

**MAT 1723HF (APM 421H1F): MATHEMATICAL CONCEPTS OF
QUANTUM MECHANICS AND QUANTUM INFORMATION
WEDNESDAYS 3–4 (BA1170), THURSDAYS 4–5 (BA6183),
FRIDAYS 6–7 (BA1170)**

I. M. SIGAL TA: LI CHEN

CONTENTS

1. Goals	2
2. Prerequisites	2
3. Syllabus	2
4. Texts	2
5. Tests and marking scheme	3
6. Modified and additional material	3
6.1. Conservation laws	3
6.2. Additional points about the spectrum	6
6.3. Angular momentum.	7
6.4. Motion in a spherically symmetric potential	10
6.5. Spin	12
6.6. Atoms and molecules and their spectra	15
6.7. The Harmonic Oscillator	20
6.8. A particle in a constant external magnetic field	23
6.9. Aharonov-Bohm Effect	25
6.10. Variational Characterization of Eigenvalues	26
6.11. Particle in a Periodic Potential	28
6.12. Quantum Statistics	30
6.13. Density Matrices	30
6.14. Stationary States	32
6.15. Quantum statistics: General framework	34
6.16. Information Reduction	36
6.17. Reduced dynamics	39
6.18. Decoherence and Thermolization	42
6.19. Irreversibility	43
6.20. Information and communication channels	44
6.21. Measuring classical and quantum information	46
6.22. Shannon's theorems and their quantum extensions	47
6.23. Spin and Statistics	48

Date: September, 2017.

6.24. Self-consistent Approximations	52
References	57

1. GOALS

The goal of this course is to explain key concepts of Quantum Mechanics and to arrive quickly to some topics which are at the forefront of active research. Examples of the latter topics are Bose-Einstein condensation and quantum information, both of which have witnessed an explosion of research in the last decade and both involve deep and beautiful mathematics.

2. PREREQUISITES

I will cover all necessary definitions beyond multivariable calculus and linear algebra, but without familiarity with elementary ordinary and partial differential equations, the course will be tough. Knowledge of elementary theory of functions and operators would be helpful. No physics background is required.

3. SYLLABUS

- Schrödinger equation
- Quantum observables
- Spectrum and evolution
- Important special cases
- Angular momentum and group of rotations
- Spin and statistics
- Atoms and molecules
- Quasiclassical asymptotics
- Adiabatic theory and geometrical phases
- Hartree-Fock theory
- Feynman and Wiener path integrals
- Density matrix
- Open systems
- Quantum entropy
- Quantum channels

4. TEXTS

[GS] S. Gustafson and I.M. Sigal: Mathematical Concepts of Quantum Mechanics, 2nd edition, Springer, 2011.

For the material which is not in the book, I will either post the lecture notes or refer to on-line material.

5. TESTS AND MARKING SCHEME

3 quizzes (October 4, 25 and November 22), midterm test (November 15), final test (December 6).

Location and time: three quizzes and midterm BA1170, 3:10-4:00pm; final BA1170, 6:10-7:00pm

All the tests will be on the material covered in the lectures. For the quizzes and midterm exam, all the problems will be taken from homework and for the final test, most of the problems.

Breakup of the grade:

Class participation (20%), three quizzes (5% + 10% + 15%), midterm (25%), final test (25%).

* * * * *

* * *

6. MODIFIED AND ADDITIONAL MATERIAL

6.1. Conservation laws. We say that an observable A (or more precisely, the physical quantity represented by this observable) is *conserved* if its average in any evolving state $\psi(t)$ is independent of t :

$$\langle A \rangle_{\psi(t)} = \langle A(t) \rangle_{\psi_0} = \langle A \rangle_{\psi_0}, \quad (6.1)$$

where $\psi_0 = \psi(0)$, $A(t) := e^{itH/\hbar} A e^{-itH/\hbar}$, and $\psi(t)$ solves the Schrödinger equation:

$$i\hbar \frac{\partial \psi}{\partial t} = H\psi. \quad (6.2)$$

Due to (6.1), an observable A is conserved if and only if $A(t)$ is constant, which is equivalent, due to (??), to A commuting with the Schrödinger operator H , i. e.

$$[A, H] = 0$$

(provided certain domain properties, which we ignore here, hold; proving that $\langle \psi, [A, H]\psi \rangle = 0$, $\forall \psi$, implies $[A, H] = 0$ requires some work).

Consider several examples. Since obviously $[H, H] = 0$, we have $\langle H \rangle_{\psi(t)} = \text{constant}$ for any solution to (6.2), which is the mean-value version of the conservation of energy.

As the second example, we take the obvious relation $[\mathbf{1}, H] = 0$, which give the conservation of the total probability: $\langle \mathbf{1} \rangle_{\psi(t)} = \|\psi(t)\|^2 = \text{constant}$ for any solution to (6.2).

Furthermore, $\frac{i}{\hbar}[x_j, H] = p_j \neq 0$, which says that the particle position is never conserved, i.e. a quantum particle cannot be localized at a point.

The relation $\frac{i}{\hbar}[p_j, H] = -\partial_{x_j} V \neq 0$ implies that p_j is conserved iff V is independent of x_j . This taken for all j 's gives the first Newton's law: a particle will be in the state of the motion with a constant velocity as long as no force is acting on it.

Homework 6.1. Prove that (a) the momentum is conserved iff the potential $V(x)$ is constant; and (b) the angular momentum is conserved iff the potential $V(x)$ is spherically symmetric, i.e. $V(Rx) = V(x)$ for any rotation R in \mathbb{R}^3 around the origin, i.e. $V(x) = W(|x|)$ for some function $W(r)$ on the interval $[0, \infty)$.

Later on, we will prove this statement an elementary group theory.

Most conservation laws come from symmetries of the quantum system in question. Continuous symmetries are often associated with one-parameter groups U_s , $s \in \mathbb{R}$, of unitary operators. We say that U_s is a symmetry if U_s maps $D(H)$ into itself and

$$\psi_t \text{ is a solution to (6.2)} \rightarrow U_s \psi_t \text{ is a solution to (6.2), } \forall s \in \mathbb{R}.$$

Let A be a generator of a one-parameter group U_s : $\partial_s U_s = -iAU_s$. Then (ignoring domain questions)

$$U_s \text{ is a symmetry of (6.2)} \rightarrow A \text{ commutes with } H.$$

Indeed, the fact that U_s is a symmetry implies that (here again we disregard domain questions and proceed formally) $i\hbar\partial_t U_s \psi_t = HU_s \psi_t$. Inverting U_s gives

$$i\hbar\partial_t \psi_t = U_s^{-1} H U_s \psi_t.$$

Differentiating the last equation with respect to s and setting $s = 0$ and $t = 0$ we arrive at $i[H, A]\psi_0 = 0$, where $\psi_0 = \psi_{t=0}$. Since this is true for any ψ_0 , we conclude that

$$[H, A] = 0,$$

i. e. A commutes with H .

Let $e_1 := (1, 0, 0)$, $e_2 := (0, 1, 0)$, $e_3 := (0, 0, 1)$. Examples of symmetry groups and their generators:

- Spatial translations: $U_s^{\text{transl}} : \psi(x) \rightarrow \psi(x + se_j)$, $s \in \mathbb{R}$, with the generator $\frac{1}{\hbar}p_j = -i\nabla_{x_j} \Rightarrow$ conservation of momentum
- Spatial rotation: $U_s^{\text{rot}} : \psi(x) \rightarrow \psi(R_s^{-1}x)$, where R_s is the contra-clockwise rotation around the j -axis by the angle $s \in [0, 2\pi)$, with the generator $\frac{1}{\hbar}L_j = (x \times (-i\nabla_x))_j \Rightarrow$ conservation of angular momentum
- Gauge invariance: $U_s^{\text{gauge}} : \psi(x) \rightarrow e^{is}\psi(x)$, $s \in \mathbb{R}$, with the generator $i \Rightarrow$ conservation of probability.

Recall that the operator $L = x \times p$ is called the *angular momentum* operator, or just angular momentum. The term is justified by conservation law above.

We summarize the discussion above as

- Time translation invariance (V is independent of t) \rightarrow conservation of energy
- Space translation invariance (V is independent of x) \rightarrow conservation of momentum
- Space rotation invariance (V is rotation invariant, i.e. is a function of $|x|$) \rightarrow conservation of angular momentum
- Gauge invariance (invariance of the equation under the transformation $\psi \rightarrow e^{i\alpha}\psi$) \rightarrow conservation of charge/probability.

Discussion. More generally, the symmetries associated with groups. In the last three examples these are the groups of translations and rotations of \mathbb{R}^3 , denoted as \mathbb{R}^3 and $O(3)$, respectively, (forming together the group of rigid motions of \mathbb{R}^3), and the group $U(1)$ of complex numbers of unit modulus, called the gauge group. (For particles with internal degrees of freedom, specifically with a spin to be considered later on, called the gauge group is $SU(n)$ of complex, unitary, $n \times n$ matrices, for an appropriate n .)

We represent these groups by unitary operators on the state space $L^2(\mathbb{R}^3)$: $U_y^{\text{transl}} : \psi(x) \rightarrow \psi(x + y)$, $y \in \mathbb{R}^3$, $U_R^{\text{rot}} : \psi(x) \rightarrow \psi(R^{-1}x)$, $R \in O(3)$, and $U_\alpha^{\text{gauge}} : \psi(x) \rightarrow e^{i\alpha}\psi(x)$, $e^{i\alpha} \in U(1)$, for spatial translations and rotations and the gauge transformations. These operators satisfy

$$U_y^{\text{transl}} U_{y'}^{\text{transl}} = U_{y+y'}^{\text{transl}} \quad \text{and} \quad U_R^{\text{rot}} U_{R'}^{\text{rot}} = U_{RR'}^{\text{rot}},$$

and similarly for U_α^{gauge} . The operators U_y^{transl} and U_R^{rot} give unitary representations of the groups of translations and rotations of \mathbb{R}^3 , respectively, (the first one commutative, the second one not). They can be written as products of one parameter groups, so that the analysis relevant for us can be reduced to the latter case.

Homework 6.2. Find the generators of translation, rotation and gauge groups, and show that under certain conditions (which?) these groups are symmetry groups of the Schrödinger equation.

Consider the group of rotations, $O(3)$, in some more detail. Note that the rotations, R , are represented by orthogonal matrices (i.e. real matrices satisfying $R^T R = \mathbf{1}$) of the determinant 1. Consider the contra-clockwise rotation, R_θ^ω , around the axis along the unit vector ω by the angle $\theta \in [0, 2\pi)$. This is one-parameter group and one can show that its generator is given by

$$\partial_\theta \Big|_{\theta=0} R_\theta^\omega x = \omega \times x.$$

(It is easy to prove this for each co-ordinate axis separately which implies the relation above.) Let $A_j := e_j \times x$. Then A_j satisfy the commutation relations

$$[A_1, A_2] = A_3, [A_2, A_3] = A_1, [A_3, A_1] = A_2. \quad (6.3)$$

Now, $\partial_\theta|_{\theta=0}\psi(R_{-\theta}^\omega x) = -\nabla\psi(x) \cdot (\omega \times x) = -\omega \cdot (x \times \nabla\psi(x))$. Hence

$$i\hbar\partial_\theta|_{\theta=0}U_{R_{-\theta}^\omega}^{\text{rot}} = -i\hbar\omega \cdot (x \times \nabla) = \omega \cdot L,$$

which proves the statement made above.

Furthermore, it is easy to see for $H := -\frac{\hbar^2}{2m}\Delta + V(x)$ that

$$U_R^{\text{rot}-1}HU_R^{\text{rot}} = -\frac{\hbar^2}{2m}\Delta + V(Rx).$$

Homework 6.3. Prove the last relation.

The last relations imply the proof of the statement in Problem 6.1.

Unlike the components of the position and momentum operators, x_j and p_j , the components of the angular momentum one do not mutually commute:

$$\frac{i}{\hbar}[L_k, L_l] = \epsilon^{klm}L_m. \quad (6.4)$$

Here ϵ^{klm} is the Levi-Chivita symbol: $\epsilon^{123} = 1$ and ϵ^{klm} changes the sign under the permutation of any two indices. This is related to the fact that U_h^{transl} and U_R^{rot} are representations of the groups of translations and rotations of \mathbb{R}^3 and, unlike the group of translations, the group of rotation is non-abelian.

Eq. (6.5) implies $\frac{1}{i\hbar}L_j$ are the generators of the group $SO(3)$ in the sense defined above. The vector space spanned by $\frac{1}{i\hbar}L_j$, equipped with the commutator relation $[\cdot, \cdot]$, is the Lie algebra $so(3)$ of the group $SO(3)$.

6.2. Additional points about the spectrum. If ψ is a normalized eigenvector of A with an eigenvalue λ , then $\langle\psi, A\psi\rangle = \lambda$, i.e. measuring of of the corresponding physical observable in the state ψ will always give the same answer, λ . Let ψ_i be several normalized eigenvectors of A , with eigenvalues λ_i . By Problem ??(2) below, they can be chosen to be mutually orthogonal. If ψ is a superposition, $\psi = \sum_i a_i\psi_i$, of ψ_i 's, then $\langle\psi, A\psi\rangle = \sum_i \lambda_i|a_i|^2$.

* * * * *

The second property of spreading sequences implies that we can choose a subsequence $\{\psi'_n\}$ so that $\text{supp } \psi'_n \subset \{|x| \geq R_n\}$, with $R_n \rightarrow \infty$. In what follows we always assume that the sequence $\{\psi_n\}$ has this property.

As an example, a spreading sequence for $-\frac{\hbar^2}{2m}\Delta$ and any $\lambda > 0$ is given by $\psi_n(x) := f(|x|/n)e^{ik \cdot x/\hbar}$, where f is any smooth function, s.t. $f(x) = 1$ for $|x| \leq 1$ and $= 0$ for $|x| \geq 2$, and $\frac{1}{2m}|k|^2 = \lambda$.

Homework 6.4. Show that the sequence constructed above is a spreading sequence for $-\frac{\hbar^2}{2m}\Delta$ and $\lambda = \frac{1}{2m}|k|^2$.

Homework 6.5. What are the essential spectra and what are possible locations of the discrete spectra of the following operators (justify your answer)

- (a) $H = -\frac{\hbar^2}{2m}\Delta - 10|x|^3 + |x|^4$,
 (b) $H = -\frac{\hbar^2}{2m}\Delta - (1 + |x|)^{-2}$.

6.3. Angular momentum. Recall that the operators $L_j = (x \times p)_j$ are called *angular momentum* (component) operators, or angular momentum observables, or just angular momenta. The term is justified by conservation law above. $\frac{1}{i\hbar}L_j$ are the generators of the group $SO(3)$ in the sense defined above. They satisfy the commutation relations

$$[L_k, L_l] = i\hbar\epsilon^{klm}L_m. \tag{6.5}$$

Here ϵ^{klm} is the Levi-Chivita symbol: $\epsilon^{123} = 1$ and ϵ^{klm} changes the sign under the permutation of any two indices.

Spectrum of the angular momentum. In this subsection we study spectral properties of the angular momentum operators,

$$L_j = (x \times p)_j, \tag{6.6}$$

where $p = -i\hbar\nabla$, as usual. Recall that $L_j = (x \times p)_j$ satisfy the commutation relations

$$[L_k, L_l] = i\hbar\epsilon^{klm}L_m, \tag{6.7}$$

where ϵ^{klm} is the Levi-Chivita symbol: $\epsilon^{123} = 1$ and ϵ^{klm} changes the sign under the permutation of any two indices.

Homework 6.6. Prove (6.7). (Hint: Use that $[p_j, x_k] = -i\hbar\delta_{ij}$.)

We define the squared magnitude of the angular momentum, $L^2 = L_1^2 + L_2^2 + L_3^2$. Note that L^2 commutes with L_k :

$$[L^2, L_k] = 0, \quad \forall k. \tag{6.8}$$

Homework 6.7. Prove (6.8). (Hint: Use that $[p_j, x_k] = -i\hbar\delta_{ij}$.)

Furthermore, L^2 is a homogeneous degree zero operator, in the sense that it commutes with the rescalings, $T_\lambda^{\text{scal}} : \psi(x) \rightarrow \lambda^{3/2}\psi(\lambda x)$, $\lambda > 0$, and therefore it does not act on the radial variable $r = |x|$.

Theorem 6.8. (a) The spectrum of the operator L^2 is purely discrete; (b) it consists of isolated eigenvalues $\lambda = \hbar^2 l(l+1)$, where $l = 0, 1, 2, \dots$, of the multiplicities $2l + 1$.

We will derive the proof from the following

Theorem 6.9. Let L_j be operators satisfying the commutation relations (6.7). Assume the operator $L^2 = L_1^2 + L_2^2 + L_3^2$ has purely discrete spectrum. Then this spectrum consists of the eigenvalues $\lambda = \hbar^2 l(l + 1)$, where $l = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \dots$, of the multiplicities $2l + 1$.

Proof. A proof of (a) is somewhat involved and we skip it here.¹ We prove statement (b). Let λ be an eigenvalue of L^2 with the eigenspace V_λ (i.e. the subspace spanned by all eigenvectors of L^2 with this eigenvalue). It follows from (6.8) that the eigenspace V_λ is invariant under L_k , i.e. $L_k V_\lambda \subset V_\lambda$. (Indeed, if $L^2 \phi_\lambda = \lambda \phi_\lambda$, then $L_3 L^2 \phi_\lambda = L^2 L_3 \phi_\lambda = \lambda L_3 \phi_\lambda$.) Hence one can show that one can choose a basis in V_λ consisting of common eigenfunctions, $\phi_{\lambda\mu}(\theta, \phi)$, for these operators:

$$L^2 \phi_{\lambda,\mu} = \lambda \phi_{\lambda,\mu}, \quad L_3 \phi_{\lambda\mu} = \mu \phi_{\lambda\mu}.$$

(By choosing an arbitrary ON basis in a finite dimensional subspace V_λ , one can reduce this problem to one for matrices.)

We still would like to use the remaining operators L_1, L_2 and so we form the combinations

$$L_\pm := L_1 \pm iL_2. \tag{6.9}$$

The virtue of the new operators are the following commutation relations, which are easily checked:

$$[L^2, L_\pm] = 0, \quad [L_+, L_-] = 2L_3, \quad [L_3, L_\pm] = \pm L_\pm. \tag{6.10}$$

Homework 6.10. *Prove (6.10). (Hint: Use that $[p_j, x_k] = -i\hbar\delta_{ij}$.)*

The last relation can be rewritten as $L_3 L_\pm = L_\pm (L_3 \pm 1)$ which shows that L_\pm are raising/lowering operators in the following sense. Apply the operators L_\pm to the eigenfunctions as $\phi_{\lambda,\mu}$ and use $L_3 \phi_{\lambda\mu} = \mu \phi_{\lambda\mu}$ and $L_3 L_\pm = L_\pm (L_3 \pm 1)$ to obtain

$$L_3 L_\pm \phi_{\lambda\mu} = L_\pm (L_3 \pm 1) \phi_{\lambda\mu} = (\mu \pm 1) \phi_{\lambda\mu}.$$

Hence we see that the operators L_\pm raise/lower eigenvalues of L_3 . More precisely, assuming the eigenfunctions as $\phi_{\lambda\mu}$ are non-degenerate (correspond to different eigenvalues), we have $L_\pm \phi_{\lambda\mu} = C \phi_{\lambda\mu \pm 1}$, for some constants C .

Next, we claim that $\mu^2 \leq \lambda$, for given λ . Indeed, we use $L^2 = L_1^2 + L_2^2 + L_3^2 \geq L_3^2$ to conclude that $\mu^2 = \langle \phi_{\lambda\mu}, L_3 \phi_{\lambda\mu} \rangle \leq \langle \phi_{\lambda\mu}, L^2 \phi_{\lambda\mu} \rangle = \lambda$. Furthermore, since $\bar{L}_3 = -L_3$, if μ is an eigenvalue of L_3 on V_λ , then so is $-\mu$.

Denote $\mu_* := \max\{|\mu| : \text{given } \lambda\}$. Then by the definition of μ_* and the raising/lowering property of the operators L_\pm , we find that

$$L_\pm \phi_{\lambda \pm \mu_*} = 0.$$

¹From the known form of the laplacian in the spherical coordinates, (r, θ, ϕ) , (6.17) shows that $L^2 = -\hbar^2 \Delta_\Omega$, where Δ_Ω is the Laplace-Beltrami operator on \mathbb{S}^2 , given in spherical coordinates (θ, ϕ) , by

$$\Delta_\Omega = \frac{1}{\sin(\theta)} \frac{\partial}{\partial \theta} (\sin(\theta) \frac{\partial}{\partial \theta}) + \frac{1}{\sin^2(\theta)} \frac{\partial^2}{\partial \phi^2}.$$

Using the Weyl criterion of essential spectrum and the fact that the sphere \mathbb{S}^2 is a compact space, one can show that Δ_Ω has purely discrete spectrum.

Next, we use $L_{\pm}L_{\mp} = L_1^2 + L_2^2 \pm \hbar L_3$, to obtain the following relation between the operators L^2 , L_{\pm} and L_3 :

$$L^2 = L_{\pm}L_{\mp} \mp \hbar L_3 + L_3^2. \quad (6.11)$$

Taking the expectation of this equation in the eigenfunctions $\phi_{\lambda, -\mu_*}$ and using that $L_- \phi_{\lambda, -\mu_*} = 0$ and the eigen-equations, we find

$$\lambda = \langle \phi_{\lambda, -\mu_*}, L^2 \phi_{\lambda, -\mu_*} \rangle = \langle \phi_{\lambda, -\mu_*}, (-\hbar L_3 + L_3^2) \phi_{\lambda, -\mu_*} \rangle = \hbar \mu_* + \mu_*^2.$$

This gives $\lambda = \mu_*(\mu_* + \hbar)$. If we denote $\mu_* = \hbar l$, then $\lambda = \hbar^2 l(l + 1)$.

Finally, since, using the raising operator L_+ , we can go from $\phi_{\lambda, -\mu_*}$ to ϕ_{λ, μ_*} in an integer number of steps, we conclude that $2\mu_*/\hbar$ is a nonnegative integer, $\mu_* = \hbar l$, $l \in \frac{1}{2}\mathbb{Z}^+$. For each l , the index m runs the values $m = -l, \dots, l$. Hence the multiplicity of the eigenvalue $\lambda = \hbar^2 l(l + 1)$ of the operator L^2 is $2l + 1$. \square

Proof of Theorem 6.8. It remains to show that only integer l 's are realized now. To this end, we note that we can solve the eigenvalue problem for L_3 easily by using spherical coordinates (r, θ, ϕ) , in which $L_3 = \hbar \partial_{\phi}$, to obtain $L_3 e^{im\phi} = \hbar m e^{im\phi}$, so that $\mu = \hbar m$, where m must be an integer.

We have shown that every eigenvalue λ of L^2 is of the form $\lambda = \hbar^2 l(l + 1)$, with $l = 0, 1, \dots$, and has the multiplicity $2l + 1$. \square

The proof allows one to construct all eigenfunctions of L^2 , by solving the equation $L_- \phi_{\lambda, -\mu_*} = 0$, for $\lambda = \hbar^2 l(l + 1)$ and $\mu_* = \hbar l$, and then using the raising operator L_+ to obtain all other eigenfunctions with the same l . (In the spherical co-ordinates, we have $L_{\pm} = \hbar e^{\pm i\phi} (\pm \partial_{\theta} + i \cot \theta \partial_{\phi})$.)

Thus we determined completely the eigenvalues of the operators L^2 and L_3 : $\lambda = \hbar^2 l(l + 1)$ and $\mu = \hbar m$, where $m, l \in \mathbb{Z}$ and $m = -l, \dots, l$. Denote the joint eigenfunction of the operators L^2 and L_3 , corresponding to the eigenvalues $\lambda = \hbar^2 l(l + 1)$ and $\mu = \hbar m$, by $Y_l^m(\theta, \phi)$:

$$L^2 Y_l^m = \hbar^2 l(l + 1) Y_l^m, \quad L_3 Y_l^m = \hbar m Y_l^m. \quad (6.12)$$

The functions $Y_l^m(\theta, \phi)$ are the celebrated *spherical harmonics*. We know it is of the form $Y_l^m(\theta, \phi) = \tilde{Y}_l^m(\theta) e^{im\phi}$, $m \in \mathbb{Z}$, where $\tilde{Y}_l^m(\theta)$ is an eigenfunction of L^2 : $L^2 \tilde{Y}_l^m(\theta) = \hbar^2 l(l + 1) \tilde{Y}_l^m(\theta)$. More precisely,

$$Y_l^k(\theta, \phi) = c_{lk} P_l^{|k|}(\cos(\theta)) e^{ik\phi} \quad (6.13)$$

where $l = 0, 1, \dots$; $k \in \{-l, -l + 1, \dots, l - 1, l\}$; c_{lk} is a constant; and the *Legendre function* P_l^k can be written as

$$P_l^k(u) = \frac{(1 - u^2)^{k/2}}{2^l l!} \left(\frac{d}{du} \right)^{l+k} (u^2 - 1)^l. \quad (6.14)$$

This completes our excursion into the spectral theory of the operator L^2 .

6.4. Motion in a spherically symmetric potential. We consider a particle moving in a potential which is spherically symmetric, i.e. it depends only on $r = |x|$, $V(|x|)$. Then the corresponding Schrödinger operator is

$$H = -\frac{\hbar^2}{2m}\Delta + V(|x|), \quad (6.15)$$

acting on the Hilbert space $L^2(\mathbb{R}^3)$.

From Subsection ??, we know that H is self-adjoint and has essential spectrum filling in the half-line $[0, \infty)$. Hence the discrete spectrum if it exists is negative with only one possible accumulation point at 0. Our goal is, using the spherical symmetry of V , to reduce the eigenvalue problem for (6.15) to a problem in the radial variable, $r = |x|$, only, i.e. an ODE problem.

Because the potential is radially-symmetric (depends only on $r = |x|$), the Schrödinger operator H commutes with rotations and, consequently, their generators, angular momenta,

$$L_j = (x \times p)_j, \quad (6.16)$$

where $p = -i\hbar\nabla$, as usual.

We define the squared magnitude of the angular momentum, $L^2 = L_1^2 + L_2^2 + L_3^2$. Note that L^2 commutes with H . Furthermore, L^2 is a homogeneous degree zero operator, in the sense that it commutes with the rescalings, $T_\lambda^{\text{scal}} : \psi(x) \rightarrow \lambda^{3/2}\psi(\lambda x)$, $\lambda > 0$, and therefore it does not act on the radial variable $r = |x|$, but, if we introduce spherical coordinates (r, θ, ϕ) , where

$$x_1 = r \sin(\theta) \cos(\phi), \quad x_2 = r \sin(\theta) \sin(\phi), \quad x_3 = r \cos(\theta),$$

$0 \leq \theta < \pi$, $0 \leq \phi < 2\pi$, only on the angles (θ, ϕ) . A straightforward computation gives

$$-\hbar^2\Delta = -\hbar^2\Delta_r + \frac{1}{r^2}L^2, \quad (6.17)$$

where Δ_r is the “radial Laplacian”, acting only on the radial variable $r = |x|$ and given by

$$\Delta_r = \frac{1}{r^2} \frac{\partial}{\partial r} r^2 \frac{\partial}{\partial r} = \frac{\partial^2}{\partial r^2} + \frac{2}{r} \frac{\partial}{\partial r}.$$

Homework 6.11. * Prove (6.17). (Hint: Compute $L^2 = L_1^2 + L_2^2 + L_3^2$ as follows. $L^2 = (x_2p_3 - x_3p_2)^2 + \text{cyclic permutations} = (x_2p_3)^2 + (x_3p_2)^2 - x_2p_3x_3p_2 - x_3p_2x_2p_3 + \text{cyclic permutations} = r^2p^2 - \sum_j x_j^2p_j^2 - \sum_{ij} p_i x_i x_j p_j + \sum_j p_j x_j^2 p_j$. This gives $L^2 = r^2p^2 + (p \cdot x)(x \cdot p)$, from which one can easily derive (6.17).)

Homework 6.12. 1) Show that in the spherical coordinates (r, θ, ϕ) the Laplacian becomes

$$\Delta = \Delta_r + \frac{1}{r^2}\Delta_\Omega$$

where Δ_r is the “radial Laplacian” given by

$$\Delta_r = \frac{\partial^2}{\partial r^2} + \frac{2}{r} \frac{\partial}{\partial r}$$

(Δ_r depends only on the radial variable), and Δ_Ω is the Laplace-Beltrami operator on \mathbb{S}^2 , given in spherical coordinates (θ, ϕ) , by

$$\Delta_\Omega = \frac{1}{\sin(\theta)} \frac{\partial}{\partial \theta} \left(\sin(\theta) \frac{\partial}{\partial \theta} \right) + \frac{1}{\sin^2(\theta)} \frac{\partial^2}{\partial \phi^2}.$$

2) Show that in the spherical coordinates (r, θ, ϕ) ,

$$L^2(f(r)g(\theta, \phi)) = -f(r)\Delta_\Omega g(\theta, \phi).$$

Homework 6.13. Prove that the family of operators $T_\lambda^{\text{scal}} : \psi(x) \rightarrow \lambda^{d/2} \psi(\lambda x)$, $\lambda > 0$, is a one-parameter unitary group on $L^2(\mathbb{R}^d)$ and find its generator.

In view of (6.17) and the fact that L^2 commutes with H , our immediate task is to find the spectrum of L^2 , which we do in the next paragraph.

We now return to our original problem - the eigenvalue problem for the spherically symmetric Schrödinger operator, (6.15). To solve this problem, we seek eigenfunctions of H in the separated-variables form

$$\psi_{n\lambda\mu}(r, \theta, \phi) = R(r)Y_l^k(\theta, \phi)$$

where Y_l^k is a spherical harmonic. Plugging this into the eigenvalue equation $H\psi = E\psi$, we obtain

$$\left(\frac{\hbar^2}{2m} \left[-\Delta_r + \frac{l(l+1)}{r^2} \right] + V(r) \right) R = ER. \tag{6.18}$$

Usually, one cannot solve this equation explicitly with exception of a very few cases. The Schrödinger operator of the hydrogen atom is one of these cases to which we now proceed.

The Hydrogen atom. Because the Coulomb potential is radially-symmetric (depends only on $r = |x|$), the results above can be applied to the Schrödinger operator of the hydrogen atom. Recall that this operator is

$$H = -\frac{\hbar^2}{2m} \Delta - e^2/|x|$$

acting on the Hilbert space $L^2(\mathbb{R}^3)$. It is a remarkable fact that we can solve the equation (6.18) explicitly and find the eigenvalues explicitly. Indeed, aside from the infinite well, the only multi-dimensional potentials for which the Schrödinger eigenvalue problem can be solved explicitly are the harmonic oscillator and the Coulomb potential.

For the Coulomb potential, the equation (6.18) reads

$$\left(\frac{\hbar^2}{2m} \left[-\Delta_r + \frac{l(l+1)}{r^2} \right] - e^2/r \right) R = ER. \tag{6.19}$$

The solutions of the ODE (6.19) are well-studied (see, eg, [LL]). Without going into details, we remark that one can show (by power-series methods) that (6.19) has square-integrable solutions only for

$$n := \frac{e^2}{\hbar} \sqrt{\frac{-m}{2E}} \in \{l+1, l+2, \dots\}.$$

The corresponding eigenfunctions, R_{nl} are of the form

$$R_{nl}(r) = \rho^l e^{-\rho/2} F_{nl}(\rho)$$

where $\rho = \frac{2me^2}{\hbar^2} r$, and F_{nl} is a polynomial.

In full, then, the solutions of the eigenvalue problem $H\psi = E\psi$ are

$$\psi(r, \theta, \phi) = R_{nl}(r) Y_l^k(\theta, \phi)$$

where

$$l = 0, 1, 2, \dots; \quad k \in \{-l, -l+1, \dots, l\}; \quad n \in \{l+1, l+2, \dots\};$$

and the eigenfunctions are

$$E (= E_n) = - \left(\frac{me^4}{2\hbar^2} \right) \frac{1}{n^2}. \quad (6.20)$$

So we see that the Hydrogen atom has an infinite number of bound states below the essential spectrum (which starts at zero), which accumulate at zero (this result can be obtained by a general technique, without solving the eigenvalue problem). The ground state energy, attained when $l = k = 0, n = 1$, is $E_1 = -me^4/2\hbar^2$. An easy count finds the degeneracy of the energy level E_n to be

$$\sum_{l=0}^{n-1} (2l+1) = n^2.$$

Finally, we note that the expression (6.20) is in agreement with the empirical formula (“Balmer series”)

$$\Delta E = R \left(\frac{1}{n_f^2} - \frac{1}{n_i^2} \right).$$

Here $1 \leq n_f < n_i$ are integers labeling the final and initial states of the atom in a radiation process, R is a constant, and ΔE is the difference of the two energy levels. This formula predates quantum mechanics, and was based on measurements of absorption and emission spectra.

6.5. Spin. Quantum mechanical particles may also have internal degrees of freedom, which have no classical counterparts. Mathematically, this means that the wave functions $\psi(x)$ have values not in \mathbb{C} but in a higher dimensional (complex) space V ($\dim_{\mathbb{C}} V > 1$). The state space in this case is $L^2(\mathbb{R}^3; V)$. For V we take

an inner product, complex, finite-dimensional space, to be specific, $V = \mathbb{C}^n$. Then we have

$$L^2(\mathbb{R}^3; V) = \underbrace{L^2(\mathbb{R}^3; \mathbb{C})}_{\text{space of external degrees of freedom}} \otimes \underbrace{V}_{\text{space of internal degrees of freedom}}.$$

Here, on the l.h.s. we have the tensor product of two Hilbert spaces, which can be thought of as the space with a basis, $\phi_i v_j$, given by products of bases elements of each factor, equipped with the corresponding inner product ($\langle \phi v, \psi w \rangle = \langle \phi, \psi \rangle \langle v, w \rangle$). It can be identified with the space of square integrable functions with values in V ($x \rightarrow \phi(x)v$).

Now, the Schrödinger equation is invariant under a larger group of gauge transformations, $U_g^{\text{gauge}} : \psi(x) \rightarrow g\psi(x)$, $g \in U(n)$, where $U(n)$ is a group of complex, unitary, $n \times n$ matrices.

Consider the simplest non-trivial case of $n = 2$ and restrict ourselves to the subgroup, $SU(n)$, of $U(n)$ of the determinant one (the special unitary group). The Lie algebra, $su(2)$, of the group $SU(2)$ is the space of anti-hermitian $A^* = -A$ matrices acting on \mathbb{C}^2 , equipped with the commutator relation. We can choose a basis, A_1, A_2, A_3 , in $su(2)$, satisfying the commutation relation

$$[A_1, A_2] = A_3, [A_2, A_3] = A_1, [A_3, A_1] = A_2, \tag{6.21}$$

which have already came up in (6.3), while considering the Lie algebra, $so(3)$, of the group of rotations $SO(3)$. In fact, the Lie algebras $su(2)$ and $so(3)$ are isomorphic, while one can find a Lie group homomorphism $SU(2) \rightarrow SO(3)$, with the kernel $\{\pm 1\}$, i.e. $SU(2)/\{\pm 1\}$ and $SO(3)$ are Lie group isomorphic (see Remarks 1) and 2) below).

If we define the hermitian matrices $S_j := -i\hbar A_j$, we we obtain

$$[S_k, S_l] = i\hbar \epsilon^{klm} S_m, \tag{6.22}$$

where, recall, ϵ^{klm} is the Levi-Chivita symbol: $\epsilon^{123} = 1$ and ϵ^{klm} changes the sign under the permutation of any two indices. These are exactly the same relations as for the operators of angular momenta, (6.5). Consequently, the observable $S = (S_1, S_2, S_3)$ is called the *spin*.

However, there is an important difference. By Theorem 6.9, the operator $S^2 := \sum_j S_j^2$ has the eigenvalues $\lambda = \hbar^2 r(r + 1)$, where $r = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \dots$, of the multiplicities $2r + 1$. Now, the half-integers are allowed. We see that, unlike the angular momentum, the spin can take the half integer values. This means that the spin observables, $S_j, j = 1, 2, 3$, acting on the eigenspace V_r , with half-integer r , cannot be realized as generators of rotations on $L^2(\mathbb{R}^3)$.

We say that the *particle has the spin r iff the internal spin space V is an eigenspace, V_r , of S^2 (in some representation) with the eigenvalue $\lambda = \hbar^2 r(r + 1)$. By Theorem 6.9, V_r is of the dimension $2r + 1$, $r = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \dots$.*

As in Theorem 6.9, it is convenient to use, as a basis in V_r , the basis of the eigenvectors of the third component, S_3 , of the spin operator. In this basis, we

write elements of V_r , as $\vec{\psi}(x, s)$, $s = -r, \dots, r$, where $\psi(x, s)$ satisfies $S_3\psi(x, s) = \hbar s\psi(x, s)$, or as vectors $\vec{\psi}(x) = (\psi_1(x), \dots, \psi_{2r+1}(x))$, where each ψ_j belongs to the familiar one-particle space $L^2(\mathbb{R}^3) = L^2(\mathbb{R}^3; \mathbb{C})$, with the identification $\psi(x, s) \leftrightarrow \psi_s(x)$. (Usually such functions are written as columns, but for typographical simplicity we write them as rows.) This identifies V_r with the space \mathbb{C}^{2r+1} and $L^2(\mathbb{R}^3; \mathbb{C}) \otimes V_r = L^2(\mathbb{R}^3; V_r)$, with $L^2(\mathbb{R}^3; \mathbb{C}) \otimes \mathbb{C}^{2r+1} = L^2(\mathbb{R}^3; \mathbb{C}^{2r+1})$.

For $r = \frac{1}{2}$, it is convenient to write S_j as $S_j = \frac{\hbar}{2}\sigma_j$, where σ_j are the Pauli matrices

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (6.23)$$

For $r = \frac{1}{2}$, the spin operators S_j act on $V_{\frac{1}{2}}$ as

$$S_1\psi(x, s) = \hbar|s|\psi(x, -s), \quad S_2\psi(x, s) = -i\hbar s\psi(x, -s), \quad S_3\psi(x, s) = \hbar s\psi(x, s).$$

It is an experimental fact that all particles belong to one of the following two groups: particles with integer spins, or *bosons*, and particles with half-integer spins, or *fermions*. (The particles we are dealing with – electrons, protons and neutrons – are fermions, with spin $\frac{1}{2}$, while photons, which we will deal with later, are bosons, with spin 1. Nuclei, though treated as point particles, are composite objects whose spin could be either integer or half-integer.)

Remarks. 1) As we mentioned above, the algebra $su(2)$ is isomorphic to the algebra of the rotation group $so(3)$. Indeed, consider rotations around the x_1 –, x_2 – and x_3 –axes, $R_1(\varphi)$, $R_2(\varphi)$ and $R_3(\varphi)$, e.g. the rotations, $R_3(\varphi)$, around the x_3 –axis is given as

$$R_3(\varphi) = \begin{pmatrix} \cos \varphi & \sin \varphi & 0 \\ -\sin \varphi & \cos \varphi & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad (6.24)$$

The generators of these rotations are given by $r_i := \partial_\varphi R_i(\varphi)|_{\varphi=0}$. Say, for the rotations, $R_3(\varphi)$, around the x_3 –axis, we have

$$r_3 = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad (6.25)$$

Homework 6.14. Find the generators, r_1 , r_2 , r_3 , of the rotations, $R_1(\varphi)$, $R_2(\varphi)$, $R_3(\varphi)$, around the x_1 –, x_2 –, x_3 –axes.

Now we can define the isomorphism, $\phi : so(3) \rightarrow su(2)$, between the algebras $so(3)$ and $su(2)$ starting with $\phi(r_i) = S_i$ and then extending to the entire $so(3)$ by linearity.

Homework 6.15. Show that the map $\phi : so(3) \rightarrow su(2)$ defined above is an algebra isomorphism.

2) Though the algebras $so(3)$ and $su(2)$ are isomorphic, the groups, $SU(2)$ and $SO(3)$, are not: There is a Lie group homomorphism, $\phi : SU(2) \rightarrow SO(3)$, s.t. $\phi^{-1}(\mathbf{1}) = \{\pm \mathbf{1}\}$. Indeed, recall that the rotation group, $O(3)$, is the collection of all rotations of the space \mathbb{R}^3 , i.e. all linear transformations of \mathbb{R}^3 preserving the Euclidean norm $|x|^2$. We define $\phi : u \rightarrow R$ by $T_u h_x = h_{Rx}$, with $x \in \mathbb{R}^3$ and the transformation T_u on traceless hermitian matrices and the traceless hermitian matrices h_x defined as $T_u h = u h u^*$ and $h_x := \sum_j x_j \sigma_j$. Using that $\det h_x = -|x|^2$, one can easily show that $\phi(u)$ is a rotation of determinant 1, i.e. $\phi(u) \in SO(3)$. Hence $\phi : SU(2) \rightarrow SO(3)$. One show easily that ϕ is a group homomorphism. Finally, since $T_u h = u h u^* = T_{-u} h$, we see that ϕ maps u and $-u$ into the same element of $SO(3)$.

The spin interacts with an external magnetic field. The energy of this interaction has the similar form of the classical energy of interaction of a charge on a circular orbit with a magnetic field, which is $-\mu \cdot B(x)$, $\mu := g \frac{e}{2m} l$, where e and m are the charge and mass of the particle, l is its angular momentum (classical) and $B(x)$ is the magnetic field. In quantum mechanics, this interaction (in the case of spin $r = \frac{1}{2}$) is

$$-\mu \cdot B(x), \quad \mu := g \frac{e}{2m} S,$$

where g is called the gyromagnetic ratio, which, based on classical mechanics, one expects to be $g = 1$, but turns out to be $g = 2$ (plus small corrections, due to creation and annihilation of photons atom out of and into the vacuum, if the electromagnetic field is quantized).

TODO: Precession of the spin in magnetic field

6.6. Atoms and molecules and their spectra. Atoms and molecules (and solids) are the most important examples of many body quantum systems. Consider a molecule with N electrons of mass m and charge $-e$, and M nuclei of masses m_j and charges $Z_j e$, $j = 1, \dots, M$. In this case, the Schrödinger operator, H_{mol} , is

$$H_{mol} = \frac{1}{2m} \sum_1^N p_j^2 + \sum_1^M \frac{1}{2m_j} q_j^2 + V(x, y) \tag{6.26}$$

acting on $L^2(\mathbb{R}^{3(N+M)})$. Here $x = (x_1, \dots, x_N)$ are the electron coordinates, $y = (y_1, \dots, y_M)$ are the nucleus coordinates, $p_j = -i\hbar \nabla_{x_j}$ is the momentum of the j -th electron, $q_j = -i\hbar \nabla_{y_j}$ is the momentum of the j -th nucleus, and

$$V(x, y) = \frac{1}{2} \sum_{i \neq j} \frac{e^2}{|x_i - x_j|} - \sum_{i,j} \frac{e^2 Z_j}{|x_i - y_j|} + \frac{1}{2} \sum_{i \neq j} \frac{e^2 Z_i Z_j}{|y_i - y_j|}$$

is the sum of Coulomb interaction potentials between the electrons (the first term on the r.h.s.), between the electrons and the nuclei (the second term), and between the nuclei (the third term).

For a neutral molecule, we have

$$\sum_{j=1}^M Z_j = N.$$

If $M = 1$, the resulting system is called an atom, or Z -atom ($Z = Z_1$).

Since nuclei are much heavier than electrons, in the leading approximation one can suppose that the nuclei are frozen at their positions. One then considers, instead of (6.26), the Schrödinger operator

$$H_N^{BO} = \frac{1}{2m} \sum_1^N p_j^2 + V(x, y)$$

on $L^2(\mathbb{R}^{3N})$, the positions $y \in \mathbb{R}^{3M}$ of the nuclei appearing as parameters. This is called the Born-Oppenheimer approximation. It plays a fundamental role in quantum chemistry, where most computations are done with the operator H_{BO} . The eigenvalues of the operator H_{BO} are functions of the coordinates, y , of the nuclei. Minimizing the lowest eigenvalue – the ground state energy – with respect to y gives the equilibrium positions of the nuclei, i.e. the shape of the molecule. In the special case of an atom (a single nucleus), we have

$$H_N^{at} = \frac{1}{2m} \sum_{i=1}^N -\frac{\hbar^2}{2m} \Delta_{x_j} + V(x), \quad (6.27)$$

where $V(x) = \frac{1}{2} \sum_{i \neq j} \frac{e^2}{|x_i - x_j|} - \sum_i \frac{e^2 Z}{|x_i|}$ (assuming that the nucleus of the atom is at the origin).

Electrons are indistinguishable (purely quantum phenomenon)

$\implies |\Psi(x_1, \dots, x_N)|^2$ is symmetric w.r.t permutations of electron coordinates

$\implies \Psi$ is multiplied by ± 1 when a pair of coordinates are interchanged.

In fact electrons are Fermions which implies a more refined symmetry condition.

The spectrum of the Schrödinger operator H_N^{at} of an atom with N electrons is given by

Theorem 6.16. (*HVZ theorem*) $\sigma_{ess}(H_N^{at}) = [\Sigma_N, \infty)$, where $\Sigma_N = \inf \sigma(H_{N-1}^{at})$.

The energy $\Sigma_N = \inf \sigma(H_{N-1}^{at})$ is called the ionization threshold. To obtain $\sigma_{ess}(H_N^{at})$ we take one of the electrons to infinity and let it move freely there. The rest of the atom is placed in the ground state, so that the energy of the atom is

$$Energy = \Sigma_N + \frac{1}{2m} |k|^2 \quad \forall k \quad (6.28)$$

where k is the momentum of the electron which is placed at infinity.

Are atoms stable? We would like to show that H_N^{at} has at least one bound state.

Theorem 6.17. For $N < Z + 1$, H_N^{at} has infinite number of eigenvalues below its ionization threshold $\Sigma_N = \inf \sigma(H_{N-1}^{at})$.

Proof. In this prove we omit the superindex *at* in H_N^{at} and write H_N . Let $\Psi_N(x_1, x_2, \dots, x_N)$ be the ground state of H_N . We use the variational principle with the test function

$$\phi = \Psi_{N-1}(x_1, \dots, x_{N-1})f(x_N), \quad (6.29)$$

where for simplicity we take a non-symmetric ϕ . Using that

$$H_N = H_{N-1} - \frac{\hbar^2}{2m} \Delta_{x_N} + I_N, \quad (6.30)$$

with

$$I_N(x) := \sum_{i=1}^{N-1} \frac{e^2}{|x_i - x_N|} - \frac{e^2 Z}{|x_N|}, \quad (6.31)$$

and using that

$$H_{N-1} \Psi_{N-1} = E_{N-1} \Psi_{N-1}, \quad (6.32)$$

we obtain

$$H_N \phi = (E_{N-1} + I_N) \phi + \Psi_{N-1} \left(-\frac{\hbar^2}{2m} \Delta_{x_N} \right) f. \quad (6.33)$$

This implies that

$$\langle \phi, H_N \phi \rangle = E_{N-1} + \langle f, -\frac{\hbar^2}{2m} \Delta f \rangle + \langle \phi, I_N \phi \rangle. \quad (6.34)$$

One can show that for any N , Ψ_N satisfies an exponential bound $|\Psi_N(x)| \leq C e^{-\alpha|x|}$ for some $\alpha > 0$. Using this bound for Ψ_{N-1} , it follows that

$$|\langle \Psi_{N-1}, (I_N(x) - I_N(x_N)) \Psi_{N-1} \rangle_{L^2(\mathbb{R}^{N-1})}| \leq (const) |x_N|^{-2}, \quad (6.35)$$

where $I_N(x_N) := I_N(x)|_{x_i=x_N \forall i}$. Observe that $I_N(x_N) = -q/|x_N|$, where $q := (Z - N + 1)e^2$. Due to the decomposition (6.30) and since $H_N \phi = \Sigma \phi$, the last inequality implies

$$\langle \phi, (H_N - \Sigma_N) \phi \rangle \leq \left\langle f, \left(-\frac{\hbar^2}{2m} \Delta_{x_N} - \frac{q}{|x_N|} + \frac{const}{|x_N|^2} \right) f \right\rangle_{L^2(\mathbb{R}^3)}.$$

We let $f \in C_0^\infty(\mathbb{R}^3)$ satisfy $\|f\| = 1$ and

$$\text{supp}(f) \subset \{x_N \mid 1 < |x_N| < 2\}.$$

Then the functions

$$f_n(x_N) = n^{-3/2} f(n^{-1} x_N), \quad n = 1, 2, 4, 8, \dots,$$

are orthonormal, and have disjoint supports. Thus the corresponding trial states $\phi_n(x) = \Psi_{N-1}(x_1, \dots, x_{N-1})f_n(x_N)$ satisfy $\langle \phi_n, H_N \phi_m \rangle = 0$ for $n \neq m$, and

$$\langle \phi_n, (H_N - \Sigma_N) \phi_n \rangle \leq -(\text{const}) \frac{q}{n} + (\text{const}) \frac{1}{n^2} < 0,$$

for some positive constants, if n is sufficiently large. Using this one can show that H_N possesses infinitely many discrete eigenvalues below the threshold Σ_N .

Taking into account statistics. Since electrons are indistinguishable and are Fermions, we should consider only test functions which are either symmetric or anti-symmetric with respect to permutation of particle co-ordinate. Hence we have to replace $\phi = \Psi_{N-1}(x_1, \dots, x_{N-1})f(x_N)$ by such a function. As a new test function we choose

$$\phi := \sum_{j=1}^N \phi^{(j)}. \quad (6.36)$$

where

$$\phi^{(j)} := \pm \frac{1}{\sqrt{N}} \Psi_{N-1}(x_1, \dots, x_{j-1} x_{j+1}, \dots, x_N) f(x_j).$$

Let $I_j := \sum_{i:i \neq j} \frac{e^2}{|x_i - x_j|} - \frac{e^2 Z}{|x_j|}$. Choose f so that

$$|\langle \phi^{(i)}, \phi^{(j)} \rangle| \ll 1 \quad \text{and} \quad |\langle \phi^{(i)}, (-\frac{\hbar^2}{2m} \Delta_{x_j} + I_j) \phi^{(j)} \rangle| \ll 1 \quad \text{for } i \neq j. \quad (6.37)$$

Namely, we take $f_\alpha(x) = \alpha^{\frac{3}{2}} f(\alpha x)$ with $\|f\| = 1$. Then for $i \neq j$ we have

$$|\langle \phi^{(i)}, \phi^{(j)} \rangle| \leq C \frac{\alpha^3}{N},$$

where $C = \sup |f|^2 \int |\bar{\Psi}_{N-1}(x_1, \dots, x_{i-1} x_{i+1}, \dots, x_N) \Psi_{N-1}(x_1, \dots, x_{j-1} x_{j+1}, \dots, x_N)| dx_1 \dots dx_N$, and, similarly,

$$|\langle \phi^{(i)}, (-\frac{\hbar^2}{2m} \Delta_{x_j} + I_j) \phi^{(j)} \rangle| \lesssim \frac{\alpha^3}{N} \quad \text{for } i \neq j.$$

By taking $\alpha \rightarrow 0$, the last two equations imply (6.37) and therefore

$$\begin{aligned} \langle \phi, H \phi \rangle &\leq E_{N-1} + \langle f_\alpha, (-\frac{\hbar^2}{2m} \Delta_x - \frac{q}{|x|}) f_\alpha \rangle + O(\frac{\alpha^3 N^2}{N}) \\ &= E_{N-1} + \alpha^2 \langle f, -\frac{\hbar^2}{2m} \Delta f \rangle - \alpha \langle f, \frac{q}{|x|} f \rangle + O(\alpha^3 N). \end{aligned}$$

Hence $\langle \phi, H \phi \rangle < E_{N-1}$, if $\alpha \ll \frac{1}{\sqrt{N}}$ and $\alpha < \frac{\langle f, \frac{q}{|x|} f \rangle}{\langle f, -\frac{\hbar^2}{2m} \Delta f \rangle}$. This proves the existence of ground state energy for $H_N \forall N$. \square

Homework: Go over the proof above and, whenever necessary, fill in the details, if needed.

A more refined estimate of $\langle \phi, I_N \phi \rangle$ (replacing (6.35)). Consider the one-electron density

$$\rho_N(y) := \int |\Psi_N(y, x_2, \dots, x_n)| dx_2 \dots dx_N. \quad (6.38)$$

Assume that for any N , $\rho_N(y)$ is spherically symmetric. Instead of (6.35) we can write

$$\langle \phi, I_N \phi \rangle = \langle f, W f \rangle, \text{ where } W(x_N) = \int I_N |\Psi_{N-1}(x_1, \dots, x_{N-1})|^2 d^{N-1}x. \quad (6.39)$$

We compute

$$W(x_N) = (N-1)e^2 \int \frac{\rho_{N-1}(y)}{|x - x_N|} dy - \frac{e^2 Z}{|x_N|}, \quad (6.40)$$

where $\rho_{N-1}(x_1) = \int |\Psi_{N-1}|^2 dx_2 \dots dx_{N-1}$. Since ρ_{N-1} is spherically symmetric, we have by Newton's theorem

$$\int_{\mathbb{R}^3} \frac{\rho_{N-1}(y) dy}{|y - x_N|} = \frac{1}{|x_N|} \int_{|y| \leq |x_N|} \rho_{N-1}(y) dy + \int_{|y| \geq |x_N|} \frac{1}{|y|} \rho_{N-1}(y) dy. \quad (6.41)$$

Using that $\int \rho_{N-1}(y) dy = 1$, this can be estimated as

$$\int_{\mathbb{R}^3} \frac{\rho_{N-1}(y) dy}{|y - x_N|} \leq \frac{1}{|x_N|} \int_{\mathbb{R}^3} \rho_{N-1}(y) dy = \frac{1}{|x_N|}. \quad (6.42)$$

(Moreover, $\rho_{N-1}(x) = O(e^{-\delta|x|}) \implies \int_{|y| \leq |x_N|} \rho_{N-1}(y) dy = 1 + O(e^{-\delta|x_N|})$.) Hence

$$W(x_N) \leq \frac{(N-1)e^2}{|x_n|} - \frac{e^2 Z}{|x_N|} = -\frac{(Z - N + 1)e^2}{|x_N|}. \quad (6.43)$$

This together with (6.30) gives the following estimate for the test function (6.29)

$$\langle H_N \rangle_\phi \leq E_{N-1} + \langle f, \left(-\frac{\hbar^2}{2m} \Delta - \frac{q}{|x_N|} \right) f \rangle.$$

where, recall, $q = (Z - N + 1)e^2$.

Remark. The accumulation of eigenvalues at Σ_N can be studied by similar arguments. Consider trial wave functions, ψ_{nm} , constructed as above, with $f_{nm}(x_N)$ a hydrogen atom eigenfunction of energy $-n^{-2}$ (in suitable units). Then one can show

$$\|(H_N - E_n)\psi_{nm}\| \leq (\text{const})n^{-\alpha}$$

for some $\alpha > 3$, where $E_n = \Sigma - n^{-2}$. This implies that H_N has groups of eigenvalues close to E_n compared to the spacing $E_{n+1} - E_n$ as $n \rightarrow \infty$ (Rydberg

states). This analysis can be easily extended to take into account the particle statistics.

Homework 6.18. *Show that the Schrödinger operator describing the Helium atom with infinitely heavy nucleus has at least one discrete eigenvalue (isolated eigenvalue of a finite multiplicity).*

6.7. The Harmonic Oscillator. The Hamiltonian of the quantum *harmonic oscillator* in r dimensions is

$$H_{\text{ho}} = -\frac{\hbar^2}{2m}\Delta + \frac{1}{2}m \sum_{i=1}^r \omega_i^2 x_i^2,$$

acting on the space $L^2(\mathbb{R}^r)$. By Theorem ??, $\sigma(H)$ consists of isolated eigenvalues, increasing to infinity. We will solve the eigenvalue problem explicitly for this operator.

Theorem 6.19. *The spectrum of H_{ho} is*

$$\sigma(H_{\text{ho}}) = \left\{ \sum_{i=1}^r \hbar\omega_i(n_i + 1/2) \mid n_i = 0, 1, 2, \dots \right\},$$

with eigenfunctions

$$\psi_0 := \prod_{i=1}^r (2\pi m \hbar \omega_i)^{-1/4} e^{-m \sum_{i=1}^r \omega_i x_i^2 / 2\hbar}, \quad n_i = 0, \quad \forall i, \quad (6.44)$$

$$\psi_{\mathbf{n}} = \prod_{i=1}^r (1/\sqrt{n_i!}) (a_i^*)^{n_i} \psi_0, \quad \mathbf{n} := (n_1, \dots, n_r), \quad (6.45)$$

with a_i^* the first order differential operators defined in (6.46).

We derive a representation of the operator H which facilitates its spectral analysis. It also prepares us for a similar technique we will encounter in the more complex situation of second quantization and quantum electrodynamics. We introduce the *creation* and *annihilation* operators

$$a_j := \frac{1}{\sqrt{2m\hbar\omega_j}}(m\omega_j x_j + ip_j) \quad \text{and} \quad a_j^* := \frac{1}{\sqrt{2m\hbar\omega_j}}(m\omega_j x_j - ip_j). \quad (6.46)$$

These operators are adjoint of each other and they satisfy the commutation relation

$$[a_i, a_j^*] = \delta_{ij}. \quad (6.47)$$

Using that $H = \sum_{i=1}^r \left(-\frac{\hbar^2}{2m} \partial_{x_i}^2 + \frac{1}{2} m \omega_i^2 x_i^2 \right)$, the Hamiltonian H can be re-written in terms of a_i and a_i^* 's as follows: on $L^2(\mathbb{R}^r)$. It can be rewritten in the form

$$H_{\text{ho}} = \sum_{i=1}^r \hbar \omega_i \left(a_i^* a_i + \frac{1}{2} \right), \quad (6.48)$$

where a_i and a_i^* are the harmonic oscillator annihilation and creation operators. We say that this expression is in *normal form* because a^* appears to the left of a . Now, we are ready to prove Theorem 6.19.

Proof of Theorem 6.19. First we find the ground state and ground state energy of H_{ho} . We define the *particle number operators*

$$N_i := a_i^* a_i,$$

so that, by the expression (6.48),

$$H_{\text{ho}} = \sum_{i=1}^r \hbar \omega_i \left(N_i + \frac{1}{2} \right),$$

Now, because a_i^* is the adjoint of a_i , N_i are non-negative $N_i \geq 0$:

$$\langle \psi, N_i \psi \rangle = \|a_i \psi\|^2 \geq 0$$

for any ψ . Therefore, ψ , with the smallest eigenvalue, $\sum_{i=1}^r \frac{1}{2} \hbar \omega_i$, (i.e. the one that minimizes the average energy, $\langle \psi, H_{\text{ho}} \psi \rangle$), satisfies $N_i \psi = 0$, $\forall i$, which holds iff $a_i \psi = 0$, $\forall i$. These equations,

$$a_i \psi = \frac{1}{\sqrt{2m\hbar\omega_i}} (m\omega_i x_i + \hbar \partial / \partial x_i) \psi = 0,$$

can be easily solved, giving the unique (up to a multiplicable constant which we denote c) family of solutions

$$\psi_{0i}(x_i) := c e^{-m\omega_i x_i^2 / 2\hbar}$$

We choose c so that ψ_{0i} is normalized to one: $\|\phi_{0i}\| = 1$. Thus (6.44) is the (normalized) ground state of H_{ho} with the ground state energy $\sum_{i=1}^r \frac{1}{2} \hbar \omega_i$.

To find the excited states we observe that (6.47) imply that the operators N_i satisfy the relations

$$N_i a_i = a_i (N_i - 1), \quad (6.49)$$

$$N_i a_i^* = a_i^* (N_i + 1). \quad (6.50)$$

The commutation relation (6.50) implies $N_i (a_i^*)^n = (a_i^*)^n (N_i + n)$, which gives $N_i a_i^* \psi_0 = a_i^* \psi_0$ and in general

$$N_i (a_i^*)^n \psi_0 = n (a_i^*)^n \psi_0.$$

Thus $\phi_{n_i} := (a_i^*)^{n_i} \psi_0$ is an eigenfunction of N_i with eigenvalue n_i and therefore $\phi_{\mathbf{n}} := \prod_{i=1}^r (a_i^*)^{n_i} \psi_0$ is an eigenfunction of H_{ho} with eigenvalue $\sum_{i=1}^r \hbar \omega_i (n_i + \frac{1}{2})$.

Homework 6.20. Show that $\|\phi_{\mathbf{n}}\|^2 = |c|^2 \prod_{i=1}^r n_i!$. Hint: write $\|\phi_{\mathbf{n}}\|^2 = \langle \psi_0, \prod_{i=1}^r (a_i)^{n_i} (a_i^*)^{n_i} \psi_0 \rangle$, then pull the a_i 's through the a_i^* 's (including the necessary commutators) until they hit ψ_0 and annihilate it.

So $\psi_{\mathbf{n}} := \prod_{i=1}^r \frac{1}{\sqrt{n_i!}} (a_i^*)^{n_i} \psi_0$, $\mathbf{n} := (n_1, \dots, n_r)$, is a normalized eigenfunction of H_{ho} with eigenvalue $\sum_{i=1}^r \hbar \omega_i (n_i + 1/2)$.

We now show that these are the only eigenfunctions. To simplify the notation we do this in dimension 1, i.e. for $r = 1$. It follows from the commutation relation (6.49) that $Na^n = a^n(N - n)$ (this is adjoint of $N(a^*)^n = (a^*)^n(N + n)$). Hence if ψ is any eigenfunction of N with eigenvalue $\lambda > 0$, then

$$Na^m \psi = (\lambda - m)a^m \psi. \quad (6.51)$$

If we choose m so that $\lambda - m < 0$ we get a contradiction to $N \geq 0$ unless $a^m \psi = 0$. Let j be the largest integer s.t. $a^j \psi \neq 0$ and $a^{j+1} \psi = aa^j \psi = 0$. This implies

$$a^j \psi = c \psi_0 \quad (6.52)$$

where $c \neq 0$ is a constant. Hence by $N\psi_0 = 0$, (6.52) and (6.51), we have $0 = Na^j \psi = (\lambda - j)a^j \psi$ and therefore $\lambda = j$. Thus ψ corresponds to the eigenvalue j . If ψ is not proportional to ψ_j , then we can choose it to satisfy $\langle \psi, \psi_j \rangle = 0$. However, by $a^j \psi = c\psi_0$ and $\psi_j = (a^*)^j \psi_0$, we have that $\langle \psi, \psi_j \rangle = \langle a^j \psi, \psi_0 \rangle = c \|\psi_0\|^2 \neq 0$, a contradiction. So we are done. \square

Remark 6.21. One can extend the last part of the proof above to show that any function $f \in L^2(\mathbb{R}^r)$ can be written as

$$f(x) = \sum_{\mathbf{n}} c_{\mathbf{n}} \psi_{\mathbf{n}} = \sum_{\mathbf{n}} c_{\mathbf{n}} \prod_{i=1}^r \frac{1}{\sqrt{n_i!}} (a_i^*)^{n_i} \psi_0.$$

Hence the set $\{\psi_{\mathbf{n}}\}$ form an orthonormal basis in $L^2(\mathbb{R}^r)$ and the space $L^2(\mathbb{R}^r)$ is isometric to the space $\mathcal{F}^{(r)} := \bigoplus_{n=0}^{\infty} \mathbb{C}_{sym}^n$ (for $r = 1$), where \mathbb{C}_{sym}^n is equal to $\{0\}$ for $n = 0$ and to $\mathbb{C}^n / \mathfrak{S}^n$ for $n \geq 1$. Here \mathfrak{S}^n is the symmetric group of permutations of n indices.

Remark 6.22. Since ψ_0 is positive and normalized, $\int \psi_0^2 = 1$, the operator $U : f \rightarrow \psi_0^{-1} f$ maps unitarily the space $L^2(\mathbb{R}^r, d^r x)$ into the space $L^2(\mathbb{R}^r, \psi_0^2 d^r x)$. Under this map the operator H_{ho} is mapped into $L := UH_{ho}U^{-1}$ acting on the space $L^2(\mathbb{R}^r, \psi_0^2 d^r x)$. We compute

$$L = \sum_{j=1}^r \hbar \left(-\frac{\hbar}{2m} \partial_{x_j} + \omega_j x_j \right) \partial_{x_j} \quad (6.53)$$

The operator L , with $\hbar = 1$, $2m = 1$, is the generator of the Ornstein-Uhlenbeck stochastic process (see [GJ]).

6.8. A particle in a constant external magnetic field. Consider a charged quantum particle moving in a constant magnetic field, B , with no electric field present. According to (??), the quantum Hamiltonian of such a system is

$$H(A) = \frac{1}{2m}(p - eA)^2,$$

acting on $L^2(\mathbb{R}^3)$, where A is the vector potential of the magnetic field B and therefore satisfies

$$\operatorname{curl} A = B. \tag{6.54}$$

Recall that the corresponding Schrödinger equation has the gauge symmetry,

$$H(A + \nabla\chi) = e^{ie\chi/\hbar}H(A)e^{-ie\chi/\hbar}. \tag{6.55}$$

We fix the gauge by choosing a special solution of equation (6.54). A possible choice for A is

$$A'(x) = \frac{1}{2}B \times x. \tag{6.56}$$

Another possibility, supposing B to be directed along the x_3 axis ($B = (0, 0, b)$) is

$$A''(x) = b(-x_2, 0, 0). \tag{6.57}$$

Homework 6.23. Check that both (6.56) and (6.57) yield the magnetic field B , and that the two are gauge-equivalent.

For the first choice of A , we chose the x_3 axis along B so that $B = (0, 0, b)$, and therefore $A'(x) = \frac{1}{2}(-x_2, x_1, 0)$.

Homework 6.24. Show that Schrödinger operator $H(A)$ in the gauge (6.56) is of the form

$$H(A') = H_2(A') + \frac{1}{2m}p_3^2, \tag{6.58}$$

$$H_2(A') := \frac{1}{2m}[p_1^2 + p_2^2 + \frac{1}{4}e^2b^2(x_1^2 + x_2^2) - ebL_3] \tag{6.59}$$

where the operator $H_2(A')$ acts only on the variables x_1 and x_2 and $L_3 = p_1x_2 - p_2x_1$ is the third component of the angular momentum operator.

The spectrum of the operator $H(A')$ can be found by the separation of variables. Hence, since the spectral properties of $\frac{1}{2m}p_3^2$ are known, it suffices to find by the spectrum of $H_2(A')$. To do this we may introduce the harmonic oscillator annihilation and creation operators, α and α^* , with

$$\alpha := (\nabla_A)_1 + i(\nabla_A)_2, \quad \alpha^* := -(\nabla_A)_1 + i(\nabla_A)_2, \tag{6.60}$$

where $\nabla_A := \nabla - i\frac{e}{\hbar}A(x)$. These operators satisfy the standard commutation relations:

$$[\alpha, \alpha^*] = 2 \operatorname{curl} A = 2b\frac{e}{\hbar}. \tag{6.61}$$

To prove this one writes α as

$$\alpha = \partial_{x_1} + i\partial_{x_2} + \frac{be}{2\hbar}(x_1 + ix_2), \quad (6.62)$$

and similarly α^* , and then computes $[\alpha, \alpha^*]$ directly. The operator $H_2(A')$ can be expressed in terms of these operators as $H_2(A') = \frac{1}{2m}(\alpha^*\alpha + b\frac{e}{\hbar})$. Hence the spectrum of $H_2(A')$ consist of the eigenvalues $\frac{\hbar e}{2m}(2bn + b), n = 0, 1, \dots$, of the infinite multiplicity.

Homework 6.25. Verify the above statements. Use them to find the spectra of $H_2(A')$ and $H(A')$.

Using the second choice for A , the corresponding Schrödinger operator is

$$H(A'') = H_2(A'') + \frac{1}{2m}p_3^2, \quad H_2(A'') := \frac{1}{2m}[(p_1 + ebx_2)^2 + p_2^2].$$

To analyze $H_2(A'')$, we apply the Fourier transform to only the first variable ($x_1 \mapsto k_1$). This results in the unitarily equivalent operator

$$\tilde{H}_2(A'') = \frac{1}{2m}p_2^2 + \frac{m\omega^2}{2}(x_2 + \frac{1}{eb}k_1)^2,$$

where $\omega = eb/m$ and k_1 acts as a multiplication operator. We remark that $\tilde{H}_2(A'')$ acts as a harmonic oscillator (centred at $-\frac{1}{eb}k_1$) in the variable x_2 , and as a multiplication operator in k_1 . In the following problem you are asked to determine the spectrum of this operator.

Homework 6.26.

- (1) Show that the energy levels of $H_2(A'')$ (called *Landau levels*) are given by

$$(n + \frac{1}{2})\hbar\omega$$

where $n = 0, 1, 2, \dots$ and $\alpha \in \mathbb{R}$.

- (2) Show that the corresponding generalized eigenfunctions are of the form

$$\psi_{n,\alpha}(x_1, x_2) = (2\pi\hbar)^{-1/2} \int e^{ix_1k_1/\hbar} f(k_1)\phi_n(x_2 + k_1/eb)dk_1$$

where ϕ_n is the n th eigenfunction of the harmonic oscillator and f is an arbitrary function.

- (3) Find the energy levels and the corresponding generalized eigenfunctions of $H(A'')$

Now, we find the eigenfunctions of the operator $H_2(A')$ given in (6.59). First we write the annihilation operator, α , given in (6.62) as

$$\alpha = 2\bar{\partial} + \frac{be}{2\hbar}z, \quad \bar{\partial} := \frac{1}{2}(\partial_{x_1} + i\partial_{x_2}), \quad z := x_1 + ix_2. \quad (6.63)$$

Next, a straightforward computation shows that

$$e^{\frac{be}{4\hbar}|z|^2} \left(\bar{\partial} + \frac{be}{4\hbar} z \right) e^{-\frac{be}{4\hbar}|z|^2} = \bar{\partial}. \quad (6.64)$$

Hence ψ_0 solves the equation $\alpha\psi_0 = 0$ iff $f(z) := e^{\frac{be}{4\hbar}|z|^2}\psi_0$ solves the equation $\bar{\partial}f = 0$. The latter equation implies that f is a holomorphic function. Hence the eigenspace of the operator $H_2(A')$ corresponding to the lowest eigenvalue $\frac{\hbar^2}{2m}b$ is $V_0 := \{f(z)e^{-\frac{be}{4\hbar}|z|^2}\}$, where f is a holomorphic function which does not grow too fast.

V_0 is the *zeroth Landau level*. Higher Landau levels are obtained by applying powers of the creation operator α^* to V_0 : $V_n := (\alpha^*)^n V_0$.

Recall that V_0 is an infinite dimensional space. A convenient basis in this space is given by using the creation and annihilation operators

$$\beta = 2\partial + \frac{be}{2\hbar}\bar{z}, \quad \beta^* = -2\bar{\partial} + \frac{be}{2\hbar}z. \quad (6.65)$$

Indeed, the equations $\alpha\psi_0 = 0$ and $\beta\psi_0 = 0$ have the solution $\psi_0 := e^{-\frac{be}{4\hbar}|z|^2}$ (unique up to multiple factors). Applying $(\beta^*)^m$ to this solution, one generates an orthogonal basis in V_0 . This a basis of eigenfunction of the angular momentum operator $L_3 = p_1x_2 - p_2x_1$. Indeed, in z variables, this operator is written as $L_3 = \hbar(z\partial - \bar{z}\bar{\partial})$, which gives $L_3(\beta^*)^m\psi_0 = m(\beta^*)^m\psi_0$.

6.9. Aharonov-Bohm Effect. Consider a charged quantum particle in 2 dimensions moving in an external electric and magnetic fields with the potentials V and A , respectively. Furthermore, we assume that V is radially symmetric, $V(x) = U(|x|)$ and A is of the form

$$A(x) := f(|x|)x^\perp/|x|^2, \quad x^\perp := (-x_2, x_1).$$

Note that $\text{div } A = 0$. Using that $(-i\hbar\nabla - eA)^2 = -\Delta + 2eA \cdot (-i\hbar\nabla) + |eA|^2$ and passing to the polar co-ordinates, (r, θ) , we find that

$$\begin{aligned} (-i\hbar\nabla - eA)^2 &= -\Delta_r - \frac{1}{r^2}(\partial_\theta^2 + 2ef(r) \cdot (-i\hbar\partial_\theta) + f^2(r)) \\ &= -\Delta_r - \frac{1}{r^2}(-i\hbar\partial_\theta + 2ef(r))^2, \end{aligned}$$

where Δ_r is the radial Laplacian, $\Delta_r := \frac{1}{r}\partial_r r \partial_r$. Using this, we obtain

$$H(A) = -\Delta_r - \frac{1}{r^2}(-i\hbar\partial_\theta + 2ef(r))^2 + U(r).$$

Expanding in the Fourier series, $\psi(r, \theta) = \sum_n \psi_m(r)e^{im\theta}$ (the separation of variables), we arrive at the eigenvalue problem for the eigenvalues of $H_m\psi_m = E\psi_m$, where $H_m := -\Delta_r - \frac{1}{r^2}(\hbar m + 2ef(r))^2 + U(r)$. **(under construction)**

6.10. Variational Characterization of Eigenvalues. We consider, for the moment, a self-adjoint operator H , acting on a Hilbert space \mathcal{H} . In this part we discuss the following important characterization of eigenvalues of H in terms of the minimization problem for the “energy” functional $\langle \psi, H\psi \rangle$.

Theorem 6.27. *The Ritz variational principle: for any $\psi \in D(H)$,*

$$\inf_{\|\psi\|=1} \langle \psi, H\psi \rangle = \inf \sigma(H).$$

Corollary 6.28. (1) *The left hand side has a minimizer if and only if $H\psi = \lambda\psi$, with $\lambda := \inf \sigma(H)$.*

(2) *If there is a ψ (called a test function) with $\|\psi\| = 1$ and*

$$\langle \psi, H\psi \rangle < \inf \sigma_{ess}(H),$$

then H has at least one eigenvalue below its essential spectrum.

Proof. Denote $E(\psi) := \langle \psi, H\psi \rangle$. If $\lambda = \inf \sigma(H)$ is an eigenvalue of H , with normalized eigenvector ψ_0 , then

$$E(\psi_0) = \lambda = \inf_{\|\psi\|=1} E(\psi),$$

and therefore ψ_0 is a minimizer of E .

On the other hand, let ψ_0 be a minimizer of E , among $\psi \in D(H)$, with $\|\psi\| = 1$, and consider the functions, $E(\psi_s)$ and $C(\psi_s)$, where $\psi_s := \psi_0 + s\xi$ and $C(\psi) := \|\psi\|^2 - 1$, depending on a single variable s . By our assumption, the function, $E(\psi_s)$ has a minimum at $s = 0$, under the constraint $C(\psi_s) = 0$. We know from the calculus that, for some $\lambda \in \mathbb{R}$ (called a *Lagrange multiplier*), ψ_0 satisfies

$$\frac{d}{d\lambda} E(\psi_s)|_{s=0} = \lambda \frac{d}{ds} C(\psi_s)|_{s=0} \quad \text{and} \quad C(\psi_0) = 0,$$

for any $\xi \in \mathcal{H}$. Using that H is symmetric, we compute $\frac{d}{d\lambda} E(\psi_s)|_{s=0} = \langle \xi, H\psi_0 \rangle$ and $\frac{d}{ds} C(\psi_s)|_{s=0} = \langle \xi, \psi_0 \rangle$, which gives the equation $\langle \xi, H\psi_0 \rangle = \lambda \langle \xi, \psi_0 \rangle$. Since this is true for all $\xi \in \mathcal{H}$, we conclude that $H\psi_0 = \lambda\psi_0$, i.e ψ_0 is an eigenvector of H and λ is the corresponding eigenvalue.

To obtain the second part of the theorem, we note that, by Theorem 6.27,

$$\inf \sigma(H) = \inf E < \langle \psi, H\psi \rangle < \inf \sigma_{ess}(H),$$

so $\lambda = \inf \sigma(H)$ must be an (isolated) eigenvalue of H . □

As an example application of this theorem, we consider the bound state problem for the Hydrogen atom. The Schrödinger operator for the Hydrogen atom is

$$H = -\frac{\hbar^2}{2m} \Delta - \frac{e^2}{|x|}$$

acting on the Hilbert space $L^2(\mathbb{R}^3)$. Take the (normalized) test function $\psi(x) = \sqrt{\mu^3/\pi}e^{-\mu|x|}$ for some $\mu > 0$ to be specified later and **compute** (passing to spherical coordinates)

$$\begin{aligned}\langle \psi, H\psi \rangle &= \frac{\hbar^2}{2m} \int |\nabla\psi|^2 - e^2 \int \frac{1}{|x|} |\psi|^2 \\ &= \frac{\hbar^2}{2m} \frac{\mu^3}{\pi} \mu^2 4\pi \int_0^\infty e^{-2\mu r} r^2 dr - e^2 \frac{\mu^3}{\pi} 4\pi \int_0^\infty e^{-2\mu r} r dr.\end{aligned}$$

To compute the above integrals, note

$$\int_0^\infty e^{-\alpha r} r^2 dr = \frac{d^2}{d\alpha^2} \int_0^\infty e^{-\alpha r} dr$$

and

$$\int_0^\infty e^{-\alpha r} r dr = -\frac{d}{d\alpha} \int_0^\infty e^{-\alpha r} dr.$$

Since $\int_0^\infty e^{-\alpha r} dr = \alpha^{-1}$, we find $\int_0^\infty e^{-\alpha r} r^2 dr = 2\alpha^{-3}$, and $\int_0^\infty e^{-\alpha r} r dr = \alpha^{-2}$. Substituting these expressions with $\alpha = 2\mu$ into the formula for $\langle \psi, H\psi \rangle$, we obtain

$$\langle \psi, H\psi \rangle = \frac{\hbar^2}{2m} \mu^2 - e^2 \mu.$$

The right hand side has a minimum at $\mu = me^2/\hbar^2$, which is equal to

$$\langle \psi, H\psi \rangle|_{\mu=me^2/\hbar^2} = -\frac{me^4}{2\hbar^2}.$$

Since $\sigma_{\text{ess}}(H) = [0, \infty)$ (according to Theorem ??), we conclude that H has negative eigenvalues, and the lowest negative eigenvalue, λ_1 , satisfies the estimate

$$\lambda_1 \leq -\frac{me^4}{2\hbar^2}.$$

As it happened, the function $\psi(x) = \sqrt{\mu^3/\pi}e^{-\mu|x|}$ for $\mu = me^2/\hbar^2$ is an exact eigenfunction for the Hydrogen atom hamiltonian and therefore the result above is in fact exact. This is just luck.

Homework 6.29. Estimate the ground state energy of $H = -\frac{\hbar^2}{2m}\Delta + |x|^4$ in three dimensions, using the variational principle with the test functions

$$\psi_\mu(x) = (8\pi)^{-1/2} \mu^{3/2} e^{-\mu|x|/2}.$$

Is the estimate found lower or upper estimate of the true ground state energy?

6.11. Particle in a Periodic Potential. We consider a particle moving in a periodic potential. A primary example of this situation is an electron moving in the potential created by ions or atoms of a solid crystal lattice. Such a particle is described by a self-adjoint Schrödinger operator

$$H = -\frac{\hbar^2}{2m}\Delta + V$$

on $L^2(\mathbb{R}^3)$, with the potential $V(x)$, having certain periodicity properties which we now explain.

First we identify the notion of physical crystal lattice with the mathematical (*Bravais*) lattice, \mathcal{L} , which is defined as a subset of \mathbb{R}^3 with the properties

- \mathcal{L} is discrete, i.e. it has no finite limit points;
- \mathcal{L} is a subgroup of the additive group of translations of \mathbb{R}^3 ;
- \mathcal{L} is not contained in any proper vector subspace of \mathbb{R}^3 .

These properties imply that \mathcal{L} is a set of points in \mathbb{R}^3 of the form

$$\mathcal{L} = \{m_1s_1 + m_2s_2 + m_3s_3 \mid m_1, m_2, m_3 \in \mathbb{Z}\}$$

for some three linearly independent vectors $s_1, s_2, s_3 \in \mathbb{R}^3$, called a *basis* of \mathcal{L} . A basis in \mathcal{L} is not unique.

We say $V(x)$ is periodic with respect to a lattice \mathcal{L} if $V(x+s) = V(x)$ for any $s \in \mathcal{L}$. This implies that the operator H commutes with the lattice translations,

$$T_s H = H T_s, \quad \forall s \in \mathcal{L}, \quad (6.66)$$

where T_s is the translation operator, given by $T_s f(x) = f(x+s)$. The latter are unitary operators satisfying $T_t T_s = T_{t+s}$.

Homework 6.30. Show that T_s are unitary operators satisfying $T_t T_s = T_{t+s}$.

Due to (6.66), the operator H can be decomposed into a direct fiber integral, as follows. We first define Ω to be the basic lattice cell of the lattice \mathcal{L} , i.e. the cell spanned by the basis vectors s_1, s_2, s_3 . Let \mathcal{L}^* be the lattice dual to \mathcal{L} , i.e. the lattice with the basis s_1^*, s_2^*, s_3^* , satisfying $s_i^* \cdot s_j = \delta_{ij}$, and let Ω^* be a basic cell in \mathcal{L}^* . The dual group to \mathcal{L} is the group consisting of all characters of \mathcal{L} , i.e., all homomorphisms from $\mathcal{L} \rightarrow U(1)$. Explicitly, for $k \in \Omega^*$, we have the character χ_k given by

$$\chi_k(t) = e^{ik \cdot t}.$$

The dual group of \mathcal{L} can be identified with the basic cell, Ω^* , of \mathcal{L}^* . Now for each $k \in \Omega^*$, we define the Hilbert space $\mathcal{H}_k := L^2(\Omega)$, and we then define the Hilbert space \mathcal{H} to be the direct fiber integral, denoted as

$$\mathcal{H} = \int_{\Omega^*}^{\oplus} \mathcal{H}_k dk$$

where dk is the standard Lebesgue measure on Ω^* , normalized so that $\int_{\Omega^*} dk = 1$, and defined as $\mathcal{H} = L^2(\Omega^*, dk; L^2(\Omega))$.

Given operators H_k , $k \in \Omega^*$, acting on $\mathcal{H}_k := L^2(\Omega)$, we define the operator $\psi_k(x) \rightarrow H_k \psi_k(x)$ on $\mathcal{H} = \int_{\Omega^*}^{\oplus} \mathcal{H}_k dk$ and denote it $\int_{\Omega^*}^{\oplus} H_k dk$. Define $U : L^2(\mathbb{R}^3) \rightarrow \mathcal{H}$ on smooth functions with compact support by the formula

$$(Uv)_k(x) = \sum_{t \in \mathcal{L}} \chi_k^{-1}(t) T_t v(x).$$

We now have the following Bloch decomposition result:

Proposition 6.31. The operator U extends uniquely to a unitary operator and

$$U H U^{-1} = \int_{\Omega^*}^{\oplus} H_k dk \quad (6.67)$$

where H_k , $k \in \Omega^*$, is the restriction of operator H to \mathcal{H}_k , with domain consisting of vectors $v \in \mathcal{H}_k \cap H^2$ satisfying the boundary conditions $T_t v(x) = \chi_k(t) v(x)$ for the basis elements $t = s_1, s_2, s_3$.

Proof. We begin by showing that U is an isometry on smooth functions with compact support. Using Fubini's theorem we calculate

$$\begin{aligned} \|Uv\|_{\mathcal{H}}^2 &= \int_{\Omega^*} \|(Uv)_k\|_{\mathcal{H}_k}^2 dk = \int_{\Omega^*} \int_{\Omega} \left| \sum_{t \in \mathcal{L}} \chi_k^{-1}(t) T_t v(x) \right|^2 dx dk \\ &= \int_{\Omega} \left(\sum_{t, s \in \mathcal{L}} T_t v(x) \overline{T_s v(x)} \int_{\Omega^*} \chi_k^{-1}(t) \chi_k(s) dk \right) dx \\ &= \int_{\Omega} \sum_{t \in \mathcal{L}} |T_t v(x)|^2 dx = \int_{\mathbb{R}^2} |v(x)|^2 dx. \end{aligned}$$

Hence $\|Uv\|_{\mathcal{H}} = \|v\|_{\mathcal{H}}$ and U extends to an isometry on all of \mathcal{H} . To show that U is in fact a unitary operator we define $U^* : \mathcal{H} \rightarrow \mathcal{H}$ by the formula

$$U^* g(x+t) = \int_{\Omega^*} \chi_k(t) g_k(x) dk,$$

for $t \in \mathcal{L}$ and $x \in \Omega$. Straightforward calculations show that U^* is the adjoint of U and that it too is an isometry, proving that U is a unitary operator.

For (6.67), we need to first show that $(Uv)_k$ is in the domain of H_k . For $(Uv)_k$ we have

$$\begin{aligned} T_t (Uv)_k(x) &= \sum_{s \in \mathcal{L}} \chi_k^{-1}(s) T_t T_s v(x) \\ &= \sum_{s \in \mathcal{L}} \chi_k^{-1}(s) T_{t+s} v(x) \\ &= \chi_k(t) \sum_{s \in \mathcal{L}} \chi_k^{-1}(t+s) T_{t+s} v(x) \\ &= \chi_k(t) (Uv)_k(x). \end{aligned}$$

Hence if $v \in D(H)$, then $Uv \in D(H_k)$. Next, we have that

$$\begin{aligned} (H_k(Uv)_k)(x) &= \sum_{t \in \mathcal{L}} \chi_k^{-1}(t) H T_t v(x) \\ &= \sum_{t \in \mathcal{L}} \chi_k^{-1}(t) T_t H v(x) \\ &= (UHv)_k(x), \end{aligned}$$

which establishes (6.67). \square

Since the resolvents of the operators H and H_k are related as $U(H-z)^{-1}U^{-1} = \int_{\Omega^*}^{\oplus} (H_k - z)^{-1} dk$, we have that

$$\sigma(H) = \bigcup_{k \in \Omega^*} \sigma(H_k). \quad (6.68)$$

By Theorem ??, the spectra of H_k are purely discrete, say, $\{\lambda_n(k)\}$. This shows that the spectrum of H is union of the sets $\{\lambda_n(k) \mid k \in \Omega^*\}$, called the bands. This is a key result in solid state physics.

6.12. Quantum Statistics. In this lecture, we address the issue of the information reduction in quantum mechanics. Namely, we would like to find out how to describe a subsystem of a larger system in terms of its own degrees of freedom. This leads us to the notion of an *open system*, whose states are described by positive, trace class operators on the L^2 state space (*density operators*). This replaces wave functions of quantum mechanics. This topic is closely related to quantum statistical mechanics. The notions of *trace* and *trace class operators*, which are extensively used in this section, are defined in Mathematical Supplement, Section ??.

6.13. Density Matrices. Consider a physical system described by a quantum Hamiltonian H acting on a Hilbert space \mathcal{H} (say, $H = -\frac{\hbar^2}{2m}\Delta + V(x)$ on $L^2(\mathbb{R}^3)$). Let $\{\psi_j\}$ be an orthonormal system (possibly, a basis) in \mathcal{H} and let $\psi = \sum a_j \psi_j$. Given an arbitrary observable A (say, position or a characteristic function of position), its average in the state ψ is given by

$$\langle A \rangle_\psi := \langle \psi, A\psi \rangle = \sum_{m,n} \bar{a}_m a_n \langle \psi_m, A\psi_n \rangle. \quad (6.69)$$

Now suppose that we know only that, for each n , the system is in the state ψ_n with a probability p_n . We thus have much less information than before. Now the average, $\langle A \rangle$, of an observable A is given by the expression

$$\langle A \rangle = \sum_n p_n \langle \psi_n, A\psi_n \rangle. \quad (6.70)$$

This corresponds to the situation when the parameters a_n in (6.69) are independent random variables with zero mean and variance $E(|a_n|^2) = p_n$. Observe that (6.70) can be written as

$$\langle A \rangle = \text{Tr}(A\rho) , \tag{6.71}$$

where $\rho = \sum_n p_n P_{\psi_n}$. Here P_ψ stands for the rank-one orthogonal projection onto the vector ψ , i. e. $P_\psi f = \langle \psi, f \rangle \psi$, or $P_\psi = |\psi\rangle\langle\psi|$ in Dirac's notation (see Sections ?? and ?? for the definition and discussion of projections and trace).

Homework 6.32. Show that $\text{Tr}(AP_\psi) = \langle \psi, A\psi \rangle$.

Note that ρ is a trace class, positive (since $p_n \geq 0$ and $P_{\psi_n} \geq 0$) operator, with trace 1: $\text{Tr}\rho = \sum_n p_n = 1$. We extrapolate from this the assumption that generalized states are given by positive, trace class operators ρ on \mathcal{H} , normalized so that $\text{Tr}\rho = 1$. Such operators are called *density matrices* or *density operators*.

If the vectors ψ_n in the expression $\rho = \sum p_n P_{\psi_n}$ evolve according to the Schrödinger equation, $i\hbar \frac{\partial \psi}{\partial t} = H\psi$, then the equation governing the state ρ is:

$$i \frac{\partial \rho}{\partial t} = \frac{1}{\hbar} [H, \rho]. \tag{6.72}$$

One takes this equation to be the *basic dynamical equation* of quantum statistical mechanics, or quantum statistics. We call it the *Landau-von Neumann equation*, or the *quantum Liouville equation*.

Homework 6.33. Derive equation (6.72) for $\rho = \sum p_n P_{\psi_n}$ with ψ_n as above, i.e. show that the solution of the equation (6.72) with the initial condition $\rho|_{t=0} =: \rho_0$ is given by $\rho = e^{-\frac{iHt}{\hbar}} \rho_0 e^{\frac{iHt}{\hbar}}$.

Homework 6.34. Show that if the initial condition $\rho|_{t=0}$ is a positive, trace class operator, then so is the solution of the equation (6.72).

If ψ_i are bound states of H (i.e. $H\psi_i = \lambda_i\psi_i$), then $\rho = \sum_i p_i P_{\psi_i}$, for any $p_i \geq 0$, $\sum p_i = 1$, is a static solutions of the equation $i \frac{\partial \rho}{\partial t} = \frac{1}{\hbar} [H, \rho]$. (*Show this.*)

An important example of density matrices are rank-one orthogonal projections $\rho = P_\psi$. Indeed, we have

$$P_\psi \geq 0 \quad \text{and} \quad \text{Tr} P_\psi = \|\psi\|^2 = 1$$

Homework 6.35. Show these relations.

Homework 6.36. Show that if ψ_i are bound states of H (i.e. $H\psi_i = \lambda_i\psi_i$), then $\rho = \sum_i p_i P_{\psi_i}$, for any $p_i \geq 0$, $\sum p_i = 1$, is a static solutions of the equation $i \frac{\partial \rho}{\partial t} = \frac{1}{\hbar} [H, \rho]$.

There is one-to-one correspondence between rays of normalized L^2 -functions and rank-one orthogonal projections: $\{e^{i\alpha}\psi\} \rightarrow P_\psi$ and any rank-one orthogonal projection P can be written as P_ψ for any normalized function $\psi \in \text{Ran } P$. Wave

functions, ψ , or rank-one projections, P_ψ , are called *pure states*. Density operators, ρ , such $\rho \neq P_\psi$ for any ψ , are called *mixed states*. Thus $p_1 P_{\psi_1} + p_2 P_{\psi_2}$ is a mixed state.

6.14. Stationary States. *Stationary*, i.e. time-independent, solutions of equation (6.72) are given by various functions, $f(H)$ (defined, say, by the formula (??)), of the quantum Hamiltonian H . Indeed, such operators, if well-defined, commute with H . However, they represent density matrices (up to normalization) only if they are positive and trace-class. The latter holds if and only if

- the functions $f(\lambda)$ are supported on the discrete spectrum of H
- if $\sigma(H)$ extends to infinity, the $f(\lambda)$ decay at infinity sufficiently fast.

As an example consider an operator H whose spectrum consists of isolated eigenvalues (of finite multiplicity) converging to ∞ . (For example, H is a Schrödinger operator with the potential $|x|^4$.) Then an operator $f(H)$ is trace class for any function f vanishing at ∞ sufficiently fast (see Section ??). Thus, for any such positive function, $f(H)$ is a density matrix (up to normalization).

To summarize, if the operator H has purely discrete spectrum, then Eqn (6.72) has an infinite-dimensional space of time-independent solutions – stationary states – which are density operators. These operators are of the form $f(H)$, where f is a positive function, decaying sufficiently fast at infinity.

However, the states above are not seen in nature if the number of particles is very large. What is seen in this case, are the states of thermal equilibrium – the (thermal) *equilibrium states*. The latter can be isolated as follows. Assume we have only one conserved quantity – the energy. Then, following the *second law of thermodynamics*, we can characterize the equilibrium states in a finite volume as states ρ which maximize the *von Neumann entropy*, $S(\rho) := -\text{Tr}(\rho \ln \rho)$, given the *internal energy* $E(\rho) := \text{Tr}(H\rho)$ and the normalization, $\text{Tr} \rho = 1$ (the ‘number of particles’):

$$\rho \text{ maximizes } S(\rho), \text{ provided } E(\rho) \text{ is fixed } (E(\rho) = E, \text{ say}) \text{ and } \text{Tr} \rho = 1. \quad (6.73)$$

The criterion above is called the *principle of maximum entropy*. This principle can be extended in an appropriate form to infinite systems.

In (6.73), ρ varies over the **convex** set, $\mathcal{S} := \{\rho \in \text{the (Banach) space of trace-class operators, } \rho \geq 0\}$. **Assuming that the maximum in (6.73) is achieved in the interior of \mathcal{S}** , one can show, using standard techniques of variational calculus (see the mathematical supplement on variational calculus of [GS]), that maximizer, if it exists, satisfies the Euler-Lagrange equation

$$E'(\rho) - TS'(\rho) - \mu \mathbf{1} = 0, \quad \text{or, explicitly, } H + T(\ln \rho + 1) - \mu = 0,$$

where T and μ are Lagrange multipliers. The latter equation can be easily solved to give the following one-parameter family of positive operators

$$\rho_T = e^{-H/T} / Z(T) \text{ , where } Z(T) := \text{Tr} e^{-H/T} \text{ ,}$$

as equilibrium states. (The parameter μ was used to fix the normalization of ρ_T .) These states are called the *Gibbs states* and T , which is the inverse of the Lagrange multiplier, is called the *temperature*. The quantity $Z(T) = \text{Tr} e^{-H/T}$ (or $Z(\beta) = \text{Tr} e^{-\beta H}$ for $\beta = 1/T$) is called the *partition function* (at temperature T , of the system described by the Hamiltonian H).

It is conjectured that in the absence of conserved quantities other than the energy, all equilibrium states of infinite systems of infinite degrees of freedom can be obtained as (weak) limits of Gibbs states.

The Lagrange multiplier theorem of variational calculus (see the mathematical supplement on variational calculus of [GS]) implies that an equilibrium state minimizes the *Helmholtz free energy*

$$F_T(\rho) := E(\rho) - TS(\rho)$$

where T , or $\beta = T^{-1}$, is the Lagrange multiplier to be found from the relation $\text{Tr}(H\rho_T) = E$.

Homework 6.37. * (will not be on the test) Show this.

By a straightforward computation, the equilibrium free energy, $F(T) := F_T(\rho_T)$, is given by

$$F(T) = -T \ln Z(T) \text{ .}$$

The next result connects Gibbs states and the free energy, to ground states and the ground state energy. Let ψ_0 be the (unique) ground state of the Hamiltonian H , and E_0 the corresponding ground state energy. Let P_ψ denote the rank-one projection onto the vector ψ . We have

Theorem 6.38 (Feynman-Kac Theorem). As $T \rightarrow 0$,

$$\rho_T \rightarrow P_{\psi_0} \quad \text{and} \quad F(T) \rightarrow E_0 \text{ .}$$

Proof. Let $E_0 < E_1 \leq E_2 \leq \dots$ be the eigenvalues of H (our standing assumption is that H has purely discrete spectrum, running off to ∞), and let ψ_0, ψ_1, \dots be corresponding orthonormal eigenstates. Then by completeness of the eigenstates, and the spectral mapping theorem (see Section ??), $\rho_T = \sum_{n=0}^{\infty} p_n P_{\psi_n}$ where $p_n = e^{-E_n/T} / Z(T)$. We can rewrite p_n as

$$p_n = e^{-(E_n - E_0)/T} / \sum_{n'=0}^{\infty} e^{-(E_{n'} - E_0)/T} \text{ .}$$

We see that $p_n \leq 1$ and as $T \rightarrow 0$

$$p_n \rightarrow \begin{cases} 1 & n = 0 \\ 0 & n \geq 1 \end{cases}$$

It follows easily that $\|\rho_T - P_{\psi_0}\| \rightarrow 0$ as $T \rightarrow 0$. Furthermore, since

$$F(T) = -T \ln \left(\sum_{n=0}^{\infty} e^{-E_n/T} \right) = E_0 - T \ln \left(1 + \sum_{n=1}^{\infty} e^{-(E_n - E_0)/T} \right),$$

we see that $F(T) \rightarrow E_0$ as $T \rightarrow 0$. \square

6.15. Quantum statistics: General framework. We formalize the theory above by making the following postulates:

- States: positive trace-class operators on \mathcal{H} (as usual, up to normalization);
- Evolution equation : $i\hbar \frac{\partial \rho}{\partial t} = [H, \rho]$;
- Observables : self-adjoint operators on \mathcal{H} ;
- Averages : $\langle A \rangle_\rho := \text{Tr}(A\rho)$.

We call the theory described above *quantum statistics*. The last two items lead to the following expressions for the probability densities for the coordinates and momenta:

- $\rho(x; x)$ - probability density for coordinate x ;
- $\hat{\rho}(k; k)$ - probability density for momentum p .

Above, $\hat{\rho}(k; k')$ is the integral kernel of the operator $\hat{\rho} := \mathcal{F}\rho\mathcal{F}^{-1}$, i.e.

$$\hat{\rho}(k; k') = (2\pi\hbar)^{-3} \int \int e^{-\frac{ik \cdot x}{\hbar}} e^{\frac{ik' \cdot x'}{\hbar}} \rho(x; x') dx dx'. \quad (6.74)$$

Homework 6.39. * (will not be on the test) Show that the integral kernel of the operator $\hat{\rho} := \mathcal{F}\rho\mathcal{F}^{-1}$ is given by (6.74).

In particular, if $\rho = P_\psi$, then

$$\begin{aligned} \rho(x; x) &= |\psi(x)|^2 \\ \hat{\rho}(k; k) &= |\hat{\psi}(k)|^2 \end{aligned}$$

as should be the case according to our interpretation.

Note that the state space here is not a linear space but a positive cone in a linear space. It can be identified with the space of all positive (normalized) linear functionals $A \rightarrow \omega(A) := \text{Tr}(A\rho)$ on the space of bounded observables. Denote the spaces of bounded observables and of trace class operators on \mathcal{H} as $L^\infty(\mathcal{H})$ and $L^1(\mathcal{H})$, respectively. There is a duality between density matrices and observables

$$\langle \rho, A \rangle = \text{Tr}(A\rho) \quad (6.75)$$

for $A \in L^\infty(\mathcal{H})$ and $\rho \in L^1(\mathcal{H})$.

Now, consider the evolution of density matrices, the von Neumann equation, $i\frac{\partial\rho}{\partial t} = \frac{1}{\hbar}[H, \rho]$, for density matrices and the Heisenberg equation, $i\frac{\partial A}{\partial t} = -\frac{1}{\hbar}[H, A]$, for observables. Denote by α_t and α_t^* corresponding flows, $\alpha_t : \rho \rightarrow \rho_t$ and $\alpha_t^* : A \rightarrow A_t$. We have shown above that they have the following explicit representations

$$\alpha_t(\rho) := e^{-\frac{iHt}{\hbar}} \rho e^{\frac{iHt}{\hbar}} \quad \text{and} \quad \alpha_t^*(A) = e^{\frac{iHt}{\hbar}} A e^{-\frac{iHt}{\hbar}}.$$

In the sense of the duality (6.75), the evolution of density matrices, α_t , and of observables, α_t^* , are dual:

$$\langle \rho, \alpha_t^*(A) \rangle = \langle \alpha_t(\rho), A \rangle.$$

Quantum mechanics is a special case of this theory, and is obtained by restricting the density operators to be rank-one orthogonal projections:

$$\begin{aligned} \{e^{i\alpha}\psi, \forall \alpha \in \mathbb{R}\} &\Leftrightarrow \rho = P_\psi \\ \langle \psi, A\psi \rangle &= \text{Tr}(AP_\psi) \\ i\hbar\frac{\partial}{\partial t}(e^{\frac{i\mu t}{\hbar}}\psi) = H(e^{\frac{i\mu t}{\hbar}}\psi), \text{ for some } \mu \in \mathbb{R}, &\Leftrightarrow i\frac{\partial\rho}{\partial t} = \frac{1}{\hbar}[H, \rho] \text{ for } \rho = P_\psi. \end{aligned}$$

Homework 6.40. Show these properties.

We show the \Leftarrow direction in the last statement. Using Dirac's notation we have that $i\frac{\partial\rho}{\partial t} = \frac{1}{\hbar}[H, \rho]$ and $\rho = P_\psi$ imply $|\chi\rangle\langle\psi| - |\psi\rangle\langle\chi| = 0$ where $\chi = i\hbar\frac{\partial\psi}{\partial t} - H\psi$, which yields that $\chi = \mu\psi$ for some real μ . This implies that the family of vectors $\tilde{\psi} := e^{\frac{i\mu t}{\hbar}}\psi$ satisfies $i\hbar\frac{\partial\tilde{\psi}}{\partial t} = H\tilde{\psi}$. \square

Recall that the states (density operators) which are rank-one operators – i.e. which are of the form $\rho = P_\psi$ for some normalized wave function ψ – are called *pure states*. They are equivalent to wave functions in the sense that they are in one-to-one correspondence with wave functions, up to a phase, and produce exactly the same expectations for arbitrary observables as the corresponding wave functions.

We have shown that quantum mechanics is a special case of quantum statistics (with $\rho \rightarrow P_\psi$). Now we show that another special case of quantum statistics is probability. Recall that the average of A in a state ρ is $\langle A \rangle_\rho = \text{Tr}(A\rho)$. We introduce the following interpretation:

- A *quantum random variable* is an observable A .
- A *quantum event* is an orthogonal projection operator P (\leftrightarrow subspace $\text{Ran } P$).
- The probability of event P in state ρ is $\text{Prob}_\rho(P) = \text{Tr}(P\rho)$.

Consider observables and orthogonal projections which are multiplication operators by measurable functions. For a projection P this means that it is the multiplication operator by a characteristic function, χ_Q , of a measurable set $Q \subseteq \mathbb{R}^3$ (i.e.

$\chi_Q(x) = \begin{cases} 1 & x \in Q \\ 0 & x \notin Q \end{cases}$. For an observable, A , which is a multiplication operator by measurable function $\xi : \mathbb{R}^3 \rightarrow \mathbb{R}$, the average is

$$\langle A \rangle_\rho = \text{Tr}(A\rho) = \int_{\mathbb{R}^3} \xi(x)\rho(x, x) dx,$$

while the probability of the event $P = \chi_Q$ is

$$\text{Prob}_\rho(P) = \text{Tr}(P\rho) = \int_{\mathbb{R}^3} \chi_Q(x)\rho(x, x) dx. \quad (6.76)$$

Thus, if we restrict ourselves to observables and projections, both of which are multiplication operators (and form a commutative subalgebra of the algebra of all observables), we obtain a standard probabilistic theory:

- The probability space $(\mathbb{R}^3, \mathbb{P})$ where $d\mathbb{P}(x) = \rho(x, x)dx$;
- Random variables which are measurable functions $\xi : \mathbb{R}^3 \rightarrow \mathbb{R}$;
- Events which are measurable subsets $Q \subseteq \mathbb{R}^3 \leftrightarrow$ characteristic functions, χ_Q .

Note that the reduced density operator $\rho = \text{Tr}_{\text{envir}}R$ (see (6.83) above) generalizes the notion of marginal distribution.

6.16. Information Reduction. Assume we are interested in measuring only properties of a subsystem, S , of a given system, e.g. averages for its various observables. Let x and y be the coordinates of the subsystem of interest, and of the rest of the total system, which we call an environment, E . Assume the total system, $S+E$, is described by a wave function, say $\psi(x, y)$. The question we would like to **address** is: is there a subsystem wave function $\varphi(x)$, such that measuring the average of any observable $A = A_x$ associated with the subsystem, in $\varphi(x)$, gives the same result as measuring it in $\psi(x, y)$; i.e. is, for any $A = A_x$,

$$\langle \psi, A\psi \rangle = \langle \varphi, A\varphi \rangle ?$$

Here the inner products on the l.h.s. and r.h.s. are in the spaces $L^2(dx, dy)$ and $L^2(dx)$. The answer is that this holds if and only if the subsystem and environment are not correlated: $\psi(x, y) = \varphi(x)\eta(y)$ for some η . If we take, for example, $\psi(x, y) = \alpha_1\varphi_1(x)\eta_1(y) + \alpha_2\varphi_2(x)\eta_2(y)$, then $\langle \psi, A\psi \rangle \neq \langle \varphi, A\varphi \rangle$ for any $\varphi(x)$.

So what does it take to describe a state of the subsystem in this case without referring to the environment? We postpone answering this question and consider another physical situation.

The answer is to do this in terms of the density operators. Namely, with any total wave function, $\psi(x, y)$, we associate the density operator, $\rho = \rho_\psi$, for our subsystem, so that

$$\langle \psi, A\psi \rangle = \text{Tr}_{\text{syst}}(A\rho) \quad (6.77)$$

for any subsystem observable A . Here Tr_{sys} is trace of the subsystem degrees of freedom. Indeed, the operator ρ is defined by its integral kernel

$$\rho(x, x') = \int \psi(x, y) \overline{\psi(x', y)} dy. \tag{6.78}$$

Homework 6.41. Check that (6.77) holds for any operator A acting on the variable x , provided ρ is given by the integral kernel (6.78).

Now, as was mentioned above, the wave function, ψ , of the total system can be associated with a density operator acting on the total system coordinates, namely the rank-one projection $R := P_\psi$, so that, as an exercise above shows,

$$\langle \psi, A\psi \rangle = \text{Tr}(AP_\psi) \tag{6.79}$$

for any operator (observable) A . With this, if A is a system observable, i.e., an operator which acts only on the variables x , then (6.77) can be rewritten as

$$\text{Tr}(AP_\psi) = \text{Tr}_{\text{sys}}(A\rho). \tag{6.80}$$

Now, assume the total system is described by a density operator R and assume that again we do observations only on the system S . Is there a density operator for the system which gives the results of these observations? The answer to this question is positive: the state of S described by the reduced density matrix $\rho = \text{Tr}_{\text{envir}}R$, obtained from R by tracing out the variables of the environment, gives the same observational values of systems observables as R .

To define the partial trace $\text{Tr}_{\text{envir}}R$, we specify state Hilbert spaces $\mathcal{H}_s = L^2(dx)$ and $\mathcal{H}_e = L^2(dy)$ of the system and environment. Here x and y the variables of the system and environment, respectively. Then the state Hilbert space of the total system (small system plus environment) is $\mathcal{H}_{\text{total}} = L^2(dxdy)$. Then the partial trace, Tr_{envir} , of R over the environment variables is defined by

$$\langle \phi, (\text{Tr}_{\text{envir}}R)\psi \rangle = \sum_i \langle \phi\chi_i, R\psi\chi_i \rangle, \tag{6.81}$$

for any $\phi, \psi \in L^2(dx)$ and for any orthonormal basis $\{\chi_j\}$ in $L^2(dy)$. Here $(\psi\chi)(x, y) := \psi(x)\chi(y)$.

Homework 6.42. * (will not be on the test) Show that the r.h.s. of (6.81) is independent of the choice of the orthonormal basis $\{\chi_j\}$.

One can also define the partial trace, $\rho = \text{Tr}_{\text{envir}}R$, in terms of the integral kernels $R(x, y, x', y')$ and $\rho(x, x')$ of R and ρ as

$$\rho(x, x') := \int R(x, y, x', y) dy. \tag{6.82}$$

This is nothing but a generalization of (6.78). Thus, the operator ρ in (6.77) is given by $\rho = \text{Tr}_{\text{envir}}P_\psi$.

Homework 6.43. * (will not be on the test) Check that the definitions of partial trace given by (6.82) and (6.81) are equivalent.

The definition above implies that for any system observable A we have

$$\mathrm{Tr}(AR) = \mathrm{Tr}_{\mathrm{sys}}(A\rho), \quad \rho = \mathrm{Tr}_{\mathrm{envir}}R. \quad (6.83)$$

To prove this we write the inner product in $L^2(dxdy)$ as $\langle \cdot \rangle = \langle \langle \cdot \rangle_{\mathrm{en}} \rangle_{\mathrm{sys}}$, i.e. first as the inner product in $L^2(dx)$ and then in $L^2(dy)$. Let $\{\phi_j\}$ and $\{\chi_j\}$ be orthonormal bases in $L^2(dx)$ and $L^2(dy)$, respectively. Then $\{\phi_i\chi_j\}$ defines an orthonormal basis in $L^2(dxdy)$ and we have

$$\begin{aligned} \mathrm{Tr}(AR) &= \sum_{ij} \langle \phi_i\chi_j, AR\phi_i\chi_j \rangle = \sum_{ij} \langle (A^*\phi_i)\chi_j, R\phi_i\chi_j \rangle \\ &= \sum_i \langle A^*\phi_i, \sum_j \langle \chi_j, R\chi_j \rangle_{\mathrm{en}} \phi_i \rangle_{\mathrm{sys}} \\ &= \sum_i \langle A^*\phi_i, (\mathrm{Tr}_{\mathrm{envir}}R)\phi_i \rangle_{\mathrm{sys}} \\ &= \sum_i \langle \phi_i, A\rho\phi_i \rangle_{\mathrm{sys}} =: \mathrm{Tr}_{\mathrm{sys}}(A\rho). \end{aligned}$$

The operator $\rho = \mathrm{Tr}_{\mathrm{envir}}R$, acting on the system state space, is called the *reduced density operator*. It has the following properties:

- 1) ρ acts on the state space of the system;
- 2) ρ is positive if R is positive;
- 3) $\mathrm{Tr}_{\mathrm{sys}}\rho = \mathrm{Tr}R$.

Homework 6.44. Show these properties.

Thus ρ is a density operator of the system. It is analogous to the conditional expectation in the probability theory.

Now, we demonstrate another way to show (6.77). If A is a system observable, i.e., an operator which acts only on the variables x , then taking the trace first with respect to the environment variables, and then over the system variables (see Mathematical Supplement, Section ??) one obtains

$$\begin{aligned} \mathrm{Tr}(AP_\psi) &= \mathrm{Tr}_{\mathrm{sys}}\mathrm{Tr}_{\mathrm{envir}}(AP_\psi) \\ &= \mathrm{Tr}_{\mathrm{sys}}[A\mathrm{Tr}_{\mathrm{envir}}(P_\psi)]. \end{aligned}$$

With the notation $\rho := \mathrm{Tr}_{\mathrm{envir}}(P_\psi)$, this gives (6.77).

Remark 6.45. In tensor product notation we write $\phi \otimes \chi$, and $A \otimes I$ for $\psi(x)\chi(y)$ and an observable A acting only on x , respectively.

We summarize our conclusions. In the situation when we are interested only in a subsystem of a given system, or when only partial information about a quantum system is available – namely, we know only that the system occupies certain states with certain probabilities – we can describe states of such a system by positive trace-class operators $\rho \geq 0$ (normalized by $\mathrm{Tr}\rho = 1$), called density operators or

density matrices, with the equation of motion given by (6.72), and averages of observables computed according to the prescription (6.71).

The totality of quantum systems to which coupling of a given system cannot be neglected is called the *environment*, while a given system, whose interaction with the environment cannot be neglected but which is described in terms of its own degrees of freedom, is called an *open system*.

In fact, every system, unless it is the entire universe, can be considered as a subsystem of a larger system. Hence, in reality no quantum system is isolated. It is coupled to nearby quantum systems and so is an open system.

6.17. Reduced dynamics. Consider a system interacting with an environment. Such a system was called open system. We found in Section 6.16, that if the total system is described by a wave function, say $\psi(x, y)$, where x and y are the coordinates of the system of interest and of the environment, respectively, or more generally by a density operator R , acting on the total system state space $L^2(dxdy)$, then the open system is described by the density operator $\rho = \text{Tr}_{\text{envir}} R$, so that for any observable $A = A_x$, associated with the system

$$\text{Tr}(AR) = \text{Tr}_{\text{syst}}(A\rho), \quad \rho = \text{Tr}_{\text{envir}} R, \quad (6.84)$$

where, recall, Tr_{envir} is the partial trace of R over the environment variables defined in (6.81). Recall that the operator $\rho = \text{Tr}_{\text{envir}} R$, acting on the system state space, is called the reduced density operator and that it has the following properties:

- 1) ρ acts on the state space of the system;
- 2) ρ is positive if R is positive;
- 3) $\text{Tr}_{\text{syst}} \rho = \text{Tr} R$.

Assume now our total system, consisting of the given system and environment, are described by the Schrödinger operators, H_{syst} and H_{envir} , acting on the corresponding state spaces $L^2(dx)$ and $L^2(dy)$, respectively. The Schrödinger operator of the total system is given, in general, by

$$H_{\text{tot}} = H_{\text{syst}} \otimes \mathbf{1}_{\text{envir}} + \mathbf{1}_{\text{syst}} \otimes H_{\text{envir}} + \lambda v, \quad (6.85)$$

where λ is a parameter, called the coupling constant and v is an operator acting on $L^2(dxdy) = L^2(dx) \otimes L^2(dy)$; λv describes the interaction of the system and the environment. The total system evolves according to the Landau-von Neumann equation

$$i \frac{\partial R_t}{\partial t} = \frac{1}{\hbar} [H_{\text{tot}}, R_t], \quad R_{t=0} = R_0, \quad (6.86)$$

The solution of this equation, with an initial condition R_0 , is given by $R_t = \alpha_t(R_0)$, where $\alpha_t(R) = e^{-\frac{iH_{\text{tot}}t}{\hbar}} R e^{\frac{iH_{\text{tot}}t}{\hbar}}$.

Now, the reduced density operator of the system at time t is given by $\rho_t := \text{Tr}_{\text{envir}} R_t$. The map $\rho_0 := \text{Tr}_{\text{envir}} R_0 \rightarrow \rho_t$ depends not only on ρ_0 and α_t but also on R_0 . In general, it is not even a linear map for fixed t . To remedy this we

consider the particular class of R_0 of the form $R_0 = \rho_0 \otimes \rho_{e0}$ for some fixed ρ_{e0} and define

$$\beta_t(\rho_0) = \text{Tr}_{\text{envir}} \alpha_t(\rho_0 \otimes \rho_{e0}). \quad (6.87)$$

The family β_t is called the *reduced evolution*. In this case, β_t is linear. It clearly depends on ρ_{e0} . What can we say about this evolution? We begin with

Theorem 6.46. The family β_t has the following properties

- 1) β_t is linear;
- 2) β_t is positive (or, positivity preserving, i.e. $\rho \geq 0 \implies \beta_t(\rho) \geq 0$);
- 3) β_t preserves the trace;
- 4) $\beta_t(\rho^*) = \beta_t(\rho)^*$.

Proof. We show only property 2), leaving the other properties as the exercise. This property follows immediately as β_t is a composition of two positivity preserving maps, Tr_{envir} and α_t . \square

Homework 6.47. Show 1) - 4).

We give a lengthy but instructive direct proof of property 2) which we will use below. In this proof we drop the subindex t , let $U := e^{-iH_{\text{tot}}t/\hbar}$ and let $\{\chi_i\}$ be orthonormal basis in the environment space $L^2(dy)$. By the definition of β and the property $\alpha(R_0) = UR_0U^*$, we have $\forall \phi, \psi \in L^2(dx)$

$$\langle \phi, \beta(\rho_0)\psi \rangle = \sum_i \langle \phi \chi_i, \alpha(\rho_0 \otimes \rho_{e0})\psi \chi_i \rangle = \sum_i \langle U^* \phi \chi_i, \rho_0 \otimes \rho_{e0} U^* \psi \chi_i \rangle. \quad (6.88)$$

Taking here $\phi = \psi$, we see the r.h.s is non-negative, provided ρ_0 is non-negative. This gives 2).

Theorem 6.48. The family β_t is of the form

$$\beta_t(\rho) = \sum_n V_{nt} \rho V_{nt}^* \quad (6.89)$$

where V_{nt} are bounded operators satisfying $\sum_n V_{nt}^* V_{nt} = \mathbf{1}$ (strong convexity).

Proof. To prove (6.89), we write the inner product in $L^2(dxdy)$ as $\langle \cdot \rangle = \langle \langle \cdot \rangle_{\text{syst}} \rangle_{\text{en}}$, i.e. first as the inner product in $L^2(dx)$ and then in $L^2(dy)$. Let χ_i be an orthonormal basis of eigenfunctions of ρ_{e0} with eigenvalues λ_j . Then $\rho_{e0} = \sum \lambda_j P_{\chi_j} = \sum \lambda_j |\chi_j\rangle \langle \chi_j|$, so that, using (6.88), we obtain

$$\begin{aligned} \langle \phi, \beta(\rho_0)\psi \rangle &= \sum_{i,j} \langle \sqrt{\lambda_j} \langle \chi_j, U^* \phi \chi_i \rangle_{\text{syst}}, \rho_0 \sqrt{\lambda_j} \langle \chi_j, U^* \psi \chi_i \rangle_{\text{syst}} \rangle_{\text{en}} \\ &= \sum_{i,j} \langle V_{ij}^* \phi, \rho_0 V_{ij}^* \psi \rangle_{\text{en}}, \end{aligned} \quad (6.90)$$

where $V_{ij}^* \phi := \sqrt{\lambda_j} \langle \chi_j, U^* \phi \chi_i \rangle_{\text{syst}}$, and therefore $\langle \phi, \beta(\rho_0) \psi \rangle = \langle \phi, \sum_{i,j} V_{ij} \rho_0 V_{ij}^* \psi \rangle$.
Now

$$\begin{aligned} \langle V_{ij} \phi, \psi \rangle_{\text{syst}} &= \langle \phi, V_{ij}^* \psi \rangle_{\text{syst}} = \sqrt{\lambda_j} \langle \phi, \langle \chi_j, U^* \psi \chi_i \rangle_{\text{en}} \rangle_{\text{syst}} \\ &= \sqrt{\lambda_j} \langle U \phi \chi_j, \psi \chi_i \rangle = \langle \sqrt{\lambda_j} \langle U \phi \chi_j, \chi_i \rangle_{\text{en}}, \psi \rangle_{\text{syst}}. \end{aligned}$$

This implies

$$V_{ij} \phi = \sqrt{\lambda_j} \langle \chi_i, U \phi \chi_j \rangle_{\text{en}},$$

which, in turn, gives

$$\begin{aligned} \sum_{i,j} V_{ij}^* V_{ij} \phi &= \sum_{i,j} V_{ij}^* \sqrt{\lambda_j} \langle \chi_i, U \phi \chi_j \rangle_{\text{en}} \\ &= \sum_{ij} \lambda_j \langle \chi_j, U^* \langle \chi_i, U \phi \chi_j \rangle_{\text{en}} \chi_i \rangle_{\text{en}} \\ &= \sum_j \lambda_j \langle \chi_j, U^* \sum_i \langle \chi_i, U \phi \chi_j \rangle_{\text{en}} \chi_i \rangle_{\text{en}}. \end{aligned}$$

Since $\sum_i \langle \chi_i, U \phi \chi_j \rangle_{\text{en}} \chi_i = U \phi \chi_j$, this gives

$$\sum_{i,j} V_{ij}^* V_{ij} \phi = \sum_j \lambda_j \langle \chi_j, U^* U \phi \chi_j \rangle_{\text{en}} = \sum_j \lambda_j \langle \chi_j, \phi \chi_j \rangle_{\text{en}}. \quad (6.91)$$

Since $\langle \chi_j, \phi \chi_j \rangle_{\text{en}} = \phi$ and $\sum \lambda_j = \text{Tr } \rho_{e0} = 1$, we have $\sum_{i,j} V_{ij}^* V_{ij} \phi = \phi$. \square

Homework 6.49. Show that (6.89) implies Theorem 6.46.

Definition 6.50. 1) Maps satisfying the conclusion of Theorem 6.48 are called *quantum (dynamical) maps* or *quantum channels*.

2) complete positivity (**define**)

3) We will call an evolution β_t satisfying the conclusions of Theorem 6.48 (*quantum*) *dissipative* evolution.

4) A reduced evolution β_t is called Markov iff

$$\beta_t \circ \beta_s = \beta_{t+s} \quad \forall t, s \geq 0. \quad (6.92)$$

For a Markov dissipative evolution β_t we define the generator by

$$\mathcal{K}(\rho) := \partial_t \beta_t(\rho)|_{t=0}, \quad (6.93)$$

so that

$$\partial_t \beta_t(\rho) = \mathcal{K}(\beta_t(\rho)). \quad (6.94)$$

Theorem 6.51. Under certain technical continuity conditions on β_t , generators of Markov dissipative evolutions are of the form

$$\mathcal{K}(\rho) = -\frac{i}{\hbar}[H, \rho] + \sum_{j=0}^{\infty} (W_j \rho W_j^* - \frac{1}{2}\{W_j^* W_j, \rho\}) \quad (6.95)$$

where H is self-adjoint, $\{A, B\} := AB + BA$ and $\sum W_j^* W_j$ converges strongly.

Recall coupling between density operators and observables:

$$\langle \rho, A \rangle = \text{Tr}_{\text{sys}}(A\rho). \quad (6.96)$$

Define the reduced evolution of observables by

$$\langle \rho, \beta_t^*(A) \rangle = \langle \beta_t(\rho), A \rangle. \quad (6.97)$$

Then, if a reduced evolution β_t satisfies the conclusions of Theorem 6.48, then its dual, β_t^* , satisfies

- 1) β_t^* is linear;
- 2) β_t^* is positivity preserving ($A \geq 0 \implies \beta_t^*(A) \geq 0$);
- 3) $\beta_t^*(\mathbf{1}) = \mathbf{1}$;
- 4) $\beta_t^*(A^*) = \beta_t^*(A)^*$
- 5) $\beta_t^*(A) = \sum_n V_{nt}^* A V_{nt}$,

where V_{nt} are bounded operators satisfying $\sum_n V_{nt}^* V_{nt} = I$ (strong convexity).

If a dissipative evolution β_t is Markov, then so is its dual: β_t^* is Markov. The latter implies that it satisfies the differential equation $\frac{\partial}{\partial t} \beta_t^* = \mathcal{L} \beta_t^*$ where $\mathcal{L} = K^*$ is the generator of Markov dissipative evolution of observables and is given by

$$\mathcal{L}(A) = \frac{i}{\hbar}[H, A] + \sum_i (W_i^* A W_i - \frac{1}{2}\{W_i^* W_i, A\}). \quad (6.98)$$

For the relation of quantum dynamical maps to completely positive maps see [T4], Statement 3.1.4 and [BR], Remarks to Section 5.3.1.

6.18. Decoherence and Thermolization. Let $\{\varphi_j\}_{j \geq 1}$ be an orthonormal basis of \mathcal{H}_{sys} diagonalizing H_{sys} . The system is said to exhibit (full) *decoherence (in the energy basis)* if the off-diagonal matrix elements, $\langle \varphi_m, \beta_t(\rho_0) \varphi_n \rangle$, of the reduced density matrix, $\beta_t(\rho_0) = \text{Tr}_{\text{envir}} \alpha_t(\rho_0 \otimes \rho_{e0})$, vanish in the limit of large times,

$$\lim_{t \rightarrow \infty} \langle \varphi_m, \beta_t(\rho_0) \varphi_n \rangle = 0, \quad (6.99)$$

whenever $m \neq n$. Denote $\bar{\rho}_t := \beta_t(\rho_0)$ and $[\bar{\rho}_t]_{m,n} := \langle \varphi_m, \bar{\rho}_t \varphi_n \rangle$. We claim then that the decoherence implies that after some time

$$\text{Tr}(A \bar{\rho}_t) \sim \sum_m p_m \langle \varphi_m, A \varphi_m \rangle, \quad (6.100)$$

where $p_m := [\bar{\rho}_t]_{m,m}$, which is a statistical sum we started with. Indeed, inserting the completeness relation, $\sum_m |\varphi_m\rangle\langle\varphi_m|$, for the basis $\{\varphi_m\}$ on both sides of A , we find

$$\mathrm{Tr}(A\bar{\rho}_t) = \sum_{m,n} [\bar{\rho}_t]_{m,n} \langle\varphi_n, A\varphi_m\rangle. \quad (6.101)$$

Eqs (6.99) and (6.101) imply the desired relation (6.100).

To elucidate the role of the environment in the decoherence, consider a total system whose evolution is given by $\Psi = \Psi_t$. Expanding this in the basis $\{\varphi_m\}$ in $L^2(dx)$, we find $\Psi = \sum_m \varphi_m \otimes \alpha_m$, for some $\alpha_i \in L^2(dy)$. Computing the reduced density operator for this state, we find $\bar{\rho}_t = \mathrm{Tr}_{\mathrm{envir}} P_\Psi = \sum_{m,n} [\bar{\rho}_t]_{m,n} |\varphi_m\rangle\langle\varphi_n|$, where $[\bar{\rho}_t]_{m,n} := \langle\varphi_m, \bar{\rho}_t \varphi_n\rangle = \langle\alpha_m, \alpha_n\rangle_{\mathrm{envir}}$. Hence the decoherence means that the environmental dynamics forces $\langle\alpha_m, \alpha_n\rangle_{\mathrm{envir}} \rightarrow 0$, whenever $m \neq n$.

According to Subsection 6.14, the equilibrium states of the total system and environment are given by $R_T := e^{-H_{\mathrm{total}}/T}/Z(T)$, $Z(T) := \mathrm{Tr} e^{-H_{\mathrm{total}}/T}$, and $\rho_{\mathrm{envir},T} := e^{-H_{\mathrm{envir}}/T}/Z_{\mathrm{envir}}(T)$, $Z_{\mathrm{envir}}(T) := \mathrm{Tr} e^{-H_{\mathrm{envir}}/T}$ (these are the Gibbs state of the total system and the environment at the temperature T). We say that the total system $\mathrm{sys} + \mathrm{envir}$ has the property of *return to equilibrium* iff

$$\lim_{t \rightarrow \infty} \mathrm{Tr}_{\mathrm{sys}+\mathrm{envir}}(\alpha_t^*(A)(\rho_0 \otimes \rho_{\mathrm{envir},T})) = \mathrm{Tr}_{\mathrm{sys}+\mathrm{envir}}(AR_T),$$

for all observables (of the joint system) A and for all initial density matrices ρ_0 on $\mathcal{H}_{\mathrm{sys}}$. If this happens then the large time limit of the reduced density matrix of such a system is given by

$$\bar{\rho}_\infty := \lim_{t \rightarrow \infty} \bar{\rho}_t = \mathrm{Tr}_{\mathrm{envir}}(R_T) = \rho_{\mathrm{sys},T} + O(\lambda),$$

since, by perturbation theory of equilibrium states, $R_T = \rho_{\mathrm{sys},T} \otimes \rho_{\mathrm{envir},T} + O(\lambda)$. The leading term of $\bar{\rho}_\infty$ (for small coupling) is just the Gibbs state of the system. In this sense, the system undergoes the process of *thermalization*.

6.19. Irreversibility. A quantum dynamical map β is said to be *irreversible* iff β is not invertible (within the class of dynamical maps). If $\beta(\rho) = U\rho U^*$, where U is unitary, then β is reversible.

How to quantify the notion of irreversibility? In classical mechanics we encounter irreversibility when we pass from Newton's equation to the Boltzmann equation. While Newton's equation is reversible ($(x(t), p(t)) \rightarrow (x(-t), -p(-t))$ is a symmetry of Newton's equation), the Boltzmann equation is irreversible: the Boltzmann entropy

$$H(f) = - \int f \log f, \quad (6.102)$$

for the particle densities, $f(x, v, t)$, which solve the Boltzmann equation, increases along the evolution (the celebrated Boltzmann H-theorem).

In quantum mechanics, we saw above that while the Schrödinger equation is reversible, the reduced evolution, say, $\rho_t = \beta_t(\rho_0)$ (or $\frac{\partial}{\partial t} \rho_t = \mathcal{K} \rho_t$, in the Markov

case) it leads to when some information is 'integrated out', is irreversible. Hence we look for an analogue of the Boltzmann entropy, and it is natural to define

$$S(\rho) = -\text{Tr}(\rho \log \rho), \quad (6.103)$$

which is nothing but the von Neumann entropy. (The operator function $\log \rho$ is defined for any $\rho \geq 0$ as the operator A s.t. $e^A = \rho$. Alternatively, if λ_j and ϕ_j are eigenfunctions and eigenvalues of ρ so that $\rho = \sum \lambda_j P_{\phi_j}$, then $S(\rho) = -\sum \lambda_j \log \lambda_j = H(\{\lambda_j\})$. Remember $0 \leq \rho \leq 1$ and therefore $0 \leq \lambda_j \leq 1$ and $S(\rho) \geq 0$.)

We list properties of $S(\rho)$:

- 1) $\rho = P_\psi$ is a pure state $\implies S(\rho) = 0$;
- 2) $S(U\rho U^*) = S(\rho)$, for unitaries U ;
- 3) $S(\sum \nu_j \rho_j) \geq \sum \nu_j S(\rho_j)$, for $\nu_j \geq 0$, $\sum \nu_j = 1$;
- 4) For any density operator, ρ_{AB} , of a composed system $A + B$,

$$S(\rho_{AB}) \leq S(\rho_A) + S(\rho_B) \quad (6.104)$$

where ρ_A and ρ_B are the marginals of ρ_{AB} :

$$\text{Tr}_B \rho_{AB} = \rho_A \text{ and } \text{Tr}_A \rho_{AB} = \rho_B.$$

(The relation 3) holds due to concavity of \log .) However, there is no H -theorem for $S(\rho)$, i.e. in general $S(\rho)$ does not decrease (or increase) under the evolution. We look for a more general object which has monotonicity properties. Such a candidate is the relative entropy:

$$S(\rho_1, \rho_2) = \text{Tr}(\rho_1(\log \rho_1 - \log \rho_2)), \quad (6.105)$$

if $\overline{\text{Ran } \rho_1} = \overline{\text{Ran } \rho_2}$ and ∞ otherwise. We have the following result whose proof can be found in [T4], Statement 3.1.12:

Theorem 6.52 (Generalized H -theorem (Lindblad)). If β is a dynamical map, then

$$S(\beta(\rho_1), \beta(\rho_2)) \leq S(\rho_1, \rho_2). \quad (6.106)$$

Note: if $\beta(\rho) = U\rho U^*$, where U is unitary, then

$$S(\beta(\rho_1), \beta(\rho_2)) = S(\rho_1, \rho_2). \quad (6.107)$$

6.20. Information and communication channels. Recall that a quantum (information) channel is a quantum dynamical map β of density operators (i.e. β is linear, completely positive and trace preserving). It is one of the key notions in quantum information theory. Quantum dynamical maps, or quantum channels, generalize the notion of classical channel. In this section we describe classical channels and their relation to quantum ones.

Classical information is represented by a probability distribution $p(x)$ ($p(x) \geq 0$ and $\int dx p(x) = 1$ or, in the discrete case, $\sum_x p(x) = 1$), or a random variable

X (with a probability distribution $p_X(x)$). (A classical message is a string of letters x chosen from an alphabet of k letters $\{a_1, \dots, a_k\}$. Letters in this message are statistically independent and each a_x occurs with an a priori probability $p(a_x)$. Binary alphabet is an alphabet with two characters.) Input/output are probability distributions p/q or random variables X/Y .

Classical information is measured in bits 0 or 1, i.e. how many bits are in a message. (There are other measures too: "nats" or "hartleys").

Classical (information) channel is a stochastic map β_{cl} of probability distributions (i.e. a map β_{cl} from probability distributions into probability distributions, $\beta_{cl} : p \rightarrow q$, which is

- 1) β_{cl} is linear;
- 2) β_{cl} is positivity preserving ($p \geq 0 \implies \beta_{cl}(p) \geq 0$);
- 3) β_{cl} preserves the total probability, $\beta_{cl}^*1 = 1$;

Let $p(y, x)$ be the integral kernel (matrix) of β_{cl} ($\beta_{cl}(p)(x) = \sum_y p(x, y)p(y)$). Then $p(x, y)$ is stochastic matrix, i.e. it satisfies

$$\sum_x p(x, y) = 1, \quad p(x, y) \geq 0. \tag{6.108}$$

If β_{cl} is a map between random variables, $\beta_{cl} : X \rightarrow Y$, then we have

$$p_X(x) = P(X = x|Y) = \sum_y P(X = x, Y = y)p_Y(y). \tag{6.109}$$

and therefore in this case, $p(x, y) := P(X = x, Y = y)$.

Recall that quantum information is represented by a density operator ρ ($\rho \geq 0$ and $\text{Tr } \rho = 1$). It is measured in qubits (quantum bits), which is a quantum 2 dimensional system (i.e. a system on \mathbb{C}^2 with basis states denoted by level states

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \text{ and } \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Realization of a qubit:

- 1) The lowest two states of an atom (with the rest of the spectrum considered as an environment.)
- 2) A particle in a double well potential.

While the classical bit takes only two values, 0 or 1, the qubit can be in any superposition of the basis states $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 1 \end{pmatrix} : \alpha \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \beta \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ with $|\alpha|^2 + |\beta|^2 = 1$.

Recall that a quantum (information) channel is a quantum dynamical map β of density operators (i.e. β is linear, completely positive and trace preserving). It is one of the key notions in quantum information theory.

Quantum dynamical maps, or quantum channels, generalize the notion of classical channel. If a quantum channel β satisfies $[\beta(\rho), \rho] = 0$ for any ρ , then it is in fact a classical one. Indeed, since β and $\beta(\rho)$ commute, they have a common orthonormal basis $\{\psi_i\}$ of eigenfunctions. So let $\rho = \sum_i p_i P_{\psi_i}$ and

$\rho' := \beta(\rho) = \sum_i p'_i P_{\psi_i}$, where $p := \{p_i\}$ and $p' := \{p'_i\}$ are the corresponding eigenvalues of the density operators ρ and ρ' , respectively. Then p and p' can be thought of probability distributions, and β , as the mapping $p \rightarrow p'$, i.e. as a classical channel. ($\beta_{\text{cl}} := P\beta P^{-1}$, where $P : \rho \rightarrow p$.)

Measurement: The measurement is defined as a map $\rho \rightarrow p(x)$, given by $p(x) := \text{Tr}(M_x \rho)$, where $\{M_x\}$ is a given collection of positive operators satisfying $\sum_x M_x = 1$, i.e. a positive operator resolution of identity. (Such a collection is called *positive operator-valued measure or POVM*.)

Preparation: The preparation is defined as a map $p \rightarrow \rho$, given by $\rho := \sum_x p(x) \rho_x$, where ρ_x is a given collection (basis), $\{\rho_x\}$, of density matrices.

In the first case we assume that the transmitter possesses a POVM $\{M_x\}$, and in the second case, a collection (basis), $\{\rho_x\}$, of density matrices. Thus we can define the following refinements:

quantum-quantum channel $\beta_{\text{qq}} : \rho \rightarrow \rho'$

classical-quantum $\beta_{\text{clq}} : p \rightarrow \rho = \sum p_x \rho_x \rightarrow \rho'$.

quantum-classical channel $\beta_{\text{qcl}} : \rho \rightarrow \rho' \rightarrow p_x = \text{Tr}(M_x \rho')$.

classical-classical channel $\beta_{\text{ccl}} : p \rightarrow \rho = \sum p_x \rho_x \rightarrow \rho' \rightarrow p_x = \text{Tr}(M_x \rho')$.

The last case is the one one expects in quantum computation. One starts with a classical information, converts it into quantum one, which is manipulated by a quantum computer or quantum channel and at the end one converts the quantum output into a classical one. In this case, the transmitter poses collection of 'basis' density matrices $\{\rho_x\}$ ($\rho_x \geq 0$, $\text{Tr} \rho_x = 1$) and receiver poses a POVM.

Problem: Quasiclassical limit of quantum channels (coherent states?)

6.21. Measuring classical and quantum information. Let a source be modeled as a probability distribution $\{p_x\}$ (random variable X with probability distribution, $p_x = P(X = x)$). To measure classical information contained in a probability distribution $p(x)$, Shannon defined the quantity (entropy)

$$H(p) = - \sum_x p(x) \log p(x) \quad \text{or} \quad H(p) = - \int dx p(x) \log p(x), \quad (6.110)$$

which he called the *Shannon's information entropy* or just entropy. It is analogous to the Boltzmann entropy of thermodynamics.

The quantum analogue of Shannon's information is the von Neumann's entropy $S(\rho) = -\text{Tr} \rho \log \rho$, which we introduced earlier (and which was defined by von Neumann 20 years earlier.) Note that, if λ_j and ϕ_j are eigenfunctions and eigenvalues of ρ so that $\rho = - \sum \lambda_j P_{\phi_j}$, then

$$S(\rho) = - \sum \lambda_j \log \lambda_j = H(\{\lambda_j\}). \quad (6.111)$$

We have the following inequalities between $H(p)$ and $S(\rho)$:

1) (Entropy of measurement) Given a density operator ρ , consider the probability distribution $p_M(x) := \text{tr}(M_x \rho)$, where $M := \{M_x\}$ is a positive operator-valued measure (POVM). Then

$$H(P_M) \geq S(\rho), \tag{6.112}$$

with equality when M_x and ρ commute. (Measurement increases entropy (randomness) and more so, more 'non-commuting' M_x and ρ are.)

2) (Entropy of preparation) $H(p) \geq S(\rho)$, where $\rho = \sum p(x)P_{\varphi_x}$, and = iff $\{\varphi_x\}$ are orthogonal.

6.22. Shannon's theorems and their quantum extensions. Two key results of classical information theory are the noiseless channel coding, or "source coding", theorem (how much can a message be compressed), and noisy channel coding, or "channel coding", theorem (at what rate can one communicate reliably over a noisy channel) (Shannon '48).

Shannon noiseless coding theorem: A source can be coded so that n symbols are sent using $\approx H(p)n$ bits and can be recovered with high probability.

To formulate the second result, we need some definitions, in particular that of the channel capacity. First we define the conditional entropy, $H(X|Y)$, of the random variable X given the random variable Y :

$$H(X = x|Y = y) = \sum_x p(x|y) \log p(x|y) \tag{6.113}$$

(so that $H(X|Y) = \sum_y H(X|Y = y)P_Y(y)$), where as usual $p(x|y) := P(X = x|Y = y)$ is the transitional probability distribution. Thus, $H(X|Y)$ is the given information about X if we know Y . Next, we introduce the mutual information entropy of X and Y :

$$I(X; Y) = H(X) - H(X|Y) = H(Y) - H(Y|X), \tag{6.114}$$

where, recall, $H(X) = H(P_X) = -\sum p_i \log p_i$ is the classical entropy. ($I(X, Y)$ says how much we know about X if Y is known.) Note that

$$I(X, Y) = H(X) + H(Y) - H(X, Y), \tag{6.115}$$

where $H(X, Y)$ is entropy of joint probability distribution of X and Y . Now, the channel capacity for a channel β_{cl} is defined as

$$C(\beta_{cl}) := \max_X I(X, \beta_{cl}(X)), \text{ or } C = \max_p I(p; \beta_{cl}(p)), \tag{6.116}$$

where X is an input random variable (resp. p is the input probability distribution).

Shannon's noisy channel theorem: $\approx nC(\beta_{cl})$ bits of information can be transmitted over a (classical) channel β_{cl} and recovered with high probability by using the channel n times (what is n ?), or more precisely, $\forall \epsilon > 0 \exists n$ and a coding scheme which takes $n(C(\beta) - \epsilon)$ bits and encodes them so that they can be transmitted by using the channel n times.

Thus with high probability, n uses of noisy channel can β_{cl} can communicate $C(\beta_{cl})n - o(n)$ bits reliably.

Now we proceed to the quantum case.

Quantum compression and transmission of information theorems

Theorem 6.53 (Quantum noiseless codimension theorem for pure states). *To obtain a good fidelity, as $n \rightarrow \infty$, the optimal compression of quantum information is $S(\rho)$ qubits per symbol. (This is the information content of $\rho \otimes \rho \dots \otimes \rho$? what is n ?)*

For mixed states we replace $S(\rho)$ by $\chi(e)$, where $e = \{p_x, \rho_x\}$ and $\chi(e)$ is the Holevo information,

$$\chi(e) := S\left(\sum p_x \rho_x\right) - \sum p_x S(\rho_x), \quad (6.117)$$

where, recall, $S(\rho) = -\text{Tr}(\rho \log \rho)$ is the von Neumann entropy. This is a generalization of the classical mutual information, $I(X, Y)$, given in (6.114).

Theorem 6.54 (Noisy channel theorem (quantum case) for cc channels). *The amount of classical information per symbol, which can be transmitted through a quantum channel, is*

$$\text{Acc}(\beta) = \max_X I(X; \hat{\beta}(X)), \quad (6.118)$$

where $\hat{\beta}(X) = \{\text{tr}(M_x \beta(\sum p_x \rho_x))\}$.

$\text{Acc}(\beta)$ is the amount of classical information which can be transmitted through a quantum channel.

Define $\chi(\beta) = \max_e \chi(\beta(e))$, where, for $e = \{p_x, \rho_x\}$,

$$\beta(e) = \beta\left(\sum p_i \rho_i\right) = \sum_{p_i} p_i \beta(\rho_i). \quad (6.119)$$

Theorem 6.55 (Quantum channel coding theorem). (*Holevo-Schumacher-Westmorland*):

$$\text{Acc}(\beta) \leq \chi(\beta). \quad (6.120)$$

6.23. Spin and Statistics. Surprising fact here is that the spin effects not only how the particles interact with the magnetic field but also how they behave socially, around each other.

Many-particle systems display a remarkable new feature of quantum physics. Unlike in classical physics, identical particles (i.e., particles with the same masses, charges and spins, or, more generally, which interact in the same way) are indistinguishable in quantum physics. Assume we have n identical particles of spin r . Classically, states of such a system are given by $(x_1, k_1, \dots, x_n, k_n)$, with the state

space, also called the *phase-space*, being $\otimes_1^n(\mathbb{R}_x^3 \times \mathbb{R}_k^3) = \mathbb{R}_x^{3n} \times \mathbb{R}_k^{3n}$. Naively, we might assume that the state space of the quantum system is

$$\otimes_1^n L^2(\mathbb{R}^3, \mathbb{C}^{2r+1}) \equiv L^2(\mathbb{R}^{3n}, \mathbb{C}^{(2r+1)^n}), \quad (6.121)$$

where the first term is the tensor product of n Hilbert spaces $L^2(\mathbb{R}^3, \mathbb{C}^{2r+1})$, defined as the Hilbert space spanned by the products of elements of orthonormal bases in each $L^2(\mathbb{R}^3, \mathbb{C}^{2r+1})$ (see Section 23.13 of [GS] for a definition for abstract Hilbert spaces).

Elements of this space can be written either as vector-functions $(\Psi_1(x_1, \dots, x_n), \dots, \Psi_{2r+1}(x_1, \dots, x_n))$, or as complex functions

$$\Psi(x_1, s_1, \dots, x_n, s_n) \in L^2(\mathbb{R}^3, \mathbb{C}^{2r+1})^{\otimes n} \equiv L^2(\mathbb{R}^{3n}, \mathbb{C}^{(2r+1)^n}) \quad (6.122)$$

where $x_j \in \mathbb{R}^3$ and $s_j \in \{1, \dots, 2r+1\}$.

The indistinguishability of the particles means that all probability distributions which can be extracted from an n -particle wave function $\Psi(x_1, s_1, \dots, x_n, s_n)$, should be symmetric with respect to permutations of the coordinates and spins, (x_j, s_j) , of identical particles. Since for bound states we can restrict ourselves to real wave functions, this is equivalent to the property that $\Psi(x_1, s_1, \dots, x_n, s_n)$ is invariant under such permutations, modulo a change of sign.

Recall that all elementary (and composite) particles are divided into two groups, the particles with half integer spins, which are called fermions (e.g. electrons, protons, and neutrons have spin $1/2$) and the particles with integer spins, which are called bosons (particles related to interactions). For bosons, the wave functions, $\Psi(x_1, s_1, \dots, x_n, s_n)$, should be symmetric with respect to permutations of the coordinates and spins of identical particles, and for fermions, antisymmetric. In particular, the state space for fermions of spin $\frac{1}{2}$ is

$$\mathcal{H}_{\text{fermi}}^{\text{total}} := \{\Psi \in L^2(\mathbb{R}^{3n}, \mathbb{C}^{2n}) \mid T_\pi \Psi = (-1)^{\#(\pi)} \Psi\} \quad (6.123)$$

where π is a permutation of the n indices,

$$\pi : (1, 2, \dots, n) \rightarrow (\pi(1), \pi(2), \dots, \pi(n)),$$

$\#(\pi)$ is the number of transpositions making up the permutation π (so $(-1)^{\#(\pi)}$ is the parity of $\pi \in S_N$). and

$$(T_\pi \Psi)(x_1, s_1, \dots, x_n, s_n) = \Psi(x_{\pi(1)}, s_{\pi(1)}, \dots, x_{\pi(n)}, s_{\pi(n)}).$$

Below we will write the space (6.123) as $\bigwedge_{i=1}^n L^2(\mathbb{R}^3, \mathbb{C}^2)$.

Since for the standard quantum-mechanical model, observables act on either spatial or the spin variables, by separation of variables, it suffices to consider wave functions of the form

$$\Psi(x_1, s_1, \dots, x_n, s_n) = \psi(x_1, \dots, x_n) \chi(s_1, \dots, s_n), \quad (6.124)$$

and therefore consider all spatial observables (including the hamiltonian) acting only on $\psi(x_1, \dots, x_n)$.

The symmetry type of $\psi(x_1, \dots, x_n)\chi(s_1, \dots, s_n)$ imposes symmetry restrictions on $\psi(x_1, \dots, x_n)$. What are those restrictions, i.e. what are the symmetry properties of these functions with respect to permutations of the coordinates?

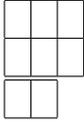
We consider fermions, which are of the main interest here. We define the Hilbert space of spatial wave functions as

$$\mathcal{H}_{\text{fermi}} := \{\psi \in L^2(\mathbb{R}^{3n}, \mathbb{C}) \mid \exists \chi : \otimes_1^n B_r \rightarrow \mathbb{C} \text{ s.t. } \psi\chi \in \mathcal{H}_{\text{fermi}}^{\text{total}}\} \quad (6.125)$$

where $B_r := \{1, \dots, 2r + 1\}$. Now, our task can be formulated as: describe explicitly the space $\mathcal{H}_{\text{fermi}}$.

To describe $\mathcal{H}_{\text{fermi}}$, one has to dip into the theory of representations of the symmetric group S_n (the group of permutations of n indices). We do not do so here, but just formulate the outcome of the theory.

Consider partitions, α , of the integer n into ordered positive integers $2r + 1 \geq \alpha_1 \geq \alpha_2 \geq \dots \geq \alpha_k \geq 1$. Denote the set of such α 's by A_r . These can be visualized as arrangements of n squares into k columns having $\alpha_1, \alpha_2, \dots, \alpha_k$ squares each, which are stuck side by side, and which are called *Young diagrams* (we retain for the them the same notation α), see the figure below.



In particular, for spin one half, $r = \frac{1}{2}$, we have one- and two-column Young diagrams. With a Young diagram α one associates a Young tableau, by filling in squares with particles.

We associate with a given Young diagram, α , the space, \mathcal{H}^α , of wave functions which are symmetric with respect to permutations of particles in the same row and antisymmetric with respect to those in the same column of some T (say, canonical one), associated with α . One can show that the subspace, \mathcal{H}^α are mutually orthogonal.

Theorem 6.56.

$$\mathcal{H}_{\text{fermi}} = \oplus_{\alpha \in A_r} \mathcal{H}^\alpha. \quad (6.126)$$

In other words, if Ψ in (6.124) is totally antisymmetric (i.e. it corresponds to the one-column Young diagram), then only functions ψ associated with Young diagrams, with $\leq 2r + 1$ columns, can appear there.

In technical terms, irreducible representations of the symmetric group S_n are in one-to-one correspondence with $\alpha \in A_r$ and on the subspace, \mathcal{H}^α , the representation of S_n is multiple to the irreducible one labeled by α . Irreducible representations of S_n can be connected, via Weyl's theory of dual pairs of groups, to irreducible representations of $SU(2)$ carried by the spin space $\mathbb{C}^{(2r+1)^n}$ and therefore determine the total spin of corresponding wave functions. Moreover, for $r = 1/2$, the total spin on \mathcal{H}^α is $\frac{1}{2}$ (the difference of the lengths of the longer and shorter columns). **(check)**

To summarize, for identical particles the state space is not $L^2(\mathbb{R}^{3n})$, but rather a subspace of it, defined by certain symmetry properties with respect to permutations of the particles.

The relation between the symmetry properties of wave functions and the spin of particles, is known as the *relation between spin and statistics*. For more detail, refer the interested reader to any of the standard books on quantum mechanics given in the references.

Example 6.57. 1) Consider an atom with N electrons of mass m and charge $-e$ and the nucleus of charge Ze . Assuming the nucleus is infinitely heavy (or passing to the center-of-mass frame), the Schrödinger operator, H_{at} , for such an atom, is

$$H_{\text{at}} = \sum_1^N \frac{1}{2m} (|p_j|^2 - \frac{e^2 Z}{|x_j|}) + \frac{1}{2} \sum_{i \neq j} \frac{e^2}{|x_i - x_j|}, \quad (6.127)$$

where x_j and $p_j = -i\hbar\nabla_{x_j}$ are the coordinate and momentum of the j -th electron, and $-\frac{e^2 Z}{|x_j|}$ and $\frac{e^2}{|x_i - x_j|}$ are the Coulomb interaction potentials between the j -th electron and the nucleus and the i -th electron. As we know, the electrons are fermions of the spin $1/2$. Hence H_{at} acts on the space $\mathcal{H}_{\text{fermi}}^N$, where we introduced the superindex N to indicate the number of particles involved.

2) Fermi/Bose gas in an external potential V : n identical particles, e.g. electrons, or atoms considered as composite particles (quantum gas), in an external potential V and with an inter-particle potential $v(x)$:

$$H_{\text{gas}} = \sum_1^n (\frac{1}{2m} |p_j|^2 + V(x_j)) + \frac{1}{2} \sum_{i \neq j} v(x_i - x_j) \quad (6.128)$$

acting on $\mathcal{H}_{\text{fermi}}$ (or on $\bigwedge_1^n (L^2(\mathbb{R}^3) \otimes \mathbb{C}^2)$) for fermions and on

$L_{\text{sym}}^2(\mathbb{R}^{3n}) = \{\psi \in L^2(\mathbb{R}^{3n}) | \psi \text{ is symmetric w.r.t permutations of coordinates } x_1, \dots, x_n \in \mathbb{R}^3\}$, for bosons.

3) Ideal quantum gas (a system of n identical particles not interacting mutually) in an external potential $V(x)$. The same as above but with $v = 0$. Hence, the Schrödinger operator for such a system is

$$H_{\text{idealgas}} = \sum_{j=1}^n (-\frac{\hbar^2}{2m} \Delta_{x_j} + V(x_j)). \quad (6.129)$$

Assume the Schrödinger operator $-\frac{\hbar^2}{2m} \Delta_x + V(x)$ has an isolated eigenvalue at the bottom of its spectrum - the ground state energy. In this case, so does H_{idealgas} and, for bosons, its ground state and ground state energy are given by $\Phi(x_1, \dots, x_n) = \phi(x_1) \dots \phi(x_n)$, and

$$E_1^{(n)} = -ne_1, \quad (6.130)$$

respectively, where $\phi(x)$ and e_1 are the ground state and the ground state energy of the Schrödinger operator $-\frac{\hbar^2}{2m}\Delta_x + V(x)$.

Homework 6.58. Prove the above statement.

On the other hand, for fermions, say of spin $\frac{1}{2}$, after separation of spin variables and for $n > 2$, the ground state of the operator \tilde{H}_{gas} cannot be given by the product $\prod_1^n \varphi_1(x_j)$, where $\varphi_1(x)$ is the ground state of the operator $-\frac{\hbar^2}{2m}\Delta_x + V(x)$. A little contemplation shows that it is given by the anti-symmetric product,

$$\bigwedge_1^{\frac{n}{2}} (\phi_j(x_{2j-1})\phi_j(x_{2j})),$$

(if n is even) of $\frac{n}{2}$ bound states, $\phi_j(x)$, of the operator H_{idealgas} , corresponding to the $\frac{n}{2}$ lowest energies, say $e_1, \dots, e_{\frac{n}{2}}$. (The case of n odd requires a slight modification.) In this case, the ground state energy, $E_1^{(n)}$, of H_{idealgas} is given by

$$E_1^{(n)} = 2 \sum_1^{\frac{n}{2}} e_j.$$

Homework 6.59. Prove the above statement.

This is, in general, a much larger number. Take for instance the Hydrogen atom ($V(x) = -\frac{e^2}{|x|}$). Then the energies E_j behave roughly as $-\frac{m(e^2Z)^2}{2\hbar^2}j^{-2}$ (see (6.20)) and $E_1^{(n)} = O(1)$, rather than $E_1^{(n)} = O(n)$ in the bosonic case. Say, the number of atoms in the room is of order $n = 10^{25}$, so we see that the difference is considerable. It is not easy to give a realistic estimate of the sum $2 \sum_1^{\frac{n}{2}} E_j$. There is considerable activity in producing such estimates (in terms of V), which are called the *Lieb-Thirring inequalities*.

Homework 6.60. Find the ground state energy of H_{idealgas} (in terms of the one-particle energies) for (a) 11 identical fermions of the spin $3/2$ on the subspace corresponding to Young diagram with with the columns of the lengths 4, 4, 2, 1; (b) 17 identical fermions of the spin $1/2$ on the subspace corresponding to Young diagram with with the columns of the lengths 10, 7.

6.24. Self-consistent Approximations. Even for a few particles the Schrödinger equation is prohibitively difficult to solve. Hence it is important to have approximations which work in various regimes. One such approximation, which has a nice unifying theme and connects to a large areas of physics and mathematics, is the one approximating solutions of n -particle Schrödinger equations by products of n one-particle functions (i.e. functions of 3 variables). This results in a single nonlinear equation in 3 variables, or several coupled such equations. The trade-off here is the number of dimensions for the nonlinearity. This method, which

goes under different names, e.g. the mean-field or self-consistent approximation, is especially effective when the number of particles, n , is sufficiently large.

We consider a system of n identical bosons or spinless fermions (i.e. with the statistics given by the single-column Young diagram). The Hamiltonian of the system of n identical particles of mass m , interacting with each other and moving in an external potential V is

$$H_n := \sum_{i=1}^n h_{x_i} + \frac{1}{2} \sum_{i \neq j} v(x_i - x_j), \quad (6.131)$$

where $h_x = -\frac{\hbar^2}{2m} \Delta_x + V(x)$ and v is the interaction potential. For bosons/spinless fermions, it acts on the state space $\mathbb{S}_1^n L^2(\mathbb{R}^d) / \wedge_1^n L^2(\mathbb{R}^d)$, $d = 1, 2, 3$. Here \mathbb{S}/\wedge is the symmetric/antisymmetric tensor product. As we know, the quantum evolution is given by the Schrödinger equation

$$i\hbar \frac{\partial \Psi}{\partial t} = H_n \Psi.$$

This is an equation in $3n + 1$ variables, x_1, \dots, x_n and t , and it is not a simple matter to understand properties of its solutions.

We would like to approximate the true QM system with one in which the particles do not interact, i.e. one described by the Schrödinger operator (6.129). To have the best approximation we have to modify the one-particle potential V making it dependent on the solution. This works well in the regime where the number of particles n is of the order v^{-1} , or the inter-particle interaction v is of the order n^{-1} . This is called the *mean-field* regime.

We give a heuristic derivation of the mean-field approximation for the Schrödinger equation above. First observe that the Schrödinger equation is the Euler-Lagrange equation for stationary points of the action functional

$$S(\Psi) := \int \{ -2\hbar \text{Im} \langle \Psi, \partial_t \Psi \rangle - \langle \Psi, H_n \Psi \rangle \} dt, \quad (6.132)$$

Now, consider the the action functional (6.132) the space (not linear!)

$$\{ \Psi := \otimes_1^n \psi \mid \psi \in H^1(\mathbb{R}^3) \}.$$

for bosons and

$$\{ \Psi := \det[\psi_i(x_j)] \mid \psi_i \in H^1(\mathbb{R}^3) \ \forall i = 1, \dots, n \}$$

for fermions. Here $(\otimes_1^n \psi)$ is the function of $3n$ variables defined by $(\otimes_1^n \psi)(x_1, \dots, x_n) := \psi(x_1) \dots \psi(x_n)$ and $[\psi_i(x_j)]$ stands for the $n \times n$ matrix with the entries indicated. (See Chapter 24 of [GS] for material on variational calculus.)

We begin with the bosons. Let $S_H(\psi) := n^{-1}S(\otimes_1^n \psi)$ ('H' stands for the Hartree). We claim that, provided $\|\psi\| = 1$, we have

$$S_H(\psi) = \int \int_{\mathbb{R}^3} \left\{ -2\hbar I m \langle \psi, \partial_t \psi \rangle - \frac{\hbar^2}{2m} |\nabla \psi|^2 - V |\psi|^2 - |\psi|^2 w * |\psi|^2 \right\} dx. \quad (6.133)$$

where $w := nv$. The regime in which $w := nv = O(1)$, or $v = O(1/n)$, is called the mean-field regime. We show this. First, we compute $\langle \Psi, \partial_t \Psi \rangle$ for $\Psi := \otimes_1^n \psi$. Denote $\psi_j := \psi(x_j)$. The relation $\partial_t \Psi = \sum_i (\prod_{j \neq i} \psi_j) \partial_t \psi_i$ gives

$$\begin{aligned} \langle \Psi, \partial_t \Psi \rangle &= \sum_i \int \prod_j \bar{\psi}_j (\prod_{j \neq i} \psi_j) \partial_t \psi_i = \sum_i \|\psi\|^{2(n-1)} \int \bar{\psi}_i \partial_t \psi_i \\ &= n \|\psi\|^{2(n-1)} \int \bar{\psi} \partial_t \psi \end{aligned}$$

Similarly, for $\Psi := \otimes_1^n \psi$, we compute

$$\begin{aligned} \langle \Psi, \sum_{i=1}^n (-\frac{\hbar^2}{2m} \Delta_{x_i} + V(x_i)) \Psi \rangle &= \sum_{i=1}^n \int (\frac{\hbar^2}{2m} |\nabla_{x_i} \Psi|^2 + V(x_i) |\Psi|^2) \\ &= \sum_{i=1}^n \|\psi\|^{2(n-1)} \int (\frac{\hbar^2}{2m} |\nabla_{x_i} \psi|^2 + V(x_i) |\psi|^2) \\ &= n \|\psi\|^{2(n-1)} \int (\frac{\hbar^2}{2m} |\nabla \psi|^2 + V(x) |\psi|^2) \end{aligned}$$

Finally, we compute the particle pair interaction terms

$$\begin{aligned} \langle \Psi, (\frac{1}{2} \sum_{i \neq j} v(x_i - x_j)) \Psi \rangle &= \frac{1}{2} \sum_{i \neq j} \int (\prod_{k \neq i, j} |\psi_k|^2) |\psi_i|^2 v(x_i - x_j) |\psi_j|^2 \\ &= \frac{1}{2} \sum_{i \neq j} \|\psi\|^{2(n-2)} \int \int |\psi_i|^2 v(x_i - x_j) |\psi_j|^2 \\ &= \frac{1}{2} n(n-1) \|\psi\|^{2(n-2)} \int \int |\psi(x)|^2 v(x - x') |\psi(x')|^2 \end{aligned}$$

Collecting the terms above and using that $v = (n-1)w$, we arrive at (6.133).

The Euler-Lagrange equation for stationary points of the action functional (6.133) considered on the first set of functions is

$$i\hbar \frac{\partial \psi}{\partial t} = (h + v * |\psi|^2) \psi. \quad (6.134)$$

This nonlinear evolution equation is called the *Hartree equation* (HE).

Homework 6.61. Prove that (6.134) is the Euler-Lagrange equation for (6.133).

Properties of (HE). The Hartree equation has the following general features

- (1) (HE) is invariant under the time-translations, $\psi(x, t) \rightarrow \psi(x, t + s)$, $s \in \mathbb{R}$, the gauge transformations,

$$\psi(x, t) \rightarrow e^{i\alpha} \psi(x, t), \quad \alpha \in \mathbb{R},$$

and, for $V = 0$, the translations, $\psi(x, t) \rightarrow \psi(x + y, t)$, $y \in \mathbb{R}^3$, and the Galilean transformations, $v \in \mathbb{R}^3$,

$$\psi(x, t) \rightarrow e^{i(mv \cdot x + \frac{mv^2 t}{2})/\hbar^2} \psi(x - vt, t),$$

and, for V spherically symmetric, the rotations, $\psi(x, t) \rightarrow \psi(Rx, t)$, $R \in O(3)$,

- (2) (HE) is a Hamiltonian system (see Section 19.1 of [GS])

As the result of the time-translational and the gauge symmetries, the energy and the number of particles functionals

$$E(\psi) := \int_{\mathbb{R}^3} \left\{ \frac{\hbar^2}{2m} |\nabla \psi|^2 + V|\psi|^2 + |\psi|^2 v * |\psi|^2 \right\} dx,$$

$$N(\psi) := \int_{\mathbb{R}^3} |\psi|^2 dx$$

are conserved. Moreover, for $V = 0$, the field momentum,

$$P(\psi) := \int \bar{\psi}(x, t) (-i\hbar \nabla_x) \psi(x, t) dx,$$

and, for V spherically symmetric, the field angular momentum,

$$L(\psi) := \int \bar{\psi}(x, t) (x \wedge (-i\hbar \nabla_x)) \psi(x, t) dx,$$

are conserved.

Homework 6.62. Prove that solutions of (6.134) conserve the energy and the number of particles given above.

The notion of bound state can be extended to the nonlinear setting as follows. The *bound states* are stationary solutions of (HFE)) of the form

$$\psi(x, t) = \phi_\mu(x) e^{i\mu t}$$

where the profile $\phi_\mu(x)$ is in $H^2(\mathbb{R}^3)$. Note that the profile $\phi_\mu(x)$ satisfies the stationary (HFE):

$$h\phi + (v * |\phi|^2)\phi = -\hbar\mu\phi \tag{6.135}$$

Thus we can think of the parameter $-\mu$ as a nonlinear eigenvalue.

A *ground state* is a bound state such that the profile $\phi_\mu(x)$ minimizes the energy for a fixed number of particles:

$$\phi_\mu \text{ minimizes } E(\psi) \quad \text{under } N(\psi) = n$$

(see Chapter 24 of [GS] which deals with variational calculus, and in particular minimization problems). Thus the nonlinear eigenvalue μ arises as a Lagrange

multiplier from this constrained minimization problem. In Statistical Mechanics μ is called the chemical potential (the energy needed to add one more particle/atom).

Remark 6.63. (1) Mathematically, the ground state can be also defined as a stationary solution with a positive (up to a constant phase factor) profile, $\psi(x, t) = \phi_\mu(x)e^{i\mu t}$ with $\phi_\mu(x) > 0$. Let $\delta(\mu) := \|\phi_\mu\|^2$. Then we have

$$\delta'(\mu) > 0 \implies \phi_\mu \text{ minimizes } E(\psi) \text{ under } N(\psi) = n.$$

(2) The Lagrange multiplier theorem in Section 24.5 of [GS] implies that the ground state profile ϕ_μ is a critical point of the functional

$$E_\mu(\psi) := E(\psi) + \hbar\mu N(\psi).$$

In fact, ϕ_μ is a minimizer of this functional under the condition $N(\psi) = n$.

The Euler-Lagrange equation for stationary points of the action functional (6.132) considered on the latter set of functions is a system of nonlinear, coupled evolution equations

$$i\hbar \frac{\partial \psi_j}{\partial t} = (h + v * \sum_i |\psi_i|^2) \psi_j - \sum_i (v * \psi_i \bar{\psi}_j) \psi_i, \quad (6.136)$$

for the unknowns ψ_1, \dots, ψ_n . This systems plays the same role for fermions as the Hartree equation does for bosons. It is called the *Hartree-Fock equations* (HFE).

Properties of (HFE). The Hartree-Fock equations have the following general features

- (1) (a) (HFE), as (HE) are invariant under the time-translations and gauge transformations, and, for $V = 0$, the spatial translations, $\psi_j(x) \rightarrow \psi_j(x + y)$, $y \in \mathbb{R}$, and the Galilean transformations, $v \in \mathbb{R}^3$, and, for V spherically symmetric, the rotations.
- (b) (HFE) is invariant under time-independent unitary transformations of $\{\psi_1, \dots, \psi_n\}$.
- (2) (HFE) are Hamiltonian systems (see Sections 24.6 and 24.7 of [GS]).

Again, similarly to (HE), as the result of the time-translational and the gauge symmetries, the energy and the number of particles functionals

$$E(\psi) := \int \left\{ \sum_i \left(\frac{\hbar^2}{2m} |\nabla \psi_i|^2 + V |\psi_i|^2 \right) + \frac{1}{2} \left(\sum_i |\psi_i|^2 \right) v * \left(\sum_i |\psi_i|^2 \right) - \frac{1}{2} \int \int v(x - y) \left| \sum_i \psi_i(x) \psi_i(y) \right|^2 \right\} dx, \quad (6.137)$$

$$N(\psi) := \sum_i \int_{\mathbb{R}^3} |\psi_i|^2 dx \quad (6.138)$$

are conserved, etc. Moreover, (HFE) conserves the inner products, $\langle \psi_i, \psi_j \rangle$, $\forall i, j$.

The item (2b) shows that the natural object for (HFE) is the subspace spanned by $\{\psi_i\}$, or the corresponding projection $\gamma := \sum_i |\psi_i\rangle\langle\psi_i|$. (HFE) can be rewritten as an equation for γ :

$$i\hbar \frac{\partial \gamma}{\partial t} = [h(\gamma), \gamma] \quad (6.139)$$

where $h(\gamma) := h + v * \rho_\gamma - ex(\gamma)$, with $\rho_\gamma(x) := \gamma(x, x) = \sum_i |\psi_i(x)|^2$ and $ex(\gamma)$ is the operator with the integral kernel $ex(\gamma)(x, y) := v(x-y)\gamma(x, y) = \sum_i (\bar{\psi}_i v * \psi_i)$. (We write $ex(\gamma)$ as $ex(\gamma) =: v\#\gamma$.)

This can be extended to arbitrary density matrices γ and leads to a new class of nonlinear differential equations. Finally, note that the energy and the number of particles in the new formulation is given by

$$E(\gamma) := \text{Tr}(h\gamma) + \frac{1}{2} \int \rho_\gamma v * \rho_\gamma - Ex(\gamma) dx, \quad (6.140)$$

$$N(\psi) := \gamma(x, x), \quad (6.141)$$

where $Ex(\gamma) := -\frac{1}{2} \int \int v(x-y) |\gamma(x, y)|^2 = \text{Tr}(\gamma v\#\gamma)$.

Homework 6.64. Prove that equations (6.136) and (6.138) can be rewritten as (6.139) and (6.141), respectively.

REFERENCES

- [BR] Bratteli, O., Robinson, D.W., Operator algebras and quantum statistical mechanics. I and II, Springer Verlag 1987
- [CFKS] H. Cycon, R. Froese, W. Kirsch, B. Simon, *Schrödinger Operators (with Applications to Quantum Mechanics and Global Geometry)*. Springer: Berlin, 1987.
- [Fo] G. Folland, *Real Analysis*. Wiley, 1984.
- [GF] I.M. Gelfand and S. Fomin, *Calculus of Variations*. Prentice-Hall, 1963.
- [GJ] J. Glimm, A. Jaffe, *Quantum Physics: a Functional Integral Point of View*. Springer-Verlag: New York, 1981.
- [GS] S. Gustafson, I.M. Sigal, *Mathematical Concepts of Quantum Mechanics*. 2nd edition Springer-Verlag, 2011.
- [Hall] Brian C. Hall, An Elementary Introduction to Groups and Representations, ArXiv, math-ph/0005032v1, May 2000.
- [HS] P. Hislop, I.M. Sigal, *Introduction to Spectral Theory*. Springer-Verlag, 1996.
- [HuS] W. Hunziker, I.M. Sigal, The quantum N -body problem. J. Math. Phys. **41** no.6 (2000) 3448-3510.
- [LL] L.D. Landau, E.M. Lifshitz, *Theoretical Physics 3. Quantum Mechanics: Non-Relativistic Theory*. Pergamon, 1977.
- [LiL] E.H. Lieb, M. Loss, *Analysis*. American Mathematical Society, 2001.

- [RSI] M. Reed, B. Simon, *Methods of Modern Mathematical Physics, Vol I. Functional Analysis*. Academic Press, 1980.
- [T4] W. Thirring, *A Course in Mathematical Physics, Vol. 4. Quantum Mechanics of Large Systems*. Springer-Verlag, 1983.