

15.11 Tensors in Quantum Mechanics

The chapter ends outside mathematics, and the arrangement is the one Chapter 14 used for Minkowski space and for the second derivative test: physics supplies a modeling assumption, the book quotes it and proves nothing about it, and everything after it is linear algebra. The assumption is that a composite quantum system is the tensor product of its parts. Two things follow. One is a phenomenon, **entanglement**, which turns out to be nothing but §02's observation that not every element of $V \otimes W$ is a simple tensor. The other is a theorem, the **Schmidt decomposition**, which turns out to be Chapter 13's singular value decomposition wearing different notation. The second is the real payoff, and it is why this section closes Part IV: it joins the two halves of the book's later chapters at a single point.

15.11.1 What is quoted and what is proved

Quantum mechanics, as a physical theory, is not derived here and cannot be. What we take from it is a short list of modeling conventions, stated once so that the reader can see exactly where the mathematics starts.

- (Q1) The states of a quantum system are the **unit vectors** of a complex inner product space V , where two unit vectors differing by a scalar of modulus 1 describe the same state.
- (Q2) If a system A has state space V and a system B has state space W , then the composite system “ A together with B ” has state space $V \otimes W$.
- (Q3) If A is in the state v and B is in the state w , then the composite is in the state $v \otimes w$.

None of (Q1)–(Q3) is a theorem, and this book does not argue for any of them. They are quoted exactly as Taylor's theorem was quoted in Chapter 14 §06 — as hypotheses that give the mathematics below something to stand on. Everything proved in this section is a statement about finite-dimensional complex inner product spaces, true whether or not (Q1)–(Q3) describe the world.

Two remarks make the list less arbitrary. First, $\otimes$ rather than $\oplus$ is visible in a dimension count: a spin has a two-dimensional state space, a **qubit**, and two spins together have a four-dimensional one, not a three-dimensional one. Of the two standard constructions, only $\otimes$ produces the observed count, since $\dim(V \otimes W) = \dim V \cdot \dim W$ by Theorem 15.2.5. Second, (Q3) very nearly forces (Q2): the assignment $(v, w) \mapsto (\text{joint state})$ is required to be bilinear, and by Definition 15.2.1 every bilinear map out of $V \times W$ factors uniquely through $V \otimes W$. The universal property is the precise statement that the joint state depends on the parts bilinearly and in no other way.

Genuine quantum mechanics uses infinite-dimensional spaces with a completeness condition. Everything below is finite-dimensional, which is where this book lives, and where qubits live too.

Before states can be unit vectors, $V \otimes W$ needs an inner product. There is only one reasonable candidate, and it exists.

Proposition 15.11.1. The Inner Product on a Tensor Product

Let V and W be finite-dimensional inner product spaces over $F = \mathbb{C}$ (or over $\mathbb{R}$). Then there is **exactly one** inner product on $V \otimes W$ satisfying

$$\langle v \otimes w, v' \otimes w' \rangle = \langle v, v' \rangle \langle w, w' \rangle$$

for all $v, v' \in V$ and $w, w' \in W$. Moreover, if $\mathcal{E}$ and $\mathcal{F}$ are orthonormal bases of V and W , then $(e_i \otimes f_j)$ is an orthonormal basis of $V \otimes W$.

Proof. Uniqueness. The simple tensors span $V \otimes W$ (Proposition 15.2.4), and an inner product is additive in the first slot and conjugate-additive in the second, so its values on all pairs of vectors are determined by its values on pairs of simple tensors.

Existence. Fix orthonormal bases $\mathcal{E} = (e_1, \dots, e_m)$ of V and $\mathcal{F} = (f_1, \dots, f_n)$ of W , which exist by Theorem 11.2.2. By Theorem 15.2.5 the mn vectors $e_i \otimes f_j$ form a basis of $V \otimes W$, so the rule

$$\left\langle \sum_{i,j} c_{ij} e_i \otimes f_j, \sum_{i,j} d_{ij} e_i \otimes f_j \right\rangle := \sum_{i,j} c_{ij} \overline{d_{ij}}$$

is well defined; it is the standard inner product of coordinate vectors, hence satisfies (IP1)–(IP3) of Definition 11.1.1. To check the product formula, write $v = \sum_i a_i e_i$ and $w = \sum_j b_j f_j$. Bilinearity of $\otimes$ gives $v \otimes w = \sum_{i,j} a_i b_j e_i \otimes f_j$, and likewise for $v' \otimes w'$, so

$$\langle v \otimes w, v' \otimes w' \rangle = \sum_{i,j} a_i b_j \overline{a'_i b'_j} = \left(\sum_i a_i \overline{a'_i} \right) \left(\sum_j b_j \overline{b'_j} \right),$$

which is $\langle v, v' \rangle \langle w, w' \rangle$ by Theorem 11.2.1. Finally, the product formula gives $\langle e_i \otimes f_j, e_k \otimes f_l \rangle = \delta_{ik} \delta_{jl}$ for **any** orthonormal $\mathcal{E}$ and $\mathcal{F}$, so the displayed construction did not depend on the pair chosen. This proves the proposition. ■

15.11.2 Entanglement is failure of simplicity

The whole phenomenon fits in the language the chapter already has. A state of the composite is a unit vector of $V \otimes W$, and a state of the form $v \otimes w$ is one in which each part has a state of its own. Section 02 proved that not every vector of a tensor product has that form; a composite state that does not describes a system whose parts have no separate states at all.

A state is entangled when it is a vector of $V \otimes W$ that is not a simple tensor.

Definition 15.11.1. Product State, Entangled State

Let V and W be finite-dimensional complex inner product spaces, and let $z \in V \otimes W$ be a unit vector. Then z is a **product state**, also called **separable**, if

$$z = v \otimes w \quad \text{for some } v \in V, w \in W,$$

and z is **entangled** otherwise. The same words are used for any nonzero $z \in V \otimes W$, the unit-length condition playing no part in the distinction.

Clause by clause. **Product state** is a synonym for “simple tensor”, the term of §02; only the context is new. The definition is relative to the **splitting** $V \otimes W$, not to the vector alone: “entangled” is a statement about a vector *and* a choice of two factors. And v and w need not be unit vectors, though they can be taken so: $\|v \otimes w\| = \|v\| \|w\|$ by [Proposition 15.11.1](#), so rescaling one factor up and the other down costs nothing.

Example. The Bell state

In $\mathbb{C}^2 \otimes \mathbb{C}^2$, write $\mathcal{E} = (e_1, e_2)$ for the standard orthonormal basis of the first factor and $\mathcal{F} = (f_1, f_2)$ for the same basis regarded as sitting in the second factor, the two names being there only to keep the factors apart. Let

$$z_0 := \frac{1}{\sqrt{2}}(e_1 \otimes f_1 + e_2 \otimes f_2).$$

Show that z_0 is a unit vector, and that it is entangled.

Solution. By Proposition 15.11.1 the four vectors $e_i \otimes f_j$ are orthonormal, so $\|z_0\|^2 = \frac{1}{2}(1 + 1) = 1$.

For entanglement there is nothing to prove that §02 has not already proved. Once the two bases are recognized as the same one, $e_1 \otimes f_1 + e_2 \otimes f_2$ is exactly the tensor of Example (A tensor that is not simple), and z_0 is a nonzero multiple of it: if $z_0 = v \otimes w$ then $e_1 \otimes f_1 + e_2 \otimes f_2 = (\sqrt{2}v) \otimes w$, which Example (A tensor that is not simple) rules out. So z_0 is entangled.

The physical reading is one paragraph long, and the book stops there. Under (Q1)–(Q3), a composite in the state z_0 is a pair of qubits of which two things are true at once: the pair has a completely specified state, and neither member has a state at all, since z_0 is not of the form $v \otimes w$. Nothing about the individual qubits is being suppressed; the postulates simply do not assign them states. That is what is meant by calling the two qubits entangled, and it is why the word attaches to a piece of linear algebra rather than to a measurement. Whether the postulates are true of the world is a question this book does not address.

Warning. A sum of simple tensors can still be simple. “Entangled” does not mean “written as a sum of more than one term”. In $\mathbb{C}^2 \otimes \mathbb{C}^2$,

$$\frac{1}{2}(e_1 \otimes f_1 + e_1 \otimes f_2 + e_2 \otimes f_1 + e_2 \otimes f_2) = \frac{e_1 + e_2}{\sqrt{2}} \otimes \frac{f_1 + f_2}{\sqrt{2}},$$

a four-term expression for a product state. Entanglement is a property of the **vector**, not of an expression for it, and deciding it needs an invariant. Supplying one is the job of the rest of this section.

Quick check. Is $z = \frac{1}{\sqrt{3}}(e_1 \otimes f_1 + e_1 \otimes f_2 + e_2 \otimes f_1)$ in $\mathbb{C}^2 \otimes \mathbb{C}^2$ entangled?

Solution. Yes. Suppose $z = v \otimes w$ with $v = a_1 e_1 + a_2 e_2$ and $w = b_1 f_1 + b_2 f_2$. Expanding, $v \otimes w = \sum_{i,j} a_i b_j e_i \otimes f_j$, and the $e_i \otimes f_j$ are a basis (Theorem 15.2.5), so comparing coefficients gives

$$a_1 b_1 = a_1 b_2 = a_2 b_1 = \frac{1}{\sqrt{3}}, \quad a_2 b_2 = 0.$$

Then $(a_1 b_1)(a_2 b_2) = 0$ while $(a_1 b_2)(a_2 b_1) = \frac{1}{3}$; but both products equal $a_1 a_2 b_1 b_2$. So no such v, w exist.

15.11.3 The Schmidt decomposition

The Quick check used a 2×2 determinant on an array of coefficients, and that is a hint about where to look. Section 03 already said what the array is: by Theorem 15.3.9 a tensor **is** a linear map, and a linear map between finite-dimensional spaces is a matrix. Chapter 13 knows exactly one thing about an arbitrary matrix, and it knows it completely.

Write out the identification for the case at hand. Fix bases $\mathcal{E} = (e_1, \dots, e_m)$ of V and $\mathcal{F} = (f_1, \dots, f_n)$ of W , and let $(\varphi_1, \dots, \varphi_n)$ be the basis of W^* dual to $\mathcal{F}$ (Theorem 5.1.3). The map

$$V \times W \rightarrow \mathcal{L}(W^*, V), \quad (v, w) \mapsto (\varphi \mapsto \varphi(w)v),$$

is bilinear in (v, w) , so by Definition 15.2.1 it induces a linear $\Theta: V \otimes W \rightarrow \mathcal{L}(W^*, V)$. It sends $e_i \otimes f_j$ to the map $\varphi_k \mapsto \delta_{jk}e_i$, whose matrix from $(\varphi_1, \dots, \varphi_n)$ to $\mathcal{E}$ is the matrix unit $\mathbf{E}_{ij}$. The matrix units are a basis of $M_{m \times n}(\mathbb{C})$ and taking the matrix of a map in fixed bases is an isomorphism (Theorem 4.1.4), so Θ carries a basis to a basis and is itself an isomorphism. It is Theorem 15.3.9 with the two factors in the other order and W identified with W^{**} (Theorem 5.4.6), and for coordinate spaces it is §02's matrix model Theorem 15.2.7; the letter is shared for that reason. What matters below is the consequence: the matrix of $\Theta(z)$ is the array of coefficients of z in the basis $(e_i \otimes f_j)$.

Theorem 15.11.2. Schmidt Decomposition

Let V and W be finite-dimensional complex inner product spaces of dimensions m and n , and let $z \in V \otimes W$ be nonzero. Then there are an integer r with $1 \leq r \leq \min(m, n)$, real numbers

$$\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0,$$

an orthonormal list $(x_1, \dots, x_r)$ in V and an orthonormal list $(y_1, \dots, y_r)$ in W such that

$$z = \sum_{k=1}^r \sigma_k x_k \otimes y_k. \quad (15.11.1)$$

Moreover r and the numbers $\sigma_1, \dots, \sigma_r$ are determined by z : they do not depend on the lists chosen, nor on any basis. Finally $\|z\|^2 = \sum_k \sigma_k^2$, so z is a unit vector exactly when $\sum_k \sigma_k^2 = 1$.

Idea. The statement is the singular value decomposition of Theorem 13.9.1 with the words changed, so the plan is to change the words back. ① Fix orthonormal bases on both sides; Theorem 15.2.5 turns z into a matrix $\mathbf{C}$ of coefficients, which by the paragraph above is the matrix of the linear map $\Theta(z)$. ② Feed $\mathbf{C}$ to the compact singular value decomposition, Theorem 13.9.3, which over $\mathbb{C}$ still returns **real positive** numbers σ_k and two orthonormal lists of complex column vectors. ③ Read the column vectors back as vectors of V and W ; orthonormal columns become orthonormal vectors because coordinates in an orthonormal basis compute inner products. For uniqueness, run ③ backwards: any expression of the shape Eq. 15.11.1 turns into a factorization $\mathbf{C} = \mathbf{X}\mathbf{D}\mathbf{Y}^\top$ with $\mathbf{X}, \mathbf{Y}$ having orthonormal columns, and then $\mathbf{C}\mathbf{C}^* = \mathbf{X}\mathbf{D}^2\mathbf{X}^*$ hands over the σ_k^2 as the nonzero eigenvalues of a matrix built from z alone.

Proof. Step 1: turn z into a matrix. Fix orthonormal bases $\mathcal{E} = (e_1, \dots, e_m)$ of V and $\mathcal{F} = (f_1, \dots, f_n)$ of W (Theorem 11.2.2). By Theorem 15.2.5 there are unique scalars $c_{ij} \in \mathbb{C}$ with $z = \sum_{i,j} c_{ij} e_i \otimes f_j$. Let $\mathbf{C} \in M_{m \times n}(\mathbb{C})$ be the matrix with entries c_{ij} ; it is the matrix of the linear map $\Theta(z)$ described above, and it is nonzero because z is.

Step 2: decompose the matrix. Put $r = \text{rank } \mathbf{C} \geq 1$. By Theorem 13.9.3 there are orthonormal lists $(\mathbf{u}_1, \dots, \mathbf{u}_r)$ in $\mathbb{C}^m$ and $(\mathbf{v}_1, \dots, \mathbf{v}_r)$ in $\mathbb{C}^n$, and the nonzero singular values $\sigma_1 \geq \dots \geq \sigma_r > 0$ of $\mathbf{C}$ (Definition 13.9.1), with

$$\mathbf{C} = \sum_{k=1}^r \sigma_k \mathbf{u}_k \mathbf{v}_k^*, \quad \text{that is} \quad c_{ij} = \sum_{k=1}^r \sigma_k (\mathbf{u}_k)_i \overline{(\mathbf{v}_k)_j}.$$

The σ_k are real and positive even though $\mathbf{C}$ is complex, since they are square roots of eigenvalues of the positive semidefinite matrix $\mathbf{C}^* \mathbf{C}$. Also $r = \text{rank } \mathbf{C} \leq \min(m, n)$.

Step 3: read the columns back. Set

$$x_k := \sum_{i=1}^m (\mathbf{u}_k)_i e_i \in V,$$

$$y_k := \sum_{j=1}^n \overline{(\mathbf{v}_k)_j} f_j \in W.$$

Substituting the formula for c_{ij} into $z = \sum_{i,j} c_{ij} e_i \otimes f_j$ and interchanging the three finite sums, bilinearity of $\otimes$ gives

$$z = \sum_{k=1}^r \sigma_k \left(\sum_i (\mathbf{u}_k)_i e_i \right) \otimes \left(\sum_j \overline{(\mathbf{v}_k)_j} f_j \right) = \sum_{k=1}^r \sigma_k x_k \otimes y_k.$$

Since $\mathcal{E}$ is orthonormal, Theorem 11.2.1 gives

$$\langle x_k, x_l \rangle = \sum_i (\mathbf{u}_k)_i \overline{(\mathbf{u}_l)_i} = \langle \mathbf{u}_k, \mathbf{u}_l \rangle = \delta_{kl},$$

and likewise, since $\mathcal{F}$ is orthonormal,

$$\langle y_k, y_l \rangle = \sum_j \overline{(\mathbf{v}_k)_j} (\mathbf{v}_l)_j = \langle \mathbf{v}_l, \mathbf{v}_k \rangle = \delta_{kl}.$$

So both lists are orthonormal, and Eq. 15.11.1 holds.

Step 4: uniqueness. Suppose $z = \sum_{k=1}^s \tau_k x'_k \otimes y'_k$ is **any** expression of the stated shape, with $\tau_1 \geq \dots \geq \tau_s > 0$ and both lists orthonormal. Keep the bases $\mathcal{E}, \mathcal{F}$ and write $x'_k = \sum_i \xi_{ik} e_i$ and $y'_k = \sum_j \eta_{jk} f_j$. Let $\mathbf{X} \in M_{m \times s}(\mathbb{C})$ have entries ξ_{ik} and $\mathbf{Y} \in M_{n \times s}(\mathbb{C})$ have entries η_{jk} . Orthonormality of the two lists says, by Theorem 11.2.1 again, that the columns of $\mathbf{X}$ and of $\mathbf{Y}$ are orthonormal, that is

$$\mathbf{X}^* \mathbf{X} = \mathbf{I}_s, \quad \mathbf{Y}^* \mathbf{Y} = \mathbf{I}_s.$$

Expanding the expression for z in the basis $(e_i \otimes f_j)$ and comparing with the unique coefficients c_{ij} of Step 1,

$$c_{ij} = \sum_{k=1}^s \tau_k \xi_{ik} \eta_{jk}, \quad \text{i.e.} \quad \mathbf{C} = \mathbf{X} \mathbf{D} \mathbf{Y}^\top,$$

where $\mathbf{D} = \text{diag}(\tau_1, \dots, \tau_s)$. Taking conjugate transposes and using $\mathbf{Y}^\top \overline{\mathbf{Y}} = \overline{\mathbf{Y}^* \mathbf{Y}} = \mathbf{I}_s$,

$$\mathbf{C}\mathbf{C}^* = \mathbf{X}\mathbf{D}(\mathbf{Y}^\top \overline{\mathbf{Y}})\mathbf{D}\mathbf{X}^* = \mathbf{X}\mathbf{D}^2\mathbf{X}^*.$$

Now read off the eigenvalues of $\mathbf{C}\mathbf{C}^* \in M_m(\mathbb{C})$. For each k , the k -th column $\mathbf{X}\mathbf{e}_k$ satisfies

$$\mathbf{C}\mathbf{C}^*(\mathbf{X}\mathbf{e}_k) = \mathbf{X}\mathbf{D}^2\mathbf{X}^*\mathbf{X}\mathbf{e}_k = \mathbf{X}\mathbf{D}^2\mathbf{e}_k = \tau_k^2(\mathbf{X}\mathbf{e}_k),$$

and these s columns are orthonormal. If instead $\mathbf{z}' \in \mathbb{C}^m$ is orthogonal to every column of $\mathbf{X}$ then $\mathbf{X}^*\mathbf{z}' = \mathbf{0}$, so $\mathbf{C}\mathbf{C}^*\mathbf{z}' = \mathbf{0}$. Since $\mathbb{C}^m = \text{col}(\mathbf{X}) \oplus \text{col}(\mathbf{X})^\perp$ by Theorem 11.3.2, putting an orthonormal basis of $\text{col}(\mathbf{X})^\perp$ beside the columns of $\mathbf{X}$ produces an orthonormal basis of $\mathbb{C}^m$ consisting of eigenvectors of $\mathbf{C}\mathbf{C}^*$. Hence the eigenvalue list of $\mathbf{C}\mathbf{C}^*$, with multiplicity, is $\tau_1^2, \dots, \tau_s^2$ together with $m - s$ zeros.

Every τ_k^2 is positive, so s is the number of nonzero eigenvalues of $\mathbf{C}\mathbf{C}^*$, and $(\tau_1^2, \dots, \tau_s^2)$ is the list of those eigenvalues in decreasing order. Both are determined by $\mathbf{C}$, which is determined by z once $\mathcal{E}$ and $\mathcal{F}$ are fixed. So any two expressions of the shape Eq. 15.11.1 have the same s and the same τ_k ; in particular they agree with the r and σ_k of Steps 2 and 3, and since the statement “ z admits Eq. 15.11.1 with these r, σ_k ” mentions no basis, the numbers cannot depend on $\mathcal{E}$ and $\mathcal{F}$ either.

Step 5: the norm. By Proposition 15.11.1, $\langle x_k \otimes y_k, x_l \otimes y_l \rangle = \langle x_k, x_l \rangle \langle y_k, y_l \rangle = \delta_{kl}$, so the r vectors $x_k \otimes y_k$ are orthonormal in $V \otimes W$. By Theorem 11.1.3, $\|z\|^2 = \sum_k \sigma_k^2$. This proves the theorem. ■

Definition 15.11.2. Schmidt Coefficients, Schmidt Rank

Let $z \in V \otimes W$ be nonzero. The numbers $\sigma_1 \geq \dots \geq \sigma_r > 0$ of Theorem 15.11.2 are the **Schmidt coefficients** of z , and the integer r is the **Schmidt rank** of z . By convention the Schmidt rank of 0 is 0.

Both are well defined precisely because of Step 4, and the proof identifies them: the Schmidt coefficients are the nonzero singular values of the coefficient matrix of z in any pair of orthonormal bases, and the Schmidt rank is the rank of that matrix. That is the invariant the warning above asked for, and it is also §02's: the smallest number of simple tensors summing to z is the rank of $\Theta(z)$, and Theorem 15.11.2 now exhibits a shortest such sum, with its terms orthogonal on both sides at once.

Corollary 15.11.3. Entanglement Is Schmidt Rank at Least Two

A nonzero $z \in V \otimes W$ is a product state if and only if its Schmidt rank is 1. Hence a state is entangled exactly when its Schmidt rank is at least 2.

Proof. ($\Leftarrow$) If $r = 1$ then $z = \sigma_1 x_1 \otimes y_1 = (\sigma_1 x_1) \otimes y_1$, a simple tensor.

($\Rightarrow$) Suppose $z = v \otimes w$. Neither factor is 0, since $z \neq 0$. Put $x_1 = v/\|v\|$ and $y_1 = w/\|w\|$, which is legitimate because $\|v\|, \|w\| > 0$. Then $z = \sigma_1 x_1 \otimes y_1$ with $\sigma_1 = \|v\|\|w\| > 0$, and the one-term lists $(x_1), (y_1)$ are orthonormal. This is an expression of the shape Eq. 15.11.1 with one term, so the uniqueness clause of Theorem 15.11.2 gives $r = 1$.



Warning. The Schmidt coefficients are determined; the vectors are not. For the Bell state z_0 of [Example \(The Bell state\)](#) the coefficient matrix is $\frac{1}{\sqrt{2}}\mathbf{I}_2$, whose singular values are $\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}$. Repeated coefficients leave the lists free, exactly as a repeated eigenvalue of $\mathbf{A}^*\mathbf{A}$ leaves $\mathbf{V}$ free in Theorem 13.9.2: for **every** orthonormal basis (x_1, x_2) of $\mathbb{C}^2$ one has $z_0 = \frac{1}{\sqrt{2}}(x_1 \otimes y_1 + x_2 \otimes y_2)$ for a suitable orthonormal (y_1, y_2) . Never speak of *the* Schmidt vectors; speak of *the* Schmidt coefficients.

Example. A Schmidt decomposition in two by three

In $\mathbb{C}^2 \otimes \mathbb{C}^3$, find the Schmidt decomposition, the Schmidt rank, and the Schmidt coefficients of

$$z = \frac{1}{2}(e_1 \otimes f_1 + e_1 \otimes f_3 + e_2 \otimes f_2 - e_2 \otimes f_3),$$

where $\mathcal{E}$ and $\mathcal{F}$ are the standard orthonormal bases.

Solution. The coefficient matrix is

$$\mathbf{C} = \frac{1}{2} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \end{pmatrix} \in M_{2 \times 3}(\mathbb{R}).$$

This is half the matrix of [Example \(A singular value decomposition of a 2 by 3 matrix\)](#), so by [Exercise \(B4: Determine whether\)](#) (b) its singular values are half of those found there: $\sigma_1 = \frac{\sqrt{3}}{2}$ and $\sigma_2 = \frac{1}{2}$. Both are nonzero, so $\text{rank } \mathbf{C} = 2$ and the Schmidt rank is 2; by [Corollary 15.11.3](#), z is entangled. As a check, $\sigma_1^2 + \sigma_2^2 = \frac{3}{4} + \frac{1}{4} = 1$, and indeed $\|z\|^2 = 4 \cdot \frac{1}{4} = 1$.

For the vectors, Theorem 13.9.3 asks for an orthonormal eigenbasis of the 3×3 matrix $\mathbf{C}^*\mathbf{C}$; as in [Example \(A singular value decomposition of a 2 by 3 matrix\)](#) it is cheaper to work with the smaller $\mathbf{C}\mathbf{C}^\top = \frac{1}{4} \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$, whose orthonormal eigenvectors are the left singular vectors $\mathbf{u}_k$. They are $\frac{1}{\sqrt{2}}(1, -1)$ for the eigenvalue $\frac{3}{4}$ and $\frac{1}{\sqrt{2}}(1, 1)$ for $\frac{1}{4}$. So $\mathbf{u}_1 = \frac{1}{\sqrt{2}}(1, -1)$ and $\mathbf{u}_2 = \frac{1}{\sqrt{2}}(1, 1)$, and the relation $\mathbf{C}^*\mathbf{u}_i = \sigma_i \mathbf{v}_i$ of Eq. 13.9.3 gives

$$\mathbf{v}_1 = \frac{\mathbf{C}^\top \mathbf{u}_1}{\sigma_1} = \frac{1}{\sqrt{6}}(1, -1, 2), \quad \mathbf{v}_2 = \frac{\mathbf{C}^\top \mathbf{u}_2}{\sigma_2} = \frac{1}{\sqrt{2}}(1, 1, 0).$$

All entries are real, so the conjugation in Step 3 of [Theorem 15.11.2](#) does nothing here, and reading the columns back gives

$$\begin{aligned} x_1 &= \frac{1}{\sqrt{2}}(e_1 - e_2), & y_1 &= \frac{1}{\sqrt{6}}(f_1 - f_2 + 2f_3), \\ x_2 &= \frac{1}{\sqrt{2}}(e_1 + e_2), & y_2 &= \frac{1}{\sqrt{2}}(f_1 + f_2), \end{aligned}$$

so that $z = \frac{\sqrt{3}}{2}x_1 \otimes y_1 + \frac{1}{2}x_2 \otimes y_2$. Multiplying out recovers z : the two terms are $\frac{1}{4}(e_1 - e_2) \otimes (f_1 - f_2 + 2f_3)$ and $\frac{1}{4}(e_1 + e_2) \otimes (f_1 + f_2)$, whose coefficients at $e_1 \otimes f_2$ and at $e_2 \otimes f_1$ cancel, leaving $\frac{1}{2}$ at $e_1 \otimes f_1$, $e_1 \otimes f_3$, and $e_2 \otimes f_2$ and $-\frac{1}{2}$ at $e_2 \otimes f_3$.

Remark. Truncating Eq. 15.11.1 after k terms is truncating a singular value decomposition, so Theorem 13.11.3 applies verbatim: the closest state of Schmidt rank at most k is the truncated sum, at distance $(\sum_{i>k} \sigma_i^2)^{1/2}$. Exercise (C2: Best approximation by a state of small Schmidt rank) works this out. The remark also marks the limit of the theorem: everything here concerns **two** factors, because a matrix has two indices. For $V \otimes W \otimes U$ a general vector admits no expression as a sum of simple tensors built from three orthonormal families; the book records that without proof.

15.11.4 Looking back at Part IV

Part IV asked one question in two ways. Chapter 14 took a single bilinear form and asked what it is up to a change of basis; Chapter 15 asked what all multilinear maps have in common, and answered by building the space that turns every one of them into a linear map.

The two chapters answer each other. Chapter 14 worked one form at a time and classified: congruence rather than similarity, complete the square, split off a hyperbolic plane, count signs rather than values. Every invariant it produced — the inertia of Theorem 14.5.1, the Witt index of Definition 14.7.4, the discriminant of Definition 14.2.2 — was a number attached to a form that survives a change of basis, and every classification theorem said that a short list of such numbers is all there is. Chapter 15 worked with all forms at once and constructed: define by universal property and then build a model, impose relations by a quotient, define a map out of a quotient upstairs first, and count a basis by counting index tuples. Where Chapter 14 diagonalized a symmetric form by completing the square, Chapter 15 produced the space $\text{Sym}^2 V$ on which that form is a single linear functional; where Chapter 14 proved that alternating is the characteristic-free strengthening of skew, Chapter 15 built $\Lambda^k V$ by killing $v \otimes v$ for exactly that reason; and where Chapter 14 could only say that a Clifford algebra is unique if it exists, Chapter 15 built one out of $T(V)$ and settled its dimension.

This last section adds a third connection, backwards rather than sideways. The Schmidt decomposition is not a new theorem; it is Theorem 13.9.3 with $V \otimes W$ written in place of $M_{m \times n}$. Part III spent a chapter on what a matrix does to two orthonormal bases, one in each of two spaces; Part IV spent a chapter on the space in which a pair of vectors, one from each of two spaces, becomes a single object. Those are the same statement read in two directions, and Theorem 15.3.9 is the sentence that says so. Asked years from now what the tensor product is for, a reader may answer that it is where a bilinear map becomes linear; the shortest true answer is that $V \otimes W$ is where the two sides of a matrix live.

15.11.5 Exercises

15.11.5.1 A. Check your understanding

Exercise. A1

- (a) State the three modeling assumptions (Q1)–(Q3) of this section, and say which of them this book proves.
- (b) Define **product state** and **entangled state**, and say what data the words depend on besides the vector itself.
- (c) State the Schmidt decomposition, with every hypothesis, and say what is unique in it and what is not.

- (d) Define the Schmidt rank, and give the criterion for entanglement in terms of it.
- (e) True or false: a state written as a sum of three simple tensors has Schmidt rank 3. Justify your answer.
- (f) Which theorem of Chapter 13 is the Schmidt decomposition, and which theorem of §03 licenses the translation?

Solution.

- (a) (Q1) states are unit vectors of a complex inner product space, up to a scalar of modulus 1; (Q2) the state space of a composite system is the tensor product of the state spaces of the parts; (Q3) if the parts are in states v and w , the composite is in the state $v \otimes w$. The book proves none of them: they are quoted from physics.
- (b) A unit vector $z \in V \otimes W$ is a product state if $z = v \otimes w$ for some $v \in V$ and $w \in W$, and entangled otherwise ([Definition 15.11.1](#)). The words depend on the **splitting**: the pair of factors V and W , not the vector alone.
- (c) [Theorem 15.11.2](#): for V, W finite-dimensional complex inner product spaces and $z \in V \otimes W$ nonzero, there are $r \geq 1$, reals $\sigma_1 \geq \cdots \geq \sigma_r > 0$, and orthonormal lists (x_k) in V and (y_k) in W with $z = \sum_k \sigma_k x_k \otimes y_k$. The number r and the σ_k are unique; the lists are not.
- (d) The Schmidt rank is that r ([Definition 15.11.2](#)). By [Corollary 15.11.3](#), z is entangled exactly when $r \geq 2$.
- (e) False. The Schmidt rank is at most 3, but need not equal 3: the warning after [Definition 15.11.1](#) displays a four-term sum whose Schmidt rank is 1.
- (f) It is Theorem 13.9.3, the compact form of the Singular Value Decomposition Theorem 13.9.1. The translation is licensed by Theorem 15.3.9, which identifies a tensor with a linear map and hence, in bases, with a matrix.

15.11.5.2 B. Practice

Exercise. B1: Determine which are entangled

Work in $\mathbb{C}^2 \otimes \mathbb{C}^2$ with the standard orthonormal bases. Determine which of the following unit vectors are entangled. Justify your answer, and for each product state exhibit the two factors.

- (a) $\frac{1}{2}(e_1 \otimes f_1 + e_1 \otimes f_2 + e_2 \otimes f_1 + e_2 \otimes f_2)$
- (b) $\frac{1}{\sqrt{2}}(e_1 \otimes f_2 - e_2 \otimes f_1)$
- (c) $\frac{1}{\sqrt{3}}(e_1 \otimes f_1 + e_1 \otimes f_2 + e_2 \otimes f_2)$
- (d) $\frac{1}{\sqrt{10}}(e_1 \otimes f_1 + 2e_1 \otimes f_2 - e_2 \otimes f_1 - 2e_2 \otimes f_2)$

Solution. By [Corollary 15.11.3](#) and the identification of the Schmidt rank with the rank of the coefficient matrix, it suffices to compute $\text{rank } \mathbf{C}$ in each case, and for a 2×2 matrix that is decided by the determinant.

- (a) $\mathbf{C} = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$, with $\det \mathbf{C} = 0$ and $\mathbf{C} \neq \mathbf{O}$, so $\text{rank } \mathbf{C} = 1$: a product state. Indeed it equals $\frac{1}{\sqrt{2}}(e_1 + e_2) \otimes \frac{1}{\sqrt{2}}(f_1 + f_2)$.

- (b) $\mathbf{C} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, with $\det \mathbf{C} = \frac{1}{2} \neq 0$, so $\text{rank } \mathbf{C} = 2$: entangled.
- (c) $\mathbf{C} = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, with $\det \mathbf{C} = \frac{1}{3} \neq 0$, so $\text{rank } \mathbf{C} = 2$: entangled.
- (d) $\mathbf{C} = \frac{1}{\sqrt{10}} \begin{pmatrix} 1 & 2 \\ -1 & -2 \end{pmatrix}$, with $\det \mathbf{C} = 0$ and $\mathbf{C} \neq \mathbf{O}$, so $\text{rank } \mathbf{C} = 1$: a product state, equal to $\frac{1}{\sqrt{2}}(e_1 - e_2) \otimes \frac{1}{\sqrt{5}}(f_1 + 2f_2)$.

Exercise. B2: A Schmidt decomposition

In $\mathbb{C}^2 \otimes \mathbb{C}^2$, let

$$z = \frac{1}{\sqrt{10}}(e_1 \otimes f_1 + 2e_1 \otimes f_2 + 2e_2 \otimes f_1 + e_2 \otimes f_2).$$

Find the Schmidt coefficients, the Schmidt rank, and a Schmidt decomposition of z , and verify that $\sum_k \sigma_k^2 = 1$.

Solution. The coefficient matrix is $\mathbf{C} = \frac{1}{\sqrt{10}} \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$, which is real symmetric. Then

$$\mathbf{C}\mathbf{C}^\top = \frac{1}{10} \begin{pmatrix} 5 & 4 \\ 4 & 5 \end{pmatrix},$$

with characteristic polynomial $x^2 - x + \frac{9}{100}$ and eigenvalues $\frac{9}{10}$ and $\frac{1}{10}$. So the Schmidt coefficients are $\sigma_1 = \frac{3}{\sqrt{10}}$ and $\sigma_2 = \frac{1}{\sqrt{10}}$, the Schmidt rank is 2, and $\sigma_1^2 + \sigma_2^2 = \frac{9}{10} + \frac{1}{10} = 1$.

Eigenvectors of $\mathbf{C}\mathbf{C}^\top$ are $\mathbf{u}_1 = \frac{1}{\sqrt{2}}(1, 1)$ for $\frac{9}{10}$ and $\mathbf{u}_2 = \frac{1}{\sqrt{2}}(1, -1)$ for $\frac{1}{10}$. Then

$$\mathbf{v}_1 = \frac{\mathbf{C}^\top \mathbf{u}_1}{\sigma_1} = \frac{1}{\sqrt{2}}(1, 1), \quad \mathbf{v}_2 = \frac{\mathbf{C}^\top \mathbf{u}_2}{\sigma_2} = \frac{1}{\sqrt{2}}(-1, 1).$$

All entries are real, so reading the columns back as in Step 3 of [Theorem 15.11.2](#),

$$\begin{aligned} x_1 &= \frac{e_1 + e_2}{\sqrt{2}}, & y_1 &= \frac{f_1 + f_2}{\sqrt{2}}, \\ x_2 &= \frac{e_1 - e_2}{\sqrt{2}}, & y_2 &= \frac{-f_1 + f_2}{\sqrt{2}}, \end{aligned}$$

and $z = \frac{3}{\sqrt{10}} x_1 \otimes y_1 + \frac{1}{\sqrt{10}} x_2 \otimes y_2$. Expanding the two terms gives $\frac{1}{2\sqrt{10}}(3 - 1)e_1 \otimes f_1$ and similarly for the other three coefficients, recovering z .

Exercise. B3: How large can the Schmidt rank be?

Let V and W be complex inner product spaces with $\dim V = m$ and $\dim W = n$.

- (a) Prove that every nonzero $z \in V \otimes W$ has Schmidt rank at most $\min(m, n)$.
- (b) Give, for each r with $1 \leq r \leq \min(m, n)$, a unit vector of Schmidt rank exactly r .

Solution.

- (a) Fix orthonormal bases and let $\mathbf{C} \in M_{m \times n}(\mathbb{C})$ be the coefficient matrix of z . By [Theorem 15.11.2](#) the Schmidt rank is $\text{rank } \mathbf{C}$, and the rank of an $m \times n$ matrix is at most $\min(m, n)$ because $\text{rank } \mathbf{C} = \dim \text{col}(\mathbf{C}) \leq m$ and $\text{rank } \mathbf{C} = \dim \text{row}(\mathbf{C}) \leq n$.

- (b) Take $z_r = \frac{1}{\sqrt{r}} \sum_{k=1}^r e_k \otimes f_k$. The lists $(e_1, \dots, e_r)$ and $(f_1, \dots, f_r)$ are orthonormal and every coefficient $1/\sqrt{r}$ is positive, so this is already an expression of the shape [Eq. 15.11.1](#); by the uniqueness clause the Schmidt rank is r . It is a unit vector since $r \cdot (1/\sqrt{r})^2 = 1$.

15.11.5.3 C. Going deeper

Exercise. C1: Acting on one part at a time

Let $S \in \mathcal{L}(V)$ and $T \in \mathcal{L}(W)$ be unitary, and let $z \in V \otimes W$ be nonzero.

- (a) Prove that $(S \otimes T)(z)$ has the same Schmidt coefficients, and hence the same Schmidt rank, as z .
- (b) Hence deduce that no operator of the form $S \otimes T$ with S, T unitary carries the Bell state z_0 of [Example \(The Bell state\)](#) to a product state.

Hint for (a): apply $S \otimes T$ to a Schmidt decomposition.

Solution.

- (a) Write $z = \sum_{k=1}^r \sigma_k x_k \otimes y_k$ as in [Theorem 15.11.2](#). By Definition 15.3.1, $(S \otimes T)(x_k \otimes y_k) = Sx_k \otimes Ty_k$, so by linearity

$$(S \otimes T)(z) = \sum_{k=1}^r \sigma_k (Sx_k) \otimes (Ty_k).$$

A unitary operator preserves inner products ([Theorem 11.7.3](#)), so $\langle Sx_k, Sx_l \rangle = \langle x_k, x_l \rangle = \delta_{kl}$ and likewise for T . Both lists are therefore orthonormal, the σ_k are still positive and decreasing, and the display is an expression of the shape [Eq. 15.11.1](#) for $(S \otimes T)(z)$. By the uniqueness clause of [Theorem 15.11.2](#), its Schmidt coefficients are $\sigma_1, \dots, \sigma_r$ and its Schmidt rank is r .

- (b) The Schmidt rank of z_0 is 2, as computed in the warning after [Corollary 15.11.3](#). By (a), $(S \otimes T)(z_0)$ also has Schmidt rank 2, and by [Corollary 15.11.3](#) a product state has Schmidt rank 1. So $(S \otimes T)(z_0)$ is entangled.

Exercise. C2: Best approximation by a state of small Schmidt rank

Let $z \in V \otimes W$ be nonzero with Schmidt decomposition $z = \sum_{k=1}^r \sigma_k x_k \otimes y_k$, and let $0 \leq k \leq r$. Put $z_{(k)} = \sum_{i=1}^k \sigma_i x_i \otimes y_i$.

- (a) Fix orthonormal bases of V and W and let $\mathbf{C}$ be the coefficient matrix of z . Prove that $\|z\| = \|\mathbf{C}\|_F$, and that the Schmidt rank of a vector equals the rank of its coefficient matrix.
- (b) Hence deduce that for every $z' \in V \otimes W$ of Schmidt rank at most k ,

$$\|z - z'\| \geq \left(\sum_{i>k} \sigma_i^2 \right)^{1/2},$$

with equality for $z' = z_{(k)}$.

Solution.

- (a) By [Proposition 15.11.1](#) the basis $(e_i \otimes f_j)$ of $V \otimes W$ is orthonormal, so $\|z\|^2 = \sum_{i,j} |c_{ij}|^2$ by [Theorem 11.9.2](#), and that sum is $\|\mathbf{C}\|_F^2$ by the definition of the Frobenius norm. For the second claim, Step 2 of [Theorem 15.11.2](#) produced the Schmidt coefficients as the nonzero singular values of $\mathbf{C}$, and [Theorem 13.9.1](#) says there are $\text{rank } \mathbf{C}$ of them.
- (b) The map $z \mapsto \mathbf{C}$ is the coordinate isomorphism attached to the basis $(e_i \otimes f_j)$ ([Corollary 2.5.5](#)), with the mn coordinates arranged in a rectangle, so it is linear and bijective; and by (a) it carries $\|\cdot\|$ to $\|\cdot\|_F$ and Schmidt rank to matrix rank. Let $\mathbf{B}$ be the coefficient matrix of z' , so $\text{rank } \mathbf{B} \leq k$. By [Theorem 13.11.3](#),

$$\|z - z'\| = \|\mathbf{C} - \mathbf{B}\|_F \geq \left(\sum_{i>k} \sigma_i^2 \right)^{1/2},$$

since the nonzero σ_i are the singular values of $\mathbf{C}$. Running Step 3 of [Theorem 15.11.2](#) backwards turns $z_{(k)}$ into the rank- k truncation $\mathbf{C}_k$ ([Definition 13.11.1](#)) of a compact singular value decomposition of $\mathbf{C}$ ([Theorem 13.9.3](#)), and [Theorem 13.11.3](#) attains equality there.

Exercise. C3: Does the field matter?

Let V and W be **real** inner product spaces of dimensions m and n , with orthonormal bases $\mathcal{E}$ and $\mathcal{F}$, and let $z \in V \otimes W$ have coefficient matrix $\mathbf{C} \in M_{m \times n}(\mathbb{R})$. Let $z_{\mathbb{C}} \in \mathbb{C}^m \otimes \mathbb{C}^n$ be the vector whose coefficient matrix in the standard orthonormal bases is the same $\mathbf{C}$.

- (a) Prove that [Theorem 15.11.2](#) holds verbatim over $\mathbb{R}$, with orthonormal lists of real vectors.
- (b) Prove that z and $z_{\mathbb{C}}$ have the same Schmidt coefficients and the same Schmidt rank.
- (c) Is the same true of the Schmidt **vectors**? Justify your answer.

Solution.

- (a) Every step of the proof runs over $F = \mathbb{R}$. [Theorem 11.2.2](#), [Theorem 15.2.5](#), [Theorem 13.9.3](#), [Theorem 11.2.1](#), and [Theorem 11.3.2](#) are all stated for $F = \mathbb{R}$ or $\mathbb{C}$, and over $\mathbb{R}$ the conjugations in Step 3 and Step 4 are the identity, so $y_k = \sum_j (\mathbf{v}_k)_j f_j$ and $\mathbf{C}\mathbf{C}^{\top} = \mathbf{X}\mathbf{D}^2\mathbf{X}^{\top}$.
- (b) By Step 4 of [Theorem 15.11.2](#), read over each field in turn, the Schmidt coefficients of z and of $z_{\mathbb{C}}$ are both times the positive square roots of the nonzero eigenvalues of $\mathbf{C}\mathbf{C}^{\top}$, and both Schmidt ranks equal $\text{rank } \mathbf{C}$. This is one and the same real symmetric matrix with one and the same characteristic polynomial, so the eigenvalues agree; and $\text{rank } \mathbf{C}$ is unchanged because row reduction over $\mathbb{R}$ is row reduction over $\mathbb{C}$.
- (c) No. The warning after [Corollary 15.11.3](#) says why: a repeated Schmidt coefficient leaves the vectors free, and over $\mathbb{C}$ there is more freedom to use. For z_0 the orthonormal pair $x_1 = \frac{1}{\sqrt{2}}(e_1 + ie_2)$, $x_2 = \frac{1}{\sqrt{2}}(e_1 - ie_2)$ occurs in a complex Schmidt decomposition —

with $y_1 = \frac{1}{\sqrt{2}}(f_1 - if_2)$ and $y_2 = \frac{1}{\sqrt{2}}(f_1 + if_2)$ — and has no real counterpart. What is invariant is the coefficient list, not the vectors.