Householder factorizations of unitary matrices

Jesús Urías*

Instituto de Física, UASLP. San Luis Potosí, SLP, México.

Abstract

A method to construct all representations of finite dimensional unitary matrices as the product of Householder reflections is given. By arbitrarily severing the state space into orthogonal subspaces, the method may, e.g., identify the entangling and single-component quantum operations that are required in the engineering of quantum states of composite (multi-partite) systems. Earlier constructions are shown to be extreme cases of the unifying scheme that is presented here.

PACS numbers: 03.65.Ud

Keywords: Representations of unitary operators – Engineering of quantum states – Quantum optics.

*jurias@ifisica.uaslp.mx

I. INTRODUCTION

The strong superposition principle of quantum mechanics tells us that every orthonormal basis represents an exhaustive test. In his book [1], Peres takes note that the principle does not tell us how to actually perform the test. This "inverse problem" is at the heart of the engineering of quantum systems. The problem has two faces. One is to write the unitary transformation $U \in U(N) \subset \mathbb{C}^{N\times N}$ that resolves quantum states into eigenstates of an observable. The other is to build the hardware realizing U.

The standard bridge from $U(N)$ to the hardware is to represent a quantum operation as the product of simpler unitary factors. In the field of quantum information it is well known [2] that any quantum operation from $U(2^n)$, acting on n-quibit states, may be factorized into single qubit operations (meaning a $U(2)$ factor acting on a particular $\mathbb{C}^2$ subspace of the joint state space) and operations entangling pairs of qubits ($U(4)$ factors acting on a particular $\mathbb{C}^2 \times \mathbb{C}^2$ subspace). The inverse problem, operatively, is to find a sequence of quantum gates producing the desired matrix in $U(N)$.

A systematic approach to the factorization problem was initiated in reference [3], wherein instructions to prepare arbitrary one-photon quantum states in multiple optical beams were supplied: $U(N)$ is represented by a product of at most $N(N-1)/2$ elementary $U(2)$ factors, each of them realizable as a Mach-Zehnder interferometer. Afterwards, in a variant [4] of the factorization method in [3] any matrix in $U(N)$ is represented as a sequences of $N-1$ factors at most. Factors in [4] are taken from $U(N)$, $U(N-1)$, ..., $U(2)$ succesively and some implemantations require, for each factor in the sequence, a number of fields and/or interactions to be applyied simultaneously on the physical system during well-controlled time intervals.

It is essential thus that factors in a sequence for $U(N)$ fit the physical nature of the system implementation. This is a feasibility problem. The physical nature (and formal description) of the set of elementary quantum operations is dictated by the application. It is, of course, far from being a unique set. Compare, for instance, the following references: [2–8].

In this article a method to construct all representations of $U(N)$ as the product of Householder factors is presented (as Theorem III.1). Our method provides a unifying scheme to attend to the feasibility problem and earlier constructions [3, 4] appear as two extreme cases of it.

To deal with the feasibility condition on unitary factors for, e.g., composite systems, the factorization method in Theorem III.1 allows us to decide freely on the sub-space for each factor in a sequence, with the aim of fitting the requirements of a particular implementation. The relevance of this adaptability of the method is illustrated by the following.

In the field of quantum information joint states are tensor products of n qubits, fact that endows the joint state space with a natural partitioning into two dimensional subspaces. In this setup, the quantum realization of the Fourier transform which is quadratic in the number of $U(2)$ factors is the natural choice since the factors combine in the form of gates acting on the "right" $\mathbb{C}^2 \times \mathbb{C}^2$ subspaces as to entangle pairs of qubits. This is a suitable choice of subspaces in the factorization of the Fourier transform for qubit systems.

While U(2) factorizations produce the longest sequences, factors have, in general, an easier implementation than $U(k > 2)$ factors. In the engineering of one-photon states that makes use of lossless optical devices [5, 6], any $U(2)$ factor is known to correspond to a Mach-Zehnder interferometer [3]. The representation of arbitrary unitary operations as sequences of $U(2)$ factors makes the M-Z interferometer a promising candidate to become the elementary building block for one-photon integrated optics (as the transistor is for electronics). Theorem III.1 provides us with a method to search, among all $U(2)$ sequences, the one having the most convenient (regarding implemantation) $U(2)$ factors.

The length of a sequence is not necessarilly an issue. By using Theorem III.1 we may produce factorizations of a Fourier transform in $U(N)$ that go from linear to quadratic in N for the number of "single" quantum operations. It is the nature of the quantum system what determines the nature of single quantum operations to choose in a sequence. Generally, a $U(k)$ factor becomes harder to relize as a piece of hardware as the value of k is increased. Theoretically, any $U(k)$ is realizable in some implementations [4] by applying k fields to the system, simultaneously and during a well controlled time interval. The fact is that such general single quantum operations, for arbitrary values of k, are difficult to implement for, e.g., trapped-ion systems and an alternative and approximate factorization scheme is, for instance, described in reference[9]. As a rule of thumb, longer factorizations involve simpler (in terms of hardware) factors and vice versa.

The article contains the following specific subjects. Given an orthonormal vector basis Y of $\mathbb{C}^N$, the idea of the factorization method is to take each of the vectors in the transformed basis UY to the corresponding vector in the original basis Y, individually, in a succession of Householder reflections along pre-selected subspaces. What we accomplish is a first characterization of $U(N)$ which we state as Lemma III.2. The next subject is to prove that the composition of all such steps for all the vectors in the orthonormal basis Y constitutes a characterization (in the form of a factorization) of $U(N)$. The formal statement is Theorem III.1.

The proof of Theorem III.1 that is presented in Section III is nothing else but a method to systematize the search of Householder factorizations of any $U \in U(N)$. A rather poor estimate we made of the number of different ways to proceed in the factorization of matrix U, for a given vector basis, shows us that it is not lesser than $N! \times B_2 \times B_3 \times \cdots \times B_{N-1}$, which is quite a number (B_n is Bell number, for which there is not a simple formula).

In the Sections to follow, we show that the strategies in [3, 8] and [4, 10] are examples of the engineering of quantum states stemming from two extreme cases of Theorem III.1. One extreme case consists of representations by the shortest sequences, having $N-1$ factors at most. The case is simple enough as to admit an explicit recording of all the factor matrices and phase factors involved in the factorizations of U. This we do in Section IV and found that there are no more than $N!$ such representations for a given $U \in U(N)$.

The second extreme case produces the longest factorizations of U into $U(2)$ factors, $N(N-1)/2$ of them at most. This restriction on Theorem III.1 is treated in Section V. For every unitary operation the method provides no more than super-factorial $N\$ \equiv N!\,(N-1)! \cdots 1$ different factorizations into $U(2)$ quantum operations. Every $U(2)$ sequence constitutes a parametrization of $U(N)$. We conclude by indicating the connection of this extreme case of our method with Murnaghan's $U(2)$ factorization [11].

II. HOUSEHOLDER REFLECTIONS

For any two vectors x and $y \in \mathbb{C}^N$, the inner product is denoted x^*y and the norm $\|x\| = \sqrt{x^*x}$. We are overloading symbol *. For a matrix U, its Hermitian conjugate is denoted U^* and the complex conjugate of $z \in \mathbb{C}$ is denoted z^*. The meaning of * follows from the context.

Householder reflections on $\mathbb{C}^N$ are defined as follows. For any vector $x \in \mathbb{C}^N$, we associate the reflection operator $[x] \in \mathbb{C}^{N\times N}$ producing the transformations

$$[x]\,y = -y \text{ for any } y \in \langle x\rangle \quad \text{and} \quad [x]\,y = y \text{ for any } y \in \langle x\rangle^{\perp},$$

where $\langle x\rangle$ is the linear span of x and $\langle x\rangle^\perp$ denotes the orthogonal complement of $\langle x\rangle$. Notice that $\texttt{[}x\texttt{]} = \texttt{[}x'\texttt{]}$ if and only if the reflecting vectors, x and x', are parallel, $x \parallel x'$. Normalization of the reflecting vector x in the definition of $\texttt{[}x\texttt{]}$ is irrelevant. Every Householder reflection $\texttt{[}x\texttt{]}$ is unitary, hermitian and idempotent, $\texttt{[}x\texttt{]}^2 = \mathbb{1}$. Every reflection $(x \neq 0)$ may be written as $\texttt{[}x\texttt{]} = \mathbb{1} - 2\mathbb{1}_{\langle x\rangle}$, where $\mathbb{1}_{\langle x\rangle}$ is the orthogonal projection onto subspace $\langle x\rangle$. For future convenience we define $\texttt{[}x\texttt{]} = \mathbb{1}$ in the case vector x is the null vector, $x = 0$. Eventhough $\mathbb{1}$ is not a reflection.

Householder reflections are the most economical unitary transformations to exchange pairs of vectors in $\mathbb{C}^N$. Given any two vectors, the Householder reflection to exchange them (including the correct phase factor) is provided by the following.

Lemma II.1 *Let vectors $x \neq y \in \mathbb{C}^N$, both x and $y \neq 0$. There is one and only one Householder reflection $\texttt{[}r\texttt{]}$ such that*

$$\texttt{[}r\texttt{]}\,x = z^*(\|x\|/\|y\|)y \ . \tag{1}$$

The unit complex number z has $\angle z = \angle x^* y$ and $r = z\|y\|x - \|x\|y$, not necessarily normalized.

Proof. The complex number z has been chosen as to have the orthogonality relation $(zx\|y\| - y\|x\|) \perp (zx\|y\| + y\|x\|)$, which may be verified by a direct calculation of the inner product

$$(zx\|y\| - y\|x\|)^*(zx\|y\| + y\|x\|) = 0 \ .$$

Then, for the given vector r we have that $\texttt{[}r\texttt{]}\,(zx\|y\| + y\|x\|) = zx\|y\| + y\|x\|$. □

Notice that the case $x \parallel y$ is not excluded in Lemma II.1. In this case the reflecting vector r is the null vector, $r = 0$, for which we have defined matrix $\texttt{[}r = 0\texttt{]}$ as the identity. This convention makes the job in (1). Thus, given any two vectors, x and y, Lemma II.1 provides the most economical unitary transformation in $\mathbb{C}^N$ that produces the exchange $z\|y\|x \leftrightarrow \|x\|y$. The economy refers to the fact that $\texttt{[}r\texttt{]}$ is the identity on $\langle x, y\rangle^\perp$, the orthogonal complement of the 2-dimensional subspace $\langle x, y\rangle$.

III. MAIN FACTORIZATION THEOREM

The transformation $U \in U(N) \subset \mathbb{C}^{N\times N}$ takes any orthonormal vector basis $Y = (y_1, \dots, y_N)$ to the basis $UY \equiv (Uy_1, \dots, Uy_N)$, which is orthonormal too. The idea of the factorization is to take each of the vectors in the transformed basis —beginning with, say, vector Uy_q— to the corresponding vector y_q in the original basis in several (or many) "partial" steps. Lemma II.1 will be adopted as the method to take such steps. From this idea we get a first characterization of unitarity stated in Lemma III.2 below. The main result in this Section states that the composition of all such steps for all the vectors in any orthonormal basis constitutes a characterization (in the form of a factorization) of unitary matrices. It is Theorem III.1, which has two extreme versions that we develop as examples in Sections IV and V.

Let us prepare the setup for the first step in the factorization of U. From basis Y select the vector y_q. There are N choices for y_q. However, to keep the exposition clear, let us say it is y_1. Then prescribe a partition $\{\mathcal{P}_i\}$ of $Y \setminus \{y_1\}$ into non-empty subsets $\mathcal{P}_i$. Let $s_1 \in \{1, \dots, N-1\}$ denote the number of elements in the partition set $\{\mathcal{P}_i\}$. Bell number B_{N-1} is the number of different choices for such a partition. Then, assign an ordering to the partition set. Let it be $\mathcal{P}_1, \dots, \mathcal{P}_{s_1}$, one out of $s_1!$ different orderings. We see that the total number of choices to start the first step are never lesser than NB_{N-1} (different orderings of the partition set are not being counted).

The partition element $\mathcal{P}_i \subset Y$ determine the subspace $\mathcal{I}_i \equiv \langle \mathcal{P}_i \rangle$ having dimension $\dim \mathcal{I}_i \equiv n_i$, $i = 1$ to s_1. The ordered set of dimension integers $(n_i)_1^{s_1}$ constitute a partition of $N-1$, $\sum_i n_i = N-1$. Every subspace $\mathcal{I}_i$ is orthogonal to $\langle y_1 \rangle$ and they are mutually orthogonal: $\mathcal{I}_i \perp \mathcal{I}_j$ whenever $i \neq j$. We have the direct sum decomposition of $\mathbb{C}^N = \langle y_1 \rangle \oplus \mathcal{I}_1 \oplus \cdots \oplus \mathcal{I}_{s_1}$.

The first step in the factorization procedure takes vector Uy_1 to vector y_1 by hopping along the subspaces $\mathcal{H}_i := \langle y_1 \rangle \oplus \mathcal{I}_i$, successively from $i = 1$ to s_1, in the prescribed order. The orthogonal projection to subspace $\mathcal{H}_i$ is denoted $\mathbb{1}_i$.

Let us proceed. First, vector Uy_1 is projected to subspace $\mathcal{H}_1$ as the vector $x_{1,1} := \mathbb{1}_1 U y_1$. Then, we make use of Lemma II.1 to exchange just the vectors $x_{1,1}$ and y_1 by means of the Householder reflection, denoted $\texttt{[1,1]}$, reflecting $\mathbb{C}^N$ along the direction of vector $r_{1,1} =$

$z_{1,1}x_{1,1} - y_1\|x_{1,1}\| \in \mathcal{H}_1$, where $z_{1,1} = e^{i\varphi_{1,1}}$ and $\varphi_{1,1} = \angle\, {x_{1,1}}^* y_1$. We have that

$$\texttt{[1,1]}\; x_{1,1} = z^*_{1,1}\, \|x_{1,1}\|\, y_1 \ . \tag{2}$$

Notice that we might get $x_{1,1} \parallel y_1$, situation that we did not exclude in Lemma II.1. It occurs if and only if the reflecting vector is the null vector, $r_{1,1} = 0$. In this case matrix $\texttt{[1,1]}$ is conveniently defined as the identity.

The important remark about the Householder matrix $\texttt{[1,1]}$ we just have computed is that

$$y^*\,\texttt{[1,1]}\,U y_1 = 0, \quad \forall y \in \mathcal{I}_1 \ . \tag{3}$$

So, for every j such that $y_j \in \mathcal{I}_1$: the $(j,1)$ entry of the matrix $V_1^{(1)} := \texttt{[1,1]}\,U$ vanishes,

$$ {y_j}^*\, V_1^{(1)} y_1 = 0 \ , \quad y_j \in \mathcal{I}_1 \ .$$

To see why (3) holds, expand vector Uy_1 as a direct sum in $\mathcal{H}_1 \oplus \mathcal{H}_1^\perp$ and write

$$\texttt{[1,1]}\; U y_1 = \texttt{[1,1]}\; \mathbb{1}_1\, U y_1 + \texttt{[1,1]}\; \mathbb{1}_{1^\perp}\, U y_1 \ .$$

The sensible point is that by definition of reflection $\texttt{[1,1]}$ in (2) we have that $\texttt{[1,1]}\,\mathbb{1}_1 U y_1 \equiv \texttt{[1,1]}\,x_{1,1} \parallel y_1 \perp \mathcal{I}_1$. While for the component $\mathcal{H}_1^\perp$ we just have that $\texttt{[1,1]}\,\mathbb{1}_{1^\perp} x = \mathbb{1}_{1^\perp} x \perp \mathcal{I}_1$ for any vector $x \in \mathbb{C}^N$. Thus, while the reflection $\texttt{[1,1]}$ is the identity on $\mathcal{H}_1^\perp$ it is not on $\mathcal{H}_1$. This situation is referred to by saying that reflection $\texttt{[1,1]}$ has a block size not greater than $n_1 + 1 \equiv \dim\mathcal{H}_1$. The actual block size depends on the basis Y. However, in any basis, the block size of $\texttt{[1,1]}$ is, by definition, never lesser than 2. Unless it happens by accident to be the identity.

Next, the same treatment is given to matrix $V_1^{(1)} = \texttt{[1,1]}\,U$. The method in Lemma II.1 is applied this time on vectors y_1 and the projection $x_{1,2} = \mathbb{1}_2 V_1^{(1)} y_1$ to subspace $\mathcal{H}_2$. The new Householder matrix exchanging vectors y_1 and $x_{1,2}$ is denoted $\texttt{[1,2]}$. The relevant yield is matrix

$$V_1^{(2)} := \texttt{[1,2]}\,V_1^{(1)} = \texttt{[1,2]}\ \texttt{[1,1]}\,U \ . \tag{4}$$

In Lemma III.1 below we will prove that it has the property

$$y^* V_1^{(2)} y_1 = 0 \ , \quad \forall y \in \mathcal{I}_1 \oplus \mathcal{I}_2 \ . \tag{5}$$

At the moment we keep going on computing Householder matrices, up to exhaust the collection of subspaces $\mathcal{H}_i$, $i = 1, \ldots, s_1$. The summing-up is matrix $V_1^{(s_1)} \equiv V_1$,

$$V_1 := \texttt{[1,}s_1\texttt{]}\ \cdots\ \texttt{[1,}i\texttt{]}\ \cdots\ \texttt{[1,1]}\ U \ . \tag{6}$$

Index i in $[1,i]$ refers to the "transit" subspace $\mathcal{H}_i$. The Householder matrix $[1,i]$ is block diagonal in the direct sum decomposition $\mathbb{C}^N = \mathcal{H}_i \oplus \mathcal{H}_i^\perp$ and has block-size at most $n_i + 1 = \dim\mathcal{H}_i$. The block of $[1,i]$ corresponding to its restriction on $\mathcal{H}_i^\perp$ is the identity.

A block-diagonal form of the product matrix V_1 in (6) is disclosed in the following.

Lemma III.1 *Matrix V_1 in (6) transforms the subspace $\langle y_1\rangle^\perp$ unitarily.*

Proof. By construction V_1 is unitary. Thus, we are to prove that $y_j^* V_1 y_1 = 0$ for every $j \neq 1$. Let us then prove that for every j (from 1 to s_1),

$$y^* V_1^{(j)} y_1 = 0 \ , \quad \forall y \in \mathcal{I}_1 \oplus \cdots \oplus \mathcal{I}_{j-1} \oplus \mathcal{I}_j \ . \tag{7}$$

For $j = 1$ we proved (3) already. Next, let $p > 1$ be given and assume (7) holds from $j = 1$ to $p-1$. We prove in two steps that (7) holds for $j = p$ too. First take a vector $y \in \mathcal{I}_1 \oplus \cdots \oplus \mathcal{I}_{p-1}$. Since $y \perp \mathcal{H}_p$ we have that $[1,p]\ y = y$ so that

$$y^* V_1^{(p)} y_1 = (\,[1,p]\ y)^* V_1^{(p-1)} y_1 = y^* V_1^{(p-1)} y_1 = 0$$

by hypothesis.

To deal with $y \in \mathcal{I}_p$ recall that $[1,p]$ is block diagonal in the decomposition $\mathbb{C}^N = \mathcal{H}_p \oplus \mathcal{H}_p^\perp$. Consider then the direct sum $V_1^{(p)} y_1 = \mathbb{I}_p V_1^{(p)} y_1 + \mathbb{I}_{p^\perp} V_1^{(p)} y_1$. Since vector y is orthogonal to $\mathcal{H}_p^\perp$ (because $\mathcal{I}_p \perp \mathcal{H}_p^\perp$) we just have to care about the component of $V_1^{(p)} y_1$ in $\mathcal{H}_p$. For it we have that

$$\mathbb{I}_p\,[1,p]\ V_1^{(p-1)} y_1 = [1,p]\ \mathbb{I}_p V_1^{(p-1)} y_1 \parallel y_1 \perp y \ ,$$

where the parallel relation holds because $x_{1,p} = \mathbb{I}_p V_1^{(p-1)} y_1$. □

Lemma III.1 tells us that the matrix V_1 we have computed as the product in (6) is block-diagonal in the direct sum decomposition $\mathbb{C}^N = \langle y_1\rangle \oplus \langle y_1\rangle^\perp$. Thus, we write

$$V_1 = \tilde{U}_1\, D_1 \tag{8}$$

where $D_1 y_1 = z_1 y_1$, with z_1 a phase factor, and D_1 is the identity on $\langle y_1\rangle^\perp$. The factor $\tilde{U}_1$ is a unitary matrix too which is the identity on $\langle y_1\rangle$ and its action on $\langle y_1\rangle^\perp$ coincides with that of V_1. In the basis Y, matrix D_1 is diagonal,

$$D_1 = \mathrm{diag}(z_1, \underbrace{1,\ldots,1}_{N-1}) \ ,$$

and $\tilde{U}_1$ is block-diagonal.

Taking $Uy_1 \mapsto y_1$ by hopping along subspaces $\mathcal{H}_1, \ldots, \mathcal{H}_{s_1}$ we have constructed matrix V_1 as the product given in (6). Inverting it gives us the "partial factorization" (associated to vector y_1) of U,

$$U = [1,1]\ [1,2] \cdots [1,s_1]\tilde{U}_1 D_1\ . \tag{9}$$

Formally, what we got is the following characterization of unitarity.

Lemma III.2 *$U \in U(N)$ if and only if the following statement holds.*

For every vector $y \in \mathbb{C}^N$ and every ordered partition $(n_i)_{i=1}^{\ell}$ of $N-1$, $1 \le \ell \le N-1$, there exist Householder reflections $\{h_i\}_{i=1}^{\ell}$ such that $Uy = zh_1 \cdots h_\ell y$. The block-size of each h_i is not greater than $n_i + 1$ and z is a phase factor.

Notice that existence of Householder factors in Lemma III.2 has been proved by providing a method to actually compute them.

Next, we go for a full factorization of U (not just the one given for vector y_1). Let us organize what we have done in the following terms. Let $H_1 := [1,1]\,[1,2] \cdots [1,s_1]$. By Lemma III.2 we have that $H_1^* U y_1 = z_1 y_1$. Then, define $U_1 \in U(N)$ such that $U_1 y_1 = z_1 y_1$ and that $U_1 = H_1^* U$ when restricted to $\langle y_1\rangle^\perp$, expressed in the form $U_1\big|\langle y_1\rangle^\perp \equiv H_1^* U\big|\langle y_1\rangle^\perp$.

Again by Lemma III.2, given the basis vector y_2, given a positive integer $s_2 \le N-2$ and given any ordered partition $(n_{2,i})_1^{s_2}$ of $N-2$ there exist Householder reflections $[2,1]$, $[2,2]$, $\ldots$, $[2,s_2]$, such that $H_2^* U_1 y_2 = z_2 y_2$, with $H_2 := [2,1]\,[2,2] \cdots [2,s_2]$ and z_2 a phase factor. Then define $U_2 \in U(N)$ such that $U_2 y_i = z_i y_i$, $i \le 2$, and that $U_2\big|\langle y_1, y_2\rangle^\perp \equiv H_2^* U_1\big|\langle y_1, y_2\rangle^\perp$. By definition we have that $H_2^* U_1 = H_2^* H_1^* U$.

When we have gone as far as the k–th basis vector y_k, have given a positive integer s_k which is not greater than $N-k$ and an ordered partition $(n_{k,i})_i^{s_k}$ of $N-k$, then Lemma III.2 tells us that there exist Householder reflections $[k,1]$, $[k,2]$, $\ldots$, $[k,s_k]$ which allow us to define $U_k \in U(N)$ such that $U_k y_i = z_i y_i$, $i \le k$ (for some z_k which is a phase factor too), and that $U_k\big|\langle y_1, \ldots, y_k\rangle^\perp \equiv H_k^* U_{k-1}\big|\langle y_1, \ldots, y_k\rangle^\perp$, with $H_k := [k,1]\,[k,2] \cdots [k,s_k]$. Just by definition, we have that $U_k = H_k^* U_{k-1} = H_k^* H_{k-1}^* \cdots H_1^* U$.

The recursive action we are describing stops at $k = N-1$ with matrix U_{N-1} being defined by the relations $U_{N-1} y_i = z_i y_i$ for $i \le N-1$ and

$$U_{N-1}\big|\langle y_1, \ldots, y_{N-1}\rangle^\perp \equiv H_{N-1}^* U_{N-2}\big|\langle y_1, \ldots, y_{N-1}\rangle^\perp =: z_N\ ,$$

which is a phase factor, by unitarity. We identify the matrix of phases

$$D \equiv U_{N-1} = \operatorname{diag}(z_1,\ \ldots,\ z_N).$$

We further have that $U_{N-1} = H^*_{N-1}U_{N-2} = H^*_{N-1}H^*_{N-2}\cdots H^*_1 U = D$.

What we just got is another characterization of unitarity.

Theorem III.1 *$U \in U(N)$ if and only if the following statement holds.*

For $k = 1$ to $N-1$, let positive integers s_k be such that $s_k \leq N-k$ and let $(n_{k,i})_{i=1}^{s_k}$ be ordered partitions of $N-k$. Then, there exist Householder reflections $[k,i] \in \mathbb{C}^{N\times N}$, $i = 1, \ldots, s_k$, such that

$$U = \underbrace{[1,1]\cdots[1,s_1-1]\,[1,s_1]}_{k=1}\ \cdots\ \underbrace{[N-2,s_{N-2}-1]\,[N-2,s_{N-2}]}_{k=N-2}\ \cdot$$
$$\underbrace{[N-1,1]}_{k=N-1}\,D\ . \qquad (10)$$

where $D = \operatorname{diag}(z_1,\ldots,z_N)$ is unitary. The block-size of each Householder factor $[k,i]$ is never greater than $n_{k,i}+1 \leq N+1-k$.

Again, notice that existence of the Householder reflections in Theorem III.1 was proved by a recursive use of Lemma III.2 which, technically, may be considered as a computing procedure. Every representation (10) involves an ordered partition for each of the following: $\{y_2,\ldots,y_N\} \supset \{y_3,\ldots,y_N\} \supset \ldots \supset \{y_{N-1},y_N\} \supset \{y_N\}$, subsets of basis Y. Thus, the set of all representations supplied by Theorem III.1 have a partial order that is induced from the relation “finer than” on the family of partition sets. The extremal elements in the partial order are treated in Sections IV and V.

Let us find a lower bound for the number of options we have for the full factorization of unitary $U \in \mathbb{C}^{N\times N}$, respect to a given vector basis Y. At every step the index for the vector to be factorized is selected ad lib. If $(q_1,\ldots,y_k,\ldots,q_N)$ is a record of the successive choices, we see that it is one of the $N!$ permutations of the indices $(1,\ldots,k,\ldots,N)$. The subspaces to be considered at every step (from 1 to $N-1$) are, successively, the linear span of the elements of some partition set of subsets of Y having cardinalities $N-1,\ldots,1$. The number of choices we have for such partition sets are, successively, the Bell numbers B_{N-1}, $B_{N-2},\ldots,1$. Without counting the ways the partition set may be ordered, we see that the number of different ways to proceed in the factorization of unitary matrix $U \in \mathbb{C}^{N\times N}$, for the given basis Y, is never lesser than $N! \times B_1 \times \cdots \times B_{N-1}$, which is quite a number. Not every choice will necessarily produce a different factorization of U.

Theorem III.1 has two extreme cases, relative to the upper bound, $\sum_k s_k$, for the number of factor matrices in (10). In the next Section IV we consider its most economical form ($s_k = 1$, that produces $N-1$ factors at most). In Section V we deal with its most expensive form ($\dim \mathcal{I}_k = 1$, that produces $N(N-1)/2$ factors at most).

IV. THE SHORTEST FACTORIZATIONS

The example we are about to consider is a very simple case of Theorem III.1. Simple enough as to admit an explicit recording of all the Householder matrices and phase factors involved in the factorization of U. In this respect we may say that the example is the simplest case of Theorem III.1.

The simplifying choice in the procedure that leads to representation (10) in Theorem III.1 is to take $s_k = 1$ at every step. Subspaces are not severed and every map $Uy_k \mapsto y_k$ is done in a single stroke. The only Householder matrix that is computed at step k is $[k, 1] \equiv [r_k]$, with $r_k \in \mathbb{C}^N$ the reflecting vector. Matrix U is represented by the product

$$U = [r_1]\,[r_2] \cdots [r_{N-1}]\,D, \tag{11}$$

which is the shortest version of (10), $N-1$ Householder factors at most. In Lemma IV.1 below we prove that, for every $i < k$, the basis vector $y_i \perp r_k$ such that $[r_k]\,y_i = y_i$. The block size of Householder factors $[r_k]$ is thus reduced at least by one from step k to the next.

A simple recursive definition of all the reflecting vectors r_i, $i = 1, \ldots, N-1$, that are involved in the factorization (11) of U is given. The projection operators $\mathbb{I}_i$ that were used in the definition of vectors $x_{k,i}$, are superfluous for the example. The single phase factors $z_{k,1}$, introduced in (2) and successive steps, are denoted ζ_k.

At step $k = 1$ introduce the quantities

$$r_1 = \zeta_1 U y_1 - y_1 \ , \quad \text{with} \quad \angle\zeta_1 = \angle(Uy_1)^* y_1 \ , \quad \text{and} \quad I_1 = [r_1]\,\zeta_1 U \ . \tag{12}$$

Then, at the following steps, $1 < k < N$, let

$$r_k = \zeta_k I_{k-1} y_k - y_k \ , \quad \text{with} \quad \angle\zeta_k = \angle(I_{k-1}y_k)^* y_k \ , \quad \text{and} \quad I_k = [r_k]\,\zeta_k I_{k-1} \ . \tag{13}$$

The for-loop ends producing the matrix

$$I_{N-1} = (\zeta_{N-1} \cdots \zeta_1)\,[r_{N-1}] \cdots [r_1]\,U. \tag{14}$$

Our claim is that the matrix D of phases in (11) is

$$D \equiv \mathrm{diag}(z_1, \ldots, z_N) = (\zeta_{N-1} \cdots \zeta_1)^* \, I_{N-1} \ , \tag{15}$$

with phases given by $z_k = (\zeta_1 \cdots \zeta_k)^*$, $k = 1$ to N (the new number ζ_N is defined later, below). Once our claim (15) is proved, the representation (11) with Householder factors given by (12)–(13) follows from (14).

To prove claim (15) we need some definitions intended to keep a proper record of the phase factors ζ_i computed at every step and then collect them all in matrix D. Let $Z_i^k := \zeta_i \cdots \zeta_k$ for $i \leq k$. Let unitary matrices F_k, $k = 1, \ldots, N-1$, be defined by their action on the vectors in the basis Y. For $k = 1$, we let $F_1 = \mathbb{1}$ while for $k > 1$,

$$F_k y_i = \begin{cases} (Z_{i+1}^k)^* \ y_i \ , & i = 1, \ldots, k-1 \\ y_i \ , & i \geq k \end{cases} \ . \tag{16}$$

Matrices F_k are diagonal in the basis Y, $F_k = \mathrm{diag}(Z_2^{k\,*}, \ldots, Z_k^{k\,*}, 1, \ldots, 1)$. By definition, $F_k^* F_k = \mathbb{1}$. Our claim (15) is a corollary of the following.

Lemma IV.1 *Matrix $F_k I_k$ is the identity on the subspace $\langle y_1, \ldots, y_k \rangle$, for each $k = 1, \ldots, N-1$.*

Proof. By definition of I_1 in (12) and by Lemma II.1 we have that $F_1 I_1$ is the identity on the subspace $\langle y_1 \rangle$ (recall that $F_1 = \mathbb{1}$). Assume next that $F_{k-1} I_{k-1}$ is the identity in the subspace $\langle y_1, \ldots, y_{k-1} \rangle$.

First we prove that $y_i \perp r_k$ for every $i < k$. A direct calculation of the scalar product ${r_k}^* y_i \equiv (\zeta_k I_{k-1} y_k - y_k)^* y_i$ proceeds as follows,

$$\begin{aligned} {r_k}^* y_i &= \zeta_k^* \, (I_{k-1} y_k)^* y_i = \zeta_k^* \, {y_k}^* (I_{k-1}^* y_i) = \zeta_k^* \, {y_k}^* (I_{k-1}^* F_{k-1}^* F_{k-1} y_i) \\ &= Z_{i+1}^{k\,*} \, {y_k}^* (I_{k-1}^* F_{k-1}^* y_i) \ , \end{aligned}$$

where we have made use of (16) and $\zeta_k^* Z_{i+1}^{k-1\,*} = Z_{i+1}^{k\,*}$. Next, observe that $I_{k-1}^* F_{k-1}^* y_i = y_i$ since, by the induction assumption, $F_{k-1} I_{k-1} y_i = y_i$ and both F_{k-1} and I_{k-1} are unitary. Thus, ${r_k}^* y_i \equiv Z_{i+1}^{k\,*} \, {y_k}^* y_i = 0$, proving that $y_i \perp r_k$ for $i < k$.

Next, we prove that $F_k I_k y_i = y_i$ for $i < k$. By definition (13) and the induction assumption it follows that

$$F_k I_k y_i = F_k \, [r_k] \, \zeta_k I_{k-1} y_i = F_k \, [r_k] \, \zeta_k F_{k-1}^* y_i = Z_{i+1}^k F_k \, [r_k] \, y_i \ .$$

We have proved that $r_k \perp y_i$ for $i < k$. Hence, $[r_k]\,y_i = y_i$ and $F_k I_k\, y_i = Z^k_{i+1} F_k\, y_i = Z^k_{i+1} Z^{k\,*}_{i+1}\, y_i$ as promised. To complete the proof, let us consider $F_k I_k\, y_k$. Directly from definitions (13) and by Lemma II.1 we find that $I_k y_k = [r_k]\,\zeta_k I_{k-1} y_k = y_k$. Then, by (16) we have that $F_k I_k\, y_k = y_k$. □

Lemma IV.1 tells us that $F_{N-1} I_{N-1}$ is block-diagonal in the direct sum $\mathbb{C}^N = \langle y_1, \ldots, y_{N-1}\rangle \oplus \langle y_N\rangle$, having the form $F_{N-1} I_{N-1} =: \mathrm{diag}(1, \ldots, 1, \zeta_N^*)$ in the basis Y. The entry ζ_N is, by unitarity, a phase factor. Using this corollary of Lemma IV.1 in (14) we identify the following matrix of phase factors,

$$\begin{aligned} D = Z_1^{N-1\,*} I_{N-1} &= F^*_{N-1} F_{N-1} I_{N-1} Z_1^{N-1\,*} \\ &= \mathrm{diag}\big(\zeta_1^*,\ (\zeta_1\zeta_2)^*,\ \ldots,\ (\zeta_1 \cdots \zeta_N)^*\big)\ , \end{aligned} \tag{17}$$

which proves our claim (15).

V. EXPLICIT $U(2)$ PARAMETRIZATIONS OF $U(N)$

The other extreme of Theorem III.1 is reached by taking $s_k = N - k$, i.e., $n_{k,i} = 1$, for each k (from 1 to $N-1$) and each i (from 1 to $N-k$). Within this choice, all subspaces in the definition of the Householder factors in representation (10) of $U \in U(N)$ in Theorem III.1 are 2-dimensional, getting (10) the following form

$$U = \underbrace{[1,1] \cdots [1, N-2]\ [1, N-1]}_{\text{step } 1}\ \cdots\ \underbrace{[N-2, 1]\ [N-2, 2]}_{\text{step } N-2}\ \underbrace{[N-1, 1]}_{\text{step } N-1}\ D\ . \tag{18}$$

The number of Householder factors is $\sum_k (N-k) = N(N-1)/2$ at most. Each reflection $[k,i]$ in (18) is the identity on $\langle y_k, y_i\rangle^\perp$, the reflecting vector r that defines $[k,i]$ lays in the two-dimensional subspace $\langle y_k, y_i\rangle$. The block-size of every $[k,i]$ is thus two (unless they happen to be the identity). In this respect we refer to factor $[k,i]$ as an elementary factor.

To fix the ideas about elementary factors in (18), for any $x \in \langle y_k, y_i\rangle$ let $r := z\|y_k\|x - \|x\|y_k$ and consider the reflection $[r]$ restricted to the 2-dimensional subspace $\langle y_k, y_i\rangle$ of $\mathbb{C}^N$. Matrix $[r]$ has eigenvalues ±1. Then, the restricted $[r]$ has $\mathrm{tr}\,[r] = 0$. For an explicit parametrization of an elementary Householder matrix $[r]$ consider x as the coordinate vector $x = (e^{i\theta}\sin\varphi, \cos\varphi) \in \langle y_k, y_i\rangle$ with $y_k = (1,0)$ and $y_i = (0,1)$. Then, by applying Lemma II.1 to x and y_k we have that $z = e^{-i\theta}$ and that $r = x - y_k = (\sin\varphi - 1, e^{-i\theta}\cos\varphi)$.

A direct calculation leads us to the following matrix for $\mathtt{[}r\mathtt{]}$, restricted to $\langle y_k, y_i\rangle$,

$$\mathtt{[}r\mathtt{]} = \begin{pmatrix} \sin\varphi & e^{i\theta}\cos\varphi \\ e^{-i\theta}\cos\varphi & -\sin\varphi \end{pmatrix} . \tag{19}$$

When $\sin\varphi = \cos\varphi = 1/\sqrt{2}$, the unitary operator $\mathtt{[}r\mathtt{]}$ in (19) is a Fourier transform between the complementary bases (y_k, y_i) and (x_1, x_2), $x_1 = (1, e^{-i\theta})/\sqrt{2}$ and $x_2 = (e^{i\theta}, -1)/\sqrt{2}$, of the 2-dimensional subspace $\langle y_k, y_i\rangle$. Every $\mathtt{[}k, i\mathtt{]}$ in (18), when restricted to $\langle y_k, y_i\rangle$, has the form in (19), up to an overall phase factor that may be absorbed in matrix D, see (18).

After (19) we see that matrix $U \in U(N)$ is parametrized in (18) by $N(N-1)/2$ angles φ and by

$$\frac{N(N-1)}{2} + N = \frac{N(N+1)}{2} \quad \text{phase factors } e^{i\theta} \text{ and } z,$$

at most. All of them make N^2 real parameters for $U(N)$.

Let us find an upper bound for the number of forms of representing U as the product of elementary factors in (18). First, the vector y_k may be chosen in 1 of $N-k+1$ forms, totaling $N!$ choices. Second, we are considering the finest partition at every step k. The choice is unique, having cardinality $N-k$ each. The number of choices for a given order of the partition set at every step is $(N-1)!$, $(N-2)!$, …, 1, successively. The number of possible factorizations of matrix $U \in U(N)$ into elementary Householder blocks is thus never greater than the superfactorial $N\$:= N! \times (N-1)! \times \cdots \times 2! \times 1! \equiv 1^N \cdot 2^{N-1} \cdot 3^{N-2} \cdots (N-1)^2 \cdot N$. This is a strict upper bound, since otherwise the dimension of $U(N)$ would be smaller than N^2.

A related parametrization of $U(N)$ is the one given by Murnaghan in [11]. Let $A \in U(N)$. Let A^D be the diagonal form of A as given by the transformation $U^*AU = A^D$, for some matrix $U \in U(N)$. Murnaghan's remark [11] is that a matrix U diagonalizing A may be represented as a product of $N(N-1)/2$ elementary factors and a matrix of phases $D = \mathbb{1}$ such that UA^DU^* is a representation of matrix A as the product of $N(N-1)$ elementary factors and a diagonal matrix of phase factors, A^D. Theorem III.1 in its version (18) provides a proof of Murnaghan's remark [11] and a method to compute the factors as well.

[1] Peres, A.: Quantum theory: concepts and methods. Kluwer Academic Publishers (1995).

[2] Barenco, A., et al.: Elementary gates for quantum computation. Phys. Rev. A 52, 3457–3467 (1995)

[3] Reck, M., Zeillinger, A., Bernstein, H.J., and Bertani, P.: Experimental realization of any discrete unitary operator. Phys. Rev. Letters 73, 58–61 (1994)

[4] Ivanov, P.A., Kyoseva, E.S., and Vitanov, N.V.: Engineering of arbitrary $U(N)$ transformations by quantum Householder reflections. Phys. Rev. A 74, 022323 (2006)

[5] Green, W.M., Rooks, M.J., Sekaric, L., and Vlasov, Y.A.: Ultra-compact, low RF power, 10 Gb/s silicon Mach-Zehnder modulator. Optics Express 15, 17106–17113 (2007)

[6] Politi, A., Cryan, M.J., Rarity, J.G., Yu, S., and O'Brien, J.L.: Silica-on-silicon waveguide quantum circuits. Science 320, 646–649 (2008)

[7] Peres,A.: Reversible logic and quantum computers. Phys. Rev. A 32, 3266–3276 (1985)

[8] Zukowski, M., Zeillinger, A., and Horne, M.A.: Realizable higher-dimensional two-particle entanglements via multiport beam splitters. Phs. Rev. A 55, 2564–2579 (1997)

[9] V. Nebendahl, H. Häffner and C.F. Roos, Optimal control of entangling operations for trapped-ion quantum computing. Phys. Rev. 79 (2009) 012312.

[10] Ivanov,P.A., Torosov, B.T., and Vitanov, N.V.: Navigation between quantum states by quantum mirrors. Phys. Rev. A 75, 012323 (2007)

[11] Murnaghan, F.D.: On a convenient system of parameters for the unitary group. Proc. N.A.S. 38, 127–129 (1952)