Teaching the mathematics of quantum entanglement using elementary mathematical tools

A. Gómez-Rodríguez
Departamento de Materia Condensada, Instituto de Física, Universidad Nacional Autónoma de México,
Coyoacán 04510, Apartado Postal 20-364, 01000, México D.F, 01000 México,
e-mail: alfredo@fisica.unam.mx

J.-L. Aragón
Departamento de Nanotecnología, Centro de Física Aplicada y Tecnología Avanzada,
Universidad Nacional Autónoma de México,
Apartado Postal 1-1010, Querétaro, 76000, México,
e-mail: aragon@fata.unam.mx

Recibido el 17 de noviembre de 2011; aceptado el 16 de enero de 2012

We address the problem of teaching the mathematics of entanglement using only elementary linear algebra. For this goal, we first discuss tensor products using only matrix multiplication and with this we discuss entanglement for pure bipartite systems of arbitrary dimensions. We show how to assess entanglement using only Gaussian methods, *i.e.* the row reduced echelon form of the familiar Gauss-Jordan algorithm for solving systems of linear equations. In this way we can present entanglement avoiding the difficulties of tensor products and without the Schmidt decomposition. Some elementary examples are provided together with MATLAB scripts. A Gaussian algorithm for the factorization of unentangled states is given.

Keywords: Entanglement; tensor products; linear algebra; Gaussian methods.

PACS: 03.65.Ud; 02.10.Ud

1. Introduction

Quantum entanglement is not only one of the most remarkable features of quantum mechanics but it is also at the core of many sci-fi-like applications of quantum mechanics such as quantum teleportation, quantum information, quantum cryptography and quantum computation [1]. Nowadays everybody talks about quantum entanglement but understanding this phenomenon requires knowledge of tensor products, a topic usually shrouded in mystery. Although several authors, such as Aczel [2], do a wonderful job explaining the concepts without mathematics, entanglement is such that it can hardly be understood in a purely verbal fashion; we need mathematics and usually not basic ones so that when invoked they are frequently just too frightening to beginners. For instance, entanglement is closely tied to the mathematical notion of tensor product. Some authors, as Isham [3], just ignore completely the task of providing an explanation of what tensor products are and declare that "The full definition of the tensor product operation is quite complex" and provide instead a "baby algebra" approach, a term due to Awodey [4].

We just wonder if to understand quantum entanglement is it then necessary to present a full fledged approach with lots of commutative diagrams and of universal properties. The answer to this question is no, and in this work we show how using elementary linear algebra (of the sort any first year undergraduate would understand), tensor products and entanglement can be presented in a reasonably rigorous way.

In Sec. 2 we introduce tensor products in a simple yet quite rigorous way; only matrix multiplication is required. In Sec. 3 we discuss entanglement. The question of factorizability or separability can be presented in terms of the most elementary techniques of linear algebra, namely, the Gauss-Jordan methods for solving linear systems of equation. Our main tool is the reduction of matrices to row echelon form. These mathematical techniques are discussed in elementary linear algebra books [5].

In the remaining sections we show how one can tell entangled states from unentangled ones and we present an algorithm for factorizing an unentangled state. No need for Schmidt or Singular Value Decompositions.

Through this presentation we restrict ourselves to pure bipartite systems but we place no restriction on the dimensions of the spaces other than they should be finite.

2. Tensor products

In advanced linear algebra textbooks it is shown that, given two vector spaces V and W it is always possible to construct a third space $V \otimes W$ called tensor product of V and W. It is also shown that, although there are many possible tensor products, they are all isomorphic, and in this sense the tensor product is unique.

According to quantum mechanics, every system S is described by means of a Hilbert space H, that is, a complex vector space with an inner product that is complete: Cauchy sequences converge in H. Assume next that our system S is composed of two subsystems S_1 and S_2. Symbolically we write $S = S_1 \cup S_2$ and quantum theory tells us that the Hilbert spaces of S, S_1 and S_2 are related by a tensor product, which is written as $H = H_1 \otimes H_2$. Consequently our first challenge is to describe tensor products.

This topic can be made easier as follows. We know that if a complex vector space with inner product (unitary space for short) has a finite dimension then it is automatically complete. Now, as every complex vector space with finite dimension is isomorphic to $\mathbb{C}^n$ we loose no generality if we take $H = \mathbb{C}^n$ and envision the vectors as column vectors, that is, as $n \times 1$ matrices. In the restricted context of spaces $\mathbb{C}^n$, tensor products are easy. A tensor product of two spaces $\mathbb{C}^m$ and $\mathbb{C}^n$ has two components:

1. A new vector space $\mathbb{C}^m \otimes \mathbb{C}^n$. In our case this new space will be simply $\mathbb{C}^{m\times n}$, the space of all $m \times n$ complex matrices.

2. A bilinear function, denoted also by the sign $\otimes$, which associates with any two vectors, $a \in \mathbb{C}^m$ and $b \in \mathbb{C}^n$, a vector $a \otimes b \in \mathbb{C}^{m\times n}$. In our case the explicit expression for the product is

$$a \otimes b = ab^T,$$

where T denotes matrix transposition .

That the function is bilinear means that:

1. For any $a, b \in \mathbb{C}^m$ and $c \in \mathbb{C}^n$ we have that

$$(a + b) \otimes c = a \otimes c + b \otimes c.$$

2. For any $a \in \mathbb{C}^m$ and $b, c \in \mathbb{C}^n$ we have that

$$a \otimes (b + c) = a \otimes b + a \otimes c.$$

3. For any number γ and vectors $a \in \mathbb{C}^m$ and $b \in \mathbb{C}^n$, we have that

$$(\gamma a) \otimes b = a \otimes (\gamma b) = \gamma (a \otimes b).$$

The reader can verify that with $a \otimes b = ab^T$ the bilinearity is a simple consequence of the most elementary properties of the matrix product. The bilinearity indicates that this "product" behaves like other products (for instance, the distributive property holds). Notice that the dimension of $\mathbb{C}^{m\times n}$ is $m \times n$, which equals the product of the dimensions of $\mathbb{C}^m$ and $\mathbb{C}^n$ ($\dim(H_1 \otimes H_2) = \dim(H_1) \times \dim(H_2)$).

At this stage the reader might wonder how is it that the tensor product of two vectors is a matrix and not a column vector. Well, matrices are also vectors in the sense that they belong to a vector space, in our case the space $\mathbb{C}^{m\times n}$ of all complex-valued $m \times n$ matrices; this space is isomorphic to the space $\mathbb{C}^{mn}$ of all the (column) vectors of length mn. Actually one can transform the matrix $g = a \otimes b = ab^T$ into a column vector just by stacking its rows (converted into vertical columns) one on top of the other, for example:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \implies \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}.$$

In the literature this is the $vect$ function, so we could transform our matrix by means of $vect\left(g^T\right)$ and the result is just the familiar Kronecker product $kron\,(a, b)$ (in MALTAB $vect$ is done with the colon operator, so $vect\,(a) = a\,(:)$). For the $vect$ and $kron$ functions the reader is referred to [6]. We could use, if we wanted, column vectors for the tensor product but, as we will see shortly, the matrix representation has significant practical advantages.

The missing ingredient for a complete definition of the tensor product is the *universal mapping property* or UMP. This is the frightening ingredient, and it will be included in the Appendix B for the curious reader (in appendix A we include some comments on commutative diagrams that are needed for an undersanding of tensor products). In short, the tensor product of two vector spaces V and W is another vector space $V \otimes W$ together with a bilinear mapping $\otimes : V \times W \to V \otimes W$ that satisfies the UMP.

3. Entanglement

Assume a bipartite system $S = S_1 \cup S_2$. A (pure) state of S is a vector $\phi \in H$ and we say that the state ϕ is separable, factorable, non-entangled or simple if ϕ can be written as $\phi = a \otimes b$ for some $a \in H_1$ and $b \in H_2$. Otherwise we say that the state is entangled. If it turns out that the state $\phi = a \otimes b$ is non-entangled, we can say that the subsystem S_1 is in state a and that the subsystem S_2 is in state b. But if the state is entangled such a separation is not possible.

Even worse, if the state is entangled there appear correlations between the properties of the subsystems that are paradoxical and counter-intuitive. Perhaps this was understood most clearly by Woody Allen, who in his book *Without Feathers* [7] described "the bizarre experience of two brothers on opposite parts of the globe, one of whom took a bath while the other suddenly got clean".

It is then of great practical importance to be able to tell if a given state is entangled or not. We will restrict ourselves to bipartite systems, but we will consider spaces of any finite dimension. In case the state is separable we will show explicitly how to produce a factorization.

4. Gaussian reduction

Recall that a given matrix A can be subjected to the so-called elementary operations:

1. Exchange any two rows of A.

2. Multiply a whole row of A (seen as a row vector in $\mathbb{R}^n$ or in $\mathbb{C}^n$) by a number $\gamma \neq 0$.

3. Add to any row a scalar multiple of any other row.

We say that a matrix is in echelon form when:

1. All rows consisting exclusively of zeros lie at the bottom of the matrix.

2. Any row that does not consist exclusively of zeros has as a first non-zero element (from left to right) the number one. This one is called *leading one* or *pivot*.

3. For every pivot, the further down it lies the more to the right it is (echelon).

If in addition, in every column with a pivot all the other elements are zero, we say that the matrix is in row reduced echelon form ($RREF$). In linear algebra text books it is shown that every real or complex matrix can be brought into an echelon form or into a $RREF$ by means of elementary operations. The echelon form is not unique but the $RREF$ is.

The rank of a matrix is the number of non-zero rows in any echelon form; it is the number of leading ones. The transit from a matrix A to its $RREF$ (represented by R) will be represented schematically as $A \to R$ and it is well known that there is a nonsingular matrix B such that $BA = R$. If $[P, Q]$ denotes the matrix P augmented with a matrix Q of the same size[i] then the matrix B can be determined by the calculation $[A, I] \to [R, B]$, where I is an identity matrix of the same size as A and the Gaussian reduction process performed on A augmented with an identity matrix will yield the $RREF$ and (as the rightmost block) the matrix B. Consequently

$$A = B^{-1}R.$$

The product of any two $n\times n$ matrices P and Q can be expressed in the so-called outer product expansion (also known as column-row expansion) as

$$PQ = \sum_{i=1}^{n} \mathrm{col}_i\,(P)\,\mathrm{row}_i\,(Q)\,,$$

where $\mathrm{col}_i\,(P)$ is the i-th column of P (viewed as a column vector, *i.e.* as an $n \times 1$ matrix) and $\mathrm{row}_i\,(Q)$ is the i-th row of Q (viewed as a row vector, *i.e.* as an $1 \times n$ matrix). If we call $p_i = \mathrm{col}_i\,(P)$ and $q_i = (\mathrm{row}_i\,(Q))^T$ then we can rewrite the product as

$$PQ = \sum_{i=1}^{n} p_i q_i^T = \sum_{i=1}^{n} p_i \otimes q_i.$$

When this is applied to the formulas given above for the $RREF$, if it turns out that the rank of A is, say, ρ then, since R will have only ρ non-zero rows, it follows that

$$A = B^{-1}R = \sum_{i=1}^{\rho} a_i b_i^T = \sum_{i=1}^{\rho} a_i \otimes b_i$$

where $a_i = \mathrm{col}_i\left(B^{-1}\right)$ and $b_i = (\mathrm{row}_i\,(R))^T$.

In particular, when $\rho = 1$

$$A = B^{-1}R = a_1 \otimes b_1,$$

thus A equals $a \otimes b$, for some vectors. We have just proved that if A is of rank one then the state is separable. The converse also holds, because if $A = cd^T$ then, for any column vector x we have that $Ax = c(d^T x)$ and the column space of A is one-dimensional (and spanned by c). This means that A has rank one. In short, the state is separable if and only if the corresponding matrix is of rank one. This holds for all values of m and n.

It is important to notice that $A = ac^T$ simply means, in terms of entries of a and c, that $A_{ij} = a_i c_j$.

5. Gauss and entanglement

The key point from the previous section is that a vector in $\mathbb{C}^{m\times n}$ is separable if and only if the rank of the corresponding matrix is one.

For this reason the proposed criterion for assessing entanglement consists in calculating the rank of a matrix. The procedure, simply stated, amounts to:

1. Reduce the matrix to an echelon form.

2. Determine its rank, just counting the number of non-zero rows in the echelon form. If there is only one such a row the state is separable, otherwise it is entangled.

But we can do a lot more that this. Given a separable state we can factor it, that is, write it as a tensor product of two vectors. In general the solution is not unique, we will give an algorithm for finding one. When A is of rank one, since $A = B^{-1}R$ then A can be factored in terms of the first column of B^{-1} and the first row of R. They can be obtained as follows:

1. from $[A, I] \to [R, B]$ we obtain R and B.

2. from $[B, I] \to [I, B^{-1}]$ we obtain B^{-1}.

This is the standard Gaussian algorithm for matrix inversion.

5.1. An example is worth a thousand words

Consider the bipartite state with $m = n = 3$ given by

$$g = \begin{bmatrix} 4 & 5 & 6 \\ 8 & 10 & 12 \\ 12 & 15 & 18 \end{bmatrix}.$$

The augmented matrices are

$$[g, I] = \begin{bmatrix} 4 & 5 & 6 & 1 & 0 & 0 \\ 8 & 10 & 12 & 0 & 1 & 0 \\ 12 & 15 & 18 & 0 & 0 & 1 \end{bmatrix},$$

$$\sim \begin{bmatrix} 1 & \frac{5}{4} & \frac{3}{2} & 0 & 0 & \frac{1}{12} \\ 0 & 0 & 0 & 1 & 0 & -\frac{1}{3} \\ 0 & 0 & 0 & 0 & 1 & -\frac{2}{3} \end{bmatrix},$$

$$[B, I] = \begin{bmatrix} 0 & 0 & \frac{1}{12} & 1 & 0 & 0 \\ 1 & 0 & -\frac{1}{3} & 0 & 1 & 0 \\ 0 & 1 & -\frac{2}{3} & 0 & 0 & 1 \end{bmatrix},$$

$$\sim \begin{bmatrix} 1 & 0 & 0 & 4 & 1 & 0 \\ 0 & 1 & 0 & 8 & 0 & 1 \\ 0 & 0 & 1 & 12 & 0 & 0 \end{bmatrix}.$$

Notice that, for consistency, we have done the full $RREF$ in both cases; however in practice any reduced form showing that the left matrix has only one non-zero row can be used.

Then

$$c = \begin{bmatrix} 1 \\ \frac{5}{4} \\ \frac{3}{2} \end{bmatrix}, \qquad a = \begin{bmatrix} 4 \\ 8 \\ 12 \end{bmatrix},$$

and $g = a \otimes c^T$.

Why textbooks never use this simple algorithm that requires no eigenvectors nor eigenvalues? Possibly the answer is related to the neglect of the concept of tensor product. Surely the reader will have noticed that, in terms of coordinates, our realization of the tensor product is nothing but the so-called *exterior product of matrices*, closely related to the Kronecker product, $a \otimes c = \text{kron}\left(a, c^T\right)$. Our point is that all this is both elementary and useful.

A simple MATLAB program for the factorization is given below.

```
function [ a,c ] = unentangled( g )
%UNENTANGLED If state es separable
%it produces the factors.
if rank(g) ~=1
    disp('the state is entangled')
    return
end
[n,m]=size(g);
al=[g,eye(n,n)];
bet=rref(al);
r=bet(:,1:n);
b=bet(:,n+1:end);
ga=[b,eye(n,n)];
del=rref(ga);
binv=del(:,n+1:end);
a=binv(:,1);
c=r(1,:);
c=c';
end
```

5.2. An even simpler procedure

We have seen in the previous section that the state is separable if and only if the corresponding matrix A has rank one and that, in this case, the matrix can be written as $A = cd^T$ for some column vectors c and d.

But if $A = cd^T$ then, for any column vector x we have that $Ax = c\left(d^T x\right)$, so c lies in the column space of A which is the row space of A^T. Similarly, since $A^T = dc^T$, for any column vector x we have that $A^T x = d\left(c^T x\right)$ so d lies in the column space of A^T which is the row space of A.

A basis for the row space of any matrix can be obtained from the $RREF$, and if the matrix is of dimension one the basis is just the first row of the $RREF$. Then a and c are proportional to the first rows of the $RREF$ of A^T and A respectively.

Thus a simplified approach would be:

1. Calculate the RREF forms of both A^T and A.
2. Extract the non zero row of each and call them r and s, respectively.
3. Form $B = r^T s$, then A is proportional to B, say $A = \gamma B$ for some number γ.
4. γ can be found in a number of ways, our choice is to use the fact that $Tr\left(A\right) = \gamma Tr\left(B\right)$ and then infer γ. This works provided the trace is not zero; in such a case one must perform an element-wise comparison between the elements of A and B. A possibility would be to form element- wise ratios taking care to exclude divisions by zero.

Consider again bipartite state with $m = n = 3$ given by

$$g = \begin{bmatrix} 4 & 5 & 6 \\ 8 & 10 & 12 \\ 12 & 15 & 18 \end{bmatrix}.$$

Performing the Gaussian reduction we see that g has the RREF given by

$$\begin{bmatrix} 1 & \frac{5}{4} & \frac{3}{2} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

so we know that g is separable (non-entangled).

The RREF for g^T is

$$\begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

so the factor a in $A = ac^T$ will be proportional to $\left[1 \;\; 2 \;\; 3\right]$ whereas c will be proportional to $\left[1 \;\; \frac{5}{4} \;\; \frac{3}{2}\right]$ or (removing fractions) to $\left[4 \;\; 5 \;\; 6\right]$ and, as a matter of fact, these are the values for the factors, as the reader is asked to verify.

A simple MATLAB program for this factorization is given below:

```
function [al,bet,factor{] = tangle(g)
% Gaussian approach to entanglement of pure
% bipartite states.

if rank(g)=1
disp('the state is entangled')
return
end
```

```
al=rref(g.');
al=al(1,:);
al=al.';
bet=rref(g);
bet=bet(1,:);
bet=bet.';
factor=trace(g)/trace(al*bet');

%factor*al*bet.' should equal g.
end
```

6. A slightly more general case

If we were told that the state is

$$\psi = \sum_{i=1}^{m} \sum_{j=1}^{n} g_{ij} v_i \otimes w_j,$$

for bases $\{v_1, v_2, \ldots, v_m\}$ and $\{w_1, w_2, \ldots, w_n\}$ and numbers g_{ij}, we could proceed as before but with the matrix g_{ij}. This can be seen explicitly from the fact that if g is separable, then $g_{ij} = a_i c_j$ and

$$\psi = \sum_{i=1}^{m} \sum_{j=1}^{n} a_i c_j v_i \otimes w_j = \left(\sum_{i=1}^{m} a_i v_i \right) \otimes \left(\sum_{j=1}^{n} c_j w_j \right).$$

7. Conclusions

We have shown how simple Gaussian reductions can be used to decide whether a given pure bipartite state is entangled or not and then also make possible to factorize a non-entangled state. Gaussian reductions are the methods first year students learn in order to handle systems of linear equations. Gaussian methods can be applied (for small sized problems) with pencil and paper alone or on the blackboard. No eigenvalues, no eigenvectors, no Schmidt coefficients (singular values) that require the use of a computer and diagonalization software.

The methods here presented have didactic value when explaining the mathematical aspects of entanglement, a compulsory subject in this era of quantum information.

Appendix

A. Commutative diagrams

We dive next into the difficult part, the one that produces panic to lecturers and students alike. The *raison d'être* of tensor products is to be able to visualize any bilinear map as if it were a linear map in some vector space. But in order to explain this we need a little bit of archery [4].

In elementary courses we see that every function F has three ingredients:

1. a domain, call it A.
2. a codomain, call it B.
3. and a rule of correspondence, *i.e.*, a formula, recipe or algorithm that associates a unique $F(a) \in B$ to every $a \in A$.

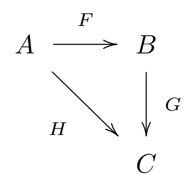
Sometimes this is written as $F{:}A \to B$ but it will prove more convenient to write $A \xrightarrow{F} B$ with identical meaning: F is a function with domain A and codomain B.

If we have two functions $A \xrightarrow{F} B$ and $B \xrightarrow{G} C$, then we can define a new function $A \xrightarrow{G \circ F} C$ by means of

$$(G \circ F)(a) = G(F(a)).$$

This new function is called the *composition* of F and G.

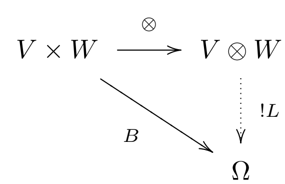
The corresponding diagram would be (with $H = G \circ F$)



So far the diagram only says that we have three functions and tells us what their domains and codomains are. But if H is the composition of F and G then we say that the diagram is commutative. It expresses the idea that going from A to B by means of F and next from B to C by means of G is exactly the same as going directly (non-stop flight) from A to C by means of H.

B. Tensor products

Theorem 1 *Let V, W and Ω be vector spaces over the same field (for instance all real or all complex), and let $B : V \times W \to \Omega$ be a bilinear function. Then there is another vector space $V \otimes W$, together with a bilinear map $\otimes : V \times W \to V \otimes W$ and a unique linear map $L : V \otimes W \to \Omega$, such that the following diagram is commutative:*



Definition 1 *The space $V \otimes W$ is called "tensor product" of the spaces V and W.*

Remark 1 *When an arrow is dashed it means that the existence of the arrow is asserted. The exclamation mark, as in $A \xrightarrow{!} C$ means that the arrow is unique. In our context, arrow, morphism, function and application mean the same.*

In short, all this means that we can replace bilinear maps by linear ones. The linear transformation L has the property

$$L\left(a \otimes b\right) = B\left(a, b\right),$$

so for vectors of the form $L\left(a \otimes b\right)$ gives exactly the same result as $B\left(a, b\right)$. But in $V \otimes W$ **there are many vectors that are not of the form** $a \otimes b$. This is the origin of the entanglement.

As an example of this consider the state represented by the matrix

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix},$$

its $RREF$ is

$$\begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix},$$

so A has rank $\rho = 2$ and can not be separable.

The construction given above with the commutative diagram is an example of what mathematicians call *universal property* or *universal mapping property* (UMP). The tensor product is universal in the sense that it allows us to represent any bilinear function as a linear mapping.

---

i. We say that a matrix A is augmented with matrix B if we form a new matrix having as columns the columns of A and B, for instance if

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} e & f \\ g & h \end{bmatrix},$$

then the augmented matrix is

$$[A, B] = \begin{bmatrix} a & b & e & f \\ c & d & g & h \end{bmatrix}.$$

1. M.A. Nielsen, I.L. Chuang, *Quantum Computation and Quantum Information*, (Cambridge University Press 2000).
2. A. Aczel, *Entanglement, the greatest mystery in Physics* (Four Walls Eight Windows, New York, 2002).
3. C. Isham, *Lectures on Quantum Theory: Mathematical and Structure Foundations* (Imperial College Press, London, 1995).
4. S. Awodey, *Category Theory* (Clarendon Press, Oxford, 2006).
5. G. Strang, *Linear Algebra and its Applications*, Fourth edition, (Wellesley-Cambridge Press, Wellesley MA, 2009).
6. C.F. Van Loan, *Journal of Computational and Applied Mathematics* **123** (2000) 85-100.
7. W. Allen, *Without Feathers* (Ballantine Books, NY, 1983).