Mathematical QM - Lecture 3

Armin Scrinzi

December 3, 2019

Contents

Literature

The arguments in this lecture are taken from
Asher Peres, Quantum Theory: Concepts and Methods

This book, in my opinion, is a very nice example of "non-mathematical physics" containing intellectually accurate and honest discussions of the conceptional foundations of quantum mechanics.

1 Unitary groups and self-adjoint operators

Stone theorem provides a one-to-one correspondence between unitary groups and self-adjoint operators. Unitarity conserves norms (=probabilities) and scalar products (=transition amplitudes) and can therefore be interpreted in terms of experimental procedures: from the point of view of physics, it is the more fundamental concept. The prominent role of self-adjointness for observables is due to their correspondence to unitary groups.

Definition: Strongly continuous one-parameter unitary group
is a family of unitary operators $U(t)$, $t \in \mathbb{R}$ with the property

$$U(t_1 + t_2) = U(t_1)U(t_2) \tag{1}$$

where

$$t_n \to t \Rightarrow U(t_n) \stackrel{s}{\to} U(t), \tag{2}$$

in the sense of strong convergence.

Theorem 1. *Let A be self-adjoint and set $U(t) = \exp(-itA)$. Then*

1. *$U(t)$ is a strongly continuous one-parameter unitary group.*
2. *The limit $\lim_{t\to 0} \frac{1}{t}[U(t)\psi - \psi]$ exists if and only if $\psi \in D(A)$ and it holds*
$$\lim_{t\to 0} \frac{1}{t}[U(t)\psi - \psi] = -iA\psi \tag{3}$$
3. *The domain is invariant under $U(t) : U(t)D(A) = D(A)$.*

As unitary groups appear as the more fundamental objects, the inverse of the above is of utmost interest:

Theorem 2. *(Stone) Let $U(t)$ be a (strongly) continuous one-parameter unitary group. Then it has the form $U(t) = \exp(-itA)$ for some self-adjoint A.*

We can here also use weak continuity for $U(t)$: as the $U(t)$ are unitary, weak continuity implies strong continuity. (Proof can be found in the literature, we have not discussed some of the rather basic tools to do that proof.)

2 Symmetries and unitary groups

Our conventional representation of symmetries such as translation by some vector $a\hat{e}$

$$(U_a\psi)(\vec{r}) = \psi(\vec{r} - a\hat{e}) \tag{4}$$

or rotation by angle α around some axis $\hat{e}$

$$(U_\alpha\psi)(\vec{r}) = \psi(R^{-1}_{\alpha\hat{e}}(\vec{r})) \tag{5}$$

are trivially unitary and also strongly continuous. $R_{\vec{\alpha}}$ is the matrix for rotations in $\mathbb{R}^3$. The group's generators are the self-adjoint operators for momentum and angular momentum.

2.1 Unitary representation of symmetries

The concept of "symmetry" is not *a priori* tied to unitary transformations. Our original concept of symmetry is that we can transform a system $\psi \to \psi' = \mathcal{T}[\psi]$ and any measurement apparatus such that the outcome of transformed measurement on the transformed system does not change. As measurements are projections we can express them as vectors in the Hilbert space and the probability for finding some result a is $\langle\psi|a\rangle\langle a|\psi\rangle = |\langle a|\psi\rangle|^2$ for any pair of vectors. This motivates to define a symmetry transformation as any transformation that leaves the modulus of all scalar products invariant. Note that no linearity assumption is made here.

Interestingly, as was realized in the 1930'ies by Eugene Wigner, for QM in Hilbert space, the only maps that qualify for symmetry transformation are either — except for a multiplication by a phase — unitary or anti-unitary, the latter meaning a map $\overline{U}$ with

$$\langle \overline{U}x|\overline{U}y\rangle = \langle y|x\rangle \tag{6}$$

Theorem 3. *Wigner Let $\mathcal{T} : u \to u' = \mathcal{T}[u]$ be a map of the Hilbert space onto itself with the property $|\langle u'|v'\rangle| = |\langle u|v\rangle|$ then $\mathcal{T}$ has the form $\mathcal{T}[u] = \phi[u]Vu$, where V is either unitary or anti-unitary and $|\phi[u]| = 1$ is an arbitrary u-dependent phase factor.*

Note that $\phi[u]$ can be any silly function, including,

$$\phi[u] = 1 \text{ for } ||u|| < 1, e^{i\alpha} \text{ else.} \tag{7}$$

The idea of the proof is by direct examination of how a given basis is mapped. This works for finite Hilbert spaces as is, needs some caution when going to separable infinite Hilbert spaces (R. Simon et al. / Physics LettersA 378 (2014) 2332). For non-separable spaces different proofs are available (see, e.g., Gy.P. Geher / Physics LettersA 378 (2014) 2054). It appears that Wigner himself did not bother to give a rigorous proof of this, which was supplied only 32 years later by Bargman (according to Geher). So we can be at ease with citing the appealingly simple finite-dimensional version. (pages 218 and 219 of Peres)

8-2. Wigner's theorem

This invariance has far-reaching implications, because of an important theorem, due to Wigner.[3] Consider a mapping of Hilbert space: $\mathbf{u} \to \mathbf{u}'$, $\mathbf{v} \to \mathbf{v}'$, and so on. The *only* thing we assume about this mapping is that

$$|\langle \mathbf{u}', \mathbf{v}'\rangle|^2 = |\langle \mathbf{u}, \mathbf{v}\rangle|^2, \qquad \forall \quad \mathbf{u}, \mathbf{v}. \tag{8.4}$$

In particular, we do not assume linearity, let alone unitarity. Wigner's theorem states that it is possible to redefine the phases of the new vectors $(\mathbf{u}', \mathbf{v}', \ldots)$ in such a way that, for any complex coefficients α and β, we have either

$$(\alpha\,\mathbf{u} + \beta\,\mathbf{v})' = \alpha\,\mathbf{u}' + \beta\,\mathbf{v}' \qquad \text{and} \qquad \langle \mathbf{u}', \mathbf{v}'\rangle = \langle \mathbf{u}, \mathbf{v}\rangle, \tag{8.5}$$

or

$$(\alpha\,\mathbf{u} + \beta\,\mathbf{v})' = \overline{\alpha}\,\mathbf{u}' + \overline{\beta}\,\mathbf{v}' \qquad \text{and} \qquad \langle \mathbf{u}', \mathbf{v}'\rangle = \langle \mathbf{v}, \mathbf{u}\rangle. \tag{8.6}$$

[3] E. P. Wigner, *Group Theory,* Academic Press, New York (1959) p. 233.

Let e_j be the vectors of an orthonormal basis, which are mapped into e'_j. The new vectors e'_j are also orthonormal, by virtue of Eq. (8.4). Consider now the set of vectors

$$\mathbf{f}_j = \mathbf{e}_1 + \mathbf{e}_j \,, \qquad j = 2, 3, \ldots, \tag{8.7}$$

which are mapped into f'_j. We have, from (8.4),

$$|\langle \mathbf{e}'_1, \mathbf{f}'_j \rangle| = |\langle \mathbf{e}_1, \mathbf{f}_j \rangle| = 1, \tag{8.8}$$

orthoganality is conserved

and

$$|\langle \mathbf{e}'_j, \mathbf{f}'_k \rangle| = |\langle \mathbf{e}_j, \mathbf{f}_k \rangle| = \delta_{jk} \,, \qquad (j > 1). \tag{8.9}$$

Therefore, for any $j > 1$, we can write

$$\mathbf{f}'_j = x_j \, \mathbf{e}'_1 + y_j \, \mathbf{e}'_j \,, \tag{8.10}$$

i.e. composed of two orthonormal vectors

where $|x_j| = |y_j| = 1$. We then redefine the phases of the transformed vectors:

$$\mathbf{f}'_j \to \mathbf{f}''_j = \mathbf{f}'_j / x_j \qquad \text{and} \qquad \mathbf{e}'_j \to \mathbf{e}''_j = y_j \, \mathbf{e}'_j / x_j \,, \tag{8.11}$$

we change T to T'', but without changing the physics
we just remove some phases

(and $e''_1 = e'_1$) so as to obtain

$$\mathbf{f}''_j = \mathbf{e}''_1 + \mathbf{e}''_j \,, \tag{8.12}$$

as in (8.7). We shall henceforth work with the new phases, and write e'_j instead of e''_j. We thus have the mapping

$$\mathbf{e}_1 + \mathbf{e}_j \quad \to \quad (\mathbf{e}_1 + \mathbf{e}_j)' = \mathbf{e}'_1 + \mathbf{e}'_j \,. \tag{8.13}$$

starts to look like linearity...

Consider now the mapping of an arbitrary vector

$$\mathbf{u} = \sum_i a_i \, \mathbf{e}_i \quad \to \quad \mathbf{u}' = \sum_i a'_i \, \mathbf{e}'_i \,. \tag{8.14}$$

We have

$$|a'_j| = |\langle \mathbf{e}'_j, \mathbf{u}' \rangle| = |\langle \mathbf{e}_j, \mathbf{u} \rangle| = |a_j|. \tag{8.15}$$

by our definition of a symmetry

Moreover,

$$\langle \mathbf{e}_1 + \mathbf{e}_j, \mathbf{u} \rangle = a_1 + a_j \qquad \text{and} \qquad \langle \mathbf{e}'_1 + \mathbf{e}'_j, \mathbf{u}' \rangle = a'_1 + a'_j \,. \tag{8.16}$$

operating in each basis seaparately

It then follows from Eqs. (8.4) and (8.13) that

$$|a_1 + a_j|^2 = |a'_1 + a'_j|^2. \tag{8.17}$$

8.15 and 8.17 will only be compatible with (anti-)linear

Together with (8.15), this gives

$$\overline{a_1}\, a_j + a_1\, \overline{a_j} = \overline{a'_1}\, a'_j + a'_1\, \overline{a'_j}\,. \tag{8.18}$$

Dividing this equation by

$$(a_1\, \overline{a_1}\, a_j\, \overline{a_j}\,)^{1/2} = (a'_1\, \overline{a'_1}\, a'_j\, \overline{a'_j}\,)^{1/2}, \tag{8.19}$$

|a'j|^2 = |aj|^2 for all j

we obtain

$$(\overline{a_1}\, a_j / a_1 \overline{a_j}\,)^{1/2} + \text{c.c.} = (\overline{a'_1}\, a'_j / a'_1 \overline{a'_j}\,)^{1/2} + \text{c.c.}\,, \tag{8.20}$$

which has the form

$$e^{i\theta} + e^{-i\theta} = e^{i\theta'} + e^{-i\theta'}, \tag{8.21}$$

i.e. cosines are equal

with two solutions, $\theta' = \pm\,\theta$. Let us consider them one after another.

Unitary mapping: If $\theta' = \theta$, we have

$$\overline{a'_1}\, a'_j / a'_1 \overline{a'_j} = \overline{a_1}\, a_j / a_1 \overline{a_j}\,. \tag{8.22}$$

Redefine the phase of u' so that $a'_1 = a_1$. We then have $a'_j / \overline{a'_j} = a_j / \overline{a_j}$ and it follows from (8.15) that $a'_j = a_j$. Therefore

$$\mathbf{u}' = \sum_i a_i\, \mathbf{e}'_i\,. \tag{8.23}$$

again, we have altered the tranformation without affecting the physics

Given another vector, $\mathbf{v} = \sum b_i \mathbf{e}_i$, we can likewise choose the phase of v' so as to have $\mathbf{v}' = \sum b_i \mathbf{e}'_i$, whence Eq. (8.5) readily follows.

Antiunitary mapping: If $\theta' = -\theta$, we have

$$\overline{a'_1}\, a'_j / a'_1 \overline{a'_j} = a_1\, \overline{a_j} / \overline{a_1}\, a_j\,. \tag{8.24}$$

Redefine the phase of u' so that $a'_1 = \overline{a_1}$. We then have $a'_j / \overline{a'_j} = \overline{a_j} / a_j$ and it follows from (8.15) that $a'_j = \overline{a_j}$. Therefore

$$\mathbf{u}' = \sum_i \overline{a_i}\, \mathbf{e}'_i\,. \tag{8.25}$$

Given another vector, $\mathbf{v} = \sum b_i \mathbf{e}_i$, we can likewise choose the phase of v' so as to have $\mathbf{v}' = \sum \overline{b_i} \mathbf{e}'_i$, which gives Eq. (8.6).

3 Time-evolution

We "know" that time-evolution forms a unitary, strongly continuous group $U(t)$ and therefore there exists (conserved) energy as a self-adjoint operator. Norm-conservation is a requirement by construction, group-property implies the possibility to continue time-evolution from any point (or revert it), continuity is a commonplace requirement in physics: we do not want sudden changes. The missing linearity, however, is not an *a priori* idea about time-evolution. We might subject the map of time-evolution to the same conditions as spatial translation and then, from the time-invariance of physics, conclude that it must be a unitary map. However, this is not legitimate: ordinary symmetries like translation merely mean to change coordinates in one way or another. In contrast time-evolution may be connecting otherwise unrelated Hilbert spaces. Time is not an operator and does not appear as one of the coordinates in our Hilbert space on which we might re-label our system and apparatus. If we mean by time-invariance that experiments started at any time show the same evolution of amplitudes, a simple time-invariant, non-linear evolution will be given next.

Less arbitrary non-linear Hamiltonians have been proposed, but could be falsified in experiments: [C. G. Shull, D. K. Atwood, J. Arthur, and M. A. Horne, Phys. Rev. Lett. 44, 765 (1980)] and [J. J. Bollinger, D. J. Heinzen, Wayne M. Itano, S. L. Gilbert, and D. J. Wineland Phys. Rev. Lett. 63, 1031, 1989] (the prominent names on these papers show that the concerns are taken seriously).

It is then interesting to see that there is a very general argument against non-linear time-evolutions: accepting all other premises of QM, only unitary time-evolution is compatible with non-increasing entropy: norm-conserving, but non-unitary evolutions invariably include states with increasing entropy.

3.1 A norm-conserving non-linear time-evolution

We have no axiom that the time-evolution must be linear or unitary. Indeed, we can construct non-linear norm-conserving maps from the Hilbert space onto itself. For example, the time-evolution

$$i\frac{d}{dt}\begin{pmatrix}\alpha\\ \beta\end{pmatrix} = (\alpha^2-\beta^2)\begin{pmatrix}0 & -i\\ i & 0\end{pmatrix}\begin{pmatrix}\alpha\\ \beta\end{pmatrix} \tag{8}$$

is manifestly non-linear, but it conserves the norm. Note that this operator is also perfectly time-invariant, i.e. an evolution starting from a given state at any time will proceed in the same way.

On can easily construct solutions for (Peres):

$$\alpha/\beta = \tanh(t-t_0) \text{ for } |\alpha| < |\beta|$$
$$\alpha/\beta = \coth(t-t_0) \text{ for } |\alpha| > |\beta|$$

In the long run, all solutions collapse to equal amplitudes for both components. That is a funny property, as it turns any vector asymptotically into the vector

$$\begin{pmatrix}\alpha(t)\\ \beta(t)\end{pmatrix} \to \frac{1}{\sqrt{2}}\begin{pmatrix}1\\ 1\end{pmatrix}. \tag{9}$$

This means that, in the remote future, all information will collapse into a single vector. Loosely associating lack of knowledge with entropy, this cause trouble for the time-evolution by a growth of entropy. Slightly more precisely, any mixed state will end up pure as

$$\rho(t) \to \frac{1}{2}\begin{pmatrix}1 & 1\end{pmatrix}\begin{pmatrix}1\\ 1\end{pmatrix}. \tag{10}$$

This problem is general for non-linear time-evolutions.

3.2 Entropy

We had introduced general "states" on a C^* algebra of observables as the normalized positive functionals $\omega : \omega(B^*B) \geq 0$. In any Hilbert space representation states appear as *density matrices*: is a positive linear operator ρ with $\mathrm{Tr}\rho = 1$. (Note that this does not represent all states, but only a subset of "normal" states **definition to be added**). As ρ positive, we have have a spectral representation. As it has finite trace, it is a fortiori compact and therefore has a purely discrete spectrum with a spectral representation as

$$\rho = \sum_i |i\rangle \rho_i \langle i| \tag{11}$$

The physical interpretation is that we have incomplete knowledge about many copies of the same system and that ρ_i is the relative frequency for finding a system picked from the ensemble in the state $|i\rangle$. It is important to remember the difference to a superposition state

$$|s\rangle = \sum_i \sqrt{\rho_i}|i\rangle : \tag{12}$$

this describes a single system that is fully determined and whose density matrix is $\rho_s = |s\rangle\langle s|$. A more detailed discussion will be given later.

Entropy is a measure about the fuzziness of our knowledge about the system, i.e. the distribution of the ρ_i between the extremes (0 entropy) = (0 fuzziness) = (perfect knowledge) = $\rho_i = \delta_{ii_0}$ = pure state and (maximal entropy) = (maximal fuzziness) = (no nontrivial knowledge) = $\rho_i = \rho_j \forall i, j$

By general desired properties as well as motivated through a probability interpretation one arrives a the von Neumann - entropy

$$S(\rho) = -\mathrm{Tr}\rho \log \rho = -\sum_i \rho_i \log \rho_i \tag{13}$$

For our discussion of the time-evolution it is only important to observe that if the density matrix were to evolve from less pure to more pure, entropy decreases, i.e. any convex function of the density matrix would do.

3.3 Only unitary time-evolution respects the second law of thermodynamics

The time-evolution of ρ is given through the time-evolution of the $|i\rangle$, i.e.

$$\rho(t) = \sum_i |i,t\rangle \rho_i \langle i,t| = \sum_i U|i,0\rangle \rho_i \langle i,0|U^\dagger = U\rho U^\dagger. \tag{14}$$

The first form is more general, but the second two forms assume unitary time-evolution.

In general, unitary transforms are the ones that leave eigenvalues of a matrix and therefor the entropy invariant. Non-unitary mappings will in general change the eigenvalues (even if the trace is conserved). We restrict the reasoning here to the finite-dimensional case. As discussed, we cannot not *a priori* require conservation of $|\langle u(t)|v(t)\rangle|^2$. Instead we will need some extra input, which will be to require non-decreasing entropy. After that one can apply Wigner's theorem.

Let us assume there is any time-evolution $u(0) \to u(t)$ and $v(0) \to v(t)$ and let us assume we have a density matrix with maximal entropy

$$\rho = \frac{1}{2} P_u + \frac{1}{2} P_v, \tag{15}$$

whose time-evolution is, of course, just given by the time-evolution $u(t)$ and $v(t)$. At any time the eigenvalues of this rank 2 matrix can be easily evaluated. Remember, though, that we do not assume orthogonality of $|\langle u(t)|v(t)\rangle| = 0$, but we do assume conservation of probability $\langle u(t)|u(t)\rangle = \langle v(t)|v(t)\rangle = 1$. Write your solution as a linear combination of $|\psi\rangle = |u\rangle c_u + |v\rangle c_v$ to obtain

$$\begin{pmatrix} \frac{1}{2} & \frac{1}{2}\langle u|v\rangle \\ \frac{1}{2}\langle v|u\rangle & \frac{1}{2} \end{pmatrix} \vec{c} = w_\pm \vec{c}. \tag{16}$$

with the eigenvalues

$$w_\pm = \frac{1}{2}(1 \pm |\langle v|u\rangle|). \tag{17}$$

Any *increase* of $|\langle v|u\rangle|^2$ leads to an unphysical decrease of entropy. To see this, one can simply calculate the derivative of the entropy. It is a direct consequence of the convexity of $-x \log x$. However, it is a good exercise to look at the qualitative behavior: for identical states $u = v$, we have a pure state, $w_+ = 1, w_- = 0$ and the entropy is zero. As we decrease the overlap between the states, the difference $w_+ - w_- = 2|\langle v|u\rangle|$ decreases monotonically towards a more and more mixed state with a limiting value at orthogonal states $\langle u|v\rangle = 0$, with corresponding increase of entropy. (With a little practice in reading 2×2 matrices, we do not even need to calculate the eigenvalues: one "knows" that the off-diagonal matrix elements contribute to an further separation of the two eigenvalues of the matrix.) I.e. the orthogonal situation produces the maximal entropy. If the second law holds, orthogonal states $|\langle v(0)|u(0)\rangle|^2$ can never become non-orthogonal during a time-evolution. Orthogonal states remain orthogonal.

That means that we can choose a *time-dependent* orthonormal basis $u_k(t)$ such that for any arbitrary state v we have

$$\sum_k \langle v(t)|u_k(t)\rangle\langle u_k(t)|v(t)\rangle = 1 \quad \forall t. \tag{18}$$

Suppose there is some state for which the "$\cos^2$" $|\langle v(t)|u_m(t)\rangle|^2$ *decreases* (rather than remaining constant), then there must be at least one term $|\langle v(t)|u_n(t)\rangle|^2$ that increases. But then we could construct a mixed state from v and u_n whose entropy spontaneously and unphysically decreases.

It follows that the time-evolution must leave moduli of all scalar products invariant, it fulfills the condition for a (continuous) symmetry and by Wigner's theorem we know it can be considered unitary.