# ($\mu$) Operator Spaces and Norms

In this mathematical digression, we introduce Hermitian-space structures for linear maps and operators, as well as several norms often used on these spaces.

Consider arbitrary linear maps between two Hermitian spaces: $M, N : \mathcal{H}_B \to \mathcal{H}_A$. The space of these linear maps is a vector space with canonical basis $|a\rangle\langle b|$, where $|a\rangle$ $(a \in A)$ and $|b\rangle$ $(b \in B)$ are the respective standard orthonormal bases of $\mathcal{H}_A$, of dimension $d_A$, and $\mathcal{H}_B$, of dimension $d_B$.

**Definition (Hilbert–Schmidt inner product).** *The* ***inner product*** *between $M$ and $N : \mathcal{H}_B \to \mathcal{H}_A$ is the form*

$$\langle M|N\rangle = \mathrm{Tr}(M^*N).$$

Like the trace, this inner product does not depend on the choice of orthonormal bases used to write $M$ and $N$. Writing $M$ and $N$ in bases $|a\rangle$ and $|b\rangle$, one obtains

$$\langle M|N\rangle = \sum_b \langle b|M^*N|b\rangle = \sum_{a,b} \langle b|M^*|a\rangle \langle a|N|b\rangle = \sum_{a,b} \langle a|M|b\rangle^* \langle a|N|b\rangle ,$$

which is exactly the inner product of the $d_A d_B$ coefficients of the two matrices. In this form, it is immediate that this is indeed an inner product.

One can check directly that this inner product is linear in the second argument: $\langle M|\alpha N + \beta N'\rangle = \alpha \langle M|N\rangle + \beta \langle M|N'\rangle$ for $\alpha, \beta \in \mathbb{C}$; conjugate-symmetric: $\langle M|N\rangle = \mathrm{Tr}(M^*N) = \mathrm{Tr}\big((M^*N)^*\big)^* = \mathrm{Tr}(N^*M)^* = \langle N|M\rangle^*$; and positive definite, since $\mathrm{Tr}(M^*M) \geqslant 0$ is the sum of the nonnegative eigenvalues of $M^*M \geqslant 0$, and this sum is zero only if all these eigenvalues are zero, that is, $M^*M = 0$, or equivalently $M^*M|\psi\rangle = 0$ for every $|\psi\rangle$, whence $\|M\psi\|^2 = \langle\psi|M^*M|\psi\rangle = 0$ and therefore $M = 0$.

Endowed with this inner product, the vector space of linear maps $\mathcal{H}_B \to \mathcal{H}_A$ is a **Hermitian space**. When $d_A = d_B$, identifying $\mathcal{H}_A = \mathcal{H}_B = \mathcal{H}$, the maps $M$ and $N$ are square matrices, that is, operators $\mathcal{H} \to \mathcal{H}$, and the vector space of these operators $\mathcal{H} \to \mathcal{H}$, endowed with the inner product $\langle M|N\rangle$, is called the **operator space**.

**Definition (Frobenius norm).** *The norm of $M$ is the Euclidean norm associated with the inner product $\langle M|N\rangle$:*

$$\|M\| = \sqrt{\langle M|M\rangle} = \sqrt{\mathrm{Tr}(M^*M)}.$$

*It is sometimes called the* ***Frobenius norm****, indicated with an "F", $\|M\|_F$, in order to distinguish it from other norms.*

By the cyclic property of the trace, it can equivalently be written as $\|M\| = \sqrt{\mathrm{Tr}(M^*M)} = \sqrt{\mathrm{Tr}(MM^*)}$. This norm is also denoted $\|M\|_2$ because, writing $M$ in an orthonormal basis $|a\rangle$, one obtains

$$\|M\| = \sqrt{\sum_{a,b} |\langle a|M|b\rangle|^2},$$

which is exactly the Euclidean norm of the $d_A d_B$ coefficients of $M$. This norm $\sqrt{\mathrm{Tr}(M^*M)}$, denoted $\|M\|_2$, should not be confused with the **trace norm** $\mathrm{Tr}\sqrt{M^*M}$, denoted $\|M\|_1$, nor with the **operator norm**, denoted $|||M|||$ or $\|M\|_\infty$; see below.

**Remark.** The canonical basis $|a\rangle\langle b|$ ($a \in A$, $b \in B$), where $|a\rangle$ and $|b\rangle$ are the respective standard orthonormal bases of $\mathcal{H}_A$ and $\mathcal{H}_B$, is also an **orthonormal basis** for the Hilbert–Schmidt inner product:

$$\langle (|a'\rangle\langle b'|)|(|a\rangle\langle b|)\rangle = \mathrm{Tr}\big((|a'\rangle\langle b'|)^*(|a\rangle\langle b|)\big) = \mathrm{Tr}(|b'\rangle\langle a'|a\rangle\langle b|) = \langle a'|a\rangle\langle b|b'\rangle\,.$$

**Remark (Orthonormality of pure states).** The orthonormality property is identical for pure states considered as vectors $|\psi\rangle \in \mathcal{H}$, or as density operators $|\psi\rangle\langle\psi|$. Indeed, by a computation similar to the one above,

$$\langle (|\phi\rangle\langle\phi|)|(|\psi\rangle\langle\psi|)\rangle = |\langle\phi|\psi\rangle|^2.$$

**Remark (Orthogonality of positive semidefinite operators).** More generally, if $M \geqslant 0$ and $N \geqslant 0$, then $\langle M|N\rangle = \mathrm{Tr}(MN) = 0$ is equivalent to $MN = NM = 0$, by the positivity property of the trace.

As in every Hermitian space, for the Hilbert–Schmidt inner product we have the **Cauchy–Schwarz inequality** $|\langle M|N\rangle| \leqslant \|M\|\|N\|$, that is,

$$|\mathrm{Tr}(M^*N)| \leqslant \sqrt{\mathrm{Tr}(M^*M)}\sqrt{\mathrm{Tr}(N^*N)},$$

with equality if and only if $M$ and $N$ are proportional, together with the triangle inequality for norms, $\|M + N\| \leqslant \|M\| + \|N\|$.

**Example.** When $M$ and $N$ are observables (Hermitian operators), one recovers the special case $\rho = \frac{\mathbb{1}}{d}$ of the Cauchy–Schwarz inequality for expectation values of observables, $|\mathrm{Tr}(MN\rho)| \leqslant \sqrt{\mathrm{Tr}(M^2\rho)}\,\sqrt{\mathrm{Tr}(N^2\rho)}$, that is,

$$|\langle MN\rangle| \leqslant \sqrt{\langle M^2\rangle}\sqrt{\langle N^2\rangle}.$$

The case of an arbitrary density operator $\rho$ is a generalization of the Cauchy–Schwarz inequality analogous to the case of random variables with respect to a given probability distribution.

The Hermitian space of linear maps $\mathcal{H}_B \to \mathcal{H}_A$ can equivalently be viewed as a Hermitian space of vectors, by means of the vectorization procedure.

**Definition (Vectorization).** *Given orthonormal bases $|a\rangle$ ($a \in A$) and $|b\rangle$ ($b \in B$) of $\mathcal{H}_A$ and $\mathcal{H}_B$, respectively, the **vectorization** of the $d_A \times d_B$ matrix*

$$M = \sum_{a,b} M_{a,b}|a\rangle\langle b|$$

*of the linear map $\mathcal{H}_B \to \mathcal{H}_A$ is the vector*

$$\mathrm{vec}(M) = \sum_{a,b} M_{a,b}|b\rangle|a\rangle \in \mathcal{H}_B \otimes \mathcal{H}_A.$$

The vectorization $\mathrm{vec}()$ is therefore the unique linear map that sends $|a\rangle\langle b|$ to $|b\rangle|a\rangle$. It depends on the choice of bases and simply consists in **stacking the columns** of $M$ into a vector of dimension $d_A d_B$. Since the coefficients are unchanged, this is a bijective map that identifies $M$ with $\mathrm{vec}(M)$, and conversely through the inverse map $\mathrm{vec}^{-1}()$, which constructs a $d_A \times d_B$ matrix column by column from a vector of length $d_A d_B$. Moreover, this vectorization preserves the inner product, since if $M = \sum_{a,b} M_{a,b}|a\rangle\langle b|$ and $N = \sum_{a,b} N_{a,b}|a\rangle\langle b|$, then

$$\langle M|N\rangle = \sum_{a,b} M^*_{a,b} N_{a,b} = \langle \mathrm{vec}(M)|\mathrm{vec}(N)\rangle .$$

It also preserves the norm, $\|M\| = \|\mathrm{vec}(M)\|$; it is an isometry. Thus we have the following result.

**Theorem.** *Vectorization is an isomorphism of Hermitian spaces.*

**Example.** The vectorizations of the Kraus matrices $K_i$ of a channel $\mathcal{W}$ are the vectors $|\psi_i\rangle$ of a decomposition $\mathbf{C}(\mathcal{W}) = \sum_i |\psi_i\rangle\langle\psi_i|$ of the Choi matrix:

$$\mathrm{vec}(K_i) = |\psi_i\rangle.$$

**Remark (Row vectorization).** One can also define the vectorization $\mathrm{vec}'$ as the unique linear map that sends $|a\rangle\langle b|$ to $|a\rangle|b\rangle$. In this case, the **rows** of $M$ are transposed and stacked as columns in the vector $\mathrm{vec}'(M) \in \mathcal{H}_A \otimes \mathcal{H}_B$. Thus $\mathrm{vec}'(M) = \mathrm{vec}(M^t)$, where $M^t$ denotes the **transpose** of $M$.

Thus, if $M = \sum_{a,b} M_{a,b}|a\rangle\langle b|$, then $\mathrm{vec}'(M) = \sum_{a,b} M_{a,b}|a\rangle|b\rangle$, with the indices $a$ and $b$ in the same order; this makes some computations easier than with $\mathrm{vec}(M)$, even though the computations are equivalent.

**Example.** The **singular value decomposition** of $M = \sum_i \sigma_i\, |i\rangle_A\langle i|_B$ is identified by $\mathrm{vec}'$ with the **Schmidt decomposition** of the row-vectorized form: $\mathrm{vec}'(M) = \sum_i \sigma_i\, |i\rangle_A|i\rangle_B$. The Schmidt rank of $\mathrm{vec}'(M)$ is the rank of the matrix $M$.

**Remark (Ket notation for a matrix).** A simplified notation for the column vectorization $\mathrm{vec}(M)$, also used for the row vectorization $\mathrm{vec}'(M)$, is the ket notation $|M\rangle$. This notation is justified by the fact that in both cases

$$\langle M|N\rangle = \sum_{a,b} M^*_{a,b} N_{a,b}$$

denotes both the Hilbert–Schmidt inner product of two matrices $\mathrm{Tr}(M^*N)$ and the ordinary inner product in the space $\mathcal{H}_B \otimes \mathcal{H}_A$ (for $\mathrm{vec}$) or $\mathcal{H}_A \otimes \mathcal{H}_B$ (for $\mathrm{vec}'$).

We now fix the notation $|M\rangle = \mathrm{vec}'(M) = \sum_{a,b} M_{a,b}|a\rangle|b\rangle$. Since $|Mb\rangle = M|b\rangle = \sum_a M_{a,b}|a\rangle$, we obtain the useful formula

$$|M\rangle = \sum_b |Mb\rangle|b\rangle.$$

Similarly, considering the transpose $M^t = \sum_{a,b} M_{a,b}|b\rangle\langle a|$, one has $|M^t a\rangle = M^t|a\rangle = \sum_b M_{a,b}|b\rangle$, whence

$$|M\rangle = \sum_a |a\rangle|M^t a\rangle.$$

We can then easily compute the action of a product operator $R\otimes S = (\mathbb{1}_A\otimes S)(R\otimes \mathbb{1}_B)$ on $|M\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$. On the one hand, $(R \otimes \mathbb{1}_B)|M\rangle = \sum_b (R \otimes \mathbb{1}_B)|Mb\rangle|b\rangle = \sum_b |RMb\rangle|b\rangle = |RM\rangle$; on the other hand, $(\mathbb{1}_A \otimes S)|M\rangle = \sum_a (\mathbb{1}_A \otimes S)|a\rangle|M^t a\rangle = \sum_a |a\rangle|SM^t a\rangle = |MS^t\rangle$. Combining the two gives the formula

$$(R \otimes S)|M\rangle = |RMS^t\rangle.$$

The inner product $\langle M|R \otimes S|N\rangle$ is therefore written as the Hilbert–Schmidt inner product $\langle M|RNS^t\rangle$, which gives the following formula.

**Theorem (Ando identity).** *For all rectangular matrices $M, N : \mathcal{H}_B \to \mathcal{H}_A$, and operators $R$ on $\mathcal{H}_A$ and $S$ on $\mathcal{H}_B$,*

$$\langle M|R \otimes S|N\rangle = \mathrm{Tr}(M^* RNS^t).$$

The space of operators on $\mathcal{H}$ is a Hermitian space with canonical orthonormal basis $|a'\rangle\langle a|$ $(a, a' \in A)$. However, these operators are not themselves Hermitian, except when $a = a'$. One can find an orthonormal basis of Hermitian operators from the canonical orthonormal basis by observing that, for $a \neq a'$, $|a'\rangle\langle a|$ decomposes into Hermitian and anti-Hermitian parts:

$$|a'\rangle\langle a| = \frac{|a'\rangle\langle a| + |a\rangle\langle a'|}{2} + i\frac{|a'\rangle\langle a| - |a\rangle\langle a'|}{2i},$$

where the two normalized operators, of norm $1$,

$$\frac{|a'\rangle\langle a| + |a\rangle\langle a'|}{\sqrt{2}} \text{ and } \frac{|a'\rangle\langle a| - |a\rangle\langle a'|}{\sqrt{2}i}$$

are Hermitian. The same operators are recovered, up to a sign for the second one, by exchanging the roles of $a$ and $a'$, and they therefore span the same space $\mathrm{Vect}(|a'\rangle\langle a|, |a\rangle\langle a'|)$. Moreover, since $|a'\rangle\langle a|$ and $|a\rangle\langle a'|$ are orthogonal $(a \neq a')$, these two Hermitian operators are also orthogonal:

$$\langle(|a'\rangle\langle a| + |a\rangle\langle a'|)|(|a'\rangle\langle a| - |a\rangle\langle a'|)\rangle = 1 - 0 + 0 - 1 = 0.$$

Thus, starting from the orthonormal basis of the $|a'\rangle\langle a|$, by replacing, whenever $a \neq a'$, each pair $|a'\rangle\langle a|$, $|a\rangle\langle a'|$ by the pair $\frac{|a'\rangle\langle a|+|a\rangle\langle a'|}{\sqrt{2}}$, $i\frac{|a'\rangle\langle a|-|a\rangle\langle a'|}{\sqrt{2}}$, one obtains an orthonormal basis of Hermitian operators.

**Theorem (Canonical Hermitian basis).** *An orthonormal basis of Hermitian operators for the operator space is formed by the operators $|a\rangle\langle a|$ and, for each unordered pair $\{a, a'\}$ with $a \neq a'$, the operators*

$$\frac{|a'\rangle\langle a| + |a\rangle\langle a'|}{\sqrt{2}} \textit{ and } \frac{i|a'\rangle\langle a| - i|a\rangle\langle a'|}{\sqrt{2}}.$$

Specifying an order on the indices, for example $k, \ell = 1, \ldots, d$ where $d = \dim \mathcal{H}$, this canonical Hermitian basis consists of the $d^2$ operators

$$\begin{cases} |k\rangle\langle k| & (k = \ell), \\ \frac{|k\rangle\langle \ell|+|\ell\rangle\langle k|}{\sqrt{2}} & (k < \ell), \\ \frac{i|k\rangle\langle \ell|-i|\ell\rangle\langle k|}{\sqrt{2}} & (k > \ell). \end{cases}$$

**Remark.** This orthonormal basis spans the entire space of arbitrary operators, which are linear combinations of these basis elements with complex coefficients. By contrast, the set of **Hermitian operators** $M = M^*$ is a **real inner-product space** with the induced inner product $\langle M|N\rangle = \mathrm{Tr}(MN)$, because if $M$ and $N$ are Hermitian, then $\alpha M + \beta N$ is Hermitian as well for all $\alpha, \beta \in \mathbb{R}$.

This real inner-product space is also spanned by the canonical Hermitian basis, but with coefficients restricted to real numbers. Indeed, since the inner product $\langle M|N\rangle = \mathrm{Tr}(MN) = \mathrm{Tr}(NM) = \langle N|M\rangle^*$ is necessarily real for Hermitian $M, N$, every Hermitian operator is a linear combination with real coefficients of the operators in the canonical Hermitian basis. One can therefore identify the real inner-product space of Hermitian operators with the space $\mathbb{R}^{d^2}$, where $d = \dim \mathcal{H}$.

**Example (Characterization of a quantum channel).** A superoperator (a linear map on operators) is uniquely determined by its action on the operators in the canonical basis $|a'\rangle\langle a|$, that is, by its Choi matrix. But it is also uniquely determined by its action on the operators of a basis of Hermitian operators; it is therefore completely determined when it is defined only for Hermitian operators.

Moreover, every Hermitian operator $M = M^*$ can be written as the difference $M = M^+ - M^-$ of positive semidefinite operators, its Hahn–Jordan decomposition. A superoperator is therefore completely determined, by linearity, when it is defined only for positive semidefinite operators, and even, still by linearity, only for density operators, that is, positive semidefinite operators with unit trace. A quantum channel $\mathcal{W}$, as a linear map, is therefore indeed determined by its action $\mathcal{W}(\rho)$ on an arbitrary density operator $\rho$.

**Remark (Decomposition into pure states).** It is sometimes convenient to express each canonical operator $|k\rangle\langle\ell|$ as a linear combination of density operators of the form $|\psi\rangle\langle\psi|$, that is, pure states. We may restrict to the case $k < \ell$. On the one hand, $|k\rangle\langle\ell|$ is the half-sum of its Hermitian and anti-Hermitian parts:

$$|k\rangle\langle\ell| = \frac{|k\rangle\langle\ell| + |\ell\rangle\langle k|}{2} - i\frac{i|k\rangle\langle\ell| - i|\ell\rangle\langle k|}{2},$$

where we recognize the canonical Hermitian operators. On the other hand, observe that

$$\begin{aligned}(|k\rangle + |\ell\rangle)(\langle k| + \langle\ell|) &= |k\rangle\langle k| + |\ell\rangle\langle\ell| + (|k\rangle\langle\ell| + |\ell\rangle\langle k|),\\ (|k\rangle + i|\ell\rangle)(\langle k| - i\langle\ell|) &= |k\rangle\langle k| + |\ell\rangle\langle\ell| - (i|k\rangle\langle\ell| - i|\ell\rangle\langle k|).\end{aligned}$$

Substituting gives

$$\begin{aligned}|k\rangle\langle\ell| = &\frac{1}{2}(|k\rangle + |\ell\rangle)(\langle k| + \langle\ell|) - \frac{1}{2}|k\rangle\langle k| - \frac{1}{2}|\ell\rangle\langle\ell|\\ &+ \frac{i}{2}(|k\rangle + i|\ell\rangle)(\langle k| - i\langle\ell|) - \frac{i}{2}|k\rangle\langle k| - \frac{i}{2}|\ell\rangle\langle\ell|.\end{aligned}$$

Writing the pure states $|+\rangle = \frac{|k\rangle+|\ell\rangle}{\sqrt{2}}$ and $|+i\rangle = \frac{|k\rangle+i|\ell\rangle}{\sqrt{2}}$, we obtain the **pure-state decomposition formula**

$$|k\rangle\langle\ell| = |+\rangle\langle+| + i \cdot |+i\rangle\langle+i| - \tfrac{1+i}{2}|k\rangle\langle k| - \tfrac{1+i}{2}|\ell\rangle\langle\ell|,$$

which shows that a channel is uniquely determined by its action on the $d^2$ pure states $|k\rangle$, $\frac{|k\rangle+|\ell\rangle}{\sqrt{2}}$, and $\frac{|k\rangle+i|\ell\rangle}{\sqrt{2}}$ $(k < \ell)$.

The Frobenius norm on the operator space can be written as $\|M\|_2 = \sqrt{\mathrm{Tr}(|M|^2)}$, where $|M| = \sqrt{M^*M}$ is the absolute value of $M$, which appears in the polar decomposition $M = U|M|$ with $U$ unitary. Since the eigenvalues of $|M|$ are the singular values $\sigma_i \geqslant 0$ of $M$, one can also identify

$$\|M\|_2 = \sqrt{\sum_i \sigma_i^2} = \|\sigma\|_2$$

as the Euclidean norm of the vector $\sigma = (\sigma_i)_i$ of singular values of $M$. The generalization to $p$-norms of order $p \geqslant 1$ suggests the following definition.

**Definition (Schatten norms).** *For every $p \geqslant 1$, the Schatten $p$-norm $\|M\|_p$ is defined by*

$$\|M\|_p = \left(\mathrm{Tr}(|M|^p)\right)^{1/p} = \left(\sum_i \sigma_i^p\right)^{1/p} = \|\sigma\|_p.$$

*This definition extends by taking the limit as $p \to +\infty$:*

$$\|M\|_\infty = \lim_{p\to+\infty} \|M\|_p = \lim_{p\to+\infty} \|\sigma\|_p = \max_i \sigma_i.$$

The $p$-norm is indeed positive definite: $\|M\|_p \geqslant 0$, with equality $\|M\|_p = 0$ if and only if $M = 0$, and positively homogeneous: $\|\alpha M\|_p = |\alpha|\|M\|_p$ for every scalar $\alpha \in \mathbb{C}$. (The third norm property, the triangle inequality, is presented below.)

**Remark.** Since $|M| = \sqrt{M^*M}$ and $|M|' = \sqrt{MM^*}$ have the same singular values $\sigma_i$, either $|M|$ or $|M|'$ can be used in the definition.

Consider finite orders $q > p \geqslant 1$. If we normalize $\sigma'_i = \frac{\sigma_i}{\|\sigma\|_p}$, then $\sigma'_i \leqslant 1$, hence $\sigma'^q_i \leqslant \sigma'^p_i$ for all $q > p$. Summing gives $\|\sigma'\|_q^q \leqslant \|\sigma'\|_p^p = 1$, that is, $\|\sigma'\|_q \leqslant 1 \iff \|\sigma\|_q \leqslant \|\sigma\|_p$: the $p$-norms $\|M\|_p = \|\sigma\|_p$ are decreasing (nonincreasing) as $p$ increases. The inequality $\|\sigma\|_\infty = \max_i \sigma_i \leqslant \|\sigma\|_p$ is immediate in the case $q = +\infty$.

Moreover, the classical Hölder inequality for finite $1 \leqslant p < q$ gives $\|\sigma\|_p^p = \sum_i \sigma_i^p \cdot 1 \leqslant \left(\sum_i \sigma_i^q\right)^{p/q} (\sum_i 1)^{1-p/q} = \|\sigma\|_q^p r^{p(1/p-1/q)}$, whence $\|\sigma\|_p \leqslant r^{1/p-1/q}\|\sigma\|_q$, where one can take $r$ to be the number of nonzero singular values, that is, the rank of $M$. Again, the inequality $\|\sigma\|_p \leqslant r^{1/p}\|\sigma\|_\infty = (r \max_i \sigma_i^p)^{1/p}$ is immediate in the case $q = +\infty$. We have obtained the following result.

**Theorem (Equivalence of $p$-norms).** *For every operator $M$ of rank $r$, and every $q > p$,*

$$\|M\|_q \leqslant \|M\|_p \leqslant r^{1/p-1/q}\|M\|_q.$$

In particular, the $p$-norm is decreasing (in the nonstrict sense) in $p \geqslant 1$. Bounding $r$ by the dimension of the space, we deduce that the Schatten $p$-norms are equivalent.

**Example.** In particular, $\|M\|_1 \leqslant \sqrt{r}\|M\|_2$ and $\|M\|_2 \leqslant \sqrt{r}\|M\|_\infty$.

In order to generalize the Cauchy–Schwarz inequality $|\langle M|N\rangle| \leqslant \|M\|_2\|N\|_2$ to $p$-norms, one has to deal with the difficulty that the positive semidefinite operators $|M|$ and $|N|$ do not commute in general and, consequently, cannot be diagonalized in the same orthonormal basis in order to apply the classical Hölder inequality $\langle\sigma|\tau\rangle \leqslant \|\sigma\|_p\|\tau\|_q$ $(\frac{1}{p} + \frac{1}{q} = 1)$ to the respective singular values $\sigma_i$ and $\tau_i$ of $M$ and $N$.

For simplicity, we restrict to the case of **operators** $M$ and $N$ (represented by square matrices) with polar decompositions $M = U|M|$ and $N = V|N|$, where $U$ and $V$ are unitary. (The case of linear maps represented by rectangular matrices can be generalized by extending the notion of polar decomposition with partial isometries $U$ and $V$.)

The operator inner product $\langle M|N\rangle$ can be written as

$$\langle M|N\rangle = \mathrm{Tr}(M^*N) = \mathrm{Tr}(|M|U^*V|N|) = \mathrm{Tr}(|M|\cdot W\cdot|N|) = \sum_i \sigma_i \langle i|W\cdot|N||i\rangle\,,$$

where $W = U^*V$ is unitary and where we have used the spectral decomposition $|M| = \sum_i \sigma_i |i\rangle\langle i|$. The classical Hölder inequality applies for $\frac{1}{p}+\frac{1}{q} = 1$ ($1 \leqslant p, q \leqslant +\infty$), and, up to exchanging the roles of $p$ and $q$, we may assume that $q$ is finite. We obtain the inequality

$$|\langle M|N\rangle| \leqslant \sum_i \sigma_i\tau_i \leqslant \|\sigma\|_p\|\tau\|_q = \|M\|_p\|\tau\|_q,$$

where the coefficients $\tau_i = |\langle i|W\cdot|N||i\rangle| = |\langle W^*i|\ |N|\ |i\rangle|$ are, in general, not equal to the singular values of $N$.

Observe that the vectors $|w_i\rangle = W^*|i\rangle$ also form an orthonormal basis, since $W^*$ is unitary. To obtain diagonal entries of $|N|$, we apply the Cauchy–Schwarz inequality

$$|\langle\phi||N||\psi\rangle| = |\langle\sqrt{|N|}\,\phi|\sqrt{|N|}\,\psi\rangle| \leqslant \|\sqrt{|N|}\phi\|\cdot\|\sqrt{|N|}\psi\| = \sqrt{\langle\phi|\ |N|\ |\phi\rangle}\sqrt{\langle\psi|\ |N|\ |\psi\rangle},$$

which gives

$$\tau_i = |\langle w_i|\ |N|\ |i\rangle| \leqslant \sqrt{\langle w_i|\ |N|\ |w_i\rangle}\sqrt{\langle i|\ |N|\ |i\rangle}.$$

Now, for every orthonormal basis $|i\rangle$, the diagonal entries $\langle i|\ |N|\ |i\rangle$ of $|N|$ written in this basis decompose over the singular values of $N$, that is, the eigenvalues $\nu_k \geqslant 0$ of $|N|$, with spectral decomposition $|N| = \sum_k \nu_k |k\rangle\langle k|$:

$$\langle i|\ |N|\ |i\rangle = \sum_k \langle i|k\rangle\langle k|i\rangle\,\nu_k = \sum_k |\langle k|i\rangle|^2\nu_k.$$

This is a convex combination of the $\nu_k$, since $\sum_k |\langle k|i\rangle|^2 = \|i\|^2 = 1$. As the function $f(x) = x^q$ ($q \geqslant 1$) is convex, the **convexity inequality** gives

$$\big(\langle i|\ |N|\ |i\rangle\big)^q \leqslant \sum_k |\langle k|i\rangle|^2\nu_k^q = \langle i|\ |N|^q\ |i\rangle\,.$$

Applying this to the two orthonormal bases $|i\rangle$ and $|w_i\rangle$ gives

$$\tau_i^q \leqslant \sqrt{\langle w_i|\ |N|^q\ |w_i\rangle}\sqrt{\langle i|\ |N|^q\ |i\rangle}.$$

Summing over $i$ and applying a variant of the Cauchy–Schwarz inequality for positive numbers, $\sum_i \sqrt{x_i y_i} \leqslant \sqrt{\sum_i x_i}\sqrt{\sum_i y_i}$ (**Cauchy's inequality**), we obtain

$$\|\tau\|_q \leqslant \big(\sqrt{\mathrm{Tr}\,|N|^q}\sqrt{\mathrm{Tr}\,|N|^q}\big)^{1/q} = \|N\|_q.$$

We have obtained the following theorem.

**Theorem (Hölder inequality for operators).** *If* $\frac{1}{p}+\frac{1}{q} = 1$, *where* $1 \leqslant p, q \leqslant +\infty$, *then*

$$|\langle M|N\rangle| \leqslant \|M\|_p\|N\|_q.$$

By analogy with the classical case, one can verify that equality holds in Hölder's inequality for finite $p, q$ if $|N|^q = c|M|^p$ (where $c > 0$) with $U = V$, because then $\|N\|_q^q = c\|M\|_p^p$ and

$$\langle M|N\rangle = \mathrm{Tr}(|M| \cdot |N|) = c^{1/q}\,\mathrm{Tr}(|M|^{p/q+1}) = c^{1/q}\|M\|_p^p = \|M\|_p\|N\|_q.$$

If one of the two orders is infinite, for example $p = 1$ and $q = \infty$, equality holds, for instance, if $|N| = c\mathbb{1}$ with $c > 0$ and the polar unitaries are aligned, because then $\|N\|_\infty = c$ and $\mathrm{Tr}(|M| \cdot |N|) = c\|M\|_1$.

In all cases, one can always choose the proportionality constant $c$ so that $\|N\|_q = 1$, and $\|M\|_p$ is then the maximum value of $|\langle M|N\rangle|$, which gives the following characterization.

**Theorem (Linearization of the $p$-norm by duality).** *If $\frac{1}{p} + \frac{1}{q} = 1$, where $1 \leqslant p, q \leqslant +\infty$, then*

$$\|M\|_p = \max_{\|N\|_q=1} |\langle M|N\rangle|.$$

Because of this property, the norms $\|\cdot\|_p$ and $\|\cdot\|_q$, with $\frac{1}{p} + \frac{1}{q} = 1$, are said to be **dual**. The triangle inequality for operator $p$-norms follows immediately, just as in the classical case of vector $p$-norms. Indeed, since $|\langle M + N|L\rangle| \leqslant |\langle M|L\rangle| + |\langle N|L\rangle|$, we have

$$\|M + N\|_p = \max_{\|L\|_q=1} |\langle M + N|L\rangle| \leqslant \max_{\|L\|_q=1} |\langle M|L\rangle| + \max_{\|L\|_q=1} |\langle N|L\rangle| = \|M\|_p + \|N\|_p,$$

which gives the **triangle inequality** for $p$-norms, or **Minkowski inequality**:

$$\|M + N\|_p \leqslant \|M\|_p + \|N\|_p.$$

**Example (Operator norm $\|M\|_\infty$).** Write $M = U|M|$, where $U$ is unitary, and write the spectral decomposition $|M| = \sum_i \sigma_i |i\rangle\langle i|$. Since $U$ preserves the norm, for every vector $|\psi\rangle$ we have

$$\|M\psi\|^2 = \|\,U|M|\,\psi\|^2 = \|\,|M|\,\psi\|^2 = \sum_i \sigma_i^2 |\langle i|\psi\rangle|^2 \leqslant (\max_i \sigma_i)^2 \|\psi\|^2,$$

which shows that $\|M\psi\| \leqslant \|M\|_\infty \|\psi\|$, with equality if $|\psi\rangle$ lies in the eigenspace of $|M|$ corresponding to a maximal singular value. Consequently, $\|M\|_\infty$ is the **operator norm** (also called the **spectral norm**) of $M$, for the usual Euclidean norm $\|\psi\| = \|\psi\|_2$ on vectors:

$$\|M\|_\infty = \max_i \sigma_i = \max_{\|\psi\|=1} \|M\psi\|,$$

which is sometimes denoted $|||M|||$ (**triple norm**).

**Example (Trace norm $\|M\|_1$).** The trace norm

$$\|M\|_1 = \mathrm{Tr}\,|M| = \mathrm{Tr}\,\sqrt{M^*M} = \mathrm{Tr}\,\sqrt{MM^*} = \sum_i \sigma_i$$

is very widely used in quantum information theory. The trace norm $\|M\|_1$ and the operator norm $\|M\|_\infty$ are dual.

Writing the polar decomposition $M = U|M|$, where $U$ is unitary, and the spectral decomposition $|M| = \sum_i \sigma_i |i\rangle\langle i|$, we have

$$\operatorname{Tr}(M) = \operatorname{Tr}(U|M|) = \sum_i \sigma_i \operatorname{Tr}(U|i\rangle\langle i|) = \sum_i \sigma_i \langle i|U|i\rangle ,$$

where the diagonal coefficients of $U$ all have modulus at most $1$, whence the **trace triangle inequality**

$$|\operatorname{Tr}(M)| \leqslant \operatorname{Tr}|M| = \|M\|_1,$$

with equality, in particular, if $U = \mathbb{1}$, that is, if $M \geqslant 0$.

**Remark.** In the special case where $M$ is Hermitian, one can recover the trace triangle inequality from the convexity inequality for the function $f(x) = |x|$: $|\operatorname{Tr}(M)| = |\sum_i \langle i|M|i\rangle| \leqslant \sum_i |\langle i|M|i\rangle| \leqslant \operatorname{Tr}(|M|)$. In this special Hermitian case, the inequality is also immediate from the Hahn–Jordan decomposition: $|\operatorname{Tr}(M)| = |\operatorname{Tr} M^+ - \operatorname{Tr} M^-| \leqslant \operatorname{Tr} M^+ + \operatorname{Tr} M^- = \operatorname{Tr}(M^+ + M^-) = \operatorname{Tr}|M|$.

We can generalize the trace triangle inequality to every unitary matrix $V$:

$$|\langle V|M\rangle| = |\operatorname{Tr}(V^* M)| \leqslant \|M\|_1,$$

because in the preceding proof, $V^*U$ is still unitary; equality holds if $V = U$. We therefore have the following characterization.

**Theorem (Characterization of the trace norm).**

$$\|M\|_1 = \max_{U \text{ unitary}} |\langle M|U\rangle|.$$

This characterization of the trace norm is stronger than the one by duality, $\|M\|_1 = \max_{\|N\|_\infty = 1} |\langle M|N\rangle|$, since every unitary $U$ satisfies $\|U\|_\infty = 1$.

**Remark.** One can also use the characterization of the trace norm to show the triangle inequality $\|M + N\|_1 \leqslant \|M\|_1 + \|N\|_1$, as we did above using duality of the $p$-norms.

**Remark.** If $\mathbf{M}$ is an operator on $\mathcal{H}_A \otimes \mathcal{H}_B$, the identity-tracing formula $\operatorname{Tr}_B\big((U^* \otimes \mathbb{1}_B)\mathbf{M}\big) = U^* \operatorname{Tr}_B \mathbf{M}$ gives, by taking the trace over $A$,

$$\langle U|\operatorname{Tr}_B \mathbf{M}\rangle = \langle U \otimes \mathbb{1}_B|\mathbf{M}\rangle .$$

Therefore

$$\begin{aligned} \|\operatorname{Tr}_B \mathbf{M}\|_1 &= \max_{U \text{ unitary}} |\langle U|\operatorname{Tr}_B \mathbf{M}\rangle| = \max_{U \text{ unitary}} |\langle U \otimes \mathbb{1}_B|\mathbf{M}\rangle| \\ &\leqslant \max_{\mathbf{U} \text{ unitary}} |\langle \mathbf{U}|\mathbf{M}\rangle| = \|\mathbf{M}\|_1. \end{aligned}$$

The trace norm decreases (does not increase) under a partial trace.