Orthogonal and Orthonormal Basis

1. Let X be a vector space over $\mathbb{K}$, $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$. We note that $\mathbb{R} \subseteq \mathbb{C}$ and the conjugation. Let $X \times X \xrightarrow{s} \mathbb{K}$ be a sesquilinear form which is hermitian; that is, $s(x,u) = \overline{s(u,x)}$. Then $s(x,u) + \overline{s(x,u)} = s(x,u) + s(u,x) \in \mathbb{R}$ for each $x, u \in X$. If now s is positive i.e. $Q(x) = s(x,x) \geq 0$ *(note that $s(x,x) = \overline{s(x,x)}$ because s is hermitian so that $s(x,x) \in \mathbb{R}$)*, let us write $a - s(x,x)$, $b = s(u,u)$, $c = s(u,u)$ so that $a \geq 0$, $b \geq 0$, $c\bar{c} = |c|^2 \geq 0$. For any $r \in \mathbb{R}$, we have $0 \leq s(x - ur\bar{c}, x - ur\bar{c})$

$= s(x,x) - s(x,u)r\bar{c} - s(u,x)rc + s(u,u)r^2c\bar{c}$

$= a - 2c\bar{c}r + br^2c\bar{c} = |c|^2br^2 - 2|c|^2r + a$

The quadratic expression $|c|^2br^2 - 2|c|^2r + a$ positive iff its discriminant $4|c|^4 - 4|c|^2ba \leq 0$ and thus we have, $if c \neq 0$, $|c|^2 \leq ba$. In case $c = 0$, this is true any way since a, b are positive real numbers. This is all situations we have $|c|^2 \leq ab$ i.e. $|s(x,u)|^2 \leq s(x,x)s(u,u)$. When s is a positive hermitian form, which is positive-definite in the sense that $s(x,x) = 0$ forces $x \in X$ to be $o \in X$, we shall denote $s(x,u)$ by $(x \mid u)$ and call it an inner product on X. In this notation, we have just proved

$$|(x \mid u)|^2 \leq \|x\|^2\|u\|^2 \tag{1}$$

This is known as the Cauchy-Schwarz inequality; note that we wrote $\|x\|^2$ for $s(x,x)$. The number is called the norm of $x \in X$.

A vector space X over $\mathbb{K}$ ($\mathbb{K} = \mathbb{R}, \mathbb{C}$) is called the inner product space if we have equipped it with a preferred inner product. *(From the theory of quadratic equations, we know that if $At^2 + Bt + C = A(t-\alpha)(t-\beta)$. Assuming without loss of generality $\alpha > \beta$, we have $(t-\alpha) > 0 \Rightarrow (t-\beta) > 0$ and $(t-\beta) < 0 \Rightarrow (t-\alpha) < 0$ so that $(t-\alpha)(t-\beta) > 0$ always holds and the expression $At^2 + Bt + c$ has the same sign as A unless t has a value lying between α and β in which case the expression has the sign opposite to that of A. If α, β are real with $\alpha = \beta$ then $At^2 + Bt + C = A(t-\alpha)^2$ has the same sign which A has. If α and β are complex then $At^2 + Bt + C = A\left[(t + \frac{B}{2A})^2 + \frac{4AC - B^2}{4A^2}\right]$ with $4AC - B^2 \geq 0$ since the roots are complex and we conclude that again the expression $At^2 + Bt + C$ has the same sign which A has.*

To sum up: Suppose A, B, C real numbers. Then *(i)* $At^2 + Bt + C \geq 0$ *for all* $t \in \mathbb{R}$ *iff* $B^2 - 4AC \leq 0$, $A > 0$ *we have used this here with* $A = |c|^2 b$, $B = -2|c|^2$, $C = a$

(ii) $At^2 + Bt + C \leq 0$ *for all* $t \in R$ *iff* $B^2 - 4AC \leq 0$, $A < 0$ *)*

We say $X \times X \xrightarrow{s} \mathbb{K}$ is sesquilinear iff $s(\lambda x, u + w\mu) = \overline{\lambda} s(x, u) + \overline{\lambda} s(x, w)\mu$ and $s(x + u, w\mu) = s(x, w) + s(u, w)\mu$ i.e. s is conjugate-linear in the first and linear in the second variable; 'sesqui' means 'one and a half'. This is the 'physicists's convention'; the 'mathematician's convention' is linear in the first and conjugate-linear in the second variable. Clearly, if s is sesquilinear in the physicist' convention, $\overline{s}$ given by $\overline{s}(x, u) := s(u, x)$ is sesquilinear in the mathematician's convention; one can adhere to either of the two. 'Conjugate-linear' is also called 'semi-linear' whence the name 'sesqui-linear'.

2. Suppose now X is an inner-product space. Then $(x \mid u) = (z \mid u)$ for each $u \in X$ means $(x - z \mid x - z) = (x \mid x) - (x \mid z) - (z \mid x) + (z \mid z) = (z \mid x) - (z \mid z) - (z \mid x) + (z \mid z) = 0$ and since $s(w, u) := (w \mid u)$ is positive-definite (so that $(w \mid w) = 0 \Rightarrow w = 0$), we get $x - z = 0$ i.e $x = z$.

Similarly, $(u \mid x) = (u \mid z)$ for each $u \in X$ means $(x - z \mid x - z) = (x \mid x) - (x \mid z) - (z \mid x) + (z \mid z)$ $= (x \mid z) - (x \mid z) - (z \mid z) + (z \mid z) = 0$ and again, since $s(w, u) := (w \mid u)$ is positive-definite, we get $x - z = 0$ i.e $x = z$.

To sum up:

If s *is positive -definite (and in particular if* s *is an inner product)* $s(u, x) = s(u, z)$ *for each* $u \in X$ *forces* $x = z$*(in particular* $(u \mid x) = (u \mid z)$ *for each* $u \in X$ *forces* $x = z$*); similarly,* $s(x, u) = s(z, u)$ *for each* $u \in X$ *forces* $x = z$*). Further,* $(x \mid u) = 0$ *for all* $u \in X$ *means* $(x \mid x) = 0$ *as well and then positive-definiteness ensures* $x = 0$

3. Since the inner product is positive, $\|x\| = +\sqrt{(x \mid x)} \geq 0$; further, since it is positive-definite, $\|x\| = +\sqrt{(x \mid x)} = 0$ forces $x = 0$. For $\lambda \in \mathbb{K}$ ($\mathbb{K} = \mathbb{C}$, or $\mathbb{R}$) we have $\|x\lambda\| = +\sqrt{(x\lambda \mid x\lambda)} = +\sqrt{\overline{\lambda}(x \mid x)\lambda} = +\sqrt{\overline{\lambda}\lambda(x \mid x)} = |\lambda|\|x\| = \|x\||\lambda|$ ($\because$ $(x \mid x) \in \mathbb{R}$). Further, $\|x+u\|^2 = (x+u \mid x+u)$ $= (x \mid x) + (x \mid u) + (u \mid x) + (u \mid u) = \|x\|^2 + 2(x \mid u) + \|u\|^2$

$\leq \|x\|^2 + 2|(x \mid u)| + \|u\|^2 \leq \|x\|^2 + 2\|x\|\|u\| + \|u\|^2 = (\|x\| + \|u\|)^2$. To sum up, we have proved *(using the Cauchy-Schwarz inequality in the last argument)*

The norm function $X \xrightarrow{x \to \|x\|} [0, \infty)$ satisfies:

(i) $\|x\| = 0 \Rightarrow x = 0$ *(of course* $\|x\| \geq 0$ *and* $\|0\| = 0$ *hold)*

(ii) $\|x\lambda\| = \|x\||\lambda| = |\lambda|\|x\|$

(iii) $\|x + u\| \leq \|x\| + \|u\|$

for all $x, u \in X$, $\lambda \in \mathbb{H}$ (or $\lambda \in \mathbb{C}$ or $\lambda \in \mathbb{R}$)

The inequality $\|x + u\| \leq \|x\| + \|u\|$ is called the triangle inequality.

4. In an inner product space, $\|u \pm x\|^2 = (u \pm x \mid u \pm x) = \|u\|^2 \pm 2(u \mid x) + \|x\|^2$ so that $\|u+x\|^2 + \|u-x\|^2 = \left[\|u\|^2 + \|x\|^2\right]$ for each $u, x \in X$. We refer to this equation as the parallelogram law.

5. We write $u \perp x$ *(read: u is orthogonal to x)* iff $(u \mid x) = 0$; this is clearly a symmetric relation *($u \perp x$ iff $x \perp u$).*

The Cauchy-Schwarz inequality $|(u \mid x)| \leq \|u\|\|x\|$ ensures that if u, x are nonzero, we have $\frac{|(u|x)_0|}{\|u\|\|x\|} \leq \frac{|(u|x)|}{\|u\|\|x\|} \leq 1$ so that

$-1 \leq \frac{(u|x)_0}{\|u\|\|x\|} \leq 1$

We define $\theta = \cos^{-1} \frac{(u|x)_0}{\|u\|\|x\|}$ to be the angle between u and x. This is called

(i) obtuse if $(u \mid x)_0 < 0$, and (ii) acute if $(u \mid x)_0 \rangle 0$; we see it is a right angle if $(u \mid x)_0 = 0$. Note that now $\|u \pm x\|^2 = \|u\|^2 \pm 2\|u\|\|x\| \cos\theta + \|u\|^2$ Clearly, if $\cos\theta = 0$ we get $\|u \pm x\|^2 = \|u\|^2 + \|x\|^2$. Moreover, $u \perp x \Rightarrow \cos\theta = 0$ but $\cos\theta = 0 \not\Rightarrow u \perp x$ unless $\mathbb{K} = \mathbb{R}$.

A set $\{x_\alpha \in X\}$ is called an orthogonal set iff $x_\alpha \perp x_\beta$ whenever $\alpha \neq \beta$. Then $\left\{\frac{x_\alpha}{\|x_\alpha\|}\right\}$ also has this property if each $x_\alpha \neq 0$ and each vector in this set has norm 1.

6. (i) A set $\{x_\alpha \in X\}$ is called an orthonormal set iff $(x_\alpha \mid x_\beta) = \delta^\alpha_\beta$ (δ^α_β is the 'Kronecker delta',

$$\delta^\alpha_\beta = \begin{cases} 1 & \text{if } \alpha = \beta \\ 0 & \text{otherwise} \end{cases}$$

.

(ii) An orthogonal set of nonzero vectors must be linearly independent

(If $\{x_\alpha\}$ is orthogonal and $x_1\lambda^1 + \cdots + x_n\lambda^n = 0$ then $0 = (x_1 \mid 0) = (x_i \mid x_1\lambda^1 + \cdots + x_n\lambda^n) =$ $(x_i \mid x_1\lambda^1 + \cdots + x_n\lambda^n) = (x_i \mid x_1)\lambda^1 + \cdots + (x_i \mid x_n)\lambda^n = \|x_i\|\lambda^i$, $(\because (x_i \mid x_j) =$

$$\delta^i_j = \begin{cases} \|x_i\| & \textit{if } i = j \\ 0 & \textit{otherwise} \end{cases}$$

; thus $\lambda^i = 0$ for $1 \leq i \leq n$ because $\|x\| \neq 0$ for any i.)

(iii) Suppose $\{x_\alpha\}$ is orthonormal and $u \perp x_\alpha$ for each α forces $u = 0$. Then we can surely find no $u \in X$ with $\|u\| = 1$ such that $u \perp x_\alpha$ for each α which means that $\{x_\alpha\}$ cannot be enlarged to a bigger orthonormal set and if $\{x_\alpha\} \cup \{u\}$ is orthonormal, u must be one of the $\{x_\alpha\}$. Thus, if we decide to call a set total orthonormal set when it is an orthonormal set which cannot be enlarged to a bigger orthonormal set with the property that $x \in X$, $x \perp u_\alpha$ for each α forces x to be 0, $\{u_\alpha\}$ is total. A total orthonormal set will also be a linearly independent set *(since every orthonormal set is linearly independent but it is not necessarily a maximal linearly independent set since it is not proved that it cannot be enlarged to a bigger linearly independent set; what is proved is that it cannot be enlarged to a bigger orthonormal set. Thus a total orthonormal set is not necessarily a Hamal basis of the vector space under consideration. Hamel basis=basis of a vector space defined earlier)* However, we have the following

Theorem 0.1. If X is a finite dimensional innerproduct space over $\mathbb{K}(\mathbb{C}$ or $\mathbb{R})$ then every total orthonromal set is a Hamel basis of X.

Proof : Suppose $\{x_0, \cdots, x_{n-1}\}$ is an orthogonal set of nonzero vectors in an innerproduct space X having the property that $(x_i \mid x) = 0$ for $0 \le i \le n-1$ forces $x = 0$. Let $x \in X$ and $z = \sum_{i=0}^{n-1} x_i(x_i \mid x) = x_0(x_0 \mid x) + \cdots + x_{n-1}(x_{n-1} \mid x)$; then $(x_i \mid z) = (x_i \mid x)$ so that $(x_i \mid z - x) = 0$ for $0 \le i \le n-1$ and we have $x = z = \sum_{i=0}^{n-1} x_i(x_i \mid x)$ which shows that x is in the vector space generated by $\{x_0, \cdots, x_{n-1}\}$. On the other hand, $\{x_0, \cdots, x_{n-1}\}$ is linearly independent *(since every orthogonal set of nonzero vectors is).* Thus $dim\, X = n$ and $\{x_0, \cdots, x_{n-1}\}$ is a Hamel basis of X.

7. In any innerproduct space X, one says that a sequence $\{x_n \in X\}$ is a Cauchy sequence iff it is possible to find an integer k such that $\|x_k - x_{n+k}\| \to 0$ as $n \to \infty$ and one says that the innerproduct space is a Hilbert space iff each Cauchy sequence is convergent in the sense that if $\{x_n \in X\}$ is a Cauchy sequence, one can find $x \in X$ with $\|x_n - x\| \to 0$ as $n \to \infty$; the vector $x \in X$ is then called the limit of $\{x_n\}$ and one writes $x_n \to x$ or $x = \lim_{n\to\infty} x_n$. Each sequence $\{x_n\}$ yields a new sequence $s_n := \sum_{j=1}^{n} x_j$ and one writes $x = \sum_{j=1}^{\infty} x_j$ iff $s_n \to x$ saying that the *series* $\sum x_j$ converges to x.

8.

Theorem 0.2. When $\{x_n\}_{n=0}^{\infty}$ is a linearly independent set in an innerproduct space X, $y_n := x_n - \sum_{j=0}^{n-1} u_j(u_j \mid x_n)$; $u_n = \frac{y_n}{\|y_n\|}$, $y_0 = x_0$ supply an orthonormal set $\{u_n\}_{n=0}^{\infty}$ such that $span\{u_0, \cdots, u_{n-1}\} =$

$span\{x_0, \cdots x_{n-1}\}$.

Proof : If $\{x_0\}$ is linearly indpendent, $y_0 = x_0 \neq 0$ an $u_0 = \frac{y_0}{\|y_0\|}$ is well defined and clearly $span\{u_0\} = span\{x_0\}$ with $\{u_0\}$ orthonormal. Assume now that $y_0, \cdots, y_{n-1}$ and $u_0, \cdots, u_{n-1}$ have been defined as above with $\{x_0, \cdots, x_{n-1}\}$ linearly independent and $span\{u_0, \cdots, u_{n-1}\} = span\{x_0, \cdots, x_{n-1}\}$, $u_0, \cdots, u_{n-1}$ orthonormal. Then $y_n = x_n - \sum_{j=0}^{n-1}(u_j \mid x_n) \neq 0$ if we have $\{x_0, \cdots, x_{n-1}, x_n\}$ linearly independent and $u_n = \frac{y_n}{\|y_n\|}$ is well defined. Further, for $0 \leq m \leq n-1$,

$$(u_m \mid u_n) = (u_m \mid x_n - \sum_{j=0}^{n-1} u_j(u_j \mid x_n))\frac{1}{\|y_n\|}$$

$$= [(u_m \mid x_n) - (u_m \mid u_m)(u_m\ x_n)]\,\frac{1}{\|y_n\|}\ (\because\ (u_m \mid u_j) = \delta_m^j\ for\, 0 \leq j \leq n-1)$$

$$= [(u_m \mid x_n) - (u_m \mid x_n)]\,\frac{1}{\|y_n\|}\ \ (\because\ (u_m \mid u_m) = 1)$$

$$= 0$$

Thus $u_n \perp u_m$ for $0 \leq m \leq n-1$ and of course $(u_n \mid u_n) = \frac{1}{\|y_n\|}(y_n \mid y_n)\frac{1}{\|y_n\|} = 1$ so that $\{u_0, \cdots, u_n\}$ is an orthonormal set. Further, $x_n = u_n\|y_n\| + \sum_{j=0}^{n-1} u_j(u_j \mid x_n) = \sum_{j=1}^{n} u_j(u_j \mid x_n)$ $(\because\ (u_n \mid x_n) = (u_n \mid y_n) = \sum_{j=0}^{n-1}(u_n \mid u_j)(u_j \mid x_n) = (u_n \mid u_n \mid y_n\|) = \|y_n\|)$ and thus $x_n \in span\{u_0, \cdots, u_n\}$ proving that $span\{u_0, \cdots, u_n\} = span\{x_0, \cdots, x_n\}$ since it is known that $span\{x_0, \cdots, x_{n-1}\} = span\{u_0, \cdots, u_{n-1}\}$. Thus, in particular, if we have a finite dimensional Hilbert space, an orthonormal Hamel basis exists for it. The process of construction an orthonormal set out of a linearly independent set outlined in the theorem is called Gram-Schmidt orthonormalization.

9.

Definition 0.1. Say a function $X \xrightarrow{T} X$ *(X an innerproduct space)* is adjointable iff there is a function $X \xrightarrow{T^+} X$ *(read T^+ as "T dagger")* such that the 'adjointness condition'

$(T^+x \mid u) = (x \mid Tu)$ for each $x, u \in X$ is satisfied; we say T^+ is the Hilbert adjoint of T

Proposition 0.3. If $X \xrightarrow{T} X$ is adjointable, the Hilbert adjoit T^+ is unique; further, both $X \xrightarrow{T} X$ and $X \xrightarrow{T^+} X$ are linear.

Proof :

(1) We first note that if $x = x'$, we have $(T^+x \mid u) = (x \mid Tu) = (x' \mid Tu) = (T^+x' \mid u)$ at each $u \in X$ so that(by positive- definiteness, para 2 page 3 above) we have $T^+x = T^+x'$. Thus the adjointness condition, if it holds, does define a function $X \xrightarrow{T^+} X$.

(2) If there are two Hilbert adjoints, say T^+ and $T^@$, for T, we have $(T^+x \mid u) = (x \mid Tu) = (T^@x \mid u)$ at each $u \in X$ and hence *(by positive-definiteness)* $T^+x = T^@x$; since this holds at each $x \in X$, we have $T^+ = T^@$. Thus there is at most one Hilbert adjoint for T.

(3) If the Hilbert adjoint T^+ exists, we have $(w \mid T(x+u\lambda)) = (T^+w \mid x+u\lambda) = (T^+w \mid x) + (T^+w \mid u)\lambda = (w \mid Tx) + (w \mid Tu)\lambda = (w \mid Tx + (Tu)\lambda)$ at each $w \in X$ so that *(by positive-definiteness again)* we have $T(x+u\lambda) = Tx + (Tu)\lambda$; this being true at each $x, u \in X$, $\lambda \in \mathbb{K}$, we see that T is *(right)*linear. Further, $(T^+(x+w\lambda) \mid u) = (x+w\lambda \mid Tu) = (x \mid Tu) + (w\lambda \mid Tu) = (x \mid Tu) + \overline{\lambda}(w \mid Tu) = (T^+x \mid u) + \overline{\lambda}(T^+w \mid u)$

$= (T^+x \mid u) + ((T^+w)\lambda \mid u) = ((T^+x + (T^+w)\lambda) \mid u)$ for each $u \in X$

Therefore*(by positive-definiteness)* we get

$T^+ = (x+w\lambda) = T^+(x) + (T^+(W))\lambda$, and this being true at each $x, w \in X$, $\lambda \in \mathbb{K}$, we conclude that $X \xrightarrow{T^+} X$ is *(right)*linear.

10. The Next question is to determine which $X \xrightarrow{T} X$ are adjointable. The question makes sense clearly for linear operators $X \xrightarrow{T} X$ only.

Suppose X is finite dimensional, say $dim\, X = n < \infty$ coorinatizing by say $(\underline{\underline{e}}, \underline{\underline{\epsilon}})$, $\underline{\underline{\epsilon}}$ is the dual basis of $\underline{\underline{e}}$, we already know that $X \times X \xrightarrow{s} \mathbb{K}$ is a hermitian sesquilinear form iff $[s_i^j]_{n\times n}$ $(s_I^j = s(e_j, e_i))$ is a hermitian matrix i.e. $s_i^j = \overline{s_j^i}$ *(verify this)*

Thus s_j^i must be real *(for $\mathbb{K} = \mathbb{C}$)*. For an inner product on X (*thus X is the finite dimensional Hilbert space)* we choose an orthonormal basis e *(which is possible by Gram-Schmidt)* and note $(x \mid u) = (\sum_{i=1}^{n-1} e_i x^i \mid \sum_{j=0}^{n-1} e_j u^j) = \sum_{i=0}^{n-1}\sum_{j=0}^{n-1} \overline{x^i}(e_i \mid e_j)u^j = \sum_{i=0}^{n-1} \overline{x^i}u^i$ ($\because$ $(e_i \mid e_j) = 0$ for $i \neq j$ $(e_i \mid e_i) = 1$)

Thus $(x \mid u) = x^*u$ where x^* is the conjugate-transpose of the column vector $x = \begin{pmatrix} x^0 \\ . \\ . \\ . \\ x^{n-1} \end{pmatrix}_{n\times 1}$

and u is the column vector $\begin{pmatrix} u^0 \\ . \\ . \\ . \\ u^{n-1} \end{pmatrix}_{n\times 1}$ *(being calculated with respect to* (e,ϵ) *).*

Therefore $(u \mid Tx) = u^*Tx = u^*(x^*T^*)^*$ (T being the matrix of T with respect to (e,ϵ) and T^* its congate-transpose)

$= (x^*T^*u)^* = (T^*u)^*x$ ($\because x^{**} = x$)

$= (T^{@}u \mid x)$

Where $X \xrightarrow{T^{@}} X$ is the operator having the matrix T^* with respect to (e,ϵ) and sending u to T^*u.

But since $(u \mid Tx) = (T^+u \mid x)$ and the Hilbert adjoint T^+ is unique, we conclude that $T^{@} = T^+$ and thus the matrix of T^+ with respect to (e,ϵ) T^*

To sum up:

Every linear operator $X \xrightarrow{T} X$ is adjointable when X is a finite dimensional Hilbert space.

This result does not hold in infinite dimensional inner product space in general *(when X is a Hilbert space and $\|T\| := \sup\limits_{\|x\|\leq 1} \|Tx\| < \infty$, it is however true. We are not concerned with these considerations in this course).*

11. Now take $\mathbb{K} = \mathbb{C}s.$ that X is a finite dimensional complex Hilbert space, say $\mathbb{C}^n$.

Let $L(X,X)$ be the space of all linear operators $X \to X$. Then we know that $L(X,X)$ has dimension n^2 and if (e,ϵ) is a coordinatization of $\mathbb{C}^2$,

$$\left\{ |e_j\rangle\langle\epsilon^i| \;\middle|\; \begin{array}{c} 0 \leq i \leq n-1 \\ 0 \leq j \leq n-1 \end{array} \right\} \text{ is a basis of } L(X,X).$$

For $X \xrightarrow{A} X$, we define $trace\,(A) := \sum\limits_{i=0}^{n-1} \langle\epsilon^i \mid Ae_i\rangle$; *(Thus if we denote the associated matrix of A by $a = [a_i^j]$, we have $trace\,(A) = \sum\limits_{i=0}^{n-1} a_i^i$; we then define the trace of a square matrix to be the sum of its diagonal elements and in view of the bijectivity $\mathbb{K}_{n\times n} \longleftrightarrow^{A\leftrightarrow a} Lim(X,X)$ (just define A by $Ax := ax$ for $x \in X$) the two definitions record the same mathematical concept).* Clearly, $L(X,X) \xrightarrow{trace} \mathbb{C}$ is a linear form *(verify it).*

If (d,δ) δ is the dual basis of d, any other coordinatization for X we use the decomposition of identity

$X \xrightarrow{id} X = X \xrightarrow{\sum |e_i\rangle\langle\epsilon^i|} X = X \xrightarrow{\sum |d_j\rangle\langle\delta^j|} X$ to find

$trace\,(A) = \sum_{i=0}^{n-1} \langle \epsilon^i \mid Ae_i\rangle = \sum_{i=0}^{n-1} \langle \epsilon^i \mid A(|de_i)\rangle$

$= \sum_{i=0}^{n-1}\sum_{j=0}^{n-1} \langle \epsilon^i \mid Ad_j\rangle\langle\delta^j \mid e_i\rangle$

$= \sum_{j=0}^{n-1}\sum_{i=0}^{n-1} \langle \delta^j \mid e_i\rangle\langle\epsilon^i \mid Ad_j\rangle$

$= \sum_{j=0}^{n-1} \langle \delta^j \mid |d(Ad_j)\rangle = \sum_{j=0}^{n-1} \langle \delta^j \mid Ad_j\rangle$

which means that the trace of A is coordinate-free *(trace (A) is the same whatever coordintization is chosen).*

12. Considering $(X, (e, \epsilon)) \xrightarrow[a,\,b]{A,\,B} (X, (e, \epsilon))$ we define $(A \mid B) := trace\,(A^+B) = trace\,(a^*b) \in \mathbb{C}$ *(since a^* is the matrix of A^+ as noted in para (10) above).* Then

(i) $(\lambda A + B \mid C + D\mu) = trace\,((\lambda a + b)^*(c + d\mu)) = trace\,((b^* + a^*\overline{\lambda})(c + d\mu))$

$= trace\,(\overline{\lambda}a^*c + b^*c + \overline{\lambda}a^*d\mu + b^*d\mu)$

$= \overline{\lambda}trace\,(a^*c) + trace\,(b^*c) + \overline{\lambda}trace\,(a^+d)\mu + trace\,(b^*d)\mu$

$= \overline{\lambda}(A \mid C) + (B \mid C) + \overline{\lambda}(A \mid D)\mu + (B \mid D)\mu$

which says that $(A \mid B)$ as defined is a sesquilinear form on $L(X, X)$

(ii) $(A \mid B) = trace\,(a^*b) = \overline{trace\,(b^*(a^*)^*)} = \overline{trace\,(b^*a)} = \overline{(B \mid A)}$ which says it is hermitian, and

(iii) $(A \mid A) = trace\,(a^*a) = \sum_{i=0}^{n-1} (a^*a)^i_i = \sum_{i=0}^{n-1}\sum_{j=0}^{n-1} (a^*)^i_j a^j_i$

$= \sum_{i=0}^{n-1}\sum_{j=0}^{n-1} \overline{a^j_i} a^j_i = \sum_{i=0}^{n-1}\sum_{j=0}^{n-1} |a^j_i|^2$ which says $(A \mid A) \geq 0$, $(A \mid A) = 0$ iff each $a^j_i = 0$, $0 \leq i \leq n-1$, $0 \leq i \leq j-1$ i.e. $a = 0$ i.e. $A = 0$ i.e. this is positivedefinite.

Thus $(A, B) \to (A \mid B) = trace\,(A^+B) = trace\,(a^*b)$ defines an inner product on $L(X, X)$ into a complex Hilbert space of dimension n^2. This inner product is called the Hilbert-Schmidt inner product on $L(X, X)$ and equipped with this inner product, $L(X, X)$ is called the Liouville space of the Hilbert space X. *(the name 'Liouville space' is used mostly in quantum mechanics; see Quantum Information: An Overview by Gregg Jaeger; spinger 2007 page 248)*

(iv) If $dim\,V = n\langle\infty$, we have essentially $V = \mathbb{C}^n$ and V^*, the dual, also has dimension n when V is an ips, let us write, given $z \in V$, $V \xrightarrow{\varphi_z} \mathbb{C}$ supplied by $\varphi_z(x) := (z \mid x)$. Then $\varphi_z(x + u\lambda) = (z \mid x + u\lambda) = (z \mid x) + (z \mid u)\lambda = \varphi_z(x) + (\varphi_z(u))\lambda$ so that $\varphi_z \in V^{tr}$ Further, $V \xrightarrow{\varphi} V^{tr}$ defined by $\varphi(z) := \varphi_z$ obeys $(\varphi(\lambda z + w))(x) = (\lambda z + w \mid x) = \overline{\lambda}(z \mid x) + (w \mid x) = \overline{\lambda}\varphi_z(x) + \varphi_w(x)$ at each $x \in X$

which means $\varphi(\lambda z + w) = \overline{\lambda}\varphi(z) + \varphi(w)$ showing that $V \xrightarrow{\varphi} V^{tr}$ is conjugate-linear. If $\varphi_z = \varphi_w$ the at each $x \in V$ we have $\varphi_z(x) = (z \mid x) = \varphi_w(x) = (w \mid x)$ which means*(by positive definiteness)* $z = w$. Thus this φ is a bijective*(injectivity would ensure this because* $dim\, V = dim\, V^{tr} \langle \infty$*)*. Explicitly, Suppose $f \in V^{tr}$ and choose an orthonormal basis $\{e_0, \cdots, c_{n-1}\}$ of V; with respect to this basis of V (and the basis $\{1\}$ of $\mathbb{C}$ the linear form $V \xrightarrow{f} \mathbb{C}$ will be given by a row vector $[a_0, \cdots\cdots, a_{n-1}] \in Mat_{1\times n}\mathbb{C}$ so that for any $x = \sum_{i=0}^{n-1} e_i x^i$ we have $f(x) = \sum_{i=0}^{n-1} f(e_i)x^i$, $f(e_i) = a_i$ so that if $a = \sum \overline{a_j} e_j$, we have $\varphi_a(x) = (a \mid x) = (\sum \overline{a_j} e_j \mid \sum e_i x^i) = \sum a_i x^i = f(x)$ at each $x \in V$ and thus $f = \varphi_a$.

To sum up:

Riesz Representation Theorem: If $dim\, V \langle \infty$, there exists a conjugate-linear bijection $V \xrightarrow{T} V^{tr}$ given by $T(z) \in V^{tr}$ computing as $(T(z))(x) = (z \mid x)$ for $z, x \in V$. Given $f \in V^{tr}$, it is

$$z = \begin{pmatrix} \overline{a_0} \\ . \\ . \\ . \\ \overline{a_{n-1}} \end{pmatrix} \in V$$

which is such that $T(z) = f$; we say z is the representor for f and this column vector is with reference to a chosen orthonormal basis $\underline{e}$; $a_i = f(e_i)$.