Draft

we.taper

October 5, 2016

Abstract

This is a draft.

Contents

1 Problems in Bernevig's Topological ...

1.1 Chapter 2

Non-Abelian Berry Transport Derive Berry curvature to the adiabatic transport of a degenerate multiplet of states separated by a gap from the excited states. (Cautious about rotation within degenerate states).

Answer: $\gamma_{mn}(t) = i\int_0^t \langle m(R(t'))|\frac{d}{dt'}|n(R(t'))\rangle\, dt'$

Approach: Assuming that the those degenerate states are labeled by $1\cdots N$. Thus we have naturally:

$$H\phi = i\hbar\frac{\partial}{\partial t}\phi,\ \phi = \sum_n A_n\psi_n \tag{1.1.1}$$

Then we have:

$$\begin{aligned} H\sum_n A_n\psi_n &= i\hbar\frac{\partial}{\partial t}\sum_n A_n(t)\psi_n(R(t)) \\ \sum_n EA_n\psi_n &= i\hbar\sum_n\left(\frac{\partial A_n(t)}{\partial t}\psi_n(R(t)) + A_n(t)\frac{\partial\psi_n(R(t))}{\partial t}\right) \\ EA_m &= i\hbar\frac{\partial A_m(t)}{\partial t} + \sum_n A_n(t)\langle m|\frac{\partial}{\partial t}|n\rangle \end{aligned} \tag{1.1.2}$$

Put in another form:

$$\sum_n\left(\delta_m^n E - \langle m|\frac{\partial}{\partial t}|n\rangle\right)A_n = i\hbar\frac{\partial A_m(t)}{\partial t}$$

In matrix form:

$$(E-P)\mathbf{A} = i\hbar\dot{\mathbf{A}} \tag{1.1.3}$$

where:

$$E = \begin{pmatrix} \cdots & & \\ & E & \\ & & \ldots \end{pmatrix} \tag{1.1.4}$$

$$P = (P_n^m) = \left(\langle m|\frac{\partial}{\partial t}|n\rangle\right) \tag{1.1.5}$$

$$A = \begin{pmatrix} A_1(t) \\ A_2(t) \\ \cdots \end{pmatrix} \tag{1.1.6}$$

Note that $\langle n|\frac{\partial}{\partial t}|m\rangle^* \neq \langle m|\frac{\partial}{\partial t}|n\rangle$, thus P may not be Hermitian.Ergo $E-P$ is Hermitian. So it is diagonalizable.

Notice that

$$0 = \frac{\partial}{\partial t}\langle m|n\rangle = \langle\frac{\partial}{\partial t}m|n\rangle + \langle m|\frac{\partial}{\partial t}n\rangle \qquad \text{(any } m,n) \tag{1.1.7}$$

temporary mathematica code:

2 Quantum Field Theory

2.1 Relatvistic Quantum Mechanics

2.1.1 2.1 Quantum Mechanics

This part sumarize the axioms of quantum mechanics. (Omitted) However, one class of important but seldom mentioned operator, the antilinear operators and the antiunitary operators are not discussed here but postponed to the next part.

2.1.2 2.2 Symmetries

This part discribe some important theorems concerning the symmetries in quantum mechanics.

Firstly, symmetries in quantum mechanics some times requires the use of antilinear and antiunitary operators.

Antilinear and Antiunitary Antilinear operators: $U(\lambda A + \upsilon B) = \lambda^* UA + \upsilon^* UB$.

Antiunitary: $(U\phi, U\psi) = (\psi, \phi) = (\phi, \psi)^*$

The adjoint operator of an antilinear operator requires special attention:

$$(\phi, A^\dagger \psi) = (A\phi, \psi)^*$$

(The usual definition will become troublesome since the LHS will be antilinear in ϕ, whereas the RHS is linear.)

Thankfully, the criterion for unitarity or antiunitarity is the same:

$$U^\dagger = U^{-1}$$

An important example is those symmetries which can be infinitesimally close to identity. In the vincinity of unity, it is:

$$U = 1 + i\epsilon t$$

where t is Hermitian and linear.

Symmetry Group and its Representation The set of all symmetries transformations obviously can have a group structure. By giving each symmetry transformation a unitary or antiunitary operator, a representation of such group is obtained. However, such a representation can be projective (projective as in projective geometry, or $\mathbb{CP}^n$), since operators acts on ket spaces, which is already projective (within a freedom of phase).

If the phase is taken into consideration, it is found that this phase is independent of the ket that the operator acts on:

$$U(T_1 T_2)\Psi = e^{i\phi} U(T_1) U(T_2) \Psi \qquad (2.1.1)$$

Here $U(T_1)/U(T_2)$ is the operator corresponding to the symmetry transformation T_1/T_2. ϕ is the phase, depends only on $T_1 T_2$.

Howver, there is one exception to this rule. It exists when the state Ψ is not the linear some of some other wave functions. That is, we can not express Ψ as $\sum_i \lambda_i \psi_i$. This seems incomprehesible to me. There are some link on page 53 between the structure of Lie group associated with this symmetry and whether the representation can be projective. [1]

Connected Lie group is of special importance in physics. However, the books description obfuscate the mathematical description of this structure. I will update this note later to remedy such discussion. The message I understand is that the representation is tightly bound to the Lie algebra. An important example is that when

$$U(T(\theta)) \approx 1 + i\theta^a t_a$$

$$[t_b, t_c] = 0$$

then

$$U(T(\theta)) = \exp(i t_a \theta^a)$$

2.1.3 2.3 Quantum Lorentz Transformations

Note: although the title suggests "quantum", this part is mostly classical.

Starting with the invariance of interval:

$$\eta_{\mu\nu} dx'^{\mu} dx'^{\nu} = \eta_{\mu\nu} dx^{\mu} dx^{\nu} \tag{2.1.2}$$

or equivalently

$$\eta_{\mu\nu} \frac{\partial x'^{\mu}}{\partial x^{\rho}} \frac{\partial x'^{\nu}}{\partial x^{\sigma}} = \eta_{\rho\sigma} \tag{2.1.3}$$

It claims that any coordinate transformation satisfying 2.1.3 must be linear.

Such linear Lorentz transformations satisfy:

$$T(\bar{\Lambda}, \bar{a}) T(\Lambda, a) = T(\bar{\Lambda}\Lambda, \bar{\Lambda}a + \bar{a}) \tag{2.1.4}$$

where $\bar{\Lambda} + \bar{a}$ and $\Lambda + a$ are two such transformations.

The Λ and Λ^{-1} can be determined by several equations that relate it with $\eta_{\mu\nu}$.

Note that:

$$\Lambda^{\rho}{}_{\nu} \neq \Lambda_{\nu}{}^{\rho}$$

And:

$$(\Lambda^{-1})^{\rho}{}_{\nu} = \Lambda_{\nu}{}^{\rho} = \eta_{\nu\mu} \eta^{\rho\sigma} \Lambda^{\mu}{}_{\sigma} \tag{2.1.5}$$

The only part about quantum mechanics here, is that the operators corresponding to the Lorentz transformations, have the property that:

$$U(\bar{\Lambda}, \bar{a}) U(\Lambda, a) = U(\bar{\Lambda}\Lambda, \bar{\Lambda}a + \bar{a}) \tag{2.1.6}$$

where a represents the translation.

[1] Though not state, it can be infered that a representation can be made not-projective by a good choice of $U(T)$, so that $\phi \equiv 0$ in equation (2.1.1) on the preceding page

The group of Lorentz transformations is very important. There is:

$$\begin{tikzcd}
\text{Poincaré Group} \\
\text{Homogeneous Lorentz group} \arrow[u, hook] \\
\text{proper orthochronous Lorentz group} \arrow[u, hook]
\end{tikzcd}$$

The important *proper orthochronous Lorentz group* consists of those Λ with $\det \Lambda = 1$ and ${\Lambda^0}_0 \geq 1$.

We have also $\mathscr{P}$, the space inversion matrix. And $\mathscr{T}$, the time-reversal matrix. The definition for both can be easily written.

A important rule is that all homogeneous Lorentz transformations can are generated by $\{$proper orthochronous, $\mathscr{P}, \mathscr{T}\}$. However this property is not proved

Digression about tensor notation It seems that phsysicists have not settled down on their notation for tensors. Weinberg using the notation ${\Lambda^\nu}_\mu \equiv \Lambda(e^\nu, e_\mu)$, so it acts on (v, w), where v is a cotangent vector and w is a tangent vector. I see that this notation helps to distinguish the type of arguments. This is unnecessary, since the upper and lower indices already fulfill this function. Another benefit is that it conveies the process of lowering of raising a index.

However, this idea that each slot in tensor indices are distinct and can be lowered and raised seperatly, is not even mentioned in some modern mathematical physics book.

Helpful link: Convention of tensor indices in Phy.SE, and Working with indices of tensors in special relativity in Phy.SE.

2.1.4 2.4 The Poincaré Algebra

I am lost in this part. The general idea seems to develop, in a infinitesimal sense, the property of a Lorentz transformation. For a infinitesimal Lorentz transformation:

$$U(1+\omega, \epsilon) = 1 + \frac{1}{2} i\omega_{\rho\sigma} J^{\rho\sigma} - i\epsilon_\rho P^\rho + \cdots \tag{2.1.7}$$

where J and P are independent of the infinitesimal value ω and ϵ.

Later P is identified as the energy-momentum operator (with P^0 being the energy operator), and J^{23}, J^{31}, J^{12} are identified with the angular momentum operator. However, the reason for this identification is not provided.

Later, it was eatablished that J and P satisfy:

$$i[J^{\mu\nu}, J^{\rho\sigma}] = \eta^{\nu\rho} J^{\mu\sigma} - \eta^{\mu\rho} J^{\nu\sigma} - \eta^{\sigma\mu} J^{\rho\nu} + \eta^{\sigma\nu} J^{\rho\mu} \tag{2.1.8}$$
$$i[P^\mu, J^{\rho\sigma}] = \eta^{\mu\rho} P^\sigma - \eta^{\mu\sigma} P^\rho \tag{2.1.9}$$
$$[P^\mu, P^\rho] = 0 \tag{2.1.10}$$

This is called the *Lie Algebra* of the Poincaré group. The relation $[P^\mu, P^\rho] = 0$ is particular interesting.

With this, the finite translation and a rotation by angle θ are expressed as:

$$U(1, a) = \exp(-iP^\mu a_\mu) \tag{2.1.11}$$

$$U(\theta, 0) = \exp(i\mathbf{J} \cdot \theta) \tag{2.1.12}$$

This part ends with a discussion on the low-velocity limit of the Lie Algebra obtained above, i.e. the Galilean algebra.

2.2 Note

I will now change to another book: Q.F.T. in a Nutshell by A. Zee.

3 Miscellnaneous Math

3.1 $\int_0^\infty \frac{\sin(x)}{x}$

From Math.SE. By Aryabhata.

I believe this can also be solved using double integrals.

It is possible (if I remember correctly) to justify switching the order of integration to give the equality:

$$\int_0^\infty \left(\int_0^\infty e^{-xy} \sin x \, dy\right) dx = \int_0^\infty \left(\int_0^\infty e^{-xy} \sin x \, dx\right) dy$$

Notice that

$$\int_0^\infty e^{-xy} \sin x \, dy = \frac{\sin x}{x}$$

This leads us to

$$\int_0^\infty \left(\frac{\sin x}{x}\right) dx = \int_0^\infty \left(\int_0^\infty e^{-xy} \sin x \, dx\right) dy$$

Now the right hand side can be found easily, using integration by parts.

$$\begin{aligned}
I = \int e^{-xy} \sin x \, dx &= -e^{-xy} \cos x - y \int e^{-xy} \cos x \, dx \\
&= -e^{-xy} \cos x - y\Big(e^{-xy} \sin x + y \int e^{-xy} \sin x \, dx\Big) \\
&= \frac{-ye^{-xy} \sin x - e^{-xy} \cos x}{1 + y^2}.
\end{aligned}$$

Thus

$$\int_0^\infty e^{-xy} \sin x \, dx = \frac{1}{1 + y^2}$$

Thus

$$\int_0^\infty \left(\frac{\sin x}{x}\right) dx = \int_0^\infty \frac{1}{1 + y^2} \, dy = \frac{\pi}{2}.$$

3.2 About $e^{-i\pi S_y/\hbar}$

(Related: Sakurai's Q.M., about (4.4.65)).

Since

$$S_y = \frac{\hbar}{2}\sigma_y = \frac{\hbar}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$$

It is easy to check that $\sigma_y^2 = \mathbb{I}$. So:

$$\begin{aligned} e^{-i\phi S_y/\hbar} = e^{-i\frac{\phi}{2}\sigma_y} &= \sum_{n=0}^{\infty} \frac{(-i\frac{\phi}{2})^n(\sigma_y)^n}{n!} \\ &= \sum_{\text{odd } n} -\frac{i(\frac{\phi}{2})^n}{n!}\sigma_y + \sum_{\text{even } n} \frac{(i\frac{\phi}{2})^n}{n!}\mathbb{I} \end{aligned}$$

Using the taylor expansions:

$$\sinh x = \sum_{n=0}^{\infty} \frac{x^{2n+1}}{(2n+1)!} \text{ (odd terms)} \tag{3.2.1}$$

$$\cosh x = \sum_{n=0}^{\infty} \frac{x^{2n}}{(2n)!} \text{ (even terms)} \tag{3.2.2}$$

We see that:

$$e^{-i\phi S_y/\hbar} = e^{-i\frac{\phi}{2}\sigma_y} = -\sigma_y \sinh i\frac{\phi}{2} + \mathbb{I}\cosh i\frac{\phi}{2}$$

$$= -i\sigma_y \sin\frac{\phi}{2} + \mathbb{I}\cos\frac{\phi}{2} \tag{3.2.3}$$

$$= -i\frac{2}{\hbar}S_y \sin\frac{\phi}{2} + \mathbb{I}\cos\frac{\phi}{2} \tag{3.2.4}$$

$$= \begin{pmatrix} \cos(\frac{\phi}{2}) & -\sin(\frac{\phi}{2}) \\ \sin(\frac{\phi}{2}) & \cos(\frac{\phi}{2}) \end{pmatrix} \tag{3.2.5}$$

So

$$e^{-i\pi S_y/\hbar} = -i\frac{2S_y}{\hbar}$$

By exactly the same argument, or simply by the symmetry ($x \leftrightarrow y$), we have:

$$e^{-i\phi S_x/\hbar} = e^{-i\frac{\phi}{2}\sigma_x} = -i\sigma_x \sin\frac{\phi}{2} + \mathbb{I}\cos\frac{\phi}{2} \tag{3.2.6}$$

$$= -i\frac{2}{\hbar}S_x \sin\frac{\phi}{2} + \mathbb{I}\cos\frac{\phi}{2} \tag{3.2.7}$$

$$= \begin{pmatrix} \cos(\frac{\phi}{2}) & -\sin(\frac{\phi}{2}) \\ \sin(\frac{\phi}{2}) & \cos(\frac{\phi}{2}) \end{pmatrix} \tag{3.2.8}$$

Also, one can easily find that:

$$e^{-i\phi S_z/\hbar} = e^{-i\frac{\phi}{2}\sigma_z} = \begin{pmatrix} e^{-i\frac{\phi}{2}} & 0 \\ 0 & e^{i\frac{\phi}{2}} \end{pmatrix} \tag{3.2.9}$$

Here a formula that encompass all the above relationship:

$$\exp(\frac{-i\vec{\sigma}\cdot\hat{\vec{n}}\phi}{2}) = \cos(\frac{\phi}{2})\mathbb{I} - i\vec{\sigma}\cdot\hat{\vec{n}}\sin(\frac{\phi}{2}) \tag{3.2.10}$$

where $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$, and $\hat{\vec{n}}$ is any unit vector.

4 Miscellnaneous Notes

4.1 Super conductor

Mean-field approach to deal with a four operator diagonalization.

Suppose we have: D^*C^*CD, then let $\delta = CD - \langle CD \rangle = CD - avg$. Then if we assume $\langle CD \rangle \neq 0$, and $\delta \approx 0$. Then we have:

$$\delta^2 \approx 0$$

i.e.:

$$((CD)^* - avg)(CD - avg) = 0 \tag{4.1.1}$$

$$D^*C^*CD = avg * (CD + D^*C^*) - avg^2 \tag{4.1.2}$$

Hence a four operator is reduced into a few of two operators. Such method could be naturally extended to treat the operator $\sum_{i,j} D_i^* C_i^* C_j D_j$.

A copper pair has the energy of:

$$\Delta = \langle C_{k\uparrow} C_{-k\downarrow} \rangle$$

To resist the flow of current carried by Copper Pair, is equivalent to destroying a pair of Copper Pair:

$$\langle C_{k\uparrow} C_{-k\downarrow} \rangle \longrightarrow C_{k\uparrow} C_{-k\downarrow}$$

This will require an additional enegy of 2Δ.

The exact meaning of "equivalent to" is as follows:

$$\begin{aligned} &\text{break a copper pair} \longrightarrow \text{scatter two electrons consecutively} \\ &\longrightarrow \text{create two electron-hole mixed type quasi-particle} \longrightarrow 2\Delta \end{aligned}$$

4.2 Why 0/0 is undefined?

If we suppose

$$\frac{0}{0} = \triangle$$

Consider the following derivation:

$$\frac{0}{0} \cdot 1 = \triangle \cdot 1 = \triangle \tag{4.2.1}$$

$$0 \cdot \frac{1}{0} = \triangle \tag{4.2.2}$$

$$\Rightarrow \triangle = 0 \tag{4.2.3}$$

This is already bad enough. And we are forced to define $\frac{1}{0}$. Let $\frac{1}{0} = \square$, which literally means $1 = 0 \cdot \square = 0$. This is disastrous.

Alternatively, we could let

1. Let $\frac{1}{0}$ be undefined.
2. Or let $\frac{1}{0} = \infty$.
3. Or, let $\frac{a}{b} \cdot c = a \cdot \frac{c}{b}$ be not true when $b = 0$.

The third idea is disastrous for algebraic manipulation. [2] The first idea is not good. Since defining $\frac{1}{0} = \infty$ turns out to be very useful in both mathematics and physics. Actully, in physics it is common practice to set $\frac{a}{0} = \pm\infty$ for any nonzero number a, where the sign of ∞ is determined by the sign of a. The second idea is okey. But then we are faced with a serious problem. We have to define $\triangle \equiv 0 \cdot \infty$

$\triangle \cdot 2 = \triangle$, What will be of $\triangle + 1$?

4.3 Preface of BSCS

BSCS: see [2]. Parallism between theories in condensed matter physics and those in particle physics.

- Anderson-Higgs Phenomenon (Paritcle), Meissner effect (C.M.P.)
- 'inflation' in Cosmology, first order phase transition
- 'cosmic strings', magnetic field vortex lines in type II superconductors
- Hadron-meson interaction, Ginzburg-Landau theory of superfluid He^3.

Same ideas on different space-time scales, different hierachical 'layers'. Strong parallism: **strongly correlated low dimensional system**

E.g.:

The problem of formation and structure of heavy particles - hadrons and mesons. The corresponding fine structure constant $\alpha_G \approx 1$.

Approaches:

1. Exact solutions
2. Reformulate complicated interacting models in such a way that they become weekly interacting. -> Bosonization.

   Spin 1/2 anisotropic Hisenberg chain $\approx$ Model of interacting fermions. (Jordan and Wigner, 1928)

Bosonization: transformation from fermions to a scalar massless bosonic field.

4.4 System of Differential Equations

This is a small note of [3].

pp. 266.

Definition 4.1. $\mathbf{x}(\mathbf{t})$ *is a vector whose elements are* $x_i(t)$. $\frac{d}{dt}$ *acts on vector* $\mathbf{x}$ *element-wise.* $\dot{\mathbf{x}}$ *is abbrevation for* $\frac{d}{dt}\mathbf{x}$

pp. 291.

Theorem 4.1 (Existence-uniqueness theorem). *There exists one, and only one, solution of the initial-value problem*

[2] Or more speicifically, it is a disaster for field theory.

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x},\ \mathbf{x}(t_0) = \mathbf{x}^0 = \begin{pmatrix} x_1^0 \\ x_2^0 \\ \cdots \end{pmatrix} \tag{4.4.1}$$

Moreover, this solution exists for $-\infty\langle t\langle\infty$.

Remark 4.1. *By this, any non-trivial solution* $\mathbf{x}(t) \neq 0$ *at any time* t. *Also notice that the elements of* $\mathbf{A}$ *are just numbers.*

Theorem 4.2. *The dimension of the space* $\mathbf{V}$ *of all solutions of the homogeneous linear system of differential equations:*

$$\frac{d\mathbf{x}}{dt} = \mathbf{A}\mathbf{x} \tag{4.4.2}$$

is n, *i.e. the dimension of vector* $\mathbf{x}$.

4.5 ODE by Arnold

sec. 14

Definition 4.2.

$$e^A = I + A + \frac{A^2}{2!} + \frac{A^3}{3!} \tag{4.5.1}$$

or

$$e^A = \lim_{n\to\infty}(I + \frac{A}{n})^n \tag{4.5.2}$$

where I *is the identity matrix.*

Equivalance of the two definition will be addressed in the Theorem on pp. 165.

Important theorems:

Theorem 4.3 (pp. 158)**.** *The series* e^A *converges for any* A *uniformly on each set* $X = \{A : ||A|| \leq a\}$, $a \in \mathbb{R}$.

Theorem 4.4 (pp. 160)**.**

$$e^{At} = H^t$$

where H^t *is the translation operator which sends every polynomial* $p(x)$ *into* $p(x+t)$.

Theorem 4.5 (pp. 163)**.**

$$\frac{d}{dt}e^{tA} = Ae^{tA}$$

Theorem 4.6 (Fundamental Theorem of the Theory of Linear Equations with Constant Coefficients)**.** *The solution of:*

$$\dot{\mathbf{x}} = A\mathbf{x} \tag{4.5.3}$$

with initial condition $\phi(0) = \mathbf{x}_0$ *is*

$$\phi(t) = e^{tA}\mathbf{x}_0 \tag{4.5.4}$$

Practically solution to

$$\dot{\mathbf{x}} = A\mathbf{x}$$

(pp. 173, Sec 17) (Assuming A is diagonalizable.)

- Find the eigenvectors $\xi_1, \cdots, \xi_n$ and eigenvalues $\lambda_1, \cdots, \lambda_n$. Use them as basis.
- Expand the initial condition in the new basis.

$$\mathbf{x}_0 = \sum_{k=1}^{n} C_k \xi_k \tag{4.5.5}$$

- Then $\phi(t) = \sum_{k=1}^{n} C_k e^{\lambda_k t} \xi_k$

4.6 Appearance of Gauge Structure in Simple Dynamical Systems

$$0 = (\eta_b, \dot{\eta_a}) = (\eta_b, \dot{U}_{ac}\psi) + (\eta_b, U_{ac}\dot{\psi}_c \tag{4.6.1}$$

4.7 Quantum Statistical Mechanics

Definition 4.3 (Time Evolution Operator)**.** *The time evolution operator $U(t, t_0)$ is defined such that*

$$|\Psi(t)\rangle = U(t, t_0)\, |\Psi(t_0)\rangle \tag{4.7.1}$$

It satisfy the relationship:

$$i\hbar \partial_t U(t, t_0) = HU(t, t_0) \tag{4.7.2}$$

This is obvious when substituting $U(t, t_0)$ into the Schrodinger Equations.

Quantum Macrostates Macrostates of the system depend on only a few the thermodynamic functions. We can form an ensemble of a large number $\mathcal{N}$ of microstates $\{\psi_\alpha\}$, corresponding too a given macrostates. The different microstates occur with probability p_α. When wen no longer have exact knowledge of the microstate of a system the system is said to be in a *mixed state*. The ensemble average of the quantum mechanical expectation value is given by:

$$\begin{aligned} \langle \bar{O} \rangle &= \sum_\alpha p_\alpha \langle \psi_\alpha | O | \psi_\alpha \rangle = \sum_{\alpha, m, n} p_\alpha \langle \psi_\alpha | m \rangle \langle m | O | n \rangle \langle n | \psi_\alpha \rangle \\ &= \sum_{m,n} \langle n | \rho | m \rangle \langle m | O | n \rangle = \mathrm{tr}(\rho O) \end{aligned} \tag{4.7.3}$$

where we have introduced the density matrix:

Definition 4.4 (Density Matrix)**.** *The density matrix $\rho(t)$ is defined as*

$$\langle n | \rho(t) | m \rangle \equiv \sum_\alpha p_\alpha \langle n | \psi_\alpha \rangle \langle \psi_\alpha | m \rangle \tag{4.7.4}$$

or

$$\rho(t) \equiv \sum_\alpha p_\alpha |\psi_\alpha\rangle \langle \psi_\alpha| \tag{4.7.5}$$

Density matrix is denoted by $\rho(t)$ by analogy of the notation for P.D.F, since ρ often represents density.

Density matrix satisfies several good properties:

- Normalized
- Hermiticity
- Positivity. For any Φ, $\langle\Phi|\rho|\Phi\rangle \geq 0$.

The time evolution of density matrix, directly obtained from Schrodinger's equation, is

$$i\hbar\frac{\partial}{\partial t}\rho = [H, \rho] \tag{4.7.6}$$

4.8 The Mathematical Theory of Communication

4.8.1 Introduction

What is information In this part, it is implied that *information* in this work does not carry the usual sense in people's daily life. The semantic aspect of a message is considered to be irrelavant, for purpose of generality of the design of communication systems.

Measure of information Then it refers to a paper by Hartley to substantiate the use of

$$S = \log(M) \tag{4.8.1}$$

as a measure of information. More specifically, we assume we have a set of possible messages. Then M is the cardinality of this set. Then S is a measure of the information produced when one message is chosen from the set. Once again, we regard all choices being equally possible.

Note that the base of logrithm in 4.8.1 is undefined. Choosing a base constituting choosing a unit of the measure. Two such measures, when calculated in different units, are related by a simple constant.

Conventionally, a base 2 is chosen. The resulted unit is called bits. If the base 10 is chosen, then the units may be called decimal digits. If the base e is chosen, then the units is called natural units.

Also, the author lists several points to illustrate the convenience of this measure.

Communication systems Next the author defines the necessary components of a *communication system*, and categorizes it into discrete systems, continuous systems and mixed systems.

4.8.2 Discrete Noiseless Systems

Discrete Noiseless Channel This part deals with another measure, the measure of the capacity of a channel to transmit information. It defines the capacity of a discrete channel as:

$$C \equiv \lim_{T\to\infty} \frac{\log N(T)}{T} \tag{4.8.2}$$

where $N(T)$ is the number of allowed signals of duration T. Several examples are given with formula for C in each particular example.

Discrete Source of Information Next, it proceeds to discuss the statistical property of the source of information. Pointing out that a statistical knowledge of the source of information can help peopole craft special protocals to reduce the required capacity of the channel, the article gradually focuses on the statistical property of sourcs. It profess that while a discrete source could be represented by a statistical source, a statistical process can also be considered a discrete source. The second claim is substatiated by several exmaples.

In one example of natural language, the article defines a *n-gram* case to produce natural language from statistical information.

Series of Approximations to English As the title suggests, this part illustrate two serial levels of steps to approximate the English language using statistical knowledge of appearance of alphabets (the first method) and words (the second example). The article claims that "a sufficiently complex stochastic process will give a satisfactory representation of a discrete source". Although I am largely against this juvenile view.

Graphical Representation of a Markoff Process Then the article mentions a graphical way to represent the aforementioned approximation process, and gives three examples on page 46.

Ergodic and Mixed Sources Now the article comes to a special type of stochastic process, ergodic processes. A rough idea of "ergodic" is given in page 45. The idea is so important that I felt compeled to present it here:

"In an ergodic process every sequence produced by the process is the same in statistical properties. Thus the letter frequencies, digram frequencies, etc., obtained from particular sequences, will, as the lengths of the sequences increase, approach definite limits independent of the particular sequence. Actually this is not true of every sequence but the set for which it is false has probability zero. Roughly the ergodic property means **statistical homogeneity**."

Next, the article claims that artificial languages given in previous examples are ergodic, because the corresponding graph does not have two properties: they does not comprise two or more *isolated parts*, and they *gcd* of the lengths of all *circuits* is one. The precise meaning is listed in page 47. Roughly, an analogy made by myself helps to catch the points. If we picture an stochastic process as a connected area, then isolated parts are its connected components, whereas the circuit are the recurrent pattterns.

Naturally , a stochastic process may exhibit a mixed behavir, in which there are several different sources $L_1, L_2, L_3, \cdots$, which are each of homogeneous, i.e. ergodic, statistical structure. This is discussed following the introduction of ergodicity in page 48.

Then the article declare that except in special cases, ergodicity is always assumed. This purpose is analogous to that of in statistical physics, to "identify averages along a sequence with averages over the ensemble of possible sequences", with "the probability of a discrepancy being zero".

Lastly, the article mentions a fact regarding the equilibrium of the system. A process is called stationary, if it satisfies a equilibrium condition:

$$P_j = \sum_i P_i \cdot P_i(j) \tag{4.8.3}$$

where P_j is the probability of being in state j, and $P_i(j)$ is the transition probability from i to j. The fact is that ergodic process is, in a sense, always stationary.

The Entropy of an Information Source This part first defines the entropy of a discrete source of finite state to be:

$$\begin{aligned} H &\equiv \sum_i P_i H_i \\ &= -\sum_{i,j} P_i p_i(j) \log(p_i(j)) \end{aligned} \tag{4.8.4}$$

It is "the entropy of the source per symbol of text". Another definition for entropy per second is also listed.

Following this definition are some theorems, which I consider to be the most essential and influential part of the whole book (although I have not yet read the whole book). They are:

Theorem 4.7 (Theorem 3 on page 55)**.** *Given any $\epsilon > 0$ and $\delta > 0$, we can find an N_0 such that the sequences of any legnth $N \geq N_0$ fall into two class:*

- *A set whose total probability is less than ϵ.*
- *The remainder, all of whose members have probabilities satisfying the inequality*

$$\left| \frac{\log p^{-1}}{N} - H \right| < \delta \tag{4.8.5}$$

"In other words we are almost certain to have $\frac{\log p^{-1}}{N}$ very close to H when N is large." For me, this reads quite like the *second law of thermaldynamics.*

4.9 Adjoint of antilinear operator

If we were to define adjoint of an antilinear operator, we might run into troubles with traditional definition. For example, suppose A is antilinear, let multiply it with $c\mathbb{1}$, where c is just some nonzero constant.

We have:

$$\langle \phi | c A \psi \rangle = c \langle \phi | A \psi \rangle \tag{4.9.1}$$

on the other hand:

$$\begin{aligned}\langle\phi|cA\psi\rangle &= \langle cA\psi|\phi\rangle^* = \langle\psi|A^\dagger c^*\phi\rangle^* \\ &= (c\,\langle\psi|A^\dagger\phi\rangle)^* = c^*\,\langle A\psi|\phi\rangle^* = c^*\,\langle\phi|A\psi\rangle \end{aligned} \tag{4.9.2}$$

A contracdiction. Then, it was proposed, in the first answer to this Math.SE post, that we define the adjoint operator differently for antilinear operator A:

$$\langle A^\dagger v|w\rangle = \langle v|Aw\rangle^* \tag{4.9.3}$$

Then the above discrepancy will not exist anymore (easily confirmed).

Here I also include the complete answer for reference:

I) First of all, one should never use the [Dirac bra-ket notation](http://en.wikipedia.org/wiki/Bra-ket_notation) (in its ultimate version where an operator acts to the right on kets and to the left on bras) to consider the definition of [adjointness](http://en.wikipedia.org/wiki/Adjoint_operator), since the notation was designed to make the adjointness property look like a mathematical triviality, which it is not. See also [this](http://physics.stackexchange.com/q/43069/2451) Phys.SE post.

II) OP's question(v1) about the existence of the adjoint of an [antilinear](http://en.wikipedia.org/wiki/Antilinear_map) operator is an interesting mathematical question, which is rarely treated in textbooks, because they usually start by assuming that operators are $\mathbb{C}$-linear.

III) Let us next recall the mathematical definition of the adjoint of a linear operator. Let there be a [Hilbert space](http://en.wikipedia.org/wiki/Hilbert_space) H over a [field](http://en.wikipedia.org/wiki/Field_%28mathematics%29) $\mathbb{F}$, which in principle could be either real or complex numbers, $\mathbb{F}=\mathbb{R}$ or $\mathbb{F}=\mathbb{C}$. Of course in quantum mechanics, $\mathbb{F}=\mathbb{C}$. In the complex case, we will use the standard physicist's convention that the [inner product/sequilinear form](http://en.wikipedia.org/wiki/Sesquilinear_form) $\langle\cdot|\cdot\rangle$ is conjugated $\mathbb{C}$-linear in the first entry, and $\mathbb{C}$-linear in the second entry.

Recall [Riesz' representation theorem](http://en.wikipedia.org/wiki/Riesz_representation_theorem): For each continuous $\mathbb{F}$-linear functional $f : H \to \mathbb{F}$ there exists a unique vector $u \in H$ such that

$$\mathrm{tag}\{1\} f(\cdot) \;=\; \langle u|\cdot\rangle.$$

Let $A : H \to H$ be a continuous[1] $\mathbb{F}$-linear operator. Let $v \in H$ be a vector. Consider the continuous $\mathbb{F}$-linear functional

$$\mathrm{tag}\{2\} f(\cdot) \;=\; \langle v|A(\cdot)\rangle.$$

The value $A^\dagger v \in H$ of the adjoint operator $A^\dagger$ at the vector $v \in H$ is by definition the unique vector $u \in H$, guaranteed by Riesz' representation theorem, such that

$$\mathrm{tag}\{3\} f(\cdot) \;=\; \langle u|\cdot\rangle.$$

In other words,

$$\text{tag}\{4\}\langle A^\dagger v|w\rangle \;=\; \langle u|w\rangle \;=\; f(w) = \langle v|Aw\rangle.$$

It is straightforward to check that the adjoint operator $A^\dagger : H \to H$ defined this way becomes an $\mathbb{F}$-linear operator as well.
IV) Finally, let us return to OP's question and consider the definition of the adjoint of an antilinear operator. The definition will rely on the complex version of Riesz' representation theorem. Let H be given a complex Hilbert space, and let $A : H \to H$ be an antilinear continuous operator. In this case, the above equations (2) and (4) should be replaced with

$$\text{tag}\{2'\}f(\cdot) \;=\; \overline{\langle v|A(\cdot)\rangle},$$

and

$$\text{tag}\{4'\}\langle A^\dagger v|w\rangle \;=\; \langle u|w\rangle \;=\; f(w) = \overline{\langle v|Aw\rangle},$$

respectively. Note that f is a $\mathbb{C}$-linear functional.
It is straightforward to check that the adjoint operator $A^\dagger : H \to H$ defined this way becomes an antilinear operator as well.
–
[1]We will ignore subtleties with [discontinuous/unbounded operators](http://en.wikipedia.org/wiki/Unbounded_operator), domains, [selfadjoint extensions](http://en.wikipedia.org/wiki/Extensions_of_symmetric_operators), etc., in this answer.

4.10 Bayes' theorem

I am quite unfamiliar with this formula, so I decided to make a note here for future reference.

This formula starts from the definition of conditional probability:

$$P(A|B) \equiv \frac{P(A\&B)}{P(B)} \tag{4.10.1}$$

this definition if quite intuitively pleasing if $P(B)$ is multiplied to the left side. Then one easily deduce that:

Theorem 4.8 (Bayes' theorem)**.**

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \tag{4.10.2}$$

We have also an extended form. Suppose the event space is partitioned into $\{A_i\}$, then we have (also easily proved):

$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{\sum_j P(B|A_j)P(A_j)} \tag{4.10.3}$$

Here's a good reading Bayes' Theorem in SEP.

4.11 About Matlab

On using A/B or A
B Matlab's matrix divide function `mrdivide`, `mldivide` , abbrevated A/B and A
B, are different and should be used with caution. One can check that:

$$A \cdot D^{-1} \text{ is equivalent to } A/D,\ D^{-1} \cdot A \text{ is equivalent to } D\backslash A.$$

Also, it is straightforward to prove that the two are equal if and only if A commutes with D.

5 Anchor

References

[1] Sakurai, J. J. Modern Quantum Mechanics, Addison Wesley.

[2] Bosonization and Strongly Correlated Systems. Cambridge. Cambridge Press Link

[3] Martin Braun. Differential Equations and Their Applications. 4ed. Springer.

[4] M. Kardar. Statistical Physics of Particles (2007)

[5] Existence of adjoint of an antilinear operator, time reversal

6 License