The Average-Value Correspondence Principle

Philip Goyal

Perimeter Institute, Waterloo, Canada

Abstract. In previous work [1], we have presented an attempt to derive the finite-dimensional abstract quantum formalism from a set of physically comprehensible assumptions. In this paper, we continue the derivation of the quantum formalism by formulating a correspondence principle, the *Average-Value Correspondence Principle*, that allows relations between measurement outcomes which are known to hold in a classical model of a system to be systematically taken over into the quantum model of the system, and by using this principle to derive many of the correspondence rules (such as operator rules, commutation relations, and Dirac's Poisson bracket rule) that are needed to apply the abstract quantum formalism to model particular physical systems.

Keywords: Quantum-classical correspondence principles, Origin of Quantum Formalism, Correspondence Rules
PACS: 03.65.-w, 03.65.Ta

INTRODUCTION

In order to obtain a quantum mechanical model for a particular physical system such as a particle moving in space, it is necessary to supplement the abstract quantum formalism (due to von Neumann and Dirac) with various rules, to which we shall refer collectively as *correspondence* rules [1], which explicitly determine the form of the operators that represent particular measurements performed on the system, or that represent particular symmetry transformations (such as displacement or rotation) of the frame of reference in which the system is being observed.

The physical origin of many of these correspondence rules is obscure. For example, the commutation relationship $[\mathsf{x}, \mathsf{p}_x] = i\hbar$ is typically obtained from Schroedinger's equation or from Dirac's Poisson bracket rule, both of whose derivations involve abstract assumptions whose physical origin is obscure.

Operator rules have been discussed in a few publications, for example in [3, 4, 5], but a derivation of the operator rules on the basis of a physical principle, taking the abstract formalism as a given, has not been successfully completed. The most recent systematic attempt to derive the commonly employed measurement and measurement–transformation commutation relations on the basis of a physical principle is found in [6, 7, 8]. However, this approach implicitly makes use of operator rules, and relies upon auxiliary assumptions (such as the assumption that certain measurement operators are unchanged by the action of particular symmetry operators) whose physical origin is unclear. In this paper, we show that, starting from the abstract quantum formalism, it is possible to derive the above-mentioned correspondence rules in a straightforward manner from a correspondence principle.

THE AVERAGE-VALUE CORRESPONDENCE PRINCIPLE

The *Average-Value Correspondence Principle* (AVCP), asserts that, in a classical experiment, if a measurement of observable A on a physical system can be implemented by an arrangement where a number of other measurements of observables $A', A'', \ldots$ are performed on the same system (perhaps performed at different times), so that the outcome of the measurement of A can be calculated as a function of the outcomes of the measurements of $A', A'', \ldots$, then we can construct a corresponding quantum experiment, where quantum measurements $\mathbf{A}, \mathbf{A}', \mathbf{A}'', \ldots$, corresponding to classical measurements of $A, A', A'', \ldots$, are performed on the same physical system, such that the same functional relation holds *on average* between the outcomes of the quantum measurements, where the average is taken over infinitely many trials of the quantum experiment.

---

[1] A systematic classification of the correspondence rules is given in [2]. Readers are referred to [2] for a more thorough discussion of the subject of the present paper.

The precise form of the corresponding quantum experiment is arrived at by carefully considering a sequence of particular examples (see Ref. [2] for details). The key difficulty one has to face is what the quantum experiment must be in the case where the operators involved are non-commuting. For example, suppose that, in the classical framework, a measurement of C at time t can be implemented by performing measurements of A and B at time t. Suppose further that, in the quantum framework, it is known that operators A and B represent measurements of A and B respectively, and that $[\mathsf{A},\mathsf{B}] \neq 0$. How should one deal with this situation? The AVCP insists that the quantum measurements are performed on separate copies of the system. On the contrary, if the operators commute, and if the quantum measurements are not performed on sub-systems of a composite system, then the measurements are to be performed on the same copy of the system. This makes some intuitive sense: when the operators commute, we are in an essentially classical situation, and so there is no harm in having the measurements performed on the same copy of the system; when the operators do not commute, we are in a distinctively non-classical situation, in which case, to avoid the complication of one measurement's results being affected by the other's and to avoid the issue of measurement ordering, one insists that the measurements are performed on separate copies of the system.

Statement of the Principle

We shall now state the general principle. An illustrative example is given in Fig. 1.

Consider a classical idealized experiment in which a system (possibly a composite system) is prepared in some state at time t_0, and is allowed to evolve in a given background. Suppose that a measurement of $A^{(m)}$ $(m \geq 2)$, performed on the system at time t_2 with outcome $a^{(m)}$, can be implemented by an arrangement where measurements of $A^{(1)}, A^{(2)}, \ldots, A^{(m-1)}$ are performed upon one copy of the system at time t_1, and their respective outcomes, denoted $a^{(1)}, a^{(2)}, \ldots, a^{(m-1)}$, are then used to compute the result $f(a^{(1)}, a^{(2)}, \ldots, a^{(m-1)})$, where f is an analytic function, so that the relation

$$a^{(m)} = f(a^{(1)}, a^{(2)}, \ldots, a^{(m-1)}) \qquad (*)$$

holds for all initial (classical) states of the system.

Consider the case where the quantum measurements $\mathbf{A}^{(1)}, \mathbf{A}^{(2)}, \ldots, \mathbf{A}^{(m)}$, with operators $\mathsf{A}^{(1)}, \mathsf{A}^{(2)}, \ldots, \mathsf{A}^{(m)}$, represent the measurements of $A^{(1)}, A^{(2)}, \ldots, A^{(m)}$, respectively. Then, consider the following idealized quantum experimental arrangement consisting of several set-ups, each consisting of identical sources and backgrounds, where, in each set-up, a copy of the system is prepared in the same initial state, v_0, at time t_0.

In one set-up, only measurement $\mathbf{A}^{(m)}$ is performed (at time t_2) and, for any i, j with $i \neq j$ and $i, j \leq m-1$, the measurements $\mathbf{A}^{(i)}, \mathbf{A}^{(j)}$ are performed (at time t_1) in:

1. the *same* set-up if both $\left[\mathsf{A}^{(i)}, \mathsf{A}^{(j)}\right] = 0$ and, if the system is composite, the measurements are performed on the same sub-system;
2. *different* set-ups if $\left[\mathsf{A}^{(i)}, \mathsf{A}^{(j)}\right] \neq 0$.

Let the outcomes of the measurements $\mathbf{A}^{(1)}, \ldots, \mathbf{A}^{(m)}$ in any given run of the experimental arrangement be denoted $a^{(1)}, \ldots, a^{(m)}$, respectively. The function $f(a^{(1)}, a^{(2)}, \ldots, a^{(m-1)})$ is defined as *simple* provided that its polynomial expansion contains no terms involving a product of eigenvalues belonging to measurements whose operators do not commute. If f is simple, then $(*)$ holds on average, the average being taken over an infinitely large number of runs of the experiment.

The above principle can be generalized in a number of ways, for example to the case where the measurements of $A^{(1)}, \ldots, A^{(m-1)}$ are not performed at the same time. However, these generalizations are unnecessary for the derivations of the usual correspondence rules of quantum theory, and are therefore not discussed here.

GENERALIZED OPERATOR RULES

We will now apply the AVCP to derive operator relations which hold when the function f takes various useful forms. In each instance of f, we shall first derive a generalized operator rule which relates the expected values of the relevant operators at *different times*. Then, taking the special case when the expected values are computed at the same time, we obtain the corresponding operator rule which relates the operators themselves.

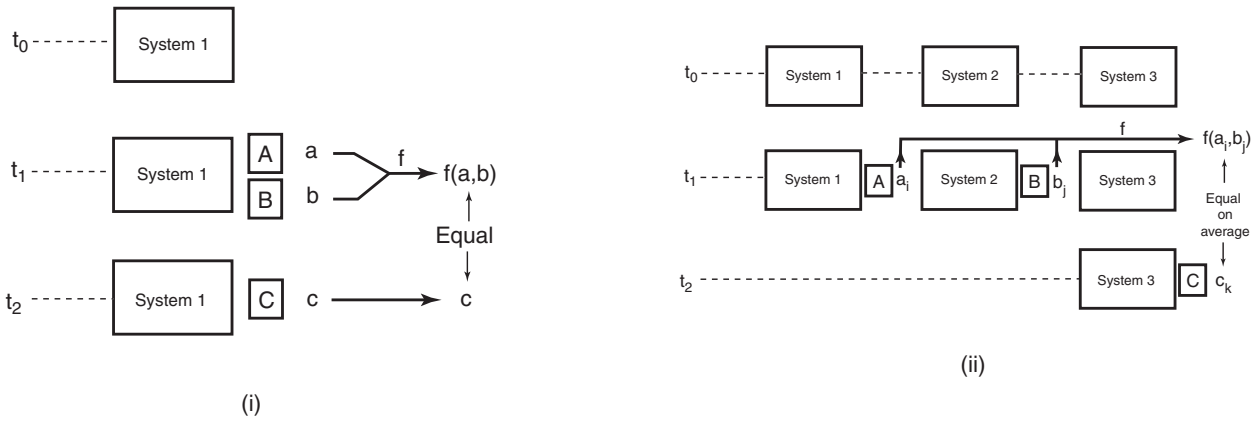

FIGURE 1. An example of the application of the AVCP. (i) A classical experiment showing the measurements of A, B and C performed at times t_A, t_B, and t_C, respectively, with outcomes denoted as a, b, and c, respectively. Here, $t_A = t_B = t_1$ and $t_C = t_2$. Suppose that one finds that the relation $c = f(a,b)$ holds for all initial states of the system. (ii) The corresponding quantum experiment. Three copies of the system are prepared in the same initial state, v_0, at time t_0, and are placed in identical backgrounds. In this example, it is assumed that the operators A and B do not commute. Hence, by the AVCP, measurements **A** and **B** are performed on different copies of the system. Measurement **C** is performed on a separate copy of the system. In any given run of the experiment, the probabilities that measurements **A**, **B** and **C** yield outcome values a_i, b_j and c_k $(i, j, k = 1, \ldots, N)$, are P_i, P'_j and P''_k, respectively. The AVCP then asserts that, provided the polynomial expansion of $f(a,b)$ contains no product terms involving a and b, the relation $\bar{c} = \overline{f(a,b)}$ holds for all initial states, v_0, of the system, where the average is taken over an infinite number of runs of the experiment; that is, $\sum_k c_k P''_k = \sum_{ij} f(a_i, b_j) P_i P'_j$ for all v_0.

We shall consider a classical experiment where a system is subject to measurements of A and B at time t_1, and to a measurement of C at time t_2. We shall suppose that a measurement of C, with outcome c, can be implemented by an arrangement in which the measurements of A and B are performed, with respective outcomes a and b, and the function $f(a,b)$ then computed, so that the relation $c = f(a,b)$ holds for all initial states of the system.

In a quantum model of the appropriate experimental arrangement, let the operators that represent these measurements be denoted A, B, and C, respectively. To simplify the presentation, we shall only consider the case where these operators have finite dimension, N; the results obtained below can be readily shown to apply in the infinite dimensional case. Let the elements of orthonormal sets of eigenvectors of A, B and C be denoted $\mathsf{v}_i, \mathsf{v}'_j$, and v''_k, respectively $(i, j, k = 1, 2, \ldots, N)$, let the corresponding eigenvalues be denoted a_i, b_j and c_k, and let the probabilities of the ith, jth and kth outcomes of measurements **A**, **B** and **C** in any given experimental arrangement be denoted by P_i, P'_j and P''_k, respectively.

Case 1. f is a function of a only.

In this case, according to the AVCP, the quantum experiment simply consists of two set-ups, involving two copies of the system, where **A** is performed on one copy at time t_1 and **C** on the other copy at time t_2. Consequently, we obtain the relation

$$\langle \mathsf{C} \rangle_{t_2} = \left\langle f(\mathsf{A}) \right\rangle_{t_1}, \tag{1}$$

which holds for all v_0. We can summarize the above result in the form of the *generalized function rule*:

$$c(t_2) = f\left(a(t_1)\right) \mapsto \langle \mathsf{C} \rangle_{t_2} = \left\langle f(\mathsf{A}) \right\rangle_{t_1} \quad \forall \mathsf{v}_0, \tag{2}$$

where, for clarity, the times at which the outcomes are obtained has been explicitly indicated. In the special case where $t = t_1 = t_2$, we obtain the usual operator rule, the *function rule*:

$$c = f(a) \mapsto \mathsf{C} = f(\mathsf{A}). \tag{3}$$

Case 2. $f(a,b) = f_1(a) + f_2(b)$

In this case, by straightforward application of the AVCP, one finds that, whether or not the operators A, B commute, and whether or not they are performed on the same sub-system in the case of a composite system, one obtains the *generalized sum rule*:

$$c(t_2) = f_1\left(a(t_1)\right) + f_2\left(b(t_1)\right) \mapsto \langle \mathsf{C} \rangle_{t_2} = \left\langle f_1(\mathsf{A}) \right\rangle_{t_1} + \left\langle f_2(\mathsf{B}) \right\rangle_{t_1} \qquad \forall \mathsf{v}_0. \tag{4}$$

In the special case where $t = t_1 = t_2$, we obtain the *sum rule*:

$$c = f_1(a) + f_2(b) \mapsto \mathsf{C} = f_1(\mathsf{A}) + f_2(\mathsf{B}). \tag{5}$$

Case 3. $f(a,b) = f_1(a) f_2(b)$

In this case, application of the AVCP yields the *generalized product rule*:

$$c(t_2) = f_1\big(a(t_1)\big) f_2\big(b(t_1)\big) \mapsto \langle \mathsf{C} \rangle_{t_2} = \left\langle f_1(\mathsf{A}) f_2(\mathsf{B}) \right\rangle_{t_1} \quad \forall \mathsf{v}_0 \qquad \text{if } [\mathsf{A}, \mathsf{B}] = 0. \tag{6}$$

In the special case where $t_1 = t_2$, we obtain the *product rule*:

$$c = f_1(a) f_2(b) \mapsto \mathsf{C} = f_1(\mathsf{A}) f_2(\mathsf{B}) \quad \text{if } [\mathsf{A}, \mathsf{B}] = 0. \tag{7}$$

Note that there is no rule in the case where A and B do not commute.

APPLICATION OF THE AVCP

Remarkably, one finds that the AVCP enables correspondence rules of each of the four types described in [2], namely operator rules, measurement commutation relations, measurement–transformation commutation relation, and transformation operators, to be obtained in a uniform manner. Consequently, one can see that the difference between these types of rules depends simply upon whether the classical relations that one is taking over into the quantum framework are relations between measurements performed at the same time (leading to the operator rules), at different times (leading to measurement commutation relations), or in different frames of reference (leading to measurement–transformation commutation relations and to the explicit forms of transformation operators). In short, from the perspective provided by the derivation, the commutation relation $[\mathsf{x}, \mathsf{p}_x] = i\hbar$ is no more elusive in its origin than fact that the classical relation $H = p_x^2/2m$ is taken over into the operator relation $\mathsf{H} = \mathsf{p}_x^2/2m$.

In this section, for reasons of space, we will pick just one simple illustrative application of the AVCP, namely the derivation of the measurement commutation relation $[\mathsf{x}, \mathsf{p}_x] = i\hbar$. The reader is referred to [2] for further applications, which include the derivation of the angular momentum commutation relations, namely $[\mathsf{L}_x, \mathsf{L}_y] = i\hbar \mathsf{L}_z$ (and cyclic permutations thereof), and Dirac's Poisson-Bracket Rule.

The position–momentum commutation relationships

Consider a particle of mass m moving in the $+x$-direction, where measurements of the x-component of position, the x-component of momentum, and the energy, are represented by the operators $\mathsf{x}, \mathsf{p}_x, \mathsf{H}$, respectively.

First, to determine the relationship of H to the operators x and p_x, we make use of the fact that the relation $H = p_x^2/2m + V(x)$ holds for all classical states (x, p_x) of the system, so that, from the function rule, it follows that $\mathsf{H} = \mathsf{p}_x^2/2m + V(\mathsf{x})$.

Next, to obtain a relation between x and p_x, we make use of the fact that, in the quantum model of the system, the expected value of x at time $t+\delta t$ can be calculated in two separate ways. First, from the definition of $\langle x\rangle_t$, the relation

$$\begin{aligned}\langle \mathsf{x}\rangle_{t+\delta t} &= \left\langle \mathsf{U}_t^\dagger(\delta t)\,\mathsf{x}\,\mathsf{U}_t(\delta t)\right\rangle_t \\ &= \left\langle \left(1+\frac{i\mathsf{H}\delta t}{\alpha}\right)\mathsf{x}\left(1-\frac{i\mathsf{H}\delta t}{\alpha}\right)\right\rangle_t + O(\delta t^2) \\ &= \langle \mathsf{x}\rangle_t + \frac{i}{\alpha}\delta t\langle \mathsf{H}\mathsf{x}-\mathsf{x}\mathsf{H}\rangle_t + O(\delta t^2) \\ &= \langle \mathsf{x}\rangle_t - \frac{i}{2m\alpha}\delta t\left\langle [\mathsf{x},\mathsf{p}_x^2]\right\rangle_t + O(\delta t^2)\end{aligned} \tag{8}$$

holds for all states, v, of the system. Second, using the generalized function rule, it follows from the classical relation $x(t+\delta t) = x(t) + p_x(t)\delta t/m + O(\delta t^2)$ that the relation

$$\langle \mathsf{x}\rangle_{t+\delta t} = \langle \mathsf{x}\rangle_t + \frac{\delta t}{m}\langle \mathsf{p}_x\rangle_t + O(\delta t^2) \tag{9}$$

holds for all v.

Equating the above two expressions for $\langle x\rangle_{t+\delta t}$, we obtain

$$\mathsf{v}^\dagger\left(\mathsf{p}_x[\mathsf{x},\mathsf{p}_x] + [\mathsf{x},\mathsf{p}_x]\,\mathsf{p}_x\right)\mathsf{v} = \mathsf{v}^\dagger 2i\alpha\mathsf{p}_x\mathsf{v} \tag{10}$$

for all v, which implies that

$$\mathsf{p}_x[\mathsf{x},\mathsf{p}_x] + [\mathsf{x},\mathsf{p}_x]\,\mathsf{p}_x = 2i\alpha\mathsf{p}_x. \tag{11}$$

Representing the operators p_x and $\mathsf{A}\equiv[\mathsf{x},\mathsf{p}_x]$ as limits of infinite-dimensional matrices, and diagonalizing p_x by a unitary transformation U, one obtains

$$\mathsf{p}'_x\mathsf{A}' + \mathsf{A}'\mathsf{p}'_x = 2i\alpha\mathsf{p}'_x, \tag{12}$$

where $\mathsf{p}'_x = \mathsf{U}^{-1}\mathsf{p}_x\mathsf{U}$ and $\mathsf{A}' = \mathsf{U}^{-1}\mathsf{A}\mathsf{U}$. Writing $\mathsf{p}'_x = p'_i\delta_{ij}$, where i and j range from $-\infty$ to $+\infty$, and multiplying out Eq. (12), we obtain

$$(p'_i + p'_j)A'_{ij} = 2i\alpha p'_i\delta_{ij}. \tag{13}$$

Now let the eigenvalues, p'_i, be chosen such that they have equal spacing δ, so that $p'_{i+1} - p'_i = \delta$, and set $p_0 = \varepsilon$ with $0 < \varepsilon < \delta$, which ensures that $(p'_i + p'_j) \neq 0$ for any i, j. In this case, it follows that $\mathsf{A} = i\alpha\mathsf{I}$, which, in the limit as δ tends to zero, yields the commutation relationship

$$[\mathsf{x},\mathsf{p}_x] = i\alpha. \tag{14}$$

Although a particular system has been used to obtain this commutation relation, one can argue that it is generally valid [2]. We note that, if one instead considers a massless particle moving at the speed of light in the positive x direction, with classical Hamiltonian given by $H = cp_x$, and so that $x(t+\delta t) = x(t) + c\delta t$, then the above computation immediately yields Eq. (14).

REFERENCES

1. P. Goyal. An information-theoretic approach to quantum theory, I: The abstract quantum formalism. quant-ph/0702124, 2007.
2. P. Goyal. An information-theoretic approach to quantum theory, II: The correspondence rules of quantum theory. quant-ph/0702149, 2007.
3. P.A.M. Dirac. *Principles of Quantum Mechanics*. Oxford Science Publications, fourth edition, 1999.
4. J. von Neumann. *Mathematical Foundations of Quantum Mechanics*. Princeton University Press, 1955.
5. H. J. Groenewold. On the principles of elementary quantum mechanics. *Physica*, 12:405–60, 1946.
6. T. F. Jordan. Why $-i\nabla$ is the momentum. *Am. J. Phys.*, 43(12), 1975.
7. T. F. Jordan. *Linear Operators for Quantum Mechanics*. John Wiley and Sons, Inc., 1969.
8. L. E. Ballentine. *Quantum Mechanics : A Modern Development*. World Scientific, 1999.