitian matrices is subject to this transpose approach. If Ω is hermitian and λ is an eigenvalue with vector \mathbf{v}

$$\Omega|v\rangle = \lambda|v\rangle \implies \langle v|\Omega|v\rangle = \lambda\|v\|^2 = (\langle v|\Omega|v\rangle)^\dagger = \lambda^\dagger\|v\|^2$$

so that $\lambda = \lambda^\dagger$. Thus **eigenvalues of hermitian matrices are real**.

Theorem. *Hermitian matrices admit an orthonormal eigenbasis (diagonalizeable)*

In a finite dimensional space, with U purely real, and in an orthonormal basis, $U \cdot U^t = I$ is orthogonality of U with respect to $\langle|\rangle$, meaning that the basis vectors defined by the columns of U are orthogonal under $\langle|\rangle$. Generalizing this condition, U is **unitary** if

$$U \cdot U^\dagger = I$$

This is analogous to conjugation of values on the complex unit circle $c \cdot \bar{c} = 1$.

Theorem. *Unitary matrices admit eigenbases, eigenvalues of unitary matrices are in S^1 , and non-degenerate eigenvectors are orthogonal*

Proof. For the first claim, suppose that $|u_i\rangle$ and $|u_j\rangle$ are two nonzero eigenvectors with eigenvalues u_i and u_j , that is

$$\begin{aligned} U|u_i\rangle &= u_i|u_i\rangle \\ U|u_j\rangle &= u_j|u_j\rangle \end{aligned}$$

Then

$$\langle u_i|u_i\rangle = u_i^* u_i \cdot \langle u_i|u_i\rangle \implies \|u_i\|^2 = 1$$

and

$$u_j^* u_i \langle u_j|u_i\rangle = \langle u_j|u_i\rangle$$

This is only possible if $u_j^* u_i = 1 \implies u_j = u_i$, or, $\langle u_j|u_i\rangle = 0$. Thus, only nondegenerate eigenvectors must be orthogonal. \square

2.1.2 Projections and Completeness: Infinite Dimensional Space

In an arbitrary basis, due to the axioms of the inner product, the inner product function can identify coordinates of all elements in the space. Assuming the inner product is known on the space, this says

$$|v\rangle = \sum_i |i\rangle \langle i|v\rangle$$

Since an object of the form $|w\rangle \langle u|$ is a linear operator (just check its action on any vector), this means

$$|v\rangle = \left(\sum_i |i\rangle \langle i| \right) |v\rangle$$

Since $|v\rangle$ is arbitrary,

$$\sum_i |i\rangle \langle i| = \mathbf{I}$$

For any basis on an inner product space $\{|\alpha\rangle\}_{\alpha \in \mathcal{A}}$, it is desirable that the statement $\sum_{\alpha} |\alpha\rangle \langle \alpha| = \mathbf{I}$, where α is used as an arbitrary index in anticipation of extending this statement to function space. Define

$$\mathbb{P}_{\alpha} = |\alpha\rangle \langle \alpha|$$

as the **projection operator** for the subspace generated by $|\alpha\rangle$ (assuming here that it is 1-dimensional). This way, the statement above is understood as

$$\mathbf{I} = \sum_{\alpha} \mathbb{P}_{\alpha}$$

This has not used orthonormality of basis vectors. If the basis is orthonormal, then for any vector, its various components, $(\mathbb{P}_{\alpha}|v\rangle) = v_{\alpha}|\alpha\rangle$ are orthogonal, i.e.

$$\langle \mathbb{P}_{\alpha}|v\rangle | \mathbb{P}_{\beta}|v\rangle \rangle = v_{\alpha}^* v_{\beta} \langle \alpha | \beta \rangle = 0$$

In fact

$$\mathbb{P}_{\alpha} \cdot \mathbb{P}_{\beta} = \mathbb{P}_{\beta} \delta_{\alpha,\beta} = |\alpha\rangle \langle \alpha | | \beta \rangle \langle \beta | = 0$$

Various coordinates and their corresponding projection operators are linearly independent in such a system. Thus, due to the inner product, and choice of basis respectively,

$$(1). \sum_{\alpha} |\alpha\rangle\langle\alpha| = \mathbf{I}$$

$$(2). \langle\alpha|\beta\rangle = \delta_{\alpha,\beta} \iff \mathbb{P}_{\alpha} \cdot \mathbb{P}_{\beta} = \delta_{\alpha,\beta} \mathbb{P}_{\beta}$$

In this form, these concepts are extended to function space. Consider first function space on discrete set of complex values with elements

$$f: \{z_i\}_{i \in S_n} \longrightarrow \mathbb{C}$$

Functions are here distinguished by taking distinct values at any of n complex numbers $\{z_i\}$. *We need n independent, orthogonal vectors (functions), to yield n independent coordinates that bijectively correspond to the independent valuations of an arbitrary f against n points.* Arbitrary is important here, since perhaps f may have some constraint on its value among certain points; a basis expressing an arbitrary vector (function) can certainly express a vector with fewer than n independent coordinates (values). The **points of the space itself, since f is valued against each of these independently** provide a suitable basis. Thus, take:

$$\mathcal{B} = \{|z_i\rangle\}$$

and require the satisfaction of (1). and (2). Orthogonality (or even just linear independence) here makes sense here since an arbitrary f is valued against these points independently, and is completely specified only by obtaining its value on all points. Finally, define coordinates (or equivalently, specify the inner product) to give the simplest possible correspondence between the coordinates of f , and its values:

$$\langle z_i | f \rangle := f(z_i)$$

These requirements imply:

$$|f\rangle = \sum_i f(z_i) |z_i\rangle$$

and for example

$$\langle f | g \rangle = \sum_i f(z_i)^* g(z_i)$$

To extend this to infinite dimension, start with the requirement to specify an inner product and basis which satisfy (1.) and (2). If the inner product is to be extended as

$$\sum_i f^*(z_i) g(z_i)$$

with $n \rightarrow \infty$ convergence is an obvious problem. It is readily observed that modifying $\langle | \rangle$ on a discrete space by a scalar, does no harm to (1) and (2). Replace the prior inner product by

$$\langle f | g \rangle = \sum_i f^*(z_i) g(z_i) \cdot \Delta$$

where $\Delta = \frac{m(X)}{n+1}$ where L is the measure of the whole space (so that Δ is a uniform weight). This reduces to a uniform step when the space is an interval, and in that case, this inner product is a proper Riemann sum. Thus, when the variable is real and the space an interval, the $n \rightarrow \infty$ limit of the above inner product is:

$$\int_{[a,b]} f(x_i)^* g(x_i) dx$$

This special case of a valid, infinite-dimensional inner product, generalizes over any complex measure space to:

$$\langle f|g\rangle = \int_X f^* g d\mu$$

The question is now, what are the basis vectors in the $n \rightarrow \infty$ limit, $|\alpha\rangle$ for an infinite dimensional space $X \subset \mathbb{C}$ or even $[a, b] \in \mathbb{R}$. Does the index stay countable? The inner product is specified in terms of the functions, and now we wish to translate its arguments to the language of vector spaces.

The prior argument about functions being independently valued at each point of the space, and the requirement of the whole space to specify an arbitrary function, is still valid. Thus

$$\mathcal{B} = \{|z\rangle : z \in X \subset \mathbb{C}\}$$

and we hope to specify the relation between coordinates and function values just as before, that is,

$$\langle z|f\rangle = f(z)$$

Most importantly, we want (1). and (2). to hold with the specification of $|z\rangle$. To have (1) and this coordinate simplicity:

$$\left(\int_X dz' |z'\rangle\langle z'| \cdot |f\rangle = I|f\rangle = |f\rangle \implies \langle w| \left(\int_X dz' |z'\rangle\langle z'| \cdot |f\rangle\right) = \int_X \langle w|z'\rangle f(z') = f(w)\right.$$

Since f is arbitrary,

$$\langle w|z'\rangle = \delta(z' - w)$$

as a distribution. Thus, this function is zero for $z' \neq w$, satisfying orthogonality, but acts as the identity distribution under convolution. Thus, when $z' = w$

$$\langle w|w\rangle \sim \infty$$

and thus, $|w\rangle$ are now not normalizeable complex valued functions on X , but distributional functions, characterized by their action on other functions. In fact, this is how all functions are viewed as vectors.

An important example of such a distribution is the **delta prime distribution**, defined by exchanging the derivative limit, with the limit to distribution on a smooth function family corresponding to the dirac delta:

$$\delta'(z - z') = \frac{d}{dz}(\delta(z - z'))$$

Consider a gaussian family of functions

$$g_\sigma(z - z') = \frac{1}{[(\sqrt{\pi})\sigma]} \exp\left[-\frac{1}{\sigma^2}(z - z')\right]$$

to compute this distribution:

$$\begin{aligned} \delta(z - z') &= \lim_{\sigma \rightarrow \infty} g_\sigma(z - z') \implies \\ \delta'(z - z') &= \lim_{\sigma \rightarrow \infty} \frac{d}{dz} g_\sigma(z - z') = \lim_{\sigma \rightarrow \infty} -\frac{2}{\sigma^2}(z - z') \frac{1}{(\sqrt{\pi})\sigma} \exp\left[-\frac{(z - z')^2}{\sigma^2}\right] \end{aligned}$$

which, is extremized at $z' \pm \frac{\sigma}{\sqrt{2}}$, positive at the first and negative at the second. This function is essentially linear between these two regions, and decays exponentially about then. Thus, in the $\sigma \rightarrow 0$ limit, this acts as a dirac delta at each of these points, and applying opposite signs to the function under convolution. With $\varepsilon = \sigma/\sqrt{2}$, one gleans from the normalization factor of $g'_\sigma(z - z')$ that

Thus *under convolution with smooth functions*

$$\delta' \equiv \delta \frac{d}{dz}$$

under convolution.

2.1.3 Linear Operators in Infinite Dimensional Hilbert Space

Now with functions expressed in the language of vector space, linear operators on function space may now be described with the same language. If $\Omega \in \mathcal{L}(\mathbb{V})$ then

$$\Omega|f\rangle = |\tilde{f}\rangle$$

Take the spatial derivative for example on \mathbb{C} ,

$$D|f\rangle = \left| \frac{d}{dz} f(z) \right\rangle$$

Since, as discussed, the coordiantes of \mathbf{D} are given by the inner prodct in any space, what are

$$\langle z|\mathbf{D}|z'\rangle = \mathbf{D}_{z,z'}$$

We know

$$\langle z|\mathbf{D}|f\rangle = \langle z|(\frac{d}{dz}f)\rangle = (\frac{d}{dz}f)(z)$$

so

Now, since

$$\begin{aligned} \langle z|D|f\rangle &= \langle z|D \int_{\mathbb{C}} dz' |z'\rangle \langle z'| |f\rangle = \int_{\mathbb{C}} dz' (\langle z|D|z'\rangle) \langle z'|f\rangle = \int_{\mathbb{C}} dz' (\langle z|D|z'\rangle) f(z') = (\frac{d}{dz}f)(z) \implies \\ \langle z|D|z'\rangle &= \delta'(z - z') \end{aligned}$$

as a distribution.

Now, as it stands D appears anti-hermitian on real variables

$$\mathbf{D}_{r,r'} = \delta'(r - r') = -\delta(r - r')^* = -\delta(r' - r) = -\mathbf{D}_{r',r}$$

and may be corrected by applying a purely complex value as with

$$K = -iD$$

For K to be hermitian, however,

$$\langle f(\mathbf{x})|K^\dagger = \langle f(\mathbf{x})|K$$

for all f . Thus, for any g

$$\langle f|K|g\rangle = \langle f|K^\dagger|g\rangle = \langle Kf|g\rangle = \langle g|Kf\rangle^* = (\langle g|K|f\rangle)^*$$

This says

$$\begin{aligned}\langle g|K|f\rangle &= \int_X \int_X dx dx' \langle g||\mathbf{x}\rangle \langle \mathbf{x}|K|\mathbf{x}'\rangle \langle \mathbf{x}'||f\rangle = \int_X \int_X g^*(\mathbf{x})(-i\delta'(\mathbf{x}-\mathbf{x}'))f(\mathbf{x}') = \\ &= \int_X g^*(\mathbf{x})(-i\frac{d}{dx}f(x) = -ig^*(\mathbf{x})f(\mathbf{x})|_{\partial\mathbf{x}} + \int_X i\frac{d}{dx}g^*(\mathbf{x})f(\mathbf{x})\end{aligned}$$

which is only equal to its adjoint if the surface term is zero.

2.1.4 Functions on Operators

If f is an analytic function on \mathbb{C} , this map can be extended to $\mathcal{L}(\mathbb{V})$ simply. Take its power series expansion about zero,

$$f = \sum_k a_k z^k$$

where $a_k = \frac{f^k(0)}{k!}$ and define

$$\begin{aligned}f: \mathcal{L}(\mathbb{V}) &\longrightarrow \mathcal{L}(\mathbb{V}) \\ f: \Omega &\rightarrow \sum_k \frac{f^k(0)}{k!} \Omega^k = \sum_k a_k \Omega^k\end{aligned}$$

Restricting to hermitian Ω guarantees that this will converge, since f is well defined on \mathbb{C} . Since this must be a basis independent map, take Ω in its eigenbasis so that it is diagonal to compute $f(\Omega)$. $\Omega = (\omega_{i,\alpha}\delta_{ij})$ where $\omega_{i,\alpha}$ are the eigenvalues with α indexing degeneracy. Then,

$$f(\Omega) = \sum_k a_k \Omega^k = (\sum_k a_k \omega_{i,\alpha}^k \delta_{ij}) = (f(\omega_{i,\alpha})\delta_{ij})$$

Now, suppose there is an operator valued map, with one (complex) parameter

$$\theta: \mathbb{C} \longrightarrow \mathcal{L}(V)$$

Taking θ in some basis, then $\frac{d}{d\lambda}(\theta)$ is just the derivative of the coordinates with respect to λ . Take the example

$$\theta(\lambda) = \exp \lambda \Omega$$

for some operator Ω . If Ω is diagonalizable, then it is easy to show that in the eigenbasis

$$\frac{d}{d\lambda}\theta(\lambda) = \Omega \exp \lambda \Omega = \Omega \theta(\lambda)$$

If the power series converges, this is easily obtained directly from the power series definition of $\exp \Omega$. More interesting still is that this differential equation

$$\frac{d}{d\lambda}(\theta(\lambda)) = \Omega \theta(\lambda)$$

can be integrated with respect to lambda

$$\theta(\lambda) = \int_0^\lambda \frac{d}{d\lambda'}(\theta(\lambda'))d\lambda' = \int_0^\lambda \Omega$$

Additional complexity arises from the fact that, unlike complex numbers, operators do not commute in general. Hence, it is not generally true that

$$\exp \alpha \Omega \cdot \exp \beta \Lambda \neq \exp (\alpha \Omega + \beta \Lambda)$$

but it is easily deduced from the definition

$$[f(\Theta), \Theta] = 0 \text{ and } [f(\Theta, g(\Theta))] = 0$$

for any well-defined f, g , and any Θ . Familiar facts of calculus, such as differentiating a product (which is now ordered), now give ordered results:

$$\frac{d}{d\lambda} (\exp \lambda \Omega \exp \lambda \Theta) = \Omega (\exp \lambda \Omega) (\exp \lambda \Theta) + (\exp \lambda \Omega) (\exp \lambda \Theta) \Theta$$

where multiplication is ordered for any functions of Θ and Ω . This additional complexity is eliminated for all functions if, as with complex numbers, $[\Theta, \Omega] = 0$

Question 1. 1. 1.6.3- *Unitary Subgroup: Show that a product of unitary operators is unitary, and that unitary operators form a Subgroup under matrix multiplication*

2. 1.6.2- *Hermitian Subgroup: Show that the sum of hermitian matrices is hermitian, and that hermitian matrices form a subgroup under addition. Observe the common operator operations on hermitian matrices, and adjust the commutator to make it closed.*

3. 1.6.4 (a) *Prove that $\det(\Omega^t) = \det(\Omega)$. and (b) prove that if U is unitary then $\det U$ is a complex number of unit modulus.*

4. *Verify that $R(\frac{1}{2}\pi \cdot \mathbf{i})$ is orthogonal by examining its matrix*

5. 1.7.1 *Prove that $\text{Tr}(\Omega) = \sum_j \Omega_{jj}$ satisfies*

$$(a) \text{Tr}(\Omega \Lambda) = \text{Tr}(\Lambda \Omega)$$

$$(b) \text{Tr}(ABC) = \text{Tr}(BCA) = \text{Tr}(CAB)$$

$$(c) \text{Tr}(\Omega) = \text{Tr}(U^\dagger \Omega U)$$

6.

Proof. 1. Suppose Ω and Λ are unitary. Then

$$(\Omega \cdot \Lambda)^\dagger = \Lambda^\dagger \Omega^\dagger$$

so that

$$(\Omega \cdot \Lambda) \cdot (\Omega \cdot \Lambda)^\dagger = \Omega (\Lambda \cdot \Lambda^\dagger) \Omega^\dagger = I$$

so that $(\Omega \Lambda)$ is unitary. This occurs because \dagger obeys the rules of a group inverse on all operators, and acts as inverse on unitary operators.

2. Hermitian operators are not a group under matrix multiplication, unless, the operators commute. For, suppose that Ω and Λ are hermitian. Then, $(\Omega \Lambda)^\dagger = \Lambda^\dagger \Omega^\dagger = \Lambda \Omega$ which is not generically equal to $\Omega \Lambda$. However, $(\Omega + \Lambda)^\dagger = \Omega + \Lambda$. The anticommutator, too is then an operation on the hermitian operators. $\{\Omega, \Lambda\}^\dagger = \{\Omega, \Lambda\}$. The commutator, by contrast, is not an operation on hermitian matrices since $[\Lambda, \Omega]^\dagger = -[\Omega, \Lambda]$. This may be corrected by complex conjugation, with $(i[\Omega, \Lambda])^\dagger = -i^*[\Omega, \Lambda] = i[\Omega, \Lambda]$.

3. Assuming n dimensions,

$$\det(\Omega) = \sum_{\sigma \in S^n} (-1)^{\text{sgn}(\sigma)} \Pi_j(\Omega)_{j\sigma(j)} = \sum_{\sigma \in S^n} (-1)^{\text{sgn}(\sigma)} \Pi_j \Omega_{\sigma(j),j}^t = \sum_{\sigma^{-1}} (-1)^{\text{sgn}(\sigma^{-1})} \Pi_j \Omega_{j,\sigma^{-1}(j)}^t =$$

$$\sum_{\tau \in S^n} (-1)^{\text{sgn}(\tau)} \Pi_j \Omega_{j,\tau(j)}^t = \det(\Omega^t)$$

4. Assuming that the determinant distributes over products, $\det(U) \det(U)^t = \det(U) \cdot (\det U)^\dagger = \det(I) = 1$ so that $\|\det U\|^2 = 1$. Thus, $\det U \in S^1$.

5.

$$R\left(\frac{\pi}{2}\right)|j\rangle = \begin{cases} |1\rangle & j = 1 \\ |3\rangle & j = 2 \\ -|2\rangle & j = 3 \end{cases}$$

This is a rotation matrix, and it clearly does not stretch space. Since it rotates only 1 plane, it preserves the relative order of the unit vectors. In each of their planes. The orientation and volume of space is the same, so that $\det R$ must be 1.

$$6. \text{Tr}(\Omega\Lambda) = \sum_{\alpha} \left(\sum_k \Omega_{\alpha k} \Lambda_{k\alpha} \right) = \sum_k \sum_{\alpha} \Lambda_{k,\alpha} \Omega_{\alpha,k} = \text{Tr}(\Lambda\Omega)$$

$$7. \text{Tr}(ABC) = \sum_{\alpha} \sum_k a_{\alpha,k} d_{k,\alpha} = \sum_{\alpha} \sum_k a_{\alpha,k} \left(\sum_{\beta} b_{k,\beta} c_{\beta,k} \right) = \sum_{\alpha,k,\beta} a_{\alpha,k} b_{k,\beta} c_{\beta,\alpha} \text{ Take } \beta \rightarrow \alpha \alpha \rightarrow k$$

and $k \rightarrow \beta$ to obtain

$$\sum_{\alpha,\beta,k} c_{\alpha,k} a_{k,\beta} b_{\beta,\alpha} = \text{Tr}(CAB)$$

Repeat to obtain

$$\sum_{\alpha,\beta,k} b_{\alpha,k} c_{k,\beta} a_{\beta,\alpha} = \text{Tr}(BCA)$$

$$8. \text{Using the above fact, } \text{Tr}(U^\dagger \Omega U) = \text{Tr}(\Omega(U \cdot U^\dagger)) = \text{Tr}(\Omega \cdot I) = \text{Tr}(\Omega)$$

□

2.2 The Postulates and Immediate Consequences

2.2.1 Single Degree of Freedom

1. The state of the particle is represented by a vector $|\Psi(t)\rangle$ in Hilbert space
2. The independent variables x and p of classical mechanics are represented by Hermitian operators X and P on hilbert space with the followign with the following coordinates in the X eigenbasis:

$$\langle x|X|x'\rangle = x\delta(x-x')$$

$$\langle x|P|x'\rangle = -i\hbar\delta'(x-x')$$

All variables corresponding to dependent variables $\omega(x,p)$ are given Hermitian operators

$$\Omega(X,P) = \omega(x \rightarrow X, p \rightarrow P)$$

3. If the particle is in a state $|\Psi\rangle$, measuring the variable correspondant to Ω gives an element of $E(\Omega)$ with probability

$$P(w) \propto \|\langle \omega | \Psi \rangle\|^2$$

and result in the change $|\Psi\rangle \rightarrow \mathbb{P}_\omega |\Psi(t)\rangle$

4. The state $|\Psi\rangle$ evolves according to

$$i\hbar \frac{d}{dt} |\Psi\rangle = H |\Psi\rangle$$

In classical mechanics, the state of a single particle had two independent degrees of freedom, x and p . Now, a general state $|\Psi\rangle$ has infinitely many components in a given basis, i.e. infinitely many degrees of freedom. States form a vector space as classical waves, since for classical waves, if $|f\rangle$ and $|g\rangle$ form displacements,

$$\alpha|f\rangle + \beta|g\rangle$$

is a possible displacement. However, the state $\alpha|f\rangle + \beta|g\rangle$ is a new state, not resembling $|f\rangle$ and $|g\rangle$. The superposition of possible states in quantum mechanics produces states which resemble the elements of the superposition.

2.3 Measurements: Evaluating a Dynamical Variable

In classical mechanics, given the state of a particle, (x,p) , the **dynamical variable** $\omega(x,p)$ is a uniquely determined real value; we say the particle has a value of $\omega(x,p)$. In quantum mechanics, dynamical variables correspond to Hermitian operators, the analogs of real numbers. What is the information of Ω given the state $|\Psi\rangle$? This is answered by postulate III.

Since Ω is Hermitian, it has real eigenvectors $E(\Omega) = \{\omega_\alpha\}$ and admits an orthonormal eigenbasis, $\mathcal{B}(\Omega) = \{|\omega_\alpha\rangle\}$. Utilizing the completeness relation afforded by the inner product, $|\Psi\rangle = (\sum_\alpha |\omega_\alpha\rangle \langle \omega_\alpha|) |\Psi\rangle$ Postulate three provides the following maximal information:

The relative probability that measuring $\omega(x,p)$ results in ω_α is:

$$\mathcal{P}(\omega_\alpha) = \|\langle \omega_\alpha | \Psi \rangle\|^2 = \langle \Psi | \omega_\alpha \rangle \langle \omega_\alpha | \Psi \rangle = \langle \Psi | \mathbb{P}_{\omega_\alpha} | \Psi \rangle = \langle \mathbb{P}_{\omega_\alpha} \Psi | \mathbb{P}_{\omega_\alpha} \Psi \rangle$$

For a non-normalizeable $|\Psi\rangle$, the proportions of relative probabilities are the total probabilistic information. If $|\Psi\rangle$ is normalizeable, the absolute probability is readily computed from the absolute probability, since $\mathcal{P}(\omega_\alpha) = \langle \Psi | \mathbb{P}_{\omega_\alpha}^2 | \Psi \rangle$

$$\mathcal{P}(\omega_\alpha) = \frac{\|\langle \omega_\alpha | \Psi \rangle\|^2}{\sum_\alpha \mathcal{P}(\omega_\alpha)} = \frac{\|\langle \omega_\alpha | \Psi \rangle\|^2}{\langle \Psi | \Psi \rangle}$$

Thus, information of Ω given Ψ is predictive but *probabalistic* regarding $E(\Omega)$, a collection of real values.

A first observations is that both Ψ and $c\Psi$, for nonzero c , obtain identical physical information, via identical probability distributions. Each physical state thus corresponds to a single dimensional subspace.

Degeneracy of Ω complicates the analysis above; suppose that for some $\omega \in E(\Omega)$ that $\dim(\mathbb{V}^\omega) > 1$. Since Ω is Hermitian, choose orthonormal eigenvectors

$$\{|\omega_\alpha\rangle\}$$

as a basis on \mathbb{V}^w . This gives a projection operator for \mathbb{V}^w which is:

$$\mathbb{P}_\omega = \sum_\alpha |\omega_\alpha\rangle\langle\omega_\alpha|$$

Clearly this should replace the statement in postulate three:

$$\mathcal{P}(\omega) = \langle\Psi|\mathbb{P}_\omega|\Psi\rangle$$

Whereas degeneracy is a complexity of eigenspaces, a complexity of the spectrum $E(\Omega) = \mathbf{W}$ is that it may be continuous, as in the case of the momentum operator. As always, an orthonormal eigenbasis $\mathcal{B}(\Omega) = \{|\omega\rangle : \omega \in \mathbf{W}\}$ can be chosen, which satisfies completeness:

$$|\Psi\rangle = \int_{\mathbf{W}} dw' |w'\rangle\langle w'| |\Psi\rangle$$

The coordinates of Ψ in this space are determined by the inner product, and may be thought of as valuation of the function on the space \mathbf{W} :

$$\Psi(w) = \langle\omega|\Psi\rangle$$

or Ψ in the \mathbf{W} space, which should be a smooth representation of Ψ . Here, $\|\langle\omega|\Psi\rangle\|^2 = \mathcal{P}(w)$ has the interpretation of a relative or absolute **probability density**, the latter in the case of normalizable Ψ . This is the likelihood of observing Ω with values in a range $\omega + d\omega$.

There is a special case where it is sensible to say the particle *has* a value of ω for some Ω . If $|\Psi\rangle \in V^\omega$ for any ω then, Postulate III says probability of observing ω is absolute. It is then sensible to say that $|\Psi\rangle$ has a value of ω for some observable Ω .

2.3.1 Multiple Dynamical Observables: Compatability

Postulate III informs what it means to evaluate a quantum dynamic variable. It revealed that, the only states for which a unique value of ω is guaranteed to be observed, are eigenstates. Since the state of a particle is immediately mapped to this eigenstate upon measurement, the measurement process acts as a filter. This analysis has not revealed, then, whether it is generally possible to take a generic state $|\Psi\rangle$, and produce states with well defined values for two dynamical variables Ω and Λ ; that is, whether or not there is a multiple filter which preserves well-definedness of certain observables.

Form an ensemble of identical states $|\Psi\rangle$, examine two dynamical variables Ω and Λ , and two possible eigenvalues (results of measurement) ω and λ . First, measure Ω , and collected the resulting states $|\omega\rangle$. Measure Λ immediately after, i.e. no time-evolution of the state $|\omega\rangle$. Collect the resulting states with $\Lambda = \lambda$. In general, the state no longer has $\Omega = \omega$, or any well defined value for Ω . The measurement process produced the following instantaneous collapse of state vectors:

$$|\Psi\rangle \rightarrow |w\rangle \rightarrow |\lambda\rangle$$

Recall from the completeness of the inner product and the existence of $\mathcal{B}(\Lambda)$

$$|w\rangle = \int_{\Lambda} d\lambda |\lambda\rangle\langle\lambda|w\rangle$$

In order that the final state $|\lambda\rangle$ has well defined values ω and λ , measuring Λ must preserve the state $|\omega\rangle$, and thus, $|\omega\rangle$ must be an eigenstate of Λ with eigenvalue λ . Thus, well defined values of multiple variables exist only occur for **simultaneous eigenstates**:

$$|\omega, \lambda\rangle$$

2.3.2 Probabilistic Interpretation of Multiple Observables

The question is then, given two operators Ω, Λ , can a simultaneous eigenbasis be constructed, so that a generic state $|\Psi\rangle$ can be expressed as a composition of components that have well defined values of two dynamical variables.

As in the finite dimensional case, if Ω and Λ are hermitian operators, a simultaneous eigenbasis can be found if $[\Omega, \Lambda] = 0$, or more importantly, $\forall \subset \mathbb{E}_0([\Omega, \Lambda])$. For a proof, see *Shankar* pp.43-46.

On the other hand, if $0 \notin E([\Omega, \Lambda])$ there are no simultaneous eigenvectors, there are no states which can have a well defined, single value of both Ω and λ .

In the first case the operators are called **compatible**, and in the second, **incompatible**.

2.3.3 The Quantum Ensemble: Cumulants and the Density Matrix

In classical theory, performing classical measurements can obtain the state of the system, (x, p) . To check classical deterministic predictions, simply measure (x_1, p_1) at some time t_1 and (x_2, p_2) at some time t_2 . This gives the full physical information about the system.

In quantum theory, quantum measurements do not obtain the state of the system before the measurement, but *after the measurement*. Since Postulate III makes only statistical predictions, individual measurements are not sufficient to evidence them. For example if the postulates predict, for a known state λ , that

$$|\lambda\rangle = \int_{\mathbf{W}} (d\omega) |\omega\rangle \langle \omega | \lambda \rangle$$

simply measuring ω on λ does not determine if $\mathcal{P}(\omega) = \|\langle \omega | \lambda \rangle\|^2$, and thus does not verify the predictions of Postulate III, and thus determine (verify) this expansion of the state $|\lambda\rangle$. Furthermore, after this measurement, $|\lambda\rangle \rightarrow |\omega\rangle$, and now the same particle cannot be used to test this prediction.

To measure $\mathcal{P}(\omega)$, Λ must be measured on a **quantum ensemble**, a large number N of particles in the same state $|\omega\rangle$. As $N \rightarrow \infty$ the proportions of particles in final states $|\omega\rangle$, giving values ω , should give $\|\langle \omega | \lambda \rangle\|^2$ with the deviation limiting to zero.

Given an operator Ω , and an ensemble of N states $|\Psi\rangle$, the state determines a mean (average) value of Ω that will be observed over the ensemble. This does not require solving the eigenvalue problem of the operator Ω , or the coordinates of $|\Psi\rangle$ in \mathbf{W} space. This is the **expectation value** of Ω as in statistics:

$$\langle \Omega \rangle = \int_{\mathbf{W}} \omega \mathcal{P}(\omega) = \int_{\mathbf{W}} \omega \|\langle \omega | \Psi \rangle\|^2 (d\omega) = \int_{\mathbf{W}} \omega \langle \Psi | \omega \rangle \langle \omega | \Psi \rangle = \int_{\mathbf{W}} \langle \Psi | \Omega | \omega \rangle \langle \omega | \Psi \rangle =$$

$$\langle \Psi | \Omega \left[\int_{\mathbb{W}} |\omega\rangle \langle \omega| (d\omega) \right] | \Psi \rangle = \langle \Psi | \Omega | \Psi \rangle$$

Ω weights each coordinate of the identity operator, before applying it to $|\Psi\rangle$. Using completeness to expand \mathbf{I} in the $\mathcal{B}(\Omega)$ basis, Ω weights each coordinate of \mathbf{I} by the eigenvalue corresponding to it. Then, $\Omega|\Psi\rangle$ is projected onto $|\Psi\rangle$ to give an integral sum of components squared times weighted probabilities. Note that $\langle \Omega \rangle = \omega \in \mathcal{E}(\Omega)$ only returns one of the eigenvalues if the state $|\Psi\rangle$ is an eigenstate in \mathbb{E}_ω .

Similarly, the standard deviation, or **uncertainty** is defined via

$$(\Delta\Omega)^2 \int_{\mathbb{W}} \mathbb{P}(\omega) (\omega - \langle \Omega \rangle)^2 = \langle \Psi | \Omega - \langle \Omega \rangle | \Psi \rangle = \langle \Omega - \langle \Omega \rangle \rangle$$

which weights each coordinate of the identity by the difference between the appropriate eigenvalue and the expectation.

These statistical functions pertain to averages over the ensemble, since an individual particle will only yield one of the accessible eigenvalues upon measurement of Ω .

The type of ensemble considered so far is called **pure**, with all N particles in an identical state $|\Psi\rangle$. An arbitrary (typical) ensembles of particles is specified by N particles (or systems), n_i of which are in some state $|i\rangle$. The probability that a random sytem is in state $|i\rangle$ is $p_i = \frac{n_i}{N}$. A general ensemble is called **mixed**, with arbitrary p_i , $i \in [m]$. This information is captured by the **density matrix**:

$$\rho = \sum_j p_j |j\rangle \langle j|$$

For a general ensemble, the statistical functions expressed before must now be reexpressed in light of the p_i . The **ensemble expectation** is

$$\langle \bar{\Omega} \rangle = \sum_j p_j \langle j | \Omega | j \rangle$$

and similarly for the **ensemble uncertainty**.

The density matrix can be used to obtain these quantities, for any hermitian operator Ω , abstractly. Supposing that the states $|i\rangle$ form an orthonormal basis:

$$\langle \bar{\Omega} \rangle = \sum_i \rho_{ii} \langle i | \Omega | i \rangle = \sum_i \langle i | \sum_j \rho_{i,j} | j \rangle \langle j | \Omega | i \rangle = \sum_i \langle i | \left[\sum_j | j \rangle (\rho \Omega)_{i,j} \right] = \text{Tr}(\rho \Omega)$$

Since $\mathcal{P}(\omega) = \|\langle \omega | \Psi \rangle\|^2 = \langle \mathbb{P}_\omega \rangle$ for a pure ensemble, the above fact shows

$$\mathcal{P}(\omega) = \text{Tr}(\rho \mathbb{P}_\omega)$$

Fact. 1. $\rho^\dagger = \rho$

2. $\text{Tr}(\rho) = 1$

3. $\rho^2 = \rho$ for a pure ensemble

4. $\rho = \frac{1}{k} I$ for a uniformly distributed ensemble

5. $\text{Tr}(\rho^2) \leq 1$

Proof. 1. $\langle j|\rho^\dagger|i\rangle = \langle i||\rho j\rangle = \rho_{jj}\langle i||j\rangle = \delta_{ij}\rho_{jj} \implies \rho^\dagger = \rho$

$$2. \text{Tr}(\rho^2) = \sum_i \langle i|\rho|i\rangle = \sum_i \rho_{ii} = \frac{1}{N} \sum_i n_i = 1$$

$$3. \text{Tr}(\rho^2) = \sum_{ii} \rho_{ii}^2 \leq (\text{Tr}(\rho))^2 = 1 \text{ by the CSIE}$$

□

2.4 The Postulates: Multiple Degrees of Freedom

1. (-)

2. Given a classical system in n cartesian coordinates $\{x_i\}_i$, there are n commuting hermitian operators $\{X_i\}_i$. Their simultaneous eigenbasis

$$\mathcal{B}(X_1, \dots, X_n) = \{|x_1, \dots, x_n\rangle\}$$

has normalization $\langle \mathbf{x}|\mathbf{x}'\rangle = \delta(\mathbf{x} - \mathbf{x}') = \prod_i \delta(x_i - x'_i)$. There are also n commuting hermitian operators $\{P_i\}$ corresponding to classical momenta p_i . These have the following presentation in the X basis:

$$\begin{aligned} \langle \mathbf{x}|X_i|\mathbf{x}'\rangle &= x_i \delta(\mathbf{x} - \mathbf{x}') \\ \langle \mathbf{x}|P_i|\mathbf{x}'\rangle &= -i\hbar \delta'_{x_i}(\mathbf{x} - \mathbf{x}') \end{aligned}$$

Dynamical variabls $\omega(X_i, P_j)$ ar operators $\Omega(X_i, P_i) = \omega(x_i \rightarrow X_i, p_i \rightarrow P_i)$.

3. (-)

4. (-)

Given an equation, once operator assignments are made in the cartesian coordinate basis, any change of coordiantes may be applied to solve the equation. This is the simplest basis to express operator assignments in. Take, for example:

$$\omega = \frac{p_1^2 + p_2^2 + p_3^2}{2m} + x_1^2 + x_2^2 + x_3^2$$

To solve the eigenvalue problem for $\Omega = \frac{1}{2m}[P_1^2 + P_2^2 + P_3^2] + X_1^2 + X_2^2 + X_3^2$

$$\Omega|\omega\rangle = \omega|\omega\rangle$$

Take $\langle x|$ on each side, and inside apply the completeness identity, to obtain:

$$\frac{-\hbar^2}{2m}[(\partial_x^2 + \partial_y^2 + \partial_z^2)] + x^2 + y^2 + z^2 \Psi_\omega(x, y, z) = \omega \Psi_\omega(x, y, z)$$

This may be transformed to spherical coordinates (r, θ, φ) to obtain

$$\frac{-\hbar^2}{2m}[\frac{1}{r^2}\partial_r(r^2\partial_r) + \frac{1}{r^2\sin(\theta)}\partial_\theta(\sin(\theta)\partial_\theta) + \frac{1}{r^2\sin^2(\theta)}\partial_\varphi^2]\Psi_\omega(r, \theta, \varphi) + r^2\Psi_\omega = \omega\Psi_\omega(r, \theta, \varphi)$$

If ω is expressed in spherical coordinates originally

$$\omega = \frac{1}{2m}[p_r^2 + \frac{p_\theta^2}{r^2} + \frac{p_\varphi^2}{r^2\sin(\theta)}] + r^2$$

there is no imediate operator assignment for p_r, p_θ, p_ψ which generates the previous equation in transformed coordinates.

2.5 The Schrodinger Equation: Generic Solutions and Simple 1-D Systems

While the prior discussion largely concerned the expression and application of Postulates I-III to a system at a given time, Postulate IV determines the evolution of this state with time.

$$i\hbar \frac{d}{dt} |\Psi(t)\rangle = \mathbf{H}(X_i, P_i) |\Psi(t)\rangle$$

where $\mathbf{H}(X_i, P_i) = \mathcal{H}(x_i \rightarrow X_i, p_i \rightarrow P_i)$ for the classical problem.

If the system were a harmonic oscillator in 1 dimension,

$$\mathcal{H}(x, p) = \frac{p^2}{2m} + \frac{m\omega^2}{2} x^2 \implies \mathbf{H} = \frac{P^2}{2m} + \frac{m\omega^2}{2} X^2$$

or in 3 dimensions

$$\mathcal{H}(x_i, p_i) = \frac{\sum_i p_i^2}{2m} + \frac{m\omega^2}{2} x^2 \implies \mathbf{H} = \frac{1}{2m} \sum_i P_i^2 + \frac{m\omega^2}{2} \sum_i X_i^2$$

Since \mathcal{H} corresponds to a classical dynamical variable, \mathbf{H} is a hermitian operator, and thus, admits an orthonormal eigenbasis $\mathcal{B}(\mathbf{H}) = \{|E\rangle\}$.

2.5.1 The Time-Independent Hamiltonian

Assuming that H is independent of time, schrodingers equation is a first order autonomous ODE in time, it is conceivable to construct a propogator $\mathbf{U}(t)$ so that

$$|\Psi(t)\rangle = \mathbf{U}(t) |\Psi(0)\rangle$$

Expanding schrodingers the equation using the eigenbasis $\mathcal{B}(\mathbf{H})$

$$i\hbar \frac{d}{dt} \int_{\mathcal{E}} (dE) |E\rangle \langle E| \Psi(t)\rangle = \mathbf{H}(X_i, P_i) \int_E (dE) |E\rangle \langle E| \Psi(t)\rangle \implies$$

$$i\hbar \frac{d}{dt} \langle E | \Psi \rangle(t) = E \langle E | \Psi \rangle(t)$$

$$\implies i\hbar \Psi(E, t) = E \Psi(E, t)$$

$$\implies \Psi(E, t) = \Psi(E, 0) \exp \frac{-i}{\hbar} Et$$

so that

$$U(t) = \int_{\mathcal{E}(\mathbf{H})} (dE) \exp \frac{-itE}{\hbar} |E\rangle \langle E|$$

Alternatively, since in $\mathcal{B}(\mathbf{H})$,

$$\langle E | \exp \frac{-it}{\hbar} \mathbf{H} | E' \rangle = (\exp \frac{-it}{\hbar} E') \delta(E - E')$$

this shows

$$U(t) = \exp \frac{-it}{\hbar} \mathbf{H}$$

The normal modes of this system are

$$|E(t)\rangle = |E\rangle \exp \frac{-iEt}{\hbar}$$

Since time is contained in the complex exponential, for any dynamical variable Ω and eigenvalue ω :

$$|E(t)\rangle = \|\langle\omega|E(t)\rangle\|^2 = \|\langle\omega|E\rangle\|^2$$

so that the physical probability distribution for any dynamical variable is independent of time; that is, the physical information afforded by them is the same at any time. The normal modes are thus called **stationary states**.

In fact, since \mathbf{H} is hermitian, the propagator $\mathbf{U}(t) = \exp \frac{-it}{\hbar} \mathbf{H}$ is a unitary operator, a "rotation" in hilbert space. Since unitary operators are inner product preserving, normalized states remain normalized: $\langle\Psi(t)|\Psi(t)\rangle = \langle\Psi(0)|U(t)^\dagger U(t)|\Psi(0)\rangle = \langle\Psi(0)|\Psi(0)\rangle$. **Small Time Dependence**

2.5.2 Example: A Free Particle

The simplest hamiltonian occurs with $V \equiv 0$ and $\mathbf{H} = \frac{P^2}{2m}$, or, a free particle. Obtaining the solution by solution by solving for $\mathcal{B}(\mathbf{H})$, consider the eigenvalue problem for \mathbf{H} .

$$\frac{P^2}{2m}|E\rangle = E|E\rangle$$

Since if $|p\rangle$ is an eigenfunction of \mathbf{P} it is also an eigenfunction of \mathbf{P}^2 , and hence of \mathbf{H} . Now, since $[\mathbf{P}, \mathbf{H}] = 0$ is clearly possible to choose a simultaneous eigenbasis for \mathbf{H} and \mathbf{P} . We begin with $\mathcal{B}(\mathbf{P})$ since this eigenbasis is already known. To observe the eigenvalues of this basis under \mathbf{H} , consider an eigenfunction $|p\rangle$ for some $p \in \text{spec}(\mathbf{P})$.

$$\frac{\mathbf{P}^2}{2m}|p\rangle = E|p\rangle \implies \frac{p^2}{2m} = E \implies p = \pm\sqrt{2mE}$$

Thus, the eigenspace of \mathbf{H} for value E is degenerate, with two orthogonal eigenfunctions $|E, \pm\rangle = |p = \pm\sqrt{2mE}\rangle$. These states corresponds to a particle with energy E moving to the left or right with momentum $\pm\sqrt{2mE}$. Since the energy eigenspace is degenerate, however, a state given by the superposition of these two momentum eigenstates is possible, wherein the energy is well defined for this state, but it $p = \pm\sqrt{2mE}$ are observed with probabilities.

To construct $\mathbf{U}(t)$, select $\mathcal{B}(\mathbf{H})$ by choosing $|E, \pm\rangle$ as the basis for each eigenspace. Since the energy eigenfunctions in each eigenspace $\mathbb{E}(\mathbf{H}, E)$ may be expressed using the eigenfunctions of the nondegenerate \mathbf{P} with the eigenvalues p , it is sufficient to index these eigenfunctions by p :

$$|E(p), \pm\rangle = |\pm p\rangle$$

Thus, the eigenvalues of \mathbf{P} are sufficient to express the eigenbasis $\mathcal{B}(\mathbf{H})$. Thus:

$$U(t) = \int_{\mathbb{S}(\mathbf{P})} (dp) |p\rangle \langle p| \exp \left[\frac{-it}{\hbar} E(p) \right]$$

2.5.3 Interpretation of the Propagator

To get a feel for the interpretation of the propagator, obtain its coordinates in the X basis by taking $\langle x | \mathbf{U}(t) | x' \rangle$. The free particle propagator just obtained is a good explicit example

$$\langle x | \mathbf{U}(t) | x' \rangle = \langle x | \int_{\mathcal{E}(\mathbf{P})} (dp \exp \left[\frac{-it}{\hbar} E(p) \right]) | p \rangle \langle p | x' \rangle = \frac{1}{2\pi\hbar} \int_{\mathcal{E}(\mathbf{P})} (dp) \exp \left[\frac{-itp^2}{2m\hbar} \right] \exp \frac{ip(x-x')}{\hbar} =$$

$$\left(\frac{2\pi\hbar it}{m}\right)^{-1/2} \exp \frac{im(x-x')^2}{2\hbar t}$$

by using quadratic factorization in the exponent. Given the state $|\Psi(t_1)\rangle$ at some time t_1 , the state at time t_2 is obtained by $|\Psi(t_2)\rangle = \mathbf{U}(t_2, t_1)|\Psi(t_1)\rangle$ which, for a time independent Hamiltonian, satisfies $\mathbf{U}(t_2, t_1) = \mathbf{U}(t_2 - t_1)$. Observing this transformation in the x basis appears as:

$$\langle x|\Psi(t_2)\rangle = \langle x|\int_{\mathcal{E}(X)} \mathbf{U}(t)|x'\rangle\langle x'|\Psi(t_2)\rangle \implies \Psi(x, t_1) = \int_{\mathcal{E}(X)} \mathbf{U}(x, x', t_2, t_1)\Psi(x', t_1)$$

Taking the case where $|\Psi(t_2)\rangle \equiv |x_0\rangle$ for some position $x_0 \in \mathcal{E}(\mathbf{X})$ shows the propagation of a state localized at a single position.

$$\Psi(x, t_2) = \int_{\mathcal{E}(\mathbf{X})} U(x, x', t_2, t_1)\delta(x' - x_0) = U(x, x_0, t_2, t_1)$$

Thus, the propagator tells the probability amplitude of a state $|x_0\rangle$, located at x_0 at time t_1 , arriving at x at time t_2 . For an arbitrary state $\Psi(x', t_1)$, the probability amplitude of the state (arriving) at x at time t_2 , is the sum over all propagation probabilities from $x' \rightarrow x$, weighted by the initial probability amplitude of the state occupying those positions. This interpretation may be given to the propagator in any basis corresponding to a dynamical variable Ω

$$\langle \omega|\Psi(t_2)\rangle = \langle \omega|\int_{\mathbf{W}} \mathbf{U}(t)|\omega'\rangle\langle \omega'|\Psi(t_1)\rangle$$

2.5.4 Particle in A 1-D Box

Now, consider a particle in a 1 dimensional system with a potential well, called a **box** or **infinite square well**

$$V(x) = \begin{cases} 0 & |x| < \frac{L}{2} \\ \infty & |x| > \frac{L}{2} \end{cases}$$

Consider the problem of solving for the eigenfunctions of \mathbf{H} again, using the x basis since this is the only immediate way to express $V(x)$. This reads:

$$\mathbf{H}|E\rangle = E|E\rangle \implies -\frac{\hbar^2}{2m} \frac{\partial^2 \Psi_E}{\partial x^2} + V(x) = E \implies \frac{\partial^2 \Psi_E}{\partial x^2} - \frac{\hbar^2}{2m}[V(x) - E] = 0$$

This problem must be solved separately in each region of space, and continuity imposed on the wave function. Suppose (though it can be shown mathematically) that the eigenfunctions for a hamiltonian \mathbf{H} where V is finite with value V_0 in the $|x| \geq \frac{L}{2}$ region, should approach the eigenfunctions of \mathbf{H} as $V_0 \rightarrow \infty$. Suppose $|E\rangle$ is an eigenfunction with $E < V_0$, since this is what all eigenvalues must satisfy in the $V_0 \rightarrow \infty$ limit anyway. Then, the solutions are immediate:

$$\frac{\partial^2 \Psi_E}{\partial x^2} - \frac{\hbar^2}{2m}[V_0 - E] = 0 \implies$$

with $k(E) = \sqrt{2m(V_0 - E)/\hbar^2}$

$$\Psi_E(x) = \alpha \exp -k(E)x + \beta \exp k(E)x$$

Since $V_0 - E > 0$, $k(E)$ is real and positive. All solutions with nonzero β must be rejected in the $x > 0$ region, and nonzero α in the $x < 0$ region, as they do not belong to real hilbert space, since

they are not normalizeable, even improperly. Taking the limit as $V_0 \rightarrow \infty$, $k(E) \rightarrow \infty$ also, for any value of E , and then

$$\Psi_E \rightarrow 0$$

Thus, the wave function must be zero in the region $|x| \geq L/2$.

Now, in the $|x| \leq L/2$ region, $V = 0$, giving the same equation as a free particle. With $k(E) = \sqrt{2mE}/\hbar$.

$$\begin{aligned} \frac{\partial^2 \Psi_E}{\partial x^2} + \frac{\hbar^2}{2m}[E] &= 0 \implies \\ \Psi_E(x) &= \alpha \exp ik(E)x + \beta \exp -ik(E)x \end{aligned}$$

In particular, $k(E)$ must be real, from the analysis of the momentum eigenvalue problem (nonreal $k(E)$ leads to unphysical momenta). Thus, $E \geq 0$. Furthermore, there are 2 constraints on Ψ_E , due to continuity at $L/2$ and $-L/2$. These may be expressed:

$$\begin{pmatrix} \exp ik(E)\frac{L}{2} & \exp -ik(E)\frac{L}{2} \\ \exp -ik(E)\frac{L}{2} & \exp ik(E)\frac{L}{2} \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Each boundary condition separately enforces:

$$(1). \alpha = -\exp[-ikL]\beta$$

But taken together, they require the linear dependence of the rows (vanishing of the determinant), so that:

$$(2). \exp ik(E)L - \exp -ik(E)L = 2i \sin(k(E)L) = 0 \implies k(E) = \frac{n\pi}{L}$$

Therefore, the solutions are

$$\Psi_E(x) = \alpha(\exp ik(E)x - \exp in\pi \exp -ik(E)x) = \begin{cases} \alpha(2i \sin(\frac{n\pi}{L}x) \in (\frac{2}{L})^{1/2} \sin(\frac{n\pi}{L}x) & n \in \mathbb{N}/2\mathbb{N} \\ \alpha(2 \cos(k(E)x) \in (\frac{2}{L})^{1/2} \cos(\frac{n\pi}{L}x) & \mathbb{N} - 2\mathbb{N} \end{cases}$$

Multiple boundary conditions result in:

$$E = \frac{\hbar^2 k(E)^2}{2m} = \frac{\hbar^2 n^2 \pi^2}{L^2 2m}$$

that is the discretization of the hamiltonian eigenvalues. Observe that $n = 0$ is not allowed since this gives

$$\Psi_{E_0} = 0$$

which is not normalizeable even to distribution. Thus, as distinct from the corresponding classical problem, the minimum physically allowed energy is nonzero, given by $E_1 > 0$. **Bound States**

The above example of a particle in a 1-dimensional box exhibited that appropriate boundary conditions on the **TISE** lead to quantization of energy, that is, discrete eigenvalues of **H**. The hamiltonian eigenstates in this case, were **bound states**, characterized by:

$$\lim_{x \rightarrow \infty} = \lim_{x \rightarrow -\infty} = \lim_{|x| \rightarrow \pm \infty} \Psi(x) = 0$$

These should occur due to the potential which prevents the particle from escaping to ∞ . As in classical mechanics, these states should occur where $V(\pm\infty) > E$ so that it must be inaccessible by the particle. In fact, it can be shown that if a particle has energy E , with

$$V_{\pm\infty} > E$$

then Ψ is clearly a bound state. It can be shown mathematically (see *Shankar* pp.160 for an intuitive argument) that if a system has energy E

$$V(\pm\infty) > E \implies \Psi_E \text{ is bound} \iff \lim_{|x| \rightarrow \pm\infty} \Psi_E = 0$$

and that the collection of **available energy levels for bound states are always discretized**.

2.5.5 Uncertainty Principle in the 1-D Box: A Nonzero Minimum Energy

Recall that, in the case of a particle in a 1-dimensional box, the minimum observable energy was not zero, as in classical mechanics for a particle at rest, but

$$\min_{n \in \mathbb{N}} \{E_n = \frac{\hbar^2 \pi^2 n^2}{2mL^2}\} = \frac{\hbar^2 \pi^2}{2mL^2}$$

What of the momentum distribution then in this minimum energy state? Recall that, in a bound state $|E_n\rangle$, the generic solution for the TISE shows that physical probability distributions obtained from eigenfunctions of \mathbf{H} are time independent. Since expectation value must be independent, $\langle \mathbf{P} \rangle$ is a time constant. Since the state $|E_n\rangle$ is bound, however, it satisfies

$$\lim_{|x| \rightarrow \infty} \Psi(x, t) = 0$$

for all time. If $\langle \mathbf{P} \rangle$ were positive or negative, the particle would, on average, drift left or right, and escape to infinity. Thus, it must be that $\langle \mathbf{P} \rangle = 0$, and thus, $\langle \Delta \mathbf{P} \rangle = \langle (\mathbf{P} - \langle \mathbf{P} \rangle)^2 \rangle = \langle \mathbf{P}^2 \rangle$.

Now examine $\langle \mathbf{H} \rangle$ to obtain a bound on the minimum energy, since $\langle \mathbf{H} \rangle = E_n$ for a bound state $|E_n\rangle$. The potential in the center of the well is zero, and any state, is zero in the x-basis for coordinates $|x| \geq \frac{L}{2}$, so that in the x basis

$$\langle \mathbf{H} \rangle = \frac{1}{2m} \langle \mathbf{P}^2 \rangle = \frac{1}{2m} (\Delta \mathbf{P})^2 \implies \Delta \mathbf{P} = \sqrt{2m \langle \mathbf{H} \rangle}$$

so that $\langle \mathbf{H} \rangle$ is constrained by the uncertainty principle to satisfy

$$(\Delta \mathbf{X}) \sqrt{2m \langle \mathbf{H} \rangle} \geq \frac{\hbar}{2} \implies \langle \mathbf{H} \rangle \geq \frac{\hbar^2}{8m(\Delta \mathbf{X})^2}$$

Now, since $|x\rangle|E_n\rangle = 0$ for $|x| \geq \frac{L}{2}$, that is, the state is confined to a box of finite length

$$(\Delta \mathbf{X}) < \frac{L}{2}$$

so that

$$\langle \mathbf{H} \rangle = E_n \geq \frac{\hbar^2}{2mL^2}$$

Though the minimum energy is π^2 higher than this lower bound, this inequality shows the minimum energy, $\min E_n$, to be nonzero.

General Theorems on the 1-D TISE

Theorem. *There is no degeneracy in 1-dimensional bound states*

Proof. Suppose that $|\Psi_E\rangle$ and $|\Theta_E\rangle$ are two eigenstates with energy $E < \lim_{V \pm \infty}$. Take the eigenvalue problem for \mathbf{H} in position coordinates as previously to obtain:

$$\begin{aligned} \frac{-\hbar^2}{2m} \frac{\partial^2}{\partial x^2} \Psi_E(x) + V(x) \Psi_E &= E \Psi_E \\ \frac{-\hbar^2}{2m} \frac{\partial^2}{\partial x^2} \Theta_E(x) + V(x) \Theta_E &= E \Theta_E \end{aligned}$$

This gives

$$\Theta \partial_x^2 \Psi - \Psi \partial_x^2 \Theta = 0$$

which, by Green's identity, is

$$\partial_x (\Psi \partial_x \Theta - \Theta \partial_x \Psi) = 0$$

So

$$\Psi \partial_x \Theta - \Theta \partial_x \Psi = c$$

for some constant c . The boundary condition $\Psi, \Theta \rightarrow 0$ as $|x| \rightarrow \infty$ shows $c \equiv 0$ so that

$$\Psi = \dots$$

□

2.6 The Uncertainty Principle and the (Semi)Classical Limit

In classical mechanics, given the state of a system (x, p) , a dynamical variable ω is a determined real value $\omega(x, p)$. In quantum mechanics, given the state of the system $|\Psi\rangle$, Postulates II says that a dynamical variable ω is a hermitian operator Ω , and Postulate III says that measurements of Ω yield elements of $\mathcal{E}(\Omega)$, with a probability of the outcome of measurement given by $\mathcal{P}(\omega) = \|\mathbb{P}_\omega \Psi\|^2$. The mean (expectation) and variance (uncertainty) of the probability distribution is completely characterized by:

$$\begin{aligned} \langle \Omega \rangle &= \langle \Psi | \Omega | \Psi \rangle \\ (\Delta \Omega)^2 &= \langle \Psi | (\Omega - \langle \Omega \rangle)^2 | \Psi \rangle \end{aligned}$$

Zero uncertainty occurs only when $|\Psi\rangle \in \mathbb{V}_\omega$ for some $\omega \in \mathcal{E}(\Omega)$.

Given two hermitian operators Ω, Λ , an arbitrary state $|\Psi\rangle$ yields uncertainties $(\Delta \Omega)$ and $(\Delta \Lambda)$. The relationship between these operators, and the outcomes of measuring both operators successively, have been explored. If the operators commute, for example, then $\Delta \Omega = \Delta \Lambda = 0$ for any simultaneous eigenstate. If the operators commutator has a null zero eigenspace, however $\mathbb{E}(0, [\Omega, \Lambda]) = \emptyset$, then $\Delta \Omega$ and $\Delta \Lambda$ cannot be simultaneously zero, since there are no simultaneous eigenstates. What if one wishes to obtain the states which simultaneously minimize the uncertainty of Ω and Λ ?

The relation between the uncertainties of two dynamical variables should generally depend on the state, and the Heisenberg uncertainty relations will provide a lower bound for $\Delta \Omega \cdot \Delta \Lambda$. Of particular interest are the cases where this lower bound is independent of the state.

The General Uncertainty Relation

Take two hermitian operators, Ω and Λ , and consider

$$(\Delta \Omega)^2 \cdot (\Delta \Lambda)^2 = \langle \Psi | (\Omega - \langle \Omega \rangle)^2 | \Psi \rangle \langle \Psi | (\Lambda - \langle \Lambda \rangle)^2 | \Psi \rangle$$

If $\hat{\Omega} = \Omega - \langle \Omega \rangle$ and $\hat{\Lambda} = \Lambda - \langle \Lambda \rangle$, then both of these are hermitian as well, and write

$$(\Delta\Omega)^2 \cdot (\Delta\Lambda)^2 = \langle \hat{\Omega}\Psi | \hat{\Omega}\Psi \rangle \cdot \langle \hat{\Lambda}\Psi | \hat{\Lambda}\Psi \rangle$$

Since the right side is a product of inner products, it is susceptible to the CSIE, so apply it

$$(\Delta\Omega)^2 \cdot (\Delta\Lambda)^2 = \langle \hat{\Omega}\Psi | \hat{\Omega}\Psi \rangle \cdot \langle \hat{\Lambda}\Psi | \hat{\Lambda}\Psi \rangle \geq \|\langle \Psi | \hat{\Omega}\hat{\Lambda} | \Psi \rangle\|^2$$

This lower bound is independent of the order of the product of $\hat{\Lambda}\hat{\Omega}$ (apply complex conjugation and \dagger), so resolve $\hat{\Omega}\hat{\Lambda}$ into its hermitian and antihermitian parts:

$$\hat{\Omega}\hat{\Lambda} = \frac{1}{2}([\Omega, \Lambda]_+ + [\Omega, \Lambda]) \implies$$

$$(\Delta\Omega)^2 \cdot (\Delta\Lambda)^2 \geq \|\langle \Psi | \hat{\Omega}\hat{\Lambda} | \Psi \rangle\|^2 = \|\langle \Psi | \frac{1}{2}([\Omega, \Lambda]_+ + [\Omega, \Lambda]) | \Psi \rangle\|^2 = \frac{1}{4}\langle [\hat{\Omega}, \hat{\Lambda}]_+ \rangle^2 + \frac{1}{4}\|\langle [\hat{\Omega}, \hat{\Lambda}] \rangle\|^2$$

since the inner product is linear, and the anticommutator is purely hermitian, giving a real expectation value, and the commutator is antihermitian, giving a purely imaginary expectation value.

This is the most **general uncertainty relation**:

$$(\Delta\Omega)^2 \cdot (\Delta\Lambda)^2 \geq \frac{1}{4}\langle [\hat{\Omega}, \hat{\Lambda}]_+ \rangle^2 + \frac{1}{4}\|\langle [\Omega, \Lambda] \rangle\|^2 \geq \frac{1}{4}\|\langle [\Omega, \Lambda] \rangle\|^2$$

by positive definiteness of the first term. The first inequality is just cauchy schwartz, and the second is positive definiteness of the anticommutator.

Now, $[\Omega, \Lambda] = i\Gamma$ where Γ is hermitian. If Ω and Λ are canonically conjugate, that is,

$$\Gamma = \hbar$$

then this fact gives the **canonically conjugate uncertainty**

$$(\Delta\Lambda)^2(\Delta\Omega)^2 \geq \frac{\hbar^2}{4}$$

as is the case for \mathbf{X} and \mathbf{P} .

2.6.1 Minimum Uncertainty Wavepacket: The Gaussian

The general uncertainty relation provides a lower bound on the joint uncertainties of Ω and Λ . It is easily observed that the state which jointly minimizes $\Delta\mathbf{X}$ and $\Delta\mathbf{P}$, called the **minimum uncertainty wavepacket**, $|\Theta\rangle$, must saturate both inequalities in the general uncertainty relation (i.e. due to Cauchy Schwarz, and positive definiteness). Precisely, $|\Theta\rangle$ must have

$$\Delta\mathbf{X}\delta\mathbf{P} = \frac{\hbar}{2} \implies$$

$$(1). \text{ Cauchy-Schwarz } c\hat{\mathbf{P}}|\Theta\rangle = \hat{\mathbf{X}}|\Theta\rangle$$

$$(2). \text{ Positive Definite } \langle [\hat{\mathbf{X}}, \hat{\mathbf{P}}] \rangle = 0$$

where expectations are computed on Θ assuming its existence.

From (1) in position coordinates,

$$c(x - \langle X \rangle)\theta(x) = (-i\hbar \frac{d}{dx} - \langle \mathbf{P} \rangle)\Theta(x) \implies$$

$$cx\Theta(x) = [-i\hbar \frac{d}{dx} - \langle \mathbf{P} \rangle]\Theta(x) \implies$$

$$\frac{i}{\hbar}[\langle \mathbf{P} \rangle + cx]\Theta(x) = \partial_x \Theta(x) \implies \Theta(x) = \alpha \exp\left[\frac{i}{\hbar}\langle \mathbf{P} \rangle x + \frac{i}{\hbar} \frac{cx^2}{2}\right]$$

after changing variables so that the origin occurs at $\langle X \rangle$.

Then, apply condition (2). and its adjoint to obtain

$$(c + \bar{c})\langle \Theta | X^2 | \theta \rangle = 0$$

so that $c = -\bar{c} \implies c = ri$ where $r \in \mathbb{R}$. Then,

$$\Theta(x) = \alpha \exp\left[\frac{i}{\hbar}\langle P \rangle x + \frac{-rx^2}{2\hbar}\right]$$

where $r > 0$ since negative r would give nonphysical Θ . This may be reexpressed using $\langle X \rangle$:

$$\Theta(x) = \alpha \exp\left[\frac{i}{\hbar}\langle P \rangle (x - \langle x \rangle) + \frac{-(x - \langle x \rangle)^2}{2\Delta^2}\right]$$

where $\Delta^2 = \frac{\hbar}{c}$. Since $\Delta, \langle \mathbf{X} \rangle$ are arbitrary and determine $\langle \mathbf{P} \rangle$ up to α , this says that **the minimum uncertainty wave packet is gaussian in position coordinates with arbitrary uncertainty and expectation.**

2.6.2 Ehrenfests Theorem: Time Evolution of Expectation Values

In the hopes of recovering classical mechanics in the appropriate limit, first consider the time evolution of expectation values of quantum dynamical variables, $\frac{d}{dt}\langle \Omega \rangle$. Since these are hermitian, and the state evolves according to Schrodingers equation, observe:

$$\frac{d}{dt}\langle \Omega \rangle = \frac{d}{dt}\langle \Psi | \Omega | \Psi \rangle = \frac{d}{dt}(\langle \Psi | \Omega | \Psi \rangle + \langle \Psi | (\frac{\partial}{\partial t} \Omega) | \Psi \rangle + \langle \Psi | (\Omega) | \Psi \rangle)$$

Schrodingers equation and its adjoint tell:

$$\frac{d}{dt}(|\Psi\rangle) = \frac{-i}{\hbar} \mathbf{H} |\Psi\rangle$$

$$\frac{d}{dt}(\langle \Psi |) = \frac{i}{\hbar} \mathbf{H}^\dagger \langle \Psi | = \frac{i}{\hbar} \mathbf{H} \langle \Psi |$$

giving **Ehrenfest's Theorem**

$$\frac{d}{dt}(\langle \Omega \rangle) = \frac{i}{\hbar} \langle \mathbf{H} \Omega \rangle - \frac{i}{\hbar} \langle \Omega \mathbf{H} \rangle + \langle \frac{d}{dt} \Omega \rangle = \frac{i}{\hbar} \langle [\mathbf{H}, \Omega] \rangle + \langle \frac{d}{dt} \Omega \rangle$$

or

$$\frac{d}{dt} \langle \Omega \rangle = \frac{i}{\hbar} \langle [\mathbf{H}, \Omega] \rangle$$

for time independent operators. Examining this result for position gives

$$\frac{d}{dt} \langle \mathbf{X} \rangle = \frac{i}{\hbar} [\mathbf{H}, \mathbf{X}] = \frac{i}{2m\hbar} \langle [\mathbf{P}^2, \mathbf{X}] \rangle = \frac{i}{2m\hbar} \langle ([\mathbf{P}, \mathbf{X}] \mathbf{P} + \mathbf{P} [\mathbf{P}, \mathbf{X}]) \rangle = \frac{i}{2m\hbar} \langle -2i\hbar \mathbf{P} \rangle = \frac{\langle \mathbf{P} \rangle}{m}$$

which says

$$m \langle \dot{\mathbf{X}} \rangle = \langle \mathbf{P} \rangle \iff \langle \dot{\mathbf{X}} \rangle = \langle \frac{\partial \mathbf{H}}{\partial \mathbf{P}} \rangle$$

which is similar to Hamilton's first equation. Writing $\frac{\partial \mathbf{H}}{\partial \mathbf{P}}$ means to compute a formal derivative on a function representable as power series. This is possible for $\mathbf{H}(\mathbf{X}, \mathbf{P})$. Repeating for momentum gives:

$$\frac{d}{dt} \langle \mathbf{P} \rangle = \frac{i}{\hbar} \langle [\mathbf{H}, \mathbf{P}] \rangle = \frac{i}{\hbar} \langle [\mathbf{V}(\mathbf{X}), \mathbf{P}] \rangle = \frac{i}{\hbar} \langle i\hbar \frac{\partial V}{\partial x}(x) \rangle = -\langle \frac{\partial \mathbf{V}}{\partial \mathbf{X}} \rangle$$

by evaluating the commutator and expectation in the position basis. This is similar to Hamilton's second equation. **The time evolution of the classical variables is exactly replaced by the time evolution of their expectation values.**

We expect, however, that these should reduce exactly to Hamilton's equations, however, when considering a macroscopic particle with a mass of $m = 1g$.

The Classical Limit: Part 1

Observe that momentum and position expectation values do in fact obey Hamilton's equations if **the mean value of functions is well approximated by the classical function evaluated at the mean.** That is, if

$$(1). \dot{x}_0 = \langle \dot{\mathbf{X}} \rangle = \langle \frac{\partial \mathbf{H}}{\partial \mathbf{P}} \rangle \simeq \frac{\partial \mathcal{H}(x_0, p_0)}{\partial p}$$

$$(2). \dot{p}_0 = \langle \dot{\mathbf{P}} \rangle = \langle \frac{\partial \mathbf{V}}{\partial \mathbf{X}} \rangle \simeq -\frac{\partial \mathcal{H}}{\partial x}(x_0)$$

Observe that this first equation is exact. Since

$$\frac{\partial \mathbf{H}}{\partial \mathbf{P}} = \frac{\langle P \rangle}{m} = \frac{p_0}{m} = \frac{\partial \mathcal{H}}{\partial p}(x_0, p_0)$$

The second, however, requires that $V'(x_0) = \frac{\partial \mathcal{H}}{\partial x}(x_0, p)$ replace $\langle \frac{\partial \mathbf{V}}{\partial \mathbf{X}} \rangle$. Consider this expectation in the position basis, and represent V in its power series about x_0 :

$$V'(x) = V'(x_0) + V''(x_0)(x - x_0) + \sum_k \frac{V^{k+1}(x_0)}{k!} (x - x_0)^k$$

Now, if all higher order terms vanished, the classical fact would hold. But these higher order corrections are a quantum mechanical fact, as the particle, unlike in classical mechanics, which is sensitive only to force at x_0 , responds to the potential at all neighboring points. However, the first order term (in V'), and all odd order terms in V' above 1, vanish under the expectation value. Thus:

$$\langle V'(x) \rangle = V'(x_0) + \langle \sum_{k \geq 1} V^{2k+1}(x_0) 2k! (x - x_0)^{2k} \rangle$$

so that **if the potential is strictly second order in \mathbf{X} , the approximation is exact.**

It is also a good approximation if the potential varies little over the "size" of the wave packet., perhaps approximated by the uncertainty. **If the force, \mathbf{V}' , varies over macroscopic scales, and the uncertainty is microscopic approximation is valid.**

In general, states with small fluctuation of the observed values of momentum and position about their mean/expectation values, i.e. states which jointly minimize $\Delta \mathbf{X}$ and $\Delta \mathbf{P}$, should provide exactness. To observe a recovery of Hamilton's equations from expectation values, then, the state over which these values are taken should resemble classical states, with definite position and momentum.

This is not quantum mechanically possible, but, there are states which jointly minimize uncertainty in both position and momentum, with arbitrary expectation values for position and momentum. These are elements of the family of $|\Theta\rangle$ observed in the previous discussion. These states were specified by $\langle \mathbf{P} \rangle = p_0$, $\langle \mathbf{X} \rangle = x_0$, and the position uncertainty Δ . These specify states from the $|\Theta\rangle$ family, so choose such a normalized state as the initial state

$$\langle x|p_0, x_0, \Delta\rangle = N \exp\left[\frac{i}{\hbar}\langle P \rangle(x - \langle x \rangle) + \frac{-(x - \langle x \rangle)^2}{2\Delta^2}\right]$$

If $\Delta \mathbf{X} \propto 10^{-13}$ then $\Delta \mathbf{P} \propto 10^{-14}$. Further, if $m = 1g$ then $\Delta \mathbf{V} \propto 10^{-14} \frac{cm}{s}$. **At classical scale**, these deviations are truly negligible, if not undetectable. Thus, the system is observed as having a nearly single valued position and momentum, which, though truly is an eigenvalue of position or momentum, is well-approximated by the expectation value.

These small uncertainties, which assured well-approximation of Hamilton's equations by mean values, however, do evolve in time. Consider the example of the free particle $\mathbf{H} = \frac{\mathbf{P}^2}{2m}$ in the initial state

$$\Theta(x', 0) = \left(\frac{1}{\pi\Delta^2}\right)^{1/4} \exp\left[\frac{ip_0}{\hbar}x' - \frac{x'^2}{2\Delta^2}\right]$$

The uncertainty of this state is $\Delta \mathbf{X} = \frac{\Delta}{\sqrt{2}}$ and $x_0 = \langle X \rangle = 0$. The free particle propagator is

$$\mathbf{U}(t) = \int_{\mathcal{E}(p)} (dp) \exp\left[-\frac{iE(p)t}{\hbar}\right] |p\rangle\langle p|$$

so that at later times

$$\Psi(x, t) = \int_{\mathcal{E}(\mathbf{X}')} \langle x|\mathbf{U}(t)|x'\rangle \langle x'|\Theta\rangle$$

Evaluating this explicitly this gives

$$\|\Psi(x)\|^2 = P(x, t) = \frac{1}{\sqrt{\pi}} \frac{1}{\sqrt{\Delta^2 + (\frac{\hbar t}{m\Delta})^2}} \exp \frac{-[x - \frac{p_0 t}{m}]^2}{\Delta^2 + (\frac{\hbar t}{m\Delta})^2}$$

Now, this shows $\langle \mathbf{P} \rangle = p_0$ for all time, and further, $\langle x \rangle = \frac{p_0 t}{m}$, the relation guaranteed by Ehrenfest's theorem. The evolution of

Energy-Time Uncertainty

3. Kardar's *Statistical Physics of Particles*: Notes and Problems

The Boltzmann Equation Scattering in a Central Force Field

First, recall that $\frac{d}{dt} \mathbf{L} = \frac{d}{dt}(\mathbf{r} \times \mathbf{p}) = \frac{d}{dt}(\mathbf{r}) \times \mathbf{p} + \mathbf{r} \times \frac{d}{dt} \mathbf{p} = \mathbf{r} \times \mathbf{F} = \mathbf{N}$. In a central force field, from a purely classical approach, $\mathbf{F}(r) = f(r)\hat{\mathbf{r}}$ where $f(r) = -\partial_r \mathcal{V}(r)$. Since the torque on the particle about the origin is $\mathbf{N} = \mathbf{r} \times \mathbf{F} = 0$, the angular momentum of the particle is clearly conserved. If the motion were not confined to a plane, the direction of \mathbf{L} would change in time. The motion is therefore confined to a plane.

Recall the result that for an s-particle density $f_s((p_1, q_1), \dots, (p_s, q_s))$

$$\frac{\partial f_s}{\partial t}(\gamma_1, \dots, \gamma_s) = \int_{\Gamma_{s+1}} \sum_{n=1}^s \partial_{q_n} V(q_n - q_{s+1}) \partial_{p_s} f_{s+1}((\gamma_1, \dots, (\gamma_s), (p_{s+1}, q_{s+1}))) + \{H_s, p_s\}$$

For $s = 1$, this is:

$$\begin{aligned} & \left[\frac{\partial}{\partial t} - [\partial_{q_1} U(q_1)] \cdot \partial_{p_1} + \frac{p_1}{m} \cdot \partial_{q_1} \right] f_1(p_1, q_1) \\ &= \int_{\Gamma_2} \partial_{q_1} V(q_1 - q_2) \cdot f_2(\gamma_1, \gamma_2) \end{aligned}$$

3 Kinetic Theory of Gases

Kinetic theory attempts to account for macroscopic behavior, the thermodynamic description of constrained macroscopic quantities, by considering a microscopic description of the system. This microscopic description, for a system of N particles say, is given by the collection of all position and momenta of these N particles at a given time, $\boldsymbol{\mu}(t) = \Pi_i(q_i(t), p_i(t))$, called the **microstate** of the system. The collection of all possible microstates is **phase space** particles as

$$\boldsymbol{\Gamma} = \Pi_{i=1}^N(q_i, p_i)$$

so that a microstate is a point in $\boldsymbol{\Gamma}$. If $\mathcal{H}(\mathbf{p}, \mathbf{q})$ is the Hamiltonian of the microstate, then $\mu(t)$ evolves in $\boldsymbol{\Gamma}$.

$$\dot{\mathbf{q}}_i = \frac{\partial \mathcal{H}}{\partial \mathbf{p}_i} \text{ and } \dot{\mathbf{p}}_i = -\frac{\partial \mathcal{H}}{\partial \mathbf{q}_i} \quad \forall i$$

Since the thermodynamic description of a system in equilibrium, its **macrostate**, \mathbf{M} , uses few state functions, while the space of possible microstates of a system is very large, there must be a many to one correspondence $\mu(t) \rightarrow \mathbf{M}$. If $d\mathcal{N}(\mathbf{p}, \mathbf{q})$ is the number of representative states of \mathbf{M} within a phase-space differential volume about (\mathbf{p}, \mathbf{q}) , define **phase space density**

$$\rho(\mathbf{p}, \mathbf{q}) = \lim_{\mathcal{N} \rightarrow \infty} \frac{d\mathcal{N}(\mathbf{p}, \mathbf{q})}{\mathcal{N}}$$

Clearly

$$\int_{\boldsymbol{\Gamma}} \rho(\mathbf{p}, \mathbf{q}) d\boldsymbol{\Gamma} = 1$$

and to compute ensemble averages for observables $\mathcal{O}(\mathbf{p}, \mathbf{q})$ one writes

$$\langle \mathcal{O} \rangle = \int_{\boldsymbol{\Gamma}} \rho(\mathbf{p}, \mathbf{q}) \mathcal{O}(\mathbf{p}, \mathbf{q})$$

If the representative correspondence specifies an exact microstate, the system is in a **pure state** whereas if the knowledge of the system is probabilistic, and the representation correspondence specifies an ensemble, it is said to be in a **mixed state**.

Liouville's Theorem $\rho(\mathbf{p}, \mathbf{q})$ behaves as an incompressible fluid, that is,

$$\frac{d}{dt} \rho = 0 \iff \frac{\partial \rho}{\partial t} = -\{p, \mathcal{H}\}$$

Proof. Consider $d\Gamma(p, q)$ states in the volume $d\Gamma(\mathbf{p}, \mathbf{q})$. After a time δt , they are in the vicinity of some point, described to first order in δt by Hamilton's equations:

$$\begin{aligned}\mathbf{q}'_i &= q_i(t) + \delta t \dot{q}_i(t) + \delta t^2 \ddot{q}_i \dots = q_i(t) + \delta t \frac{\partial \mathcal{H}(t)}{\partial \mathbf{p}_i} \\ \mathbf{p}'_i &= p_i(t) + \delta t \dot{p}_i(t) + \delta t^2 \ddot{p}_i(t) = p_i(t) + \delta t \left(-\frac{\partial \mathcal{H}(t)}{\partial \mathbf{q}_i} \right)\end{aligned}$$

Thus

$$\begin{aligned}dq'_i &= dq_i(t) + (\delta t) \partial_{q_i} \partial_{\mathcal{H}(t)} p_i dq \\ dp'_i &= dp_i(t) + (\delta t) \partial_{p_i} \partial_{\mathcal{H}(t)} p_i\end{aligned}$$

so that since $d\Gamma' = \Pi_i d\Gamma'_i$ and

$$dq'_i \cdot dp'_i = dq_i(t) \cdot dp_i(t) + \delta t \left[\frac{\partial^2 \mathcal{H}(t)}{\partial \mathbf{q}_i \partial \mathbf{p}_i} - \frac{\partial^2 \mathcal{H}(t)}{\partial \mathbf{q}_i \partial \mathbf{p}_i} \right] + \mathcal{O}(\delta t^2) = dq_i \cdot dp_i$$

to first order in δt . This shows that $d\Gamma'_i = d\Gamma_i$ and hence $d\Gamma = d\Gamma'$. Thus, $\rho(\mathbf{p}, \mathbf{q}) = \rho(\mathbf{p}', \mathbf{q}')$. Along each given microstate, then, the probability density is invariant as it evolves with time. This may be understood as ρ behaving as the density of an incompressible fluid. In summary:

$$\rho(\mathbf{p}, \mathbf{q} + \delta t \dot{\mathbf{p}}, \mathbf{q}, t + \delta t) = \rho \implies \frac{d}{dt} \rho = 0$$

and

$$\begin{aligned}\frac{d}{dt} \rho &= \frac{\partial \rho}{\partial t} + \sum_{\alpha=1}^N \frac{\partial \rho}{\partial \mathbf{p}_\alpha} \dot{\mathbf{p}}_\alpha + \frac{\partial \rho}{\partial \mathbf{q}_\alpha} \dot{\mathbf{q}}_\alpha = \frac{\partial \rho}{\partial t} + \{\rho, \mathcal{H}\} = 0 \implies \\ &\implies \frac{\partial \rho}{\partial t} = -\{\rho, \mathcal{H}\}\end{aligned}$$

□

3.1 Consequences of Liouville's Theorem

1. Time reversal symmetry of ρ :

2.

$$\begin{aligned}\frac{d}{dt} \langle \mathcal{O} \rangle &= \int_{\Gamma} \frac{\partial \rho}{\partial t}(\mathbf{p}, \mathbf{q}) \mathcal{O}(\mathbf{p}, \mathbf{q}) = \int_{\Gamma} -\{p, \mathcal{H}\} \cdot \mathcal{O} = \\ &= \int_{\Gamma} \sum_{\alpha} \left[-\frac{\partial \mathcal{O}}{\partial \mathbf{p}_\alpha} \cdot \frac{\partial \mathcal{H}}{\partial \mathbf{q}_\alpha} + \frac{\partial \mathcal{O}}{\partial \mathbf{q}_\alpha} \cdot \frac{\partial \mathcal{H}}{\partial \mathbf{p}_\alpha} \right] + \mathcal{O} \cdot \left(\frac{\partial^2 \mathcal{H}}{\partial q_\alpha \partial p_\alpha} - \frac{\partial^2 \mathcal{H}}{\partial p_\alpha \partial q_\alpha} \right) \langle \{ \mathcal{O}, \mathcal{H} \} \rangle\end{aligned}$$

via integration by parts

3. If \mathbf{M} is an equilibrium state, then the ensemble corresponding to it must have:

$$\frac{d}{dt} \langle \mathcal{O} \rangle = 0$$

for all observables \mathcal{O} . This is achieved with $\frac{\partial \rho}{\partial t} = -\{\rho, \mathcal{H}\} = 0$. This will occur, for example, if $\rho_{eq}(\mathbf{p}, \mathbf{q}) \equiv \rho(\mathcal{H}(\mathbf{p}, \mathbf{q}))$. In this case ρ will be constant on surfaces of constant energy. If there are other conserved quantities, satisfying $\{L_n, \mathcal{H}\} = 0$, then any density of the form:

$$\rho(\mathbf{p}, \mathbf{q}) = \rho(\mathcal{H}, L_i)$$

will be stationary.

4. Equal probability of accessible states

3.2 BBKGY Hierarchy - Many Particle Density Evolution

The full phase space density function contains much more information than needed to describe equilibrium properties. In particular, it may provide information that any of N particles in the system has achieved momentum and position (p, q) . Assuming that the probabilities are invariant under identical particle exchange, this clearly defines the **one particle density**

$$f(p_0, q_0) = \sum_{i=1}^N \int_{\Gamma} \delta(p_i - p_0, q_i - q_0) \rho(\mathbf{p}, \mathbf{q}) = N \int_{\Gamma} \delta(p_1 - p_0) \rho(\mathbf{p}, \mathbf{q})$$

Then, for an s -particle density $f_s((p_1, q_1), \dots, (p_s, q_s))$

$$f_s(\gamma_1, \dots, \gamma_s) = \frac{N!}{(N-s)!} \int_{\Gamma - \Gamma^s} \rho(p_{s+1}, q_{s+1}, \dots, p_N, q_N) = \frac{N!}{(N-s)!} p_s(\gamma_1, \dots, \gamma_s)$$

where p_s is the probability of s particular particles achieving these values, not any s of N identical particles.

Now, supposing that these particles are subject to the following simple Hamiltonian, including only a two-body interaction and external potential:

$$\mathcal{H}(\mathbf{p}, \mathbf{q}) = \sum_i \frac{p_i^2}{2m} + U(q_i) + \frac{1}{2} \sum_{i,j} V(q_i - q_j)$$

To evaluate its time evolution, divide this Hamiltonian as follows

$$\begin{aligned} H &= H_s + H_{N-s} + H_I \\ H_s &= \sum_{i \leq s} \frac{p_i^2}{2m} + U(q_i) + \frac{1}{2} \sum_{j \leq s} V(q_i - q_j) \\ H_{N-s} &= \sum_{i \geq s} \frac{p_i^2}{2m} + U(q_i) + \frac{1}{2} \sum_{j \geq s} V(q_i - q_j) \\ H_I &= \sum_{i \leq s, s < j} V(q_i - q_j) \end{aligned}$$

and observe

$$\frac{\partial p_s}{\partial t}(\gamma_1, \dots, \gamma_s) = \int_{\Gamma^{N-s}} \frac{\partial p}{\partial t} = - \int_{\Gamma^s} \{p, \mathcal{H}_s + \mathcal{H}_{N-s} + \mathcal{H}_I\} = -\{p_s, \mathcal{H}_s\} + - \int_{\Gamma^{N-s}} \{p, \mathcal{H}_{N-s}\} + \{p, \mathcal{H}_I\} =$$

Now,

$$\{p_s, \mathcal{H}_s\} = \sum_{\alpha > s} (\partial_{q_\alpha} p \cdot \frac{p_\alpha}{m} - \partial_{p_\alpha} p \cdot F_\alpha)$$

each term of which is zero under integration of $d\Gamma^\alpha$ via integration by parts.

Thus, only the interaction term remains under the integral:

$$\frac{\partial p_s}{\partial t} = -\{p_s, \mathcal{H}_s\} - \int_{\Gamma^{N-s}} \{p, \mathcal{H}_I\}$$

which after considering integration by parts again, shows

$$\begin{aligned}
- \int_{\Gamma^{N-s}} \{p, \mathcal{H}_I\} &= \int_{\Gamma^{N-s}} \sum_{j \leq s} \frac{\partial p}{\partial p_j} \left(\sum_{i > s} \frac{\partial V(q_j - q_i)}{\partial q_j} \right) = (N-s) \int_{\Gamma^{N-s}} \sum_{j \leq s} \frac{\partial p}{\partial p_j} \cdot \frac{\partial V(q_j - q_{s+1})}{\partial q_j} = \\
\sum_{j \leq s} (N-s) \int_{\Gamma^{s+1}} \frac{\partial V(q_j - q_{s+1})}{\partial q_j} \frac{\partial}{\partial p_j} \int_{\Gamma^{N-(s+1)}} p &= (N-s) \sum_{j \leq s} \int_{\Gamma^{s+1}} \frac{\partial V(q_j - q_{s+1})}{\partial q_j} \frac{\partial}{\partial p_j} p_{s+1}
\end{aligned}$$

by symmetry. Replacing p_s with f_s

$$\frac{\partial f_s}{\partial t} = -\{f_s, \mathcal{H}_s\} + \sum_{j \leq s} \int_{\Gamma^{s+1}} \frac{\partial V(q_j - q_{s+1})}{\partial q_j} \frac{\partial}{\partial p_j} f_{s+1}$$

The first collection of terms in this sum are called **streaming terms**, while the second collection are known as **collision terms**.

This shows f_s to depend on f_{s+1} , a system of equations at least as complicated for solving the entire phase space density ρ . To terminate this hierarchy of equations, a physically motivated approximation must be made. One such approximation yields the following.

3.3 The Boltzmann Equation

For $s = 1$, the BBKGY hierarchy reads

$$\begin{aligned}
\left[\frac{\partial}{\partial t} - [\partial_{q_1} U(q_1)] \cdot \partial_{p_1} + \frac{p_1}{m} \cdot \partial_{q_1} \right] f_1(p_1, q_1) \\
= \int_{\Gamma_2} \partial_{q_1} V(q_1 - q_2) \cdot \partial_{p_1} f_2(\gamma_1, \gamma_2)
\end{aligned}$$

and $s = 2$

$$\begin{aligned}
\left[\frac{\partial}{\partial t} - (\partial_{q_1} U(q_1) \cdot \partial_{p_1} + \partial_{q_2} U(q_2) \partial_{p_2} + \frac{\partial V(q)}{\partial q} \cdot \partial_{p_1} + -\frac{\partial V(q)}{\partial q} \partial_{p_2}) + \frac{p_1}{m} \partial_{q_1} + \frac{p_2}{m} \partial_{q_2} \right] f_1(p_1, q_1) \\
= \int_{\Gamma^3} \partial_{q_1} V(q_1 - q_3) \partial_{p_1} f_3 + \partial_{q_2} V(q_2 - q_3) \cdot \partial_{p_2} f(\gamma_1, \gamma_2)
\end{aligned}$$

By a coarse graining of the interparticle interaction, one extracts length and time scales from the external and interaction potentials, including **collision duration** τ_c and **mean free time** τ_x . In the dilute limit, for short range interactions,

$$\frac{\tau_c}{\tau_x} \ll 1$$

This justifies taking the interactions terms in $s = 2$ to zero, implying that the two body density evolves as an isolated two particles system, giving the Boltzmann equation.

3.3.1 Time Reversibility and The H-Theorem

Since the exact evolution of many body densities is time reversible, the steady-state solutions for the full phase space density $\rho = p_N$ are not attractors of generic non-equilibrium densities. If, however, p_1 is governed by Boltzmann's equation, it will approach an equilibrium density non-reversibly, as proved in the following theorem:

The H-Theorem If $f_1(p_1, q_1, t)$ satisfies the Boltzmann equation, then $\frac{d}{dt}H \leq 0$ with

$$H(t) = \int_{\Gamma^1} f_1(p_1, q_1, t) \cdot \ln f_1(p_1, q_1, t)$$

Observe that $\frac{d}{dt}H(t) = \int_{\Gamma^1} \frac{\partial f_1}{\partial t} (\ln f_1 + 1) = \int_{\Gamma^1} \frac{\partial f_1}{\partial t} \ln f_1$ so that, using the Boltzmann equation

$$\frac{\partial f_1}{\partial t} = -\left(\frac{\partial f}{\partial q_1} \frac{p_1}{m} - \frac{\partial f}{\partial p_1} \frac{\partial U(q_1)}{\partial q_1}\right) - \int_{\Gamma_p^2 \times S^2} \left|\frac{\partial \sigma}{\partial \Omega}\right| \cdot |v_1 - v_2| \cdot [f_1(p_1, q_1) \cdot f(p_2, q_2) - f_1(p'_1, q_1) f(p'_2, q_2)]$$

$$\frac{d}{dt}H(t) = \int_{\Gamma^1} \ln f_1 \left[-\left(\frac{\partial f}{\partial q_1} \frac{p_1}{m} - \frac{\partial f}{\partial p_1} \frac{\partial U(q_1)}{\partial q_1}\right) - \int_{\Gamma_p^2 \times S^2} \left|\frac{\partial \sigma}{\partial \Omega}\right| \cdot |v_1 - v_2| \cdot [f_1(p_1, q_1) \cdot f(p_2, q_2) - f_1(p'_1, q_1) f(p'_2, q_2)] \right]$$

3.4 Conservation Laws

Locally conserved quantities are unchanged by the two-body collisions, satisfying:

$$\chi(p_1, q, t) + \chi(p_2, q, t) = \chi(p'_1, q, t) + \chi(p'_2, q, t)$$

Define a **local particle density** and **local expectation value**

$$\eta(q) = \int_{\Gamma_p^1} f_1(p, q, t) dp$$

$$\langle \mathcal{O} \rangle(q, t) = \frac{1}{\eta(q)} \int_{\Gamma_p^1} f_1(p, q, t) \mathcal{O}(p, q) dp$$

Define the function

$$J(\chi)(q) := \int_{\Gamma^1(p)} \chi(q, p) \frac{d}{dt}^{coll.} f_1(p, q) dp$$

and observe that

$$J(\chi)(q) \equiv 0$$

.

$$J(\chi)(q) = - \int_{\Gamma_p^1} \chi(p_1, q) \cdot \int_{\Gamma_p^2 \times S^2} dp d\Omega |\sigma(\Omega)| \left| \frac{p_1 - p_2}{m} \right| (f_1(p_1, q) f(p_2, q) - f(p'_1, q) f(p'_2, q))$$

Since interchanging p_1 and p_2 does not affect the integrand, this gives:

$$J(\chi)(q) = - \int_{\Gamma_p^1 \oplus \Gamma_2 \oplus S^2} \frac{1}{2} (\chi(p_1, q) + \chi(p_2, q)) dp d\Omega |\sigma(\Omega)| \frac{p_1 - p_2}{m} |(f_1(p_1, q)f(p_2, q) - f(p'_1, q)f(p'_2, q))|$$

Finally, since (p_1, p_2, Ω) to (p'_1, p'_2, Ω') is a legitimate coordinate transformation with unit jacobian (due to incompressibility of phase space density) $J(\chi)(q) =$

$$- \int_{\Gamma_p^1 \oplus \Gamma_p^2 \oplus S^2} \frac{1}{2} [\frac{1}{2} [\chi(p_1, q) + \chi(p_2, q) - (\chi(p'_1, q) + \chi(p'_2, q))] dp d\Omega |\sigma(\Omega)| \frac{p_1 - p_2}{m} |(f_1(p_1, q)f(p_2, q) - f(p'_1, q)f(p'_2, q))| = 0$$

Since, due to the Boltzmann equation

$$[\partial_t + (\frac{p}{m} \partial_q + F^U(q) \partial_p)] f_1(p, q) = - \frac{d^{coll}}{dt} f_1(p, q)$$

this shows

$$J(\chi)(q) = \int_{\Gamma_p^1} \chi(p, q) [\partial_t + (\frac{p}{m} \partial_q + F^U(q) \partial_p)] f_1(p, q) = - \int_{\Gamma_p^1} \chi(p, q) \frac{d^{coll}}{dt} f_1(p, q) \implies$$

$$\int_{\Gamma_p^1} [\partial_t + (\frac{p}{m} \partial_q + F^U(q) \partial_p)] (\chi(p, q) f_1(p, q)) - f_1 [\partial_t + (\frac{p}{m} \partial_q + F^U(q) \partial_p)] (\chi(p, q)) = 0$$

via integration by parts. Since the integrand must be zero, this gives the following conservation law, componentwise:

$$\partial_t (\eta(q) \cdot \langle \chi \rangle(q)) + \partial_q (\eta(q) \langle \frac{p}{m} \chi \rangle(q)) - [\eta(q) \langle \partial_t \chi \rangle(q) + \eta(q) \langle \partial_q \frac{p}{m} \chi \rangle(q) + \eta(q) F^U(q) \langle \partial_p \chi \rangle] = 0$$

This yields the hydrodynamic equations.

3.5 The Hydrodynamic Equations

Letting χ count particle number, so that it reads 1 for a position and momentum gives

(1).

$$\partial_t \eta(q) + \partial_q (\eta(q) \cdot u) = 0$$

where $u = \langle \frac{q}{m} \rangle(q)$ is the local velocity. This shows that the time evolution of the local density is due to the velocity of particles at that position, i.e., a particle current.

(2). Since any linear function of p is conserved in a collision, take $\chi(q, p) = \mathbf{c}_\alpha = \mathbf{p}_\alpha - \mathbf{u}_\alpha$ to obtain 9 equations:

$$d_t (\eta(q) \langle c_\alpha \rangle) + \partial_\beta (\eta(q) \langle \frac{p_\beta}{m} c_\alpha \rangle) - [\eta(q) \langle \partial_t c_\alpha \rangle + \eta(q) \langle \partial_\beta \frac{p_\beta}{m} c_\alpha \rangle + \eta(q) F_\beta^U(q) \langle \partial_{p_\beta} c_\alpha \rangle] = 0$$

Now, the local expectation of c_α is zero, and

$$\int_{\Gamma_p^1} f_1 \partial_t \frac{p}{m} (dp) = 0 \implies$$

$$\partial_\beta[\eta(q)\langle(c_\beta + u_\beta)c_\alpha\rangle] + \eta(q)\partial_t u_\alpha = -u_\beta\eta(q)\partial_\beta u_\alpha + \eta(q)\frac{F_\beta^U}{m} \implies$$

which rearranges to

$$\eta(q)\partial_t u_\alpha + u_\beta\eta(q)\partial_\beta u_\alpha = -\partial_\beta[\eta(q)\langle(c_\beta + u_\beta)c_\alpha\rangle] + \eta(q)\frac{F_\beta^U}{m} = -\partial_\beta[\eta(q)\langle c_\beta \cdot c_\alpha\rangle] + \eta(q)\frac{F_\beta^U}{m}$$

Define $P_{\alpha,\beta}(q) = m \cdot \eta(q)\langle c_\alpha c_\beta \rangle$ so that this equation is:

$$\begin{aligned} \eta(q)\partial_t u_\alpha + u_\beta\eta(q)\partial_\beta u_\alpha &= -\frac{\partial_\beta}{m}[\eta(q)P_{\beta,\alpha}] + \eta(q)\frac{F_\beta^U}{m} \implies \\ m\partial_t u_\alpha + mu_\beta\partial_\beta u_\alpha &= -\frac{1}{\eta(q)}\partial_\beta P_{\alpha,\beta}(q) + F_\beta^U(q) \end{aligned}$$

The lefthandside should be $m\frac{d}{dt}u$, the mass times acceleration of an element of the fluid, which should be equal to the net force on that element. The net force includes contributions from the pressure tensor, as well as the external force on the liquid.

(3). Define **average local kinetic energy** via

$$\varepsilon := \langle \frac{mc^2}{2} \rangle = \langle \frac{p^2}{2m} - p \cdot u + \frac{u^2}{2m} \rangle$$

With $\chi = \frac{mc^2}{2}$

$$\partial_t(\eta(q)\langle \frac{mc^2}{2} \rangle) + \partial_\alpha[\eta(q)\langle \frac{p_\alpha}{m} \frac{mc^2}{2} \rangle] - \eta(q)\langle \partial_t \frac{c^2 m}{2} \rangle - \eta(q)\langle \partial_\alpha(\frac{p^\alpha}{m} \frac{mc^2}{2}) \rangle - \eta(q)\langle F_\alpha^U(q)\partial_{p_\alpha} \frac{mc^2}{2} \rangle =$$

$$\partial_t(\eta(q))\varepsilon + \eta(q)\partial_t \varepsilon + (\partial_\alpha \varepsilon)\langle (c_\alpha + u_\alpha) \frac{mc^2}{2} \rangle + \varepsilon(q)\partial_\alpha(\langle (c_\alpha + u_\alpha) \frac{mc^2}{2} \rangle) + \eta(q)m(\partial_t u_\beta)\langle c_\beta \rangle +$$

$$(\partial_\alpha u_\beta)\eta(q)m\langle \frac{(u_\alpha + c_\alpha)}{m} c_\beta \rangle + \eta(q)F_\alpha^U m\langle c_\alpha \rangle =$$

$$\partial_t(\eta(q))\varepsilon + \eta(q)\partial_t \varepsilon + (\partial_\alpha \varepsilon)\langle (c_\alpha + u_\alpha) \frac{mc^2}{2} \rangle + \varepsilon(q)\partial_\alpha(\langle (c_\alpha + u_\alpha) \frac{mc^2}{2} \rangle) + (\partial_\alpha u_\beta)\eta(q)m\langle \frac{(u_\alpha + c_\alpha)}{m} c_\beta \rangle =$$

$$\partial_t(\eta(q))\varepsilon + \eta(q)\partial_t \varepsilon + (\partial_\alpha \varepsilon)\langle (c_\alpha + u_\alpha) \frac{mc^2}{2} \rangle + \varepsilon(q)\partial_\alpha(\langle (c_\alpha + u_\alpha) \frac{mc^2}{2} \rangle) + (\partial_\alpha u_\beta)\eta(q)m\langle \frac{(u_\alpha + c_\alpha)}{m} c_\beta \rangle =$$

$$\partial_t(\eta(q))\varepsilon + \eta(q)\partial_t \varepsilon + (\partial_\alpha \varepsilon)\langle (c_\alpha + u_\alpha) \frac{mc^2}{2} \rangle + \varepsilon(q)\partial_\alpha(\langle u_\alpha \frac{mc^2}{2} \rangle) + \varepsilon(q)\partial_\alpha(\langle c_\alpha \frac{mc^2}{2} \rangle) + m\eta(q)(\partial_\alpha u_\beta)\langle \frac{c_\alpha c_\beta}{m} \rangle =$$

$$\partial_t(\eta(q))\varepsilon + \eta(q)\partial_t \varepsilon + (\partial_\alpha \varepsilon)\langle (c_\alpha + u_\alpha) \frac{mc^2}{2} \rangle + \eta(q)m\varepsilon(q)\partial_\alpha(\langle u_\alpha \frac{mc^2}{2} \rangle) + \varepsilon(q)\partial_\alpha h_\alpha + (\partial_\alpha u_\beta)P_{\alpha,\beta} = 0$$

where **h** is the **local heat flux** and $u_{\alpha,\beta}$ is the **rate of strain tensor**

$$\mathbf{h} = \frac{\eta(q)m}{2}\langle c_\alpha c^2 \rangle$$

$$u_{\alpha,\beta} = \frac{1}{2}(\partial_\beta u_\alpha + \partial_\alpha u_\beta)$$

$$\begin{aligned}\varepsilon \partial_t \eta(q) + \eta(q) \partial_t \varepsilon + \varepsilon \partial_\alpha (\eta u_\alpha) + \eta(q) u_\alpha \partial_\alpha \varepsilon + \partial_\alpha h_\alpha + (u_{\alpha,\beta}) P_{\alpha,\beta} &= 0 \\ \partial_t \varepsilon + u_\alpha \partial_\alpha \varepsilon &= -\frac{1}{\eta(q)} \partial_\alpha h_\alpha + -\frac{1}{\eta(q)} P_{\alpha,\beta} u_{\alpha,\beta}\end{aligned}$$

In summary, the **Hydrodynamic Equations** are:

$$\begin{aligned}(1) \quad & \partial_t \eta(q) + \partial_q \cdot (\eta u) = 0 \\ (2) \quad & m \partial_t u_{q_\alpha} + m u_\beta \partial_\beta u_{q_\alpha} = -\frac{1}{\eta(q)} \partial_{q_\beta} P_{\alpha,\beta}(q) + F_{q_\beta}^U(q) \\ (4) \quad & \partial_t \varepsilon + u_\alpha \partial_{q_\alpha} \varepsilon = -\frac{1}{\eta(q)} [\partial_{q_\alpha} h_\alpha + P_{\alpha,\beta} u_{\alpha,\beta}]\end{aligned}$$

Expressions for $P_{\alpha,\beta}$ and h_α are necessary to solve for ε, η , and u .

Homework 2

Question 2. Show that for any observable $\mathcal{O}(\mathbf{p}, \mathbf{q}(\mu))$ that

1.

$$\frac{d}{dt} \mathcal{O} = \{\mathcal{O}, \mathcal{H}(\mathbf{p}, \mathbf{q})\}$$

2. If $\langle \{\mathcal{O}, \mathcal{H}\} \rangle = 0$ for any \mathcal{O} , show that $\{\mathcal{H}, p\} = 0$.

Proof. 1.

$$\frac{d}{dt} \mathcal{O} = \sum_\alpha \frac{\partial \mathcal{O}}{\partial p_\alpha} \cdot \dot{p}_\alpha + \frac{\partial \mathcal{O}}{\partial q_\alpha} \cdot \dot{q}_\alpha = \sum_\alpha -\frac{\partial \mathcal{O}}{\partial p_\alpha} \frac{\partial \mathcal{H}}{\partial q_\alpha} + \frac{\partial \mathcal{O}}{\partial q_\alpha} \frac{\partial \mathcal{H}}{\partial p_\alpha} = \{\mathcal{O}, \mathcal{H}\}$$

since from Hamiltons equations

$$-\frac{\partial \mathcal{H}}{\partial q_\alpha} = \dot{p}_\alpha \text{ and } \dot{q}_\alpha = \frac{\partial \mathcal{H}}{\partial p_\alpha}$$

2. $\langle \{\mathcal{O}, \mathcal{H}\} \rangle = \int_\Gamma \mathcal{O}(\mathbf{p}, \mathbf{q}) \{\mathcal{H}, p\}$ so that if the expectation is zero for all \mathcal{O} , then $\{\mathcal{H}, p\} = 0$. □

Question 3. 3.8. Zeroth-Order Hydrodynamics

Proof. 1.

$$P_{\alpha,\beta}^0 = \eta(q) k_B T \delta_{\alpha,\beta} \text{ and } h_\alpha^0 = 0 \text{ and } \varepsilon = \frac{3}{2} k_B T$$

2. Simply substitute the above into the hydrodynamic equations, observing that $T(q, t)$.

3. Using the differential operator $D^\alpha \partial_t + u_\alpha \partial_\alpha$ the first and third equation read

$$\begin{aligned}(1) \quad & D^\alpha (\eta) = -\eta u_{\alpha,\alpha} \implies D^\alpha (\ln \eta) = -u_{\alpha,\alpha} \\ (3) \quad & \frac{3}{2} k_B D^\alpha (T) = -k_B T u_{\alpha,\alpha} \implies \frac{3}{2} D^\alpha \ln T = -u_{\alpha,\alpha}\end{aligned}$$

so that

$$D^\alpha (\ln(\eta) - \frac{3}{2} \ln T) = 0 \implies D^\alpha \ln(\eta T^{-\frac{3}{2}}) = 0$$

4.

$$H(t) = \int_{\Gamma_q} d^3q \eta \cdot \left[\ln\left(\left(\frac{1}{\pi 2m k_B T}\right)^{3/2} \eta\right) - \frac{3}{2} \right] = \int_{\Gamma_q} d^3q \cdot \eta \cdot \left[\ln(\eta T^{-3/2}) - \frac{3}{2} \ln(\pi 2m \pi k_B) - \frac{3}{2} \right]$$

5.

$$\begin{aligned} \frac{d}{dt} H(t) &= \int_{\Gamma_q} d^3q [(\partial_t \eta) \ln(\eta T^{-3/2})] + \eta [\partial_t \ln(\eta T^{-3/2})] = \int_{\Gamma_q} d^3q [-\partial_\alpha (\eta u_\alpha) \ln(\eta T^{-3/2})] + -u_\alpha \eta \partial_\alpha \ln(\eta T^{-3/2}) \\ &= - \int_{\Gamma_q} d^3q \partial_\alpha (\eta u_\alpha \ln(\eta T^{-3/2})) = 0 \end{aligned}$$

since this is a full derivative.

6.

□

4 Classical Statistical Mechanics

4.1 The Microcanonical Ensemble

As in the kinetic theory of gases, an ensemble of microstates $\{\mu_\alpha\}$ corresponding to a particular **macrostate** is examined; in the absence of heat or work a macrostate is

$$M := (E, \mathbf{x})$$

This collection of all correspondant microstates $\{\mu_\alpha\}$ is called the **microcanonical ensemble**.

One intends to obtain $p_M(\mu)$, the probabilities over points in phase space correspondant to the macrostate, for an ensemble corresponding to an equilibrium state. Each microstates evolve in time according to its Hamiltonian, $\mathcal{H}(\mu)$, and since the microcanonical ensemble $\{\mu\}$ is chosen correspondant to \mathbf{M} , and \mathcal{H} conserves total energy, microstates in the ensemble are constrained to $\mathcal{H}(\mu) = E$. Define $\Omega(\mathbf{M})$ to be the volume of this surface in phase space. In the absence of other conserved quantities, then:

Boltzmann's Assumption

$$p_M(\mu) = \frac{1}{\Omega(\mathbf{M})} \begin{cases} 1 & \mathcal{H}(\mu) = E \\ 0 & \text{else} \end{cases}$$

i.e. $p_M(\mu)$ is uniform across $\Omega(\mathbf{M})$.

Defining Entropy

$$S(E, \mathbf{x}) := k_B \ln \Omega(\mathbf{M})$$

Compare this to the the probability distribution entropy corresponding to the phase space probability distribution p_M over the continuous random variable μ :

$$\mathcal{S}(p_M) = - \int_{\Gamma} p_M \ln(p_M) = \int_{\Omega} \frac{1}{\Omega(\mathbf{M})} \ln\left(\frac{1}{\Omega(\mathbf{M})}\right) = \ln(\Omega(\mathbf{M}))$$

Therefore, $S(E, \mathbf{x}) = k_B S(p_M)$, where k_B is appended to give entropy the proper dimensions $[S(\mathbf{M})] = \frac{[E]}{[T]}$ as in classical thermodynamics.

4.2 Laws of Thermodynamics

1. 0th Law: Transitivity of Thermal Equilibrium, Existence of Empirical Temperature

Consider bringing two isolated systems \mathbf{M}_1 and \mathbf{M}_2 into contact, allowing heat exchange, but not work. The combined macrostate \mathbf{M} has a thermodynamic coordiante of total energy $E = E_1 + E_2$. The states in its microcanonical ensemble correspond to joint microstates of the two component systems, representable as: $\mu = \mu_1 \otimes \mu_2 \implies \Gamma_{\text{total}} = \Gamma_1 \otimes \Gamma_2$. This definition makes the following probabalistic fact clear for **isolated systems**

$$\Omega(\mathbf{M}) = \Omega(\mathbf{M}_1) \cdot \Omega(\mathbf{M}_2)$$

Since

$$\mathcal{H}(\mu) = \mathcal{H}(\mu_1) + \mathcal{H}(\mu_2) = E$$

for these systems in contact, μ are constrained only to have $\mathcal{H}(\mu_1) + \mathcal{H}(\mu_2) = E$. Therefore, there is a 1-parameter collection of pairs of surfaces (choose, say, E_1) satisfy $\mathcal{H}(\mu) = E$. Therefore, for **two systems exchanging heat, but mechanically isolated**

$$\begin{aligned} \Omega(\mathbf{M}(E)) &= \int_{\mathbb{R}} \Omega(\mathbf{M}_2(E_1)) \cdot \Omega(\mathbf{M}_2(E - E_1)) = \\ &= \int_{\mathbb{R}} \exp \frac{1}{k_B} [S(\Omega(E_1, \mathbf{x}_1)) + S(\Omega(E_2, \mathbf{x}_2))] \end{aligned}$$

Now, supposing due to the extensivity of entropy, that the integral is approximately equal to the maximum value of the integrand, one obtains

$$S(\Omega(E, \mathbf{x})) = S(\Omega(E_1, \mathbf{x}_1)) + S(\Omega(E_2, \mathbf{x}_2))$$

If E_1^* is the value at which the integrand is maximized, then:

$$\partial_{E_1}|_{x_1} S(E_1^*) - \partial_{E_2} S(E - E^*) = 0$$

In conclusion, near $E_1 = E^*$ and $E_2 = E - E^*$, there is an exponentially larger number of microstates, or, the volume Ω is exponentially larger.

The two systems come into contact with initial energies $E_1(0)$ and $E_2(0)$, exchange energy, and explore a new space of microstates. Once the system achieves equilibrium, so that the assumptoin of equal apriori probabilities is again valid, the volume of surfaces in Γ with $E_1 = E^*$ and $E_2 = E - E_1^*$ has the greatest volume, and so, the probability of finding microstates in this region is the greatest. If this distribution of microstates bears on the likeliness of finding E_1 and E_2 at a particular value, it is clear that these are the likeliest values.

At equilibrium, this equation is satisfied, providing a relation between two functions of state which is achieved at equilibrium. These functions of state may be chosen as empirical temperatures; for consistency with classical results,

$$\partial_E|_x S(\Omega(E, \mathbf{x})) = \frac{1}{T}$$

is chosen.

2. We want to express dE and dQ utilizing S , so that in an adiabatic process, one sees $dE = dW$. For reversible process, $dS = 0$. Thus, for a reversible, adiabatic process where mechanical work $d\mathbf{x} \cdot \mathbf{J}$ is done on the system, $dS = 0$:

$$dS = (\partial_E S \cdot \mathbf{J} + \partial_{\mathbf{x}} S) d\mathbf{x} \implies \frac{1}{T} \cdot \mathbf{J} = \partial_{\mathbf{x}} S$$

and thus

$$\frac{\mathbf{J}}{T} = -\partial_{\mathbf{x}} S$$

This way, since

$$dS = \frac{\partial S}{\partial E} dE + \frac{\partial S}{\partial \mathbf{x}} d\mathbf{x} = \frac{1}{T} dE + -\frac{\mathbf{J}}{T} \cdot d\mathbf{x} \implies$$

Thus,

$$TdS + \mathbf{J} \cdot d\mathbf{x} = dE$$

and **heat** is identified as $dQ = TdS$.

3. The definition of the equilibrium point in the joint system is characterized in the text by
 - (a) The preponderance of available microstates in the microcanonical ensemble occurring near energies E_1^* and E_2^*
 - (b) $\Omega(E) \approx \Omega(E_1^*) \cdot \Omega(E_2^*)$ due to extensivity of entropy.

$$\Omega(E) = \int_{\mathbb{R}} \exp \frac{1}{k_B} [S(\Omega(E_1, \mathbf{x}_1)) + S(\Omega(E_2, \mathbf{x}_2))]$$

This implies that the integral is exponential in particle number, and therefore may be approximated by the integrand maximum.

N increases $\implies \Gamma$ increases in volume

\implies surface $\mathcal{H}(\boldsymbol{\mu}) = E$ increases in volume $= \Omega(\mathbf{M}(E, \mathbf{x})) \uparrow \implies S(\Omega) \uparrow$

- (c) The maximum of the integrand occurs at (E_1^*, E_2^*) satisfying:

$$\frac{\partial S_1}{\partial E_1}|_{x_1}(E_1^*) = \frac{\partial S_2}{\partial E_2}|_{x_2}(E_2^*)$$

so that this characterizes equilibrium.

- (d) Therefore, near the equilibrium state, there is the greatest number of microstates in the microcanonical ensemble. *This makes evolution toward equilibrium probable*

Formally,

$$\Omega(E_1^*) \cdot \Omega(E_2^*) - \Omega(E_1)\Omega(E_2) \geq 0$$

for any (E_1, E_2) or

$$S(E_1^*) + S(E_2^*) - (S(E_1) + S(E_2)) = \delta S \geq 0$$

By the same analysis, moving towards a joint macrostate which is more densely populated always increases entropy, so $\delta S \geq 0$. In the exchange of energy as two systems are brought into contact

$$dS = \left(\frac{\partial S}{\partial E_1} - \frac{\partial S}{\partial E_2} \right) dE_1 = \left(\frac{1}{T_1} - \frac{1}{T_2} \right) dE_1 \geq 0$$

Thus, dE_1 *always flows towards the colder body* (**Clausius Theorem**)

4. Since (E_1^*, E_2^*) is a maximum for $S_1 + S_2$, it must be that:

$$\partial_{E_1}^2 S_1 + \partial_{E_2}^2 S_2 \leq 0$$

at this point

This is similarly true of changes due to position at this energy, so that $\partial_{x^{i,j}} S \leq 0$ at this point. This renders $D^2 S$ positive definite.

4.3 The Ideal Gas

Recall that microstates of a gas of N particles can be expressed as $\mu = \{\mathbf{p}_i, \mathbf{q}_i\}$, an element of a $6N$ -dimensional phase space Γ . Consider first $\mathcal{H}(\mu)$ with **no interactions** and a **box potential**, i.e.:

$$\mathcal{H}(\mu) = \sum_{i=1}^N \frac{\mathbf{p}_i^2}{2m} + U(\mathbf{q}_i)$$

The macrostate is specified by energy, volume, and number of particle: $\mathbf{M}(E, V, N)$ and

$$p_M(\mu) = \frac{1}{\Omega(E, V, N)} \begin{cases} 1 & \text{all } \mathbf{q}_i \in \text{box}, \sum_{i=1}^N \mathbf{p}_i^2 = 2mE \\ 0 & \text{else} \end{cases}$$

This gives a surface in phase space consisting of a box in position space, of volume V^N and a sphere in N -dimensional momentum space with radius $\sqrt{2mE}$:

$$\Omega(E, V, N) = V^N \cdot \beta(N)(2mE)^{\frac{N-1}{2}}$$

where $\beta(n)$ gives the area of the unit sphere in N dimensions, and $R = \sqrt{2mE}$ is the radius of the phase space sphere to which the momentum is constrained. By considering the Gaussian integral in two different ways, $\beta(N)$ is found to be $\beta(N) = \frac{\pi^{(N)/2}}{(\frac{N}{2}-1)!}$. Then,

$$S(E, V, N) = \ln(\Omega(E, V, N)) \xrightarrow{N \rightarrow \infty: \text{Stirling, } N \gg \ln(E), \ln(N)} Nk_B \ln \left[V \cdot \left(\frac{4\pi emE}{3N} \right)^{\frac{3}{2}} \right]$$

From entropy, energy and the equation of state are recovered:

$$\frac{1}{T} = \frac{\partial S}{\partial E} \implies Nk_B \frac{1}{V \cdot \left(\frac{4\pi emE}{3N} \right)^{3/2}} V \cdot \left(\frac{4\pi emE}{3N} \right)^{\frac{1}{2}} \cdot \frac{3}{2} \frac{4\pi em}{3N} \implies \frac{1}{T} = \frac{3}{2} Nk_B \frac{1}{E}$$

$$E = \frac{3}{2} N k_B T$$

and

$$\frac{P}{T} = \frac{\partial S}{\partial V} = \frac{N k_B}{V} \implies PV = N k_B T$$

To find the unconditional probability of a particular particle having momentum \mathbf{p}_1 , take:

$$p(\mathbf{p}_1) = \int_{\Gamma} d^3 q_1 \Pi_i d\Gamma^i p(\{q_i, p_i\}) = \frac{\Omega(E - \frac{\mathbf{p}_1^2}{2m}, N-1, V)}{\Omega(E, V, N)} = \dots$$

[Boltzmann Distribution]

4.3.1 Mixing of Gasses

First, observe that the entropy for an ideal gas may be written

$$S(\Omega(E, V, N)) = N k_B [\ln V + \ln(\frac{E}{N} \frac{4\pi m}{3})^{\frac{3}{2}}] = N k_B [\ln V + \sigma(\sqrt{2mE}/N)]$$

where $\sigma(\sqrt{2mE})$ is the contribution of momentum to entropy (since $\sqrt{2mE}$ is the magnitude of the total momentum).

Consider two gases at *the same temperature* T , occupying volumes V_1 and V_2 , which are allowed to mix. The initial entropy is:

$$S_1(E, V_1, N_1) + S_2(E, V_2, N_2) = N_1 k_B (\ln(V_1) + \sigma_1) + N_2 k_B (\ln(V_2) + \sigma_2)$$

and the final entropy, since $V_1 = V$ and $V_2 = V$, is

$$S_1(E, V, N_2) + S_2(E, V, N_1) = (N_1 + N_2) k_B \ln(V) + N_1 k_B \sigma_1 + N_2 k_B \sigma_2$$

The temperature does not change, since

$$T_f = \frac{2}{3 k_B} \frac{1}{(N_1 + N_2)} (E) = \frac{2}{3} \frac{1}{N_1 + N_2} (\frac{3}{2} k_B (N_1 + N_2) T_i) = T_i$$

and so the momentum contribution to each gas does not change, since energy depends only on temperature.

Thus, the entropy change due to mixing is:

$$\Delta_{\text{mix}} S = N_1 k_B \ln(V/V_1) + N_2 k_B \ln(V/V_2) = -[N_1 k_B \ln(V_1/V) + N_2 k_B \ln(V_2/V)]$$

This leads to **Gibb's Paradox**:

Take two identical gases with the same density (N_i/V_i) occupying volumes V_1 and V_2 . Remove the partition between these volumes yields no change in the state of the joint system, so that there should be no change in entropy. Yet $\Delta_{\text{mix}} S > 0$.

Since all particles are identical, there is no way of distinguishing the $N!$ permutations of any given microstate (any two permutations of a given microstate are macroscopically indistinguishable),

the volume element on phase space has been overcounted in volume by $N!$ (since it must be used to obtain the number of distinct *configurations* of the system). For a space on N identical particles

$$d\Gamma_N = \frac{1}{N!} \Pi_i d\Gamma^i$$

$$\implies \Omega(E, V, N) = \frac{1}{N!} V^N \cdot \beta(N) (2mE)^{\frac{N-1}{2}}$$

This correction could be modified to $\frac{1}{N_1!N_2!}$ for a space on N particles of two distinct types. Thus the **correct entropy** is

$$S(E, V, N) \xrightarrow{N \rightarrow \infty} Nk_B [\ln(V) - \ln(N) + \ln(e)] + \sigma = Nk_B \left[\ln \frac{Ve}{N} + \sigma \right]$$

and the entropy of mixing identical particles is:

$$\Delta_{\text{mix}} S = [-(N_1 + N_2)k_B \ln(n)] - [N_1k_B \ln(n) + N_2k_B \ln(n)] = 0$$

since $N_1/V_1 = N_2/V_2 = \frac{N_1+N_2}{V_1+V_2} = n$.

Observe thus that this new entropy is extensive:

$$S(\lambda E, \lambda, V, \lambda N) = \lambda Nk_B \ln \left[\frac{\lambda V e}{\lambda N} + \sigma(\sqrt{\lambda E / \lambda N}) \right] = \lambda S(E, V, N)$$

This correction does not affect the earlier results recovered with the uncorrected entropy, including the equation of state and energy-temperature relation.

Remark.

In quantum statistical mechanics, the factor of $\frac{1}{N!}$ in the volume element of phase space occurs naturally by considering quantum microstates. This theory additionally provides the appropriate measurement of action, \hbar , which can be used to make the phase space volume element dimensionless; as it stands, $[pq] = J \cdot s = \text{action}$. The final volume element on phase space of N identical particles is thus:

$$d\Gamma = \frac{1}{N! \hbar^{3N}} \Pi_i d\Gamma^i$$

4.4 The Canonical Ensemble

In the microcanonical ensemble, a macrostate is specified by energy and mechanical work coordinates; the temperature is thus determined by: $\frac{\partial S}{\partial E}|_{\mathbf{x}} = \frac{1}{T}$. Since energy and temperature are both functions of state and properly related, one may specify macrostates in terms of temperature, and then deduce the energy. This is the perspective of the canonical ensemble, where a **macrostate** is

$$\mathbf{M} := (T, \mathbf{x})$$

Note that the mechanical coordinates are specified, so that no mechanical work is allowed. The theoretical construction of \mathbf{M} is that, the system S is maintained at temperature T by heat exchange with a reservoir R , sufficiently large that heat exchange with S does not modify its temperature. Since heat capacity is extensive, a sufficiently large reservoir R would approximately undergo no temperature change as a result of heat input. To obtain the probability over microstates, observe that the system $\Gamma_S \oplus \Gamma_R$ has a microcanonical ensemble corresponding to its total energy $E_{\text{total}} \gg E_S$.

This has a phase space volume $\Omega(\mathbf{M}(E_{\text{total}}))$ given by the constraint $\mathcal{H}(\mu) = \mathcal{H}_S(\mu_S) + \mathcal{H}(\mu_R) = E_{\text{total}}$ density of

$$p(\mu_S \times \mu_R) = \frac{1}{\Omega(\mathbf{M}(E_{\text{total}}))} \begin{cases} 1 & \mathcal{H}_S(\mu_S) + \mathcal{H}(\mu_R) = E_{\text{total}} \\ 0 & \text{else} \end{cases}$$

Thus, the unconditional probability for a microstate of S is:

$$p_{\mathbf{M}(T, \mathbf{x})}(\mu_S) = \sum_{\mu_R} p(\mu_S \times \mu_R)$$

The probability of joint microstates is only nonzero when $\mathcal{H}(\mu_S \times \mu_R) = E_{\text{total}}$, so that this corresponds to:

$$\begin{aligned} p_{\mathbf{M}(T, \mathbf{x})}(\mu_S) &= \sum_{\mu_R: \mathcal{H}_R(\mu_R) = E_{\text{total}} - \mathcal{H}_S(\mu_S)} \frac{1}{\Omega(\mathbf{M}(E_{\text{total}}))} \\ &= \frac{\Omega(E - \mathcal{H}_S(\mu_S))}{\Omega(E_{\text{total}})} \cong \Omega(E - \mathcal{H}_S(\mu_S)) = \exp \frac{1}{k_B} S(E - \mathcal{H}_S(\mu_S)) \end{aligned}$$

Since the total energy is assumed much greater than the energy of the system, this allows the following by first order Taylor expansion about E:

$$S_R(E - \mathcal{H}_S(\mu_S)) = S(E) - \frac{\partial S}{\partial E} \cdot \mathcal{H}_S(\mu_S) \implies$$

$$(*) \ p_{\mathbf{M}(T, \mathbf{x})}(\mu_S) = \exp \frac{S_B}{k_B} \cdot \exp \left[-\frac{\mathcal{H}_S(\mu_S)}{k_B T} \right] \cong \frac{\exp \left[-\mathcal{H}_S(\mu_S)/k_B T \right]}{Z} = \frac{\exp \left[-\beta(T) \mathcal{H}_S(\mu_S) \right]}{Z}$$

where \cong is as an absolute probability distribution $\beta(T) = \frac{1}{k_B T}$. The normalization of this probability is called the **Partition Function**:

$$Z(T, \mathbf{x}) = \sum_{\mu} \exp \left[-\frac{\mathcal{H}_S(\mu)}{k_B T} \right]$$

where the subscript S can be dropped, since all dependency on R has been eliminated by applying the relative energy approximation.

Immediate consequences of this derivation include the following:

4.4.1 Energy in the Canonical Ensemble

- Note that the energy of microstates of S was not constrained in computing the probability density of ensemble states. Rather, the total energy between a reservoir, and the system, was constrained, and the energy E of S was allowed to vary by heat exchange. Since all dependence on R has been eliminated, energy will also be a random variable now among allowed microstates of S. Observe

$$\begin{aligned} p(\mathcal{E}) &= \sum_{\mu} p(\mu) \cdot \delta(\mathcal{H}(\mu) - \mathcal{E}) = \frac{1}{Z} \exp[-\beta(T)\mathcal{E}] \sum_{\mu} \delta(\mathcal{H}(\mu) - E) = \frac{\exp[-\beta(T)\mathcal{E}]}{Z} \Omega(E) \\ &= \frac{\exp[-\beta(T)\mathcal{E} + \frac{1}{k_B} S(\mathcal{E})]}{Z} = \frac{\exp[-\beta F(\mathcal{E})]}{Z} \end{aligned}$$

where $F(\mathcal{E}) = \mathcal{E} - TS(\mathcal{E})$ is the **Helmholtz Free Energy** function.

From wish to compute the Helmholtz Free Energy as a function of state.

One approach is using energy as the distribution variable in lieu of μ , obtaining the following approximation:

$$Z = \sum_{\mu} \exp[-\beta(T)\mathcal{H}(\mu)] = \sum_{\mathcal{E}} \exp[-\beta(T)F(\mathcal{E})] \approx \exp[-\beta(T)F(E^*)]$$

where E^* minimizes F , and hence maximizes $p(\mathcal{E})$. This suggests the identification

$$F = -k_B T \ln(Z)$$

Consider further the average energy:

$$\langle \mathcal{H}(\mu) \rangle = \sum_{\mu} \mathcal{H}(\mu) \frac{\exp[-\beta(T)\mathcal{H}(\mu)]}{Z} = \langle -\partial_{\beta} p(\mu) \rangle = -\partial_{\beta} \ln Z$$

By classical thermodynamics, $E = F + TS = F - T \frac{\partial F}{\partial T}|_x = -T^2 \partial_T (\frac{F}{T}) = \frac{\partial \beta(T) F}{\partial \beta(T)}$

and it is required that in equilibrium $\langle H(\mu) \rangle = E$, it is clear that the macrostate approach also suggests:

$$F = -\frac{1}{\beta} \ln Z = -k_B T \ln(Z)$$

The first identity comes from taking the most probable energy, E^* , and the second comes from taking the average energy, $E = \langle \mathcal{H}(\mu) \rangle$. The variance $\langle \mathcal{H}(\mu)^2 \rangle_c$ tells how near these values are.

Observing that $\ln Z(\beta)$ is the generating function for the cumulants of \mathcal{H} ,

$$\langle H^2 \rangle_C = -\partial_{\beta} \langle \mathcal{H}(\mu) \rangle = -\frac{\partial \langle \mathcal{H} \rangle}{\partial \frac{1}{k_B T}} = T^2 k_B \frac{\partial \langle H \rangle}{\partial T} |_{\mathbf{x}} = T^2 k_B C_{\mathbf{x}}$$

and since $C_{\mathbf{x}}$ is an extensive quantity, this shows

$$\langle H^2 \rangle_C \propto N \implies \sqrt{\langle H^2 \rangle_C} \propto \sqrt{N}$$

Since energy is extensive, this means that the relative error in treating the most probable quantity for the average energy is:

$$\frac{\sqrt{\langle H^2 \rangle_C}}{\langle H \rangle} \propto \frac{\sqrt{N}}{N} \rightarrow 0$$

in the thermodynamic limit. Thus, is is justified by the thermodynamic limit, from the macrocanonical ensemble, that;

$$F = -\frac{1}{\beta} \ln Z(\beta)$$

Since all cumulants are in fact proportional to N by analysis of the generating function, one may approximate $p(\mathcal{E})$ as having only two cumulants. Thus:

$$p(\mathcal{E}) \xrightarrow{N \rightarrow \infty} \frac{1}{\sqrt{2\pi k_B T C_{\mathbf{x}}}} \exp\left[-\frac{(\mathcal{E} - \langle H \rangle)^2}{2k_B T C_{\mathbf{x}}}\right]$$

The **canonical entropy** can be computed directly from the probability distribution, and its classical relation to F recovered:

$$S(T, \mathbf{x}) = -k_B \langle \ln p_{\mu} \rangle = -k_B \langle -\beta \mathcal{H}(\mu) - \ln Z \rangle = \frac{E}{T} - \frac{F}{T}$$

4.4.2 The Canonical Ensemble: Example Systems

1. Two state system: A two state system of N impurities with macrostate $\mathbf{M} = (T, \mathbf{x})$ has Hamiltonian $\mathcal{H}(\mathbf{n}) = \epsilon \sum_{i=1}^N n_i$. The probability density among microstates is $p(\mathbf{n}) = \frac{1}{Z} \exp[-\beta \epsilon \sum_{i=1}^N n_i]$ and the partition function

$$Z(T, \mathbf{n}) = \sum_{\mathbf{n}} \exp[-\beta \epsilon \sum_{i=1}^N n_i] = (\exp[-\beta(T)\epsilon \cdot 0] + \exp[\beta(T)\epsilon]) \sum_{n_2, \dots, n_N} \exp[-\beta(T)\epsilon \sum_{i=1}^N n_i]$$

and continuing inductively all the way through n_N

$$Z(T, \mathbf{n}) = (1 + \exp(-\beta\epsilon))^N$$

To obtain the well defined average internal energy, one may obtain S and $F(T, \mathbf{n})$ so that

$$E = F(T, \mathbf{n}) + TS(T, \mathbf{n})$$

where $S(T, \mathbf{n}) = -\frac{\partial F}{\partial T}|_N$.

Compute the free energy from $Z(T, \mathbf{n})$:

$$F(T, N) = -k_B T \ln Z(T, N) = -k_B T N \ln[1 + \exp(-\epsilon\beta)]$$

Thus

$$S = -\frac{\partial F}{\partial T}|_N = (k_B N \ln[1 + \exp(-\epsilon\beta)]) + k_B T N \frac{\frac{\epsilon}{k_B T^2} \exp(-\epsilon\beta(T))}{1 + \exp(-\epsilon\beta(T))} = -\frac{F(T, N)}{T} + \frac{N\epsilon}{T} \frac{\exp(-\epsilon\beta(T))}{1 + \exp(-\epsilon\beta(T))}$$

The internal energy is thus obtained from S

$$E = F + TS = \frac{N\epsilon \exp(-\epsilon\beta(T))}{1 + \exp[-\epsilon\beta]}$$

This can also be obtained directly from $Z(T, N)$ since

$$E = -\frac{\partial Z}{\partial \beta} = \frac{N\epsilon \exp[-\beta\epsilon]}{1 + \exp[-\beta\epsilon]}$$

2. An Ideal Gas: An ideal gas of N identical particles has a macrostate $\mathbf{M} = (T, V, N)$. The joint probability density for microstates $\mu \in \Gamma$ is given by:

$$p_{T,V,N}(\mu) = \frac{1}{Z} \exp[-\beta \sum_{i=1}^N \frac{p_i^2}{2m}] \cdot \prod_i \chi_{\{q_i \in \text{box}\}} \implies$$

$$Z(T, V, N) = \int_{\Gamma^Q \oplus \Gamma^P} \left(\frac{1}{(h)^{3N} N!} d\Gamma \right) \exp(-\beta \sum_{i=1}^N \frac{p_i^2}{2m}) \cdot \chi_{V \subset \Gamma^Q}$$

$$\frac{V^N}{h^{3N} N!} \cdot \left(\frac{2m\pi}{\beta} \right)^{3N/2} = \frac{V^N}{N!} \left(\frac{2m\pi k_B T}{h^2} \right)^{3N/2} = \frac{1}{N!} \left(\frac{V}{\lambda(T)^3} \right)^N$$

where $\lambda(T) = \frac{\hbar}{\sqrt{2m\pi k_B T}}$ is a characteristic length scale associated with \hbar .

Again compute F via:

$$F(T, V, N) = -k_B T \ln Z = -k_B T (-\ln(N!) + N[\ln V - 3/2 \ln \lambda(T)^2]) =$$

$$k_B T (N \ln(N) - N) - k_B T N \ln V + k_B T N 3/2 \ln \lambda(T)^2 = -k_B T N (\ln \frac{Ve}{N} + \frac{3}{2} \ln \frac{1}{\lambda(T)^2})$$

in the large N limit.

This gives an entropy

$$S(T, V, N) = -T \frac{\partial F}{\partial T} \Big|_{N, V} =$$

3.

Homework 4

Question 4. 4.3 Curie Susceptibility

Proof. (1).

$$Z(T, B) = \sum_{\mu} \exp [\beta(T) B \mu \sum_{i=1}^N m_i] = \sum_{m_i} \sum_{m_2, \dots, m_N} \exp [\beta(T) B \mu \sum_{i=1}^N m_i] =$$

$$\sum_j \exp [\beta(T) B \mu m_{1j}] \sum_{m_2, \dots, m_N} \exp [\beta(T) B \mu \sum_{i>2} m_{1j}] = (\sum_j \exp [\beta(T) B \mu]^{m_{1j}}) \sum_{m_2, \dots, m_N}$$

Now, with $\exp \beta(T) B \mu = z$, the first factor may be written as:

$$\sum_{j=0}^s z^j + \sum_{j=0}^s \frac{1}{z} - 1 = \frac{z^{s+1} - z^{-s}}{z - 1}$$

Continuing inductively

$$Z(T, B) = \left(\frac{z^{s+1} - z^{-s}}{z - 1} \right)^N = \left(z^{1/2} \frac{z^{s+1/2} - z^{-(s+1/2)}}{z - 1} \right)^N = \frac{z^{s+1/2} - z^{-(s+1/2)}}{z^{1/2} - z^{-1/2}} =$$

$$\left(\frac{\sinh[(s+1/2)\beta(T)B\mu]}{\sinh[1/2\beta(T)B\mu]} \right)^N$$

(2). The Gibbs free energy satisfies:

$$G(T, B) = -k_B T \ln Z(T, B) = -k_B T N [\ln \sinh(\alpha B) - \ln(\sinh \gamma B)]$$

Since the approximation calls for terms of at most order 4, expand sinh to this order. Thus

$$G(T, B) = -k_B T N [\ln[\alpha B + \frac{1}{3!}(\alpha B)^3 + \mathcal{O}(B^5)] - \ln[\gamma B + \frac{1}{3!}(\gamma B)^3 + \mathcal{O}(B^5)]]$$

Now, near $z = 0$, $\ln(1+z) = z - \frac{1}{2!}z^2 + \frac{1}{3!}z^3 + \mathcal{O}(z^4)$. Further,

$$\ln[\alpha B + \frac{1}{3!}(\alpha B)^3 + \mathcal{O}(B^5)] - \ln[\gamma B + \frac{1}{3!}(\gamma B)^3 + \mathcal{O}(B^5)] = \ln\left(\frac{\alpha}{\gamma}\right) + \ln\left(1 + \frac{1}{6}(\alpha B)^2\right) - \ln\left(1 + \frac{1}{6}(\gamma B)^2\right)$$

Since the argument of the expansion of ln is already second order in B, one need only compute the expansion of ln to its first order to obtain a sum with terms $< \mathcal{O}(B^4)$. Thus

$$G(T, B) = -Nk_B T [\ln(2s + 1) + \frac{B^2 \mu^2 \beta^2}{6} ((s + 1/2)^2 - \frac{1}{2})] = G(T, 0) - \frac{NB^2 \mu^2}{6k_B T} (s^2 + s) + \mathcal{O}(B^4)$$

(3).

Since $-\frac{\partial}{\partial B} G(T, B) = \langle M_z \rangle$ this requires

$$-\frac{\partial^2}{\partial B} G(T, B)|_{B=0} = \frac{\partial}{\partial B} \langle M_z \rangle|_{B=0} = \chi$$

Thus

$$\chi = \frac{N\mu^2}{3k_B T}$$

(4).

First, since

$$C_B = \frac{\partial}{\partial T} |_B H(T, B) = \frac{\partial}{\partial T} |_B \langle \mathcal{H}(\mu) - M \cdot B \rangle = -\frac{\partial}{\partial \beta} \ln Z(T, B) =$$

Since

$$C|_x = \frac{\partial}{\partial T} E|_M = \frac{\partial}{\partial T} [H + B \langle M \rangle] = H - B \frac{\partial H}{\partial B}$$

□

4.5 Introduction to Quantum Statistical Mechanics