Quantum measurement theory for particle oscillations

Charis Anastopoulos
Department of Physics, University of Patras, 26500 Patras, Greece

Ntina Savvidou
Theoretical Physics Group, Imperial College, SW7 2BZ, London, UK

A fundamental principle of quantum theory, clearly manifested in the two-slit experiment, is that for any alternatives that cannot be distinguished by measurement physical predictions are obtained by summation of their amplitudes. In particle oscillation experiments, a particle’s time of detection is not directly measured, consequently, the detection probability should involve the summation over amplitudes corresponding to different detection times. However, in contrast to the principle above, standard treatments involve summation over probabilities rather than amplitudes; this implicitly assumes the existence of a decohering mechanism. In this work, we construct the detection probabilities for particle oscillations by summation over amplitudes, corresponding to different detection times. The resulting wavelength of particle oscillations differs from the standard expression by a factor of two. Moreover, we predict a dependence of the oscillation wavelength on the threshold of the decay process used for detection.

PACS numbers: 03.65.Ta, 14.60.Pq, 03.65.Xp

The two-slit and related experiments demonstrate a striking feature of quantum mechanics: for any two alternatives that cannot be distinguished by a measurement scheme, their joint probability is obtained by summing their contributions at the level of probability amplitudes; therefore interference terms appear. If, however, a macroscopic distinction of these alternatives is possible, the interference terms are suppressed and the joint probability is the sum of the individual probabilities.

Let us examine the importance of the above elementary point for the correct quantum treatment of particle oscillations [1]. The detection time of the oscillating particles is not directly measured. Instead, we measure the number of detection events, at distance $L$ from the source. The question then arises: how, starting from the probability amplitude $A_\alpha(t, x)$—for the detection of a flavor $\alpha$, at a point $x$, at sharply defined moment of time $t$—, we can obtain the probability $p_\alpha(L)$ that a particle of flavor $\alpha$ will be detected at distance $L$ from the source.

A common approach is to substitute in $A_\alpha(t, x)$ the time $t$ with the classical time of arrival at $L$, namely $L/c$ for neutrinos, and to write the corresponding probability as $p_\alpha(L) \sim |A_\alpha(L, L)|^2$. The usual setting for this approach is the textbook plane-wave treatment of neutrino oscillations, where

$$A_\alpha(t, x) = \sum_i U_{\beta\alpha}^* U_{\alpha i} e^{ip_i x - iE_i t}. \tag{1}$$

Here, $i$ labels the mass eigenstates of neutrinos, $U_{\beta\alpha}$ is the mixing matrix, $\beta$ is the initial flavor, $p_i$ are the momenta in the different mass-eigenstates and the corresponding energies $E_i = \sqrt{m_i^2 + p_i^2}$. The substitution $t = L/c$ in Eq. (1) and the evaluation of the detection probability yields the standard expression for the oscillation wavenumber $k_{ij} = (m_i^2 - m_j^2)/(2p)$, for $p_i = p + O(m_i)$.

This method involves an unsatisfactory mixing of classical and quantum concepts and this problem persists if we consider realistic wave packets rather than plane waves. More importantly, the method is ambiguous in the following way: different mass eigenstates correspond to different velocities $v_i = p_i/E_i$; if we evaluate each mass-component at the corresponding arrival times $t_i = L/v_i$, then we obtain an oscillation wave-number $k_{ij} = (m_i^2 - m_j^2)/p$ for $p_i = p + O(m_i)$, which is twice the standard result [11].

Furthermore, the standard method assumes that the detection time should be computed in terms of an ‘average’ velocity for all components of the wave packet—see, for example, Refs. [1, 3]. This assumption involves the substitution of a coherent superposition by an altogether different state that cannot be done in an invariant way in quantum mechanics. Any choice of an ‘average velocity’ for a wave packet is arbitrary and unjustified at a fundamental level.

An improved derivation of the standard expression for the oscillation wavelength proceeds by calculating the time average of probabilities over detection time, i.e., defining $p_\alpha(L) \sim \int_0^T dt |A_\alpha(t, L)|^2$, where $T$ is the integration time for the experiment. However, a summation of probabilities over detection time is not justified by the rules of quantum theory. As the two-slit experiment indicates, alternatives that cannot be distinguished by measurement—in this case, detections of a flavor $\alpha$ at different moments of time—are to be summed over at the level of amplitudes.

The summation over amplitudes, defined at different moments of time, is rather intricate because there is no time operator in quantum theory. Hence the methods that apply to other observables are not directly applicable here. The method we present in this letter has
the following advantages: (i) it fully implements the basic principle, that alternatives that are not distinguished by measurement are to be summed at the level of amplitudes; (ii) it incorporates a genuinely quantum treatment of detection time, according to the quantum theory of measurement \( \hat{A} \); (iii) it employs a general formula, Eq. (3), that is derived using only the rules of quantum theory. Eq. (3) is valid for any setup in which the detection time is not measurable, not only to particle oscillations. The end result is an expression for the detection probability as a function of distance; it leads to the non-standard formula for the oscillation wave-number \( k_{ij} = (m_i^2 - m_j^2)/p \) that was mentioned previously.

In what follows, we describe in more detail our approach and the derivation of the results above. We employ a method developed in [2], for the construction of probabilities from amplitudes that are defined at different moments of time. It has been applied to various problems, such as probabilities for time-of-arrival, tunneling-time and for non-exponential decays. This method contains ideas from the decoherent histories approach to quantum mechanics [8] and it has many similarities to the Srinivas-Davies photo-detection theory [3].

The first step in our study is the derivation for a general formula for the probability of detection outcomes, when the time of detection is not observable. (The presentation here is simplified: for elaboration on the finer points, see Ref. [2].) Let \( \mathcal{H} \) be the Hilbert space of a quantum system, and \( \hat{H} \) the Hamiltonian operator. In order to describe an event—such as a particle detection—we introduce a projection operator \( \hat{P} \) that corresponds to states accessible only if the event has taken place. For example, if we detect a particle by monitoring a specific decay process, then \( \hat{P} \) is a projector onto the states of all product particles.

The Hilbert space \( \mathcal{H} \) splits into two subspaces corresponding to the projectors \( \hat{P} \) and \( \hat{Q} = 1 - \hat{P} \). Let \( \hat{P}_\lambda \) be the projection operators corresponding to different values \( \lambda \) of an observable that can be measured only if a detection event has occurred. For example, \( \hat{P}_\lambda \) may correspond to a coarse-grained position variable for a product particle. The set of projectors \( \hat{P}_\lambda \) is exclusive (\( \hat{P}_\lambda \hat{P}_\nu = 0 \) if \( \lambda \neq \nu \)) and exhaustive, provided a detection has occurred; i.e., \( \sum \hat{P}_\lambda = \hat{P} \).

Starting from the initial state of a quantum system \( |\psi_0 \rangle \), we construct the amplitude \( |\psi; t, \lambda \rangle \), at a final time \( T \), that arises if a detection event took place within the interval \( [t, t + \delta t] \subset [0, T] \) and the outcome of the measurement was \( \lambda \). By assumption, no detection took place before time \( t \). Hence, we evolve \( |\psi_0 \rangle \) with \( \hat{S}_t = \lim_{N \to \infty} (Q e^{-i\hat{H}_t/N} Q)^N \), i.e., with the restricted propagator in the eigenspace of the projector \( Q \) to no-detection states. In the interval \( [t, t + \delta t] \) the state evolves according to the full unitary evolution \( e^{-i\hat{H} \delta t} \). Then, we act by a projection \( \hat{P}_\lambda \) that selects the measurement outcome \( \lambda \). Next, the state evolves unitarily until time \( T \) because there is no other event that needs to be taken into account. At the limit of small \( \delta t \), the successive operations above yield \( |\psi; t, \lambda \rangle = -i \delta t e^{-i\hat{H}(T-t)} \hat{P}_\lambda \hat{H} \hat{S}_t |\psi_0 \rangle \).

We must emphasize here an important physical distinction on the role of time pertaining to particle oscillations. The time of detection \( t \) is not identical to the evolution parameter of Schrödinger’s equation. Instead, it is a dynamical variable that determines the moment that a physical event has taken place [10]. The construction of the amplitude \( |\psi; t, \lambda \rangle \) above highlights this distinction: the detection time \( t \) is distinct from the time \( T \) at which the amplitude is evaluated.

Furthermore, the amplitude \( |\psi; t, \lambda \rangle \) is proportional to \( \delta t \), hence it defines a density with respect to time. The integration of \( |\psi; t, \lambda \rangle \) over \( t \) is, therefore, well-defined in this scheme. In contrast, quantum theory provides no natural definition of integration of single-time probabilities over time [7]. The amplitude \( |\psi; \lambda \rangle \) corresponding to a detection with value \( \lambda \), at any (unspecified) time \( t \in [0, T] \), is

\[
|\psi; \lambda \rangle = -i \int_0^T dt \ e^{-i\hat{H}(T-t)} \hat{P}_\lambda \hat{H} \hat{S}_t |\psi_0 \rangle := \hat{C}_\lambda |\psi_0 \rangle. \tag{2}
\]

Hence, the probability \( p(\lambda) \) that a detection with outcome \( \lambda \) occurred at some time in \( [0, T] \) is

\[
p(\lambda) = \langle \psi; \lambda | \psi; \lambda \rangle = \langle \psi_0 | \hat{C}_\lambda^\dagger \hat{C}_\lambda |\psi_0 \rangle = \int_0^T dt \int_0^T dt' \ Tr(e^{i\hat{H}(t-t')} \hat{P}_\lambda \hat{H} \hat{S}_t^\dagger \hat{P}_0 \hat{S}_{t'} \hat{H} \hat{P}_\lambda), \tag{3}
\]

where \( \hat{P}_0 = |\psi_0 \rangle \langle \psi_0 | \). The probability measure Eq. (3) is positive, linear, and normalized when the probability of no-detection \( 1 - \int_0^T dt \ p(\lambda) \) is included.

If the Hamiltonian is of the form \( \hat{H} = \hat{H}_0 + \hat{H}_I \), where \( [\hat{H}_0, \hat{Q}] = 0 \) and \( \hat{H}_I \) a small perturbation, then, to leading order in the perturbation,

\[
p(\lambda) = \int_0^T dt \int_0^T dt' Tr(\hat{U}_{t'-t} \hat{P}_\lambda \hat{H}_I \hat{U}_t \hat{P}_0 \hat{S}_{t'} \hat{H} \hat{P}_\lambda), \tag{4}
\]

where \( \hat{U}_t = e^{-i\hat{H}_0 t} \).

Eqs. (3)(4) are general operator expressions valid for any system. Next we apply them to the case of particle oscillations. Since the oscillating particles are detected by means of their decay products, the precise treatment of the detection process involves the use of quantum field theory. We will not specify the type of oscillating particles so that our results will be valid for both neutrino and neutral boson oscillations.

To this end, let us denote the oscillating particles as \( A \) and consider that their detection involves the process \( A + B_m \to D_n \), where \( B_m \) and \( D_n \) are particles labeled by indices \( m \) and \( n \). The Hilbert space \( \mathcal{H} \) of the total system is a tensor product \( \mathcal{H}_A \otimes \mathcal{H}_r \), where \( \mathcal{H}_A \) is the
(bosonic or fermionic) Fock space $\mathcal{F}(H_{1A})$. The single-particle Hilbert space $H_{1A}$ is a direct sum $\bigoplus_i H_i$ of mass eigenspaces $H_i$. $H_r$ is the Hilbert space for the degrees of freedom corresponding to the $B_m$ and $D_n$ particles. It is decomposed as $H_0 \otimes H_{\text{prod}}$: $H_0$ is the subspace of states prior to the decay $A + B_m \rightarrow D_n$ and $H_{\text{prod}}$ is the subspace corresponding to states of the decay products [12]. Since we assume that the measurements are carried out to the product particles, then the projection operators $P_\lambda$ in Eq. (3) are of the form $1 \otimes \Pi_\lambda$, where $\Pi_\lambda$ projects into a subspace of $H_{\text{prod}}$.

We assume a Hamiltonian of the form $\hat{H} = \hat{H}_A \otimes 1 + 1 \otimes \hat{H}_r + \hat{H}_I$, where $\hat{H}_A$ is the Hamiltonian for the $A$ particles, $\hat{H}_r$ is the Hamiltonian for the $B_m$ and $D_n$ particles, and $\hat{H}_I$ is the interaction Hamiltonian. For simplicity, we assume that any particles $B_m$ that are present prior to detection are stationary. Hence, the restriction of $\hat{H}_r$ to $H_0$ is a constant, which can be taken equal to zero. The restriction of $\hat{H}_r$ on $H_{\text{prod}}$ equals $\epsilon_{rh} + \sum_i (\sqrt{M_{Bm}^2 + \mathbf{p}_n^2} - M_{Dn})$, where $\epsilon_{rh} = \sum_i M_{Dn} - \sum_i M_{Bm}$ is the threshold energy of the $A$ particles for the process $A + B_m \rightarrow D_n$. In the above, $M_{Bm}$ and $M_{Dn}$ are the masses of the particles $B_m$ and $D_n$, respectively and $\mathbf{p}_n$ are the momentum operators for the $D_n$ particles.

We consider an effective interaction Hamiltonian

$$\hat{H}_I = \sum_i \int d^3x \left[ b_i(x) \hat{J}_i^+(x) + b_i^\dagger(x) \hat{U}_{i\alpha} \hat{J}_\alpha^+(x) \right],$$

where $b_i, b_i^\dagger$ are annihilation and creator operators on $H_A$, $i$ labels mass eigenstates, $\hat{J}_i^+$ ($\hat{J}_i$) are current operators of flavor $\alpha$ defined on $H_r$, and $U_{i\alpha}$ is the mixing matrix [12]. The current operator $\hat{J}_\alpha^+$ involves products of annihilation operators for the $B$ particles and creation operators for the $D$ particles. Since no $A$ particles are created during the detection process, the initial state $|\phi_0\rangle$ in $H_r$ must satisfy $\hat{J}_\alpha^+(x)|\phi_0\rangle = 0$. Note that if the detection involves a scattering process rather than a decay, an interaction Hamiltonian quadratic to the field of the $A$ particles should be used, instead of (5).

Let the initial state on $H_A$ be a single-particle state $|\psi_0\rangle = \sum_i \int d^3x \langle b_i(x)|\psi_0(x)\rangle |0\rangle_A$, where $|0\rangle_A$ is the vacuum of $H_A$. Since $[\hat{H}_0, \hat{Q}] = 0$, Eq. (4) applies. Hence, the probability that a decay through flavor $\alpha$ has happened at some time in $[0, T]$ and that the value $\lambda$ for an observable of the product particles has been found equals

$$p_\alpha(\lambda) = \sum_{ij} \int_0^T dt \int_0^T dt' \int d^3x d^3x' \langle x', t' | \psi_i^*(x, t) | \chi_i \rangle \langle \chi_i | \psi_j(x', t) | \lambda \rangle \times U_{\alpha\bar{\alpha}} U_{j\alpha}^* R_\lambda^\alpha(x, x', t - t'),$$

where $\psi_i(x, t)$ is the evolution of $\psi_0(x)$ under the Hamiltonian for a single $A$-particle and

$$R_\lambda^\alpha(x, x', t - t') = \langle \phi_0 | \hat{J}_\alpha^+(x') \hat{P}_\lambda e^{-i \hat{H}_r (t' - t)} \hat{P}_\lambda \hat{J}_\alpha^+(x) | \phi_0 \rangle,$$

We next consider the measurement of position $X$, of one of the product particles, with accuracy of order $\delta$. The operators $\hat{P}_\lambda$ in (7) can be substituted by a Gaussian approximate projector for position

$$\hat{P}_X = \int d^3x e^{-\frac{\langle X^2 \rangle}{2\sigma^2}} |X\rangle \langle X| \otimes 1,$$

where the tensor product with unity refers to the remaining degrees of freedom in $H_{\text{prod}}$.

The vector $\hat{J}_\alpha^+(x)|\phi_0\rangle$ refers to the state of the product particles for decays that have taken place in a neighborhood of $x$. Hence, the operators $\hat{P}_X$ determine the locus of the decay event within an accuracy of order $\delta$, if $\hat{P}_X \hat{J}_\alpha^+(x)|\phi_0\rangle \approx 0$, for $|x - \bar{x}| \gg \delta$. Heuristically, the condition above would be satisfied for macroscopic values of $\delta$ much larger than any length parameters characterizing the interaction. At this level of coarse-graining, the localization of one product particle at the decay time essentially determines the localization of all other product particles. In general, the localization scale $\delta$ is macroscopic, but it has to be much smaller than the scale of variation in the wave function of the $A$ particles, or else no particle oscillations would be observable. Hence, to a first approximation, we can substitute $\psi_i(x, t)$ by $\psi_i(x, t)$ in Eq. (4). Taking these considerations into account, Eq. (7) becomes

$$R_\lambda^\alpha(x, x', t - t') \approx K \delta^3(X, x) \delta^3(X, x') e^{i \epsilon_{rh}(t' - t)} F(t' - t),$$

where $K > 0$ is a constant. The function $F(s)$ is obtained from the propagator of $\hat{H}_r$ and it equals $\prod_n \int d^3p e^{-i(\sqrt{M_{Dn}^2 + \mathbf{p}_n^2} - M_{Dn})x - \delta^3 \mathbf{p}_n^2}; n$ runs over all product particles $D_n$. If $M_{Dn} \delta >> 1$, the saddle-point approximation applies and

$$F(s) = \prod_n \left( \frac{M_{Dn}}{2\pi i (s - i M_{Dn} \delta^2/2)} \right)^{3/2}.$$
where $\sigma$ is the spread of the initial wave-packet and $\beta$ the initial flavor. Then,
\[
\mathcal{A}_\alpha(t,x) = \sum_i U^*_{\alpha i} U_{\alpha i} \frac{4\pi\sigma^2}{\pi}^{1/4} e^{-\frac{(x-v_i)^2}{2\sigma^2}} + (\pi\sigma^2)^{1/4},
\] (13)
where $E_i(p) = \sqrt{m_i^2 + p^2}$ and $\Gamma_i(p)$ are the decay rates in the different mass eigenstates. We employ a commonly used approximation: we expand $E_i(p)$ to first order in $p - p_i$ and $\Gamma_i(p)$ to zero-th order in $p - p_i$, i.e.,
\[
E_i(p) = E_i + v_i(p - p_i); \quad \Gamma_i(p) = \Gamma_i,
\] (14)
where $E_i = E_i(p_i)$, $v_i = p_i/E_i$, and $\Gamma_i = \Gamma_i(p_i)$ [14]. Then,
\[
\mathcal{A}_\alpha(t,x) = \sum_i U^*_{\alpha i} U_{\alpha i} e^{-\frac{(x-v_i)^2}{2\sigma^2}} + (\pi\sigma^2)^{1/4} e^{-\frac{(x-v_i)^2}{2\sigma^2}} + iE_i t - \Gamma_i t \] (15)
Substituting (15) in Eq. (11) and letting $T \rightarrow \infty$, we obtain the probability density for the detection of flavor $\alpha$ at $x = L$
\[
p_{\alpha}(L) = K' \left( \sum_i S_i e^{-\frac{(x-v_i)^2}{2\sigma^2}} + \sum_{i<j} 2 e^{-\frac{(x-v_i)^2 + (x-v_j)^2}{2\sigma^2}} \right) L \]
\[
\times Re(T_{ij} e^{ik_i^L}),
\] (16)
where $K' > 0$ is a redefined normalization constant, $S_i \sim |U^*_{\alpha i} U_{\alpha i}|^2$, $T_{ij} \sim U^*_{\beta i} U_{\alpha i} U_{\beta j} U^*_{\alpha j}$ (their precise form is not necessary for the arguments) and
\[
k_{ij} = \frac{E_i - \epsilon_{th}}{v_i} - \frac{E_j - \epsilon_{th}}{v_j} - (p_i - p_j).
\] (17)
If $\epsilon_{th}/E_i << 1$ then $k_{ij} = m_i^2/p_i - m_j^2/p_j$, i.e., for $p_i = p + O(m_i)$, $k_{ij}$ is twice the standard result $(m_i^2 - m_j^2)/(2p)$.

It is important to emphasize that this result does not depend on any *ad hoc* assumptions about the initial state. It follows from Eq. (3), which is valid for a large class of measurements, and it provides the probabilities for alternatives when the detection time is not measured. Moreover, Eq. (17) is insensitive to the approximations we made in this paper. It holds for *any* calculation, in which the summation over detection time is performed at the level of amplitudes. This is the reason why it is discerned even in the simplified plane wave description.

If we do not assume that $\epsilon_{th}/E_i << 1$, Eq. (17) becomes $k_{ij} = (1 - \frac{1}{2\sigma^2})(m_i^2 - m_j^2)/E$, in the ultra-relativistic limit. The oscillation wavelength carries a strong dependence on the threshold energy. This dependence does not arise at all in the standard treatment and it is not an artifact of any approximation. In particular, the presence of an energy threshold is equivalently described as a constant potential $V = -\epsilon_{th}$, in which the particle propagates prior to detection. The incorporation of this potential to the Hamiltonian would lead to Eq. (16) even in the simplified plane wave description.

As a final remark, we note that the kernel $F(t-t')$ in Eq. (11) turns out to cause significant suppression of interferences in detection time for the amplitude (15). One might, however, inquire whether it is possible that in a different model one might obtain $F(t-t') \sim \delta(t-t')$, so that $p_{\alpha}(L) \sim \int_0^{T} dt |\mathcal{A}(t,L)|^2$; the standard expression for the oscillation wavelength would then follow instead of (17). However, this would require the presence of an unusual and highly efficient decohering mechanism for the detection time that would be effective in timescales shorter than the shortest one appearing in (15), namely, $E^{-1}$. Moreover, such a mechanism would have to be *extrinsic* to the physics of particle oscillations, and it would have to be postulated *ad hoc*.

* Electronic address: anastop@physics.upatras.gr
* Electronic address: ntina@imperial.ac.uk

[1] M. Gell-Mann and A. Pais, Phys. Rev. 97, 1387 (1955); B. Pontecorvo, JETP 33, 599 (1957); B. Kayser, Phys. Rev. D24, 110 (1981); C. Giunti, C. W. Kim, J. A. Lee and U. W. Lee, Phys. Rev. D48, 4310 (1993). See also the reviews [2, 3] and references therein.
[2] M. Zralek, Acta Phys. Polon. B29, 3925 (1998).
[3] M. Beuthe, Phys. Rept. 375, 105 (2003).
[4] C. Giunti, Phys. Scr. 67, 29 (2003).
[5] H. Lipkin, Phys. Lett. B 642, 366 (2006).
[6] E. B. Davies, *Quantum theory of open systems* (Academic Press, 1976); P. Busch, P. J. Lahti and P. Mittelstaedt, *The Quantum Theory of Measurement*, (Springer, 1996).
[7] C. Anastopoulos and N. Savvidou, J. Math. Phys. 47, 122106 (2006); 48, 032106 (2007); 49, 022101 (2008); C. Anastopoulos, J. Math. Phys. 49, 022103 (2008).
[8] Of direct relevance is the notion of "spacetime coarse-graining"; see, J. B. Hartle, Phys. Rev. D44, 3173 (1991); in Proceedings of the 1992 Les Houches School, Gravitation and Quantisation, 1993.
[9] M. D. Srinivas and E. B. Davies, J. Mod. Opt. 28, 981 (1981); 29, 235 (1982).
[10] For an elaboration of this distinction, see, N. Savvidou, in *Approaches to Quantum Gravity*, ed. D. Oriti (Cambridge University Press, 2009), and references therein.
[11] For derivations of non-standard expressions for the oscillation wavelength, see Ref. 3 and references therein.
[12] For example, if $K_L \rightarrow 3\pi^0$, $\mathcal{H}_0$ is the Fock space for $\pi^0$ particles, $\mathcal{H}_0$ is the $\pi^0$ vacuum subspace and $\mathcal{H}_\text{prod}$ is the subspace of $H_r$ with non-zero number of $\pi^0$ particles.
[13] For neutral bosons, the mixing matrix, in general, depends on the boson’s momentum; our treatment can be straightforwardly generalized to cover this case.
[14] Keeping higher order terms in the expansion would provide a better approximation that would also incorporate the effects of wave packet dispersion.