Classical optics representation of the quantum mechanical translation operator via ABCD matrices

Marco Ornigotti and Andrea Aiello

1 Max Planck Institute for the Science of Light, Günther-Scharowsky-Strasse 1/Bau24, D-91058 Erlangen, Germany
2 Institute for Optics, Information and Photonics, University of Erlangen-Nuernberg, Staudtstrasse 7/B2, D-91058 Erlangen, Germany

E-mail: marco.ornigotti@mpl.mpg.de

Received 6 May 2013, accepted for publication 11 June 2013
Published 27 June 2013
Online at stacks.iop.org/JOpt/15/075715

Abstract

The ABCD matrix formalism describing paraxial propagation of optical beams across linear systems is generalized to arbitrary beam trajectories. As a by-product of this study, a one-to-one correspondence between the extended ABCD matrix formalism presented here and the quantum mechanical translation operator is established.

Keywords: geometrical optics, matrix theory

1. Introduction

Rays of light propagate along rectilinear trajectories in air. Therefore, at the generic position $x$ a ray, which may be represented by the linear function $f(x) = a + bx$, is completely determined by a sole pair of numbers: $f(x)$ and $f'(x)$. Such a pair may be represented in a vector-like form as follows:

$$\begin{bmatrix} f(x) \\ f'(x) \end{bmatrix}. \quad (1)$$

The simple linear relation existing between $f(x_1)$ and $f(x_2)$ at two arbitrarily chosen positions $x_1$ and $x_2 = x_1 + L$, with $L > 0$, is usually written in optics textbooks [1–4] in the following matrix form:

$$\begin{bmatrix} f(x_2) \\ f'(x_2) \end{bmatrix} = \begin{bmatrix} 1 & L \\ 0 & 1 \end{bmatrix} \begin{bmatrix} f(x_1) \\ f'(x_1) \end{bmatrix}. \quad (2)$$

The $2 \times 2$ matrix in equation (2) is known as the ABCD matrix of the optical system (free space, in the present case) and fully characterizes the propagation of rays of light in it.

The literature is rich in examples of ABCD matrices for more complicated optical systems such as lenses, planes and curved dielectric interfaces, mirrors, inhomogeneous media with a quadratic index profile etc and combinations thereof [5, 6]. For complex inhomogeneous media, the trajectory of a ray of light is more complicated than a straight line and cannot be represented by a linear function, except for a very short distance $x_2 - x_1 \approx \delta x_1 \ll 1$.

In this case, a simple approach like the one given by the ABCD matrices fails to be efficient and one has to embrace a more complicated method capable of dealing with the inhomogeneities of the medium [7, 8]. For example, in [8] the problem of ray propagation in inhomogeneous media is studied by using a Hamiltonian approach, thus solving the equations of motion of the ray itself. This approach is therefore dynamic, i.e. it takes into account not only the trajectory from a kinematic point of view, but also analyzes its dynamic aspects, answering the question of why the trajectory is like that.

Contrary to this dynamic approach, in this paper we focus attention only on the kinematic nature of the trajectory, i.e. we are not interested in the cause of motion, but only in the description of the trajectory itself. To this aim, we note that formally the ABCD matrix approach can still be used if an appropriate generalization of this method is constructed by observing that the ABCD matrix formalism is nothing but the consequence of a linearization of the trajectory of a ray of light around the initial point $x_1$ [9, 10], namely a Taylor expansion truncated up to and including first-order
terms. Such a linearization procedure, however, is physically meaningful only for \( x_2 \) close enough to \( x_1 \): \( x_2 = x_1 + \delta x_1 \). But (a) what if \( x_2 \) is no longer close to \( x_1 \)? (b) Does the first-order Taylor expansion break down? (c) If so, can such an expansion be suitably extended? (d) If higher order terms must be retained, what is their physical meaning?

To answer these questions, we intend to proceed by a two-step reasoning. Firstly, in section 2, we put this linearization procedure on rigorous basis, showing that a \( 2 \times 2 \) ABCD matrix is a principal sub-matrix \(^3\) of an effective \( \infty \times \infty \) matrix describing the full nonlinear dynamics of a curvilinear ray of light. In section 3 we then discuss the physical meaning of the proposed linearization scheme, pointing out that the generalized ABCD matrix is nothing but a physical representation of the well-known quantum translation operator \( e^{L(d/dx)} \) in one-dimensional quantum mechanics.

2. Generalized ABCD matrices for non-rectilinear light propagation

To begin with, let us first re-derive equation (2) for an arbitrary linear function \( f(x) \) that now we write in the following manner:

\[
f(x) = a_0 + a_1 x = a_0' + a_1' (x - x_i),
\]

where \( x_i \) is an arbitrary point belonging to the domain of the function \( f(x) \). If we choose \( x_i = 0 \) then we retrieve the previous expression \( f(x) = a + bx \) with \( a_0 = a \) and \( a_1 = b \). However, for \( x_i \neq 0 \), the last equality in equation (3) gives:

\[
\begin{align*}
 a_0' &= a_0 + a_1 x_i = f(x_i), \\
 a_1' &= a_1 = f'(x_i),
\end{align*}
\]

(4a)

(4b)

where, for the sake of simplicity, we have introduced the notation \( a_k' = a_k(x_i) \). Since the point \( x_i \) is arbitrarily chosen, we can pick out a different point \( x_j = x_i + L \) and write:

\[
f(x) = a_0' + a_1' (x - x_i) = a_0'' + a_1'' (x - x_j).
\]

(5)

By equating the factors with the same powers of \( x \) at the second and third terms in the equation above, we obtain \( a_0'' = a_0' + (x_j - x_i)a_1' \) and \( a_1'' = a_1' \). This relation can be rewritten in the following matrix form:

\[
\begin{bmatrix}
 a_0'' \\
 a_1''
\end{bmatrix} = 
\begin{bmatrix}
 x_j - x_i \\
 0 \\
 1
\end{bmatrix}
\begin{bmatrix}
 a_0' \\
 a_1'
\end{bmatrix}.
\]

(6)

This result is equivalent to the one given in equation (2), if we identify \( a_0' = f(x_2) \), \( a_1' = f'(x_2) \), \( a_0'' = f(x_1) \) and \( a_1'' = f'(x_1) \). Obviously, the formal derivation of equation (2) via steps (3–6) is highly redundant for the linear-function case. However, it has the virtue of being generalizable to the case of non-rectilinear ray propagation.

Now, in order to describe a ray that propagates in an arbitrary inhomogeneous medium, we need a generic smooth nonlinear function \( f(x) \) which can be expanded in a Taylor series around \( x = 0 \) as follows:

\[
f(x) = a_0 + a_1 x + a_2 x^2 + \cdots.
\]

(7)

For any \( x_i \in \mathbb{R} \) we can write \( x = x - x_i + x_i \) and insert this relation into equation (7) to obtain

\[
f(x) = a_0 + a_1 (x - x_i) + a_2 (x - x_i)^2 + \cdots
\]

\[= \sum_{n=0}^{\infty} a_n' (x - x_i)^n,
\]

(8)

where the \( a_n' \) coefficients are given by:

\[
a_n' = \frac{1}{n!} \frac{d^n f(x)}{dx^n}igg|_{x=x_i} = \sum_{k=n}^{\infty} \binom{k}{n} a_k x_i^{k-n}.
\]

(9a)

(9b)

Now we can repeat the same reasoning that led to equation (8), but with a different expansion point \( x_j = x_i + L \), and write the following equality:

\[
\sum_{n=0}^{\infty} a_n' (x - x_i)^n = \sum_{n=0}^{\infty} a_n'' (x - x_j)^n,
\]

(10)

which simply states the independence of \( f(x) \) from the expansion points \( x_i \) and \( x_j \). By expanding both sides of this equation with the help of Newton’s binomial formula one obtains:

\[
\sum_{n=0}^{\infty} a_n' \sum_{k=0}^{n} \binom{n}{k} x^k (x_i)^{n-k} = \sum_{n=0}^{\infty} a_n'' \sum_{k=0}^{n} \binom{n}{k} x^k (x_j)^{n-k}.
\]

(11)

This expression can be turned into a recursive relation by equating terms with the same power of \( x \). Then, for \( k = 0 \) we have:

\[
\sum_{n=0}^{\infty} (-1)^n a_n' x^n = \sum_{n=0}^{\infty} (-1)^n a_n'' x^n,
\]

(12)

which can be rewritten, after isolating the \( n = 0 \) term, as:

\[
a_0' = a_0'' + \sum_{n=1}^{\infty} (-1)^n (a_n' x^n - a_n'' x^n).
\]

(13)

For \( k = 1 \) the same operation yields:

\[
a_1' = a_1'' + \sum_{n=2}^{\infty} (-1)^{n-1} (a_n' x^{n-1} - a_n'' x^{n-1}).
\]

(14)

This procedure can be iterated for arbitrary values of \( k \), thus generating the following recursive relation:

\[
a_k' = a_k'' + \sum_{n=k+1}^{\infty} \binom{n}{k} (-1)^{n-k} (a_n' x^{n-k} - a_n'' x^{n-k}),
\]

(15)

with \( k = 0, 1, \ldots, n \). The equation above can be seen as a linear algebraic system relating the variables \( a_n' \) to the quantities \( a_n'' \). This result can then be written in matrix form.
as follows:
\[
\begin{bmatrix}
 b_0(0) & b_1(0) & b_2(0) & \ldots \\
 0 & b_1(1) & b_2(1) & \ldots \\
 0 & 0 & b_1(2) & \ldots \\
 \vdots & \vdots & \vdots & \ddots
\end{bmatrix} \begin{bmatrix}
 a_0 \\
 a_1 \\
 a_2 \\
 \vdots
\end{bmatrix} =
\begin{bmatrix}
 c_0 \\
 c_1 \\
 c_2 \\
 \vdots
\end{bmatrix},
\]
where \( a^i = (a_0^i, a_1^i, \ldots), a^j = (a_0^j, a_1^j, \ldots) \) and \( b_n^k = \binom{n}{k} (-1)^{n-k} x_j^{n-k} \). If we call \( B(j) \) the matrix on the left-side of the previous equation and \( B(i) \) the one on the right-side, equation (16) can be written in the compact form:
\[
B(j)a^i = B(i)a^j,
\]
where again the shorthand notations \( B(i) = B(x_i) \) and \( B(j) = B(x_j) \) are used for the sake of clarity. Solving for \( a^j \) by multiplying on the left both sides of the previous equation by \( B(j)^{-1} \) and defining \( A = B^{-1}(j)B(i) \), we can write the relation between the vectors \( a^i \) and \( a^j \) as
\[
A = A(L),
\]
where \( L = x_j - x_i \). Note that this matrix contains the usual (i.e. linear) ABCD matrix defined in equation (2) as the first \( 2 \times 2 \) principal sub-matrix. This sub-matrix verifies the following equality:
\[
\begin{bmatrix}
 1 & L & L^2 & L^3 & \ldots \\
 0 & 1 & 2L & 3L^2 & \ldots \\
 0 & 0 & 1 & 3L & \ldots \\
 0 & 0 & 0 & 1 & \ldots \\
 \vdots & \vdots & \vdots & \vdots & \ddots
\end{bmatrix} =
\begin{bmatrix}
 1 & L & L_1 + L_2 \\
 0 & 1 & 0 & 0 \end{bmatrix},
\]
as can be checked by a straightforward calculation. Similarly, for the \( 3 \times 3 \) principal sub-matrix one obtains:
\[
\begin{bmatrix}
 1 & L_1 & L_2 & L_1 + L_2 & \ldots \\
 0 & 0 & 0 & 0 & \ddots
\end{bmatrix} =
\begin{bmatrix}
 1 & L & L_1 + L_2 & (L_1 + L_2)^2/2 & \ldots \\
 0 & 0 & 0 & 0 & \ddots
\end{bmatrix}.
\]

### 3. Connection with the translation operator

The composition properties of the various sub-matrices, as given for the linear and quadratic order by equations (19) and (20) respectively, have a straightforward physical meaning: they illustrate the fact that propagation across two consecutive distances \( L_1 \) and \( L_2 \) can be described as a single propagation along the distance \( L_1 + L_2 \). From a mathematical point of view, this is a signature of the semigroup property of our generalized ABCD matrices [4]. With the help of a suitable mathematical software for algebraic manipulation, it is not difficult to verify via explicit \( N \times N \) matrix multiplications, that equation (20) is valid for arbitrary \( N \). Thus, by iteration, one can easily convince oneself that the matrix \( A \) satisfies the following relation [2]:
\[
\prod_{n=1}^{N} A(L_n) = A \left( \sum_{n=1}^{N} L_n \right).
\]

The physical implications of this relation are immediately understood: the propagation of the function through the total distance \( L_1 + L_2 + \ldots + L_N \) can be achieved by consecutive propagation across the distances \( L_1, L_2, \ldots, L_N \).

This analogy is not accidental. A closer inspection of equation (18) reveals in fact that this equation tells us how the value of the function \( f(x) \) at a point \( x_j \) can be calculated knowing the value of the same function at a point \( x_i < x_j \). With this in mind, we can calculate the derivative of \( f(x) \) as follows:
\[
\frac{df}{dx} = \lim_{\Delta x \to 0} \frac{f(x + \Delta x) - f(x)}{\Delta x} = \lim_{x \to x_j} \left( \frac{a^j - a^i}{x_j - x_i} \right).
\]

where we have chosen \( \Delta x = x_j - x_i \equiv L \) in order to represent \( f(x + \Delta x) \) as \( a^j \) and \( f(x) \) as \( a^i \). Note that this does not cause any loss of generality, since the definition of the derivative involves only the concept of neighboring points and, as discussed previously, the quantities \( a^i \) and \( a^j \) represent the value of the function \( f(x) \) in two arbitrary neighboring points. Note, moreover, that in the last equality we used equation (18) to write \( a^j \) as a function of \( a^i \). Here, \( D \) is the matrix representation of the differential operator \( \frac{d}{dx} \) [12]
\[
D = \lim_{L \to 0} \left( A - I \right) = \begin{bmatrix}
 0 & 1 & 0 & 0 & \ldots \\
 0 & 0 & 2 & 0 & \ldots \\
 0 & 0 & 0 & 3 & \ldots \\
 \vdots & \vdots & \vdots & \vdots & \ddots
\end{bmatrix}.
\]

At this point, it is not difficult to see, via an explicit calculation, that the generalized ABCD matrix is related to the differential operator by the following formula:
\[
A(L) = \sum_{k=0}^{\infty} (L D)^k \frac{1}{k!} = e^{LD}.
\]

Equation (24) gives therefore an actual physical representation of the well-known translation operator \( e^{L(d/dx)} \) [13], such that:
\[
e^{L(d/dx)} f(x) \big|_{x=0} = \sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} L^k = f(L).
\]
In our case, in fact, the action of the \( A \) matrix completely defines the vector \( F(x_2) \equiv a' \) at the point \( x_2 \) knowing the expression of \( F(x_1) \equiv a' \) at the point \( x_1 \), i.e.

\[
F(x_2) = e^{LD}F(x_1) = A(L)F(x_1). \tag{26}
\]

4. Conclusions

In this work we show how to describe the trajectory of a beam in an arbitrary inhomogeneous medium using a suitable generalization of the ABCD matrix formalism. We noticed that in such media: (a) when \( x_2 \) is no longer close to \( x_1 \), the first-order Taylor expansion of the beam trajectory (namely, the usual ABCD matrix formalism) is no longer adequate and (b) the contributions of the higher orders terms of the expansion must be taken into account. In order to find an answer to question (c), we proposed here a suitable extension of the well-known formalism of the ABCD matrix to describe an arbitrary beam trajectory consisting of the construction of an \( \infty \times \infty \) generalized ABCD matrix, whose principal \( 2 \times 2 \) matrix (corresponding to the first-order Taylor expansion of the beam trajectory) coincides with the usual ABCD matrix. Last, but not least, we also answered question (d). The higher order terms that appear here, in fact, have a simple physical meaning: they properly take into account the effects of the medium inhomogeneities on the beam trajectory. We have also shown that by taking the continuous limit of the generalized ABCD matrix, a one-to-one correspondence with the quantum mechanical translation operator in one dimension is easily established. We can interpret this correspondence as a simple relation between a physical medium, which is responsible for the generic curved trajectory of the beam, and the translation operator.

References

[1] Born M and Wolf E 1999 Principles of Optics 7th edn (Oxford: Pergamon)
[2] Hecht E 2002 Optics 4th edn (Reading, MA: Addison Wesley)
[3] Pedrotti F L and Pedrotti L S 1993 Introduction to Optics 2nd edn (Englewood Cliffs, NJ: Prentice Hall)
[4] Gerrard A and Burch J M 2012 Introduction to Matrix Methods in Optics (New York: Dover)
[5] Yariv A 1989 Quantum Electronics (New York: Wiley)
[6] Siegman A E 1986 Lasers (Cambridge: University Science Books)
[7] Landau L D and Lifshitz E M 1985 Electrodynamics of Continuous Media 2nd edn (Oxford: Pergamon)
[8] Kravtov Y A and Orlov Y I 1990 Geometrical Optics of Inhomogeneous Medium (Berlin: Springer)
[9] Stöckmann H J 2007 Quantum Chaos: An Introduction 1st edn (Cambridge: Cambridge University Press)
[10] Gaspard P 2005 Chaos, Scattering and Statistical Mechanics 1st edn (Cambridge: Cambridge University Press)
[11] Horn R A and Johnson C R 1994 Topics in Matrix Analysis (Cambridge: Cambridge University Press)
[12] Beker H 1998 Special polynomials by matrix algebra Am. J. Phys. 66 9813
[13] Ziman J M 1995 Elements of Advanced Quantum Theory (Cambridge: Cambridge University Press)