General Features of a QM Theory.

Postulates for a Relativistic QM

In seeking a relativistic generalization of QM, Dirac adopted the following features of nonrelativistic QM as <u>postulates</u> for the new theory:

$\boxed{1}$ All physical <u>observables</u> are represented by linear, Hermitian <u>operators</u>. E.g. $p_k \rightarrow (\hbar/i)\,\partial/\partial q_k$ is the k^{th} comp$^{\underline{t}}$ of (canonical) momentum for coordinate q_k.

$\boxed{2}$ There exists a <u>wavefcn</u> $\psi = \psi(q,s,t)$ which gives all possible information on the state of the quantum system, w/ t = time, q = coordinates $(\mathbf{r}, \mathbf{p})$ for the classical degrees of freedom, s = coordinates (spin, parity, etc) for additional degrees of freedom. $|\psi|^2 = \psi^*\psi \geqslant 0$ is finite, and is proportional to the <u>probability</u> of the system having coordinates q, s at time t.

$\boxed{3}$ The system is in an <u>eigenstate</u> ψ_n of an operator Ω if: $\Omega\psi_n = \omega_n\psi_n$, where ω_n = cnst is the $n^{\underline{th}}$ <u>eigenvalue</u> of Ω. If Ω is Hermitian, ω_n is a real #.

$\boxed{4}$ An arbitrary state ψ can be expanded as: $\underline{\psi = \sum_n a_n \psi_n}$, w/ the $\{\psi_n\}$ an orthonormal and complete set of eigenfcns† for the system (as defined by an appropriate set of commuting operators Ω_k). The <u>probability</u> that the system will be found in eigenstate ψ_n is $|a_n|^2$ (when ψ is normalized: $\langle\psi|\psi\rangle = 1$).

$\boxed{5}$ If $\psi = \sum_n a_n \psi_n$, with $\Omega\psi_n = \omega_n\psi_n$, then a <u>measurement</u> of the observable Ω in the state ψ will result in the eigenvalue ω_n with probability $|a_n|^2$. The <u>average</u> of a large number of measurements of Ω for state ψ will be:

$$\boxed{\langle\Omega\rangle = \sum_s \int dq\, \psi^*(q,s,t)\,\Omega\,\psi(q,s,t) = \sum_n \omega_n |a_n|^2.} \quad (1)$$

(next page)

---

† "Orthonormal" means: $\sum_s \int dq\, \psi_m^*(q,s,t)\,\psi_n(q,s,t) = \delta_{mn}$.

"Complete" means: $\sum_n \psi_n^*(q',s',t)\,\psi_n(q,s,t) = \delta_{ss'}\,\delta(q-q')$.

Dirac's linearity postulate. Remarks on Lorentz notation used here.

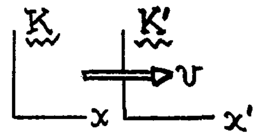
[6] The equation-of-motion for the system's state fcn ψ is of form: $i\hbar\frac{\partial\psi}{\partial t}=\mathcal{H}\psi$, where the system Hamiltonian $\mathcal{H}$ is a <u>linear</u>, Hermitian operator. Requiring $\mathcal{H}$ to be linear preserves a superposition principle for ψ (see [4] above). And a Hermitian $\mathcal{H}$ allows conservation of total probability, as...

$$\frac{\partial}{\partial t}\left[\sum_s\int dq\,\psi^*\psi\right]=\sum_s\int dq\left[(\partial\psi^*/\partial t)\psi+\psi^*(\partial\psi/\partial t)\right]$$

$$=\frac{i}{\hbar}\sum_s\int dq\left[(\mathcal{H}\psi)^*\psi-\psi^*(\mathcal{H}\psi)\right]=0. \qquad (2)$$

Dirac's postulates [1]-[6] contain the essential features of any "reasonable" theory of QM. As we shall see, satisfying #[6] is the principal challenge in constructing a relativistic QM theory.

ASIDE Lorentz notation for following notes.



1) For a Lorentz transformation Λ between cd systems K & K', i.e. for the position 4-vector $x\rightarrow x'=\Lambda x$, we adopt the following <u>conventions</u>:

1. Greek indices run from 1 to 4; Roman indices from 1 to 3.
2. Contravariant & covariant notation is <u>not</u> used. The 4-vector position is just $x=(x_\mu)=(x_1,x_2,x_3,x_4)$, and x^μ does <u>not</u> appear in the theory.
3. Sum over repeated indices, i.e. $x_\mu^2=x_\mu x_\mu=\sum_{\mu=1}^{4}x_\mu^2$.
4. Choose metric $(g_{\mu\nu})\equiv 1$. Then 4-vectors have <u>imaginary</u> 4th (timelike) components. E.g.: $x_\mu=(x_1,x_2,x_3,ict)$, w/ $(x_1,x_2,x_3)=\mathbf{r}$, and Minkowski length $x_\mu^2=r^2-(ct)^2$. Generally: $A_\mu B_\mu=A_k B_k-|A_4||B_4|$. (3)

These conventions are "old-fashioned" [comp. w/ Jackson's "Classical Electrodynamics" (Wiley, 2nd ed., 1975), where $(g_{\mu\nu})=\begin{pmatrix}1 & & 0\\ & -1 & \\ 0 & & -1 & \\ & & & -1\end{pmatrix}$, $A_\mu B^\mu=A_0B^0-A_kB^k$, etc.], but they are the conventions used by Davydov in his Ch. VIII on relativistic QM, and by Sakurai in his text "Advanced QM" (Addison-Wesley, 1967). Saku-

$\boxed{\text{ASIDE}}$ **Lorentz notation (cont'd)**

rai is emphatic in pointing out (see p. 6) that "these complications (viz. a metric tensor $g_{\mu\nu} \neq 1$, and contra & covariant vector notation) are absolutely unnecessary in the special theory of relativity."

2) Except for the appearance of $i = \sqrt{-1}$, the general properties of Lorentz transforms do not change with the conventions in Eq. (3) above. E.g.

A. Lorentz transform is : $x \to x' = \Lambda x$, i.e. $x'_\mu = \Lambda_{\mu\nu} x_\nu$, $(\Lambda_{\mu\nu}) = 4\times4$ matrix.

B. x_μ^2 is invariant $\leftrightarrow$ $\Lambda_{\mu\lambda}\Lambda_{\mu\nu} = \delta_{\mu\nu}$; $\Lambda_{\mu\nu}$ is an orthogonal matrix.

C. Λ^{-1} (inverse) $= \tilde{\Lambda}$ (transpose), i.e. $(\Lambda^{-1})_{\mu\nu} = (\tilde{\Lambda})_{\mu\nu} = \Lambda_{\nu\mu}$.

D. $\det\Lambda = \pm1$. Only $\det\Lambda = +1$ "proper" transforms are commonly used.¶

E. Since $x'_k = \Lambda_{k\mu} x_\mu$ must be pure real, while $x'_4 = \Lambda_{4\mu} x_\mu$ is pure imaginary, then the elements $\Lambda_{k\ell}$ & Λ_{44} are real, while Λ_{k4} & Λ_{4k} are imaginary.

i.e. $\left(\begin{array}{c|c} \Lambda_{k\ell} \text{ real} & \Lambda_{4k} \text{ imag.} \\ \hline \Lambda_{k4} \text{ imag.} & \Lambda_{44} \text{ real} \end{array}\right) = \Lambda$, in general. $\left.\begin{array}{l}\text{For velocity} \\ \text{boost along} \\ x_3\text{-axis}\end{array}\right\}$ $\Lambda = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \gamma & +i\beta\gamma \\ 0 & 0 & -i\beta\gamma & \gamma \end{pmatrix}$. (4)

F. Any vector A_μ that transforms like x_μ, i.e. $A_\mu \to A'_\mu = \Lambda_{\mu\nu} A_\nu$, is a 4-vector, and it has an invariant length : $A_\mu^2 = A_k^2 - |A_4|^2$. As well, the scalar product between two such 4-vectors : $A_\mu B_\mu = A_k B_k - |A_4||B_4|$ is invariant.

G. The gradient operator $\partial/\partial x_\mu = (\partial/\partial x_1, \partial/\partial x_2, \partial/\partial x_3, -\frac{i}{c}\partial/\partial t)$ is a 4-vector, since : $\partial/\partial x'_\mu = (\partial x_\nu/\partial x'_\mu)\,\partial/\partial x_\nu = \Lambda^{-1}_{\nu\mu}\,\partial/\partial x_\nu = \Lambda_{\mu\nu}\,\partial/\partial x_\nu$, transforms properly. The wave operator : $(\partial/\partial x_\mu)^2 = \nabla^2 - \frac{1}{c^2}(\partial/\partial t)^2$, is Lorentz invariant.

Etc. None of the content of special relativity is changed, just the notation is different. We will use this new (old) notation as a diversion. We remark that even Jackson used this notation in the first edition (1962) of his text.

¶ This rules out space & time inversions ($\mathbf{r} \to (-)\mathbf{r}$ & $t \to (-)t$) carried by Λ itself.

Dirac's form for a relativistic Hamiltonian.

The Dirac Equation: Derivation & Basic Properties

1) Recall (from p. fs 16 of class notes) that the Klein-Gordon Eqtn generated a probability density $\rho = -(\hbar/mc^2)\,\mathrm{Im}[\psi^*(\partial\psi/\partial t)]$ which could be either $\pm$ ve, and which needed initial values of both $\psi(t_0)$ and $\dot{\psi}(t_0)$ to fix its evolution. These difficulties are inevitably connected with using a wave equation like $(i\hbar\,\partial/\partial t)^2\psi = \mathcal{H}^2\psi$, quadratic in both $\partial/\partial t$ and the Hamiltonian $\mathcal{H}$. To avoid such difficulties, Dirac sought to write a wave equation that was linear in both the time & space operators, i.e. a Schrödinger-like form:

$$\underline{(i\hbar\,\partial/\partial t)\psi = \mathcal{H}\psi}, \quad \text{w/ } \mathcal{H} \text{ linear in space derivatives } \partial/\partial x_k. \tag{5}$$

Such an equation determines $\psi(t)$ from $\psi(t_0)$ alone $\left(\psi(t) = [e^{-\frac{i}{\hbar}(t-t_0)\mathcal{H}}]\psi(t_0)\right)$, and also permits a superposition principle similar to that of Schrödinger theory.

On grounds that in a relativistic theory time & space coordinates must occur on an equal footing, the Dirac $\mathcal{H}$ in Eq. (5) must be linear in the particle's momentum $\mathbf{p} = (p_k) = (-i\hbar\,\partial/\partial x_k)$. Also, $\mathcal{H}$ must recover the particle's rest energy mc^2 when $\mathbf{p} = 0$ (and in absence of external fields). Then for a free particle, $\mathcal{H}$ must be of the general form....

$$\left[\begin{array}{l} \mathcal{H} = c\boldsymbol{\alpha}\cdot\mathbf{p} + \beta mc^2 = c\alpha_k p_k + \beta mc^2, \\ \text{w/ } \boldsymbol{\alpha} = (\alpha_1, \alpha_2, \alpha_3) \;\&\; \beta, \text{ are 4 quantities } \int \begin{array}{l}\text{dimensionless, and}\\ \text{independent of } \mathbf{r} \,\&\, t.\end{array}\end{array}\right. \tag{6}$$

---

2) Conditions specifying α & β follow from requiring that for a stationary state of energy E, the particle obeys the relativistic energy-momentum relation:

$$\left[\begin{array}{l} \psi(\mathbf{r},t) = [e^{-\frac{i}{\hbar}t\mathcal{H}}]\psi(\mathbf{r},0) = \psi(\mathbf{r},0)\,e^{-\frac{i}{\hbar}Et}, \text{ for energy } E \\ \text{w/ } E^2 = c^2\mathbf{p}^2 + (mc^2)^2 = \mathcal{H}^2, \\ \text{i.e./ } c^2(p_k^2 + (mc)^2) = c^2(\alpha_k p_k + \beta mc)^2. \end{array}\right. \tag{7}$$

Representation of Dirac's matrices (α, β).

Consider Eq. (7) to be an identity and work out details...

$$\rightarrow p_k^2 + (mc)^2 = \alpha_k^2 p_k^2 + \beta^2 (mc)^2 + \begin{cases} (\alpha_1\alpha_2 + \alpha_2\alpha_1) p_1 p_2 + (\alpha_1\alpha_3 + \alpha_3\alpha_1) p_1 p_3 + \\ (\alpha_2\alpha_3 + \alpha_3\alpha_2) p_2 p_3 + \sum_k (\alpha_k\beta + \beta\alpha_k) mc p_k \end{cases}$$

So//

$$\left.\begin{array}{l} \alpha_k^2 = \beta^2 = 1, \text{ for } k = 1,2,3; \\ \alpha_k\alpha_l + \alpha_l\alpha_k = 0, \; l \neq k; \;\; \alpha_k\beta + \beta\alpha_k = 0, \text{ for } k = 1,2,3. \end{array}\right\} \quad (8)$$

More succinctly, define $\alpha_4 = \beta$, and write Eqs. (8) as

$$\underline{\underline{\alpha_\mu\alpha_\nu + \alpha_\nu\alpha_\mu = 2\delta_{\mu\nu}}}, \quad \text{for } \mu \,\&\, \nu = 1,2,3,4. \quad (9)$$

Since the α_μ anti-commute, they must be matrices, at least. And, so that $\mathcal{H}$ in Eq. (6) be Hermitian (with $\mathbf{p} = -i\hbar\nabla$ already so), the α_μ must be <u>Hermitian matrices</u>. The anti-commutation rule of Eq. (9) reminds us of the rule for the Pauli matrices for spin ½, i.e.

$$\left[\begin{array}{l} \sigma_k\sigma_l + \sigma_l\sigma_k = 2\delta_{kl}, \quad \text{for } k \,\&\, l = 1,2,3; \\ \text{w/}\;\; \boldsymbol{\sigma} = (\sigma_k) = \left\{ \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \right\}. \end{array}\right. \quad (10)$$

So try: $(\alpha_k, \beta) = (\sigma_k, 1)$. This doesn't work because for $\beta = 1$, we cannot satisfy $\alpha_k\beta + \beta\alpha_k = 0$ [per Eq. (8)] unless $\alpha_k \equiv 0$. The simplest choice of four independent α_μ which obey the rule of Eq. (9) is...

$$\underline{\underline{\boldsymbol{\alpha} = \begin{pmatrix} 0 & \boldsymbol{\sigma} \\ \boldsymbol{\sigma} & 0 \end{pmatrix}, \quad \alpha_4 = \beta = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}}} \quad \begin{array}{l} \boldsymbol{\sigma} = 2\times 2 \text{ Pauli matrices}, \\ 1 = 2\times 2 \text{ identity } \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \end{array} \quad (11)$$

This is Dirac's original representation of the α_μ; it is not unique.

---

3) Among other things, the fact that Dirac's matrices α_μ are 4×4 means that Dirac's wavefcn ψ must have 4 components... it is a 4-component "<u>spinor</u>" $\psi = (\psi_1, \psi_2, \psi_3, \psi_4)$, usually written as a column. <u>NOTE</u>: A 4-comp.