Inner Product Space

1. Recall that X is called a vector space over a field $\mathbb{F}$ when X is an abelian group equipped with an action $X\times\mathbb{F}\xrightarrow{(x,y)\to x\lambda} X$ satisfying expected identities like $(x+u)\lambda = x\lambda+u\lambda$, $x(\lambda+\mu)=x\lambda+x\mu$ etc; we also can, and do, write $x\lambda\in X$ as λx *(this is not a problem since $\mathbb{F}$ is commutative)*

1.1 Further, given two vector spaces X and Y *(over the same field $\mathbb{F}$)* a linear transformation $X\xrightarrow{A} Y$ is a function such that $A(x+u\lambda)=A(x)+A(u)\lambda$ for all $x,u\in X$, $\lambda\in\mathbb{F}$. For $X\xrightarrow{A,B} Y$, we define $\lambda A+B$ by $(\lambda A+B)(x)=\lambda A(x)+B(x)$ at each $x\in X$. Then the collection $L(X,Y)$ of all linear transformations $X\to Y$ becomes a vector space. Verify this fact. What is $0\in L(X,Y)$? Is $X\xrightarrow{idX} X$ in $L(X,X)$?

1.2 A linear transformation $X\xrightarrow{A} Y$ is called an isomorphism iff it is a bijection *(prove that then $Y\xrightarrow{A^{-1}} X$ is also and further, $A^{-1}A=id_X$, $AA^{-1}=id_Y$).* If there is an isomorphism $X\xrightarrow{A} Y$, we write $X\cong Y$ and regard the two vector spaces X and Y as 'the same', identifying $x\in X$ with $A(x)\in Y$.

1.2.1 Prove that $L(\mathbb{F},X)\cong X$.

Hint:

Define $L(\mathbb{F},X)\xrightarrow{f} X$ by $f(A):=A(1)$. Then $f(\lambda A+B)=(\lambda A+B)(1)=\lambda A(1)+B(1)=\lambda f(A)+f(B)$ so that f is a linear transformation.

Next, for $x\in X$, define $X\xrightarrow{g} L(\mathbb{F},X)$ by writing $(g(x))(\mu):=\mu x$. Then $(g(x))(\lambda\mu+\nu)=(\lambda\mu+\nu)(x)=\lambda(\mu x)+\nu x=\lambda g(x)(\mu)+g(x)(\nu)$ so that $g(x)$ is indeed a linear transformation $\mathbb{F}\xrightarrow{X}$ and we do have $X\xrightarrow{g} L(\mathbb{F},X)$. Further, we have $(g(\lambda x+u))(\mu)=\mu(\lambda x)+\nu x=\lambda\mu x+\mu u=(\lambda g(x)+g(u))(\mu)$ at each $\mu\in\mathbb{F}$ so that the two functions $g(\lambda x+u)$ and $\lambda g(x)+g(u)$ are the same and we get $g(\lambda x+u)=\lambda g(x)+g(u)$. This means g is a linear transformation. Then $((gf)(A))(\lambda)=(g(A(1)))(\lambda)=\lambda A(1)=A(\lambda.1)=A(\lambda)$ at each $\lambda\in\mathbb{F}$ so that $gf(A)=A$ at each $A\in L(\mathbb{F},X)$. Thus we have $gf=id_{L(\mathbb{F},X)}$. Also, $(fg(x))=(g(x))(1)=1.x=x$ at each $x\in X$ which means $fg=id_X$. This proves that $X\cong L(\mathbb{F},X)$ as desired.

Remark 0.1. In view of this, we can think of a vector $x \in X$ as a linear transformation $\mathbb{F} \xrightarrow{x} X$ and a scalar $\lambda \in \mathbb{F}$ can be regarded as a linear transformation $\mathbb{F} \xrightarrow{\lambda} \mathbb{F}$. Viewed like this, $x\lambda \in X$ is simply the composition $\mathbb{F} \xrightarrow{\lambda} \mathbb{F} \xrightarrow{x} X$. That this can be also thought as λx requires a more complicated argument which is beyond our scope.

1.3 The vector space $L(X, \mathbb{F})$ will be written as X' and called the transpose*(sometimes, the dual)* of X. An element $\varphi \in X'$ will be called a *(linear)* form *(sometimes, functional)* on X. We shall write $\varphi(x) \in \mathbb{F}$ as $< \varphi \mid x >$ *(this is the Dirac notation)*. A function $X' \times X \xrightarrow{\beta} \mathbb{F}$ satisfying $\beta(\lambda\varphi + \alpha, x + u\mu) = \lambda\beta(\varphi, x) + \beta(\alpha, x) + \lambda(\varphi, u)\mu + \beta(\alpha, u)\mu$ for all $\varphi, \alpha \in X'$, $x, u \in X$, $\lambda, \mu \in \mathbb{F}$ *(thus it is simply a linear form in each variable separately)* will be called bilinear.

1.3.1 Show that $X' \times X \xrightarrow{<-1->} \mathbb{F}$ is bilinear.

Hint:

$< \lambda\varphi + \alpha \mid x + u\mu >= \lambda < \varphi \mid x > + < \alpha \mid x > +\lambda < \varphi \mid u > \mu + < \alpha \mid u > \mu$. *(To see this, simply note that $(\lambda\varphi + \alpha)(x + u\mu)$ is the LHS and φ, α are linear forms on X).*

1.3.2 For a linear transformation $X \xrightarrow{A} Y$, define $Y' \xrightarrow{A'} X'$ by writing $< A'(\psi) \mid x >:=< \psi \mid A(x)$ for each $\psi \in Y'$, $x \in X$ and show that this is well defined linear transformation.

Hint:

$\psi = \varsigma$ means that at each $x \in X$ we have $(A'(\psi))(x) =< A'(\psi) \mid x >=< \psi \mid A(x) >=< \varsigma \mid A(x) >=< A'(\varsigma) \mid x >= (A'(\varsigma))(x)$ i.e. $A'(\psi) = A'(\varsigma)$. Thus a single ψ does not have two candidates for $A'(\psi)$. Since $< A'(\psi) \mid x+u\lambda >=< \psi \mid A(x+u\lambda) >=< \psi \mid Ax > + < \psi \mid Au > \lambda(\because$ A is a linear tr) $=< A'(\psi) \mid x > + < A'(\psi) \mid u > \lambda$, we conclude that $A'(\psi) \in X'$. Since $< A'(\lambda\psi + \varsigma) \mid x >=< \lambda\psi + \varsigma \mid Ax >= \lambda < \psi \mid Ax > + < \varsigma \mid Ax >= \lambda < A'(\psi) \mid x > + < A'(\varsigma) \mid x >$ we conclude that at each $x \in X$ we have $(A'(\lambda\psi + \varsigma))(x) = \lambda(A'(\psi))(x) + (A'(\varsigma)(x)$ so that $A'(\lambda\psi + \varsigma) = \lambda A'(\psi) + A'(\varsigma)$; this says A' is linear.

1.3.3 Write $< \varphi \mid x >$ as $< \Gamma(x) \mid \varphi >$ and extract a linear transformation $X \xrightarrow{\Gamma} X''$ from this.

Hint:

If $x = u$, we have $< \Gamma(x) \mid \varphi >=< \varphi \mid x >=< \varphi \mid u >=< \Gamma(u) \mid \varphi >$ at each $\varphi \in X'$ so that $\Gamma(x) = \Gamma(u)$ and we have no two candidates for $\Gamma(x)$, given a single $x \in X$. Since $< \Gamma(x + u\lambda) \mid \varphi >=< \varphi \mid (x + u\lambda) >=< \varphi \mid x > + < \varphi \mid u > \lambda =< \Gamma x \mid \varphi > +\Gamma(u)\lambda$, we have

$X \xrightarrow{\Gamma} X''$ *linear (since* $\Gamma(x) \in X''$ *is clear:* $<\Gamma(x) \mid \lambda\varphi + \varsigma> = <\lambda\varphi + \varsigma \mid x> = \lambda <\varphi \mid x> + <\varsigma \mid x> = \lambda <\Gamma(x) \mid \varphi> + <\Gamma(x) \mid \varsigma>$*).*

1.4 If $S = \{x_i \mid i \in I\}$ is any indexed collection of vectors in X, we say that $x = \sum_{i \in I} x_i \lambda_i$ is a linear combination of vectors from S; in this expression, it is understood that all but finitely may $\lambda_i \in \mathbb{F}$ are zero so that $x = \sum_{i \in I} x_i \lambda_i$is a finite sum and thus we have $x \in X$. In particular, we agree that $O = \sum_{i \in \phi} x_i \lambda_i$, i.e. the vacuous linear combination is the zero vector. Also, $\lambda_i = 0$ for all $i \in I$ supplies the 'trivial combination' and yields the zero vector.

1.4.1 The set S is linearly dependent if the zero vector can be expressed as a non trivial linear combination of distinct elements from S; S is called linearly independent if it is not linearly dependent. In particular, the empty set is linearly independent. Clearly then, S is linearly independent iff there is at most one way to write x as $x = \sum_i x_i \lambda_i$ *(indeed,* $\sum x_i \lambda_i = \sum x_i \mu_i$ *means* $O = \sum x_i(\lambda_i - \mu_i)$ *which means* $\lambda_i - \mu_i = 0$ *for each* i *since* S *is linearly independent).*

1.4.2 The collection $\{x \in X \mid \sum_{i \in I} x_i \lambda_i, x_i \in S\}$ is clearly a subspace of X *(prove it!)* and is called the subspace spanned by S; we shall denote this subspace by $<< S >>$ A collection $S = \{x_i \mid i \in I\}$ will be called a basis of X iff $<< S >> = X$ and S is linearly independent i.e.S is a basis of X iff each $x \in X$ can be written in exactly one way as a linear combination $x = \sum x_i \lambda_i$, $x_i \in S$, $\lambda_i \in \mathbb{F}$, only finitely many λ_i being nonzero of course.

We shall supply proofs of the following in an appendix to this handout; here we accept them and feel free to employ them.

(a) Every vector space has a basis

(b) The number of elements in any basis of a vector space X is the same; this number is called the dimension of X and we write it as $dimX$. In particular, we say that X is finite-dimensional iff dim $X < \infty$. We agree that ϕ is a basis of the vector space $\{0\}$ and thus there is exactly this space which has dimension 0.

(c) If $\{X_i\}$ is a collection of vector spaces, their product set $X = \prod X_i$ is turned into a vector space by writing $(x + y)_i := x_i + y_i$, $(\lambda x)_i = (x_i \lambda) = (x\lambda)_i = (\lambda x_i)$ (where by $x = (x_i)$ we mean the element $x \in X$ whose i-th coordinate is $x_i \in X_i$) and the projections $X \xrightarrow{proj_i} X_i$ given by $proj_i(x) = x_i$ are linear. Further, their sum, the disjoint union, can be turned into a vector space $\bigoplus X_i$ by requiring the injections $X_i \xrightarrow{inj_i} \bigoplus X_i$ given by $inj_i(x_i) = x_i$ to be linear.

We are interested only in the finite version of these cases and shall not enter into the some what intricate definitions, arguments, and differentials involved. Suffice it to say that

(i) If P and Q are two subspaces of X, $P \cap Q$ is of course a subspace*(prove it)* but $P \cap Q$ may not be; the subspace $<< P \cap Q >>$ is written $P \oplus Q$ when $P \cup Q = \{0\}$ is known; this is the only case that interests us.

We say $P \oplus Q$ is the direct sum of P and Q; its elements are $p + q$ for uniquely given $p \in P$, $q \in Q$ and $P \cap Q = \{0\}$ is of course true. One thinks of $p \in P$ as $p+0 \in P \oplus Q$ and $q \in Q$ as $0+q \in P \oplus Q$. Further, $P \oplus Q \to P$ given by $p + q \to p$ and $P \oplus Q \to Q$ given by $p + q \to q$ can be regarded as projections permitting us to think of $P \oplus Q$ as the product of P and Q while $p \to p + 0$ and $q \to 0 + q$ supply injections $Q \to P \oplus Q$, $Q \to P \oplus Q$ permitting us to think of $P \oplus Q$ as the "sum" of P and Q.

(ii) What has been said for two subspaces P and Q in *(i)* above, remains valid for a finite sum $X_1 \oplus \cdots \oplus X_n$ i.e. each $x \in X_1 \oplus \cdots \oplus X_n$ can be written uniquely as $x = x_1 + \cdots + x_n$, $x_i \in X_i$, $X_i \cap X_j = \{0\}$ if $i \neq j$ and there are injections $X_i \to X_1 \oplus \cdots X_n$ given by $x_i \to 0 + \cdots + x_i + 0 + \cdots$ as well as projection given by $x_1 + \cdots + x_n \to x_i$, $1 \leq i \leq n$. When $X = X_1 \oplus \cdots \oplus X_n$, the RHS is said to be a decomposition of X into $X_1, \cdots, X_n$.

We do not include *(c)* in our appendix and this will be not discussed further in this course. But note the following to help understanding the ideals that have come up so far.

Example 0.1. Consider $X = \sum_{i=1}^{n} X_i$, each $X_i = \mathbb{F}$. Thus $X = \mathbb{F}^n := \mathbb{F} \oplus \cdots \oplus \mathbb{F}$ (n copies). Its elements will be written as column vectors

$$x = \begin{pmatrix} x^1 \\ . \\ . \\ . \\ x^n \end{pmatrix} = \begin{pmatrix} x^1 \\ 0 \\ . \\ . \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ x^2 \\ 0 \\ . \\ 0 \end{pmatrix} + \cdots + \begin{pmatrix} 0 \\ . \\ . \\ . \\ x^n \end{pmatrix} = \sum_{i=1}^{n} e_i x^i, x^i \in \mathbb{F},\ e_i = \begin{pmatrix} 0 \\ . \\ 1 \\ 0 \\ 0 \end{pmatrix} \to i\text{-th}$$

row. This has then $e = \{e_i\}_{i=1}^n$ as a basis*(prove this by showing that e is linearly independent and spans $\mathbb{F}^n$)*. If V has dimension $n < \infty$, choose a basis $\{b_1, \cdots, b_n\}$ for V*(which must exist and have exactly n distinct elements by (a) and (b) above)*. Define $\mathbb{F}^n \xrightarrow{A} V$ by $A(e_i) = b_i$ so that for $x = \sum_{i=1}^{n} e_i x^i$, we have $Ax := \sum_{i=1}^{n} A(e_i)x^i$*(by linearity of A which is agreed upon)*$= \sum_{i=1}^{n} b_i x^i$. Then since any $v \in V$ is $v = \sum_{i=1}^{n} b_i v^i$ for exactly one n-tuple$\{v^1, \cdots, v^n\}$ of scalars*(= elements of* $\mathbb{F}$*)*, we have $A^{-1}(v) := \sum_{i=1}^{n} A^{-1}(b_i)v^i = \sum_{i=1}^{n} e_i v^i$. Show that A^{-1} is linear and $AA^{-1} = id_V$ and $A^{-1}A = id_{\mathbb{F}^n}$.

This means any vector space v of dimension $n < \infty$ is isomorphic to $\mathbb{F}^n$.

For instance, consider the vector space $P_2 = \{a_0 + a_1\theta + a_2\theta^2 | a_0, a_1, a_2 \in \mathbb{F}\}$ consisting of all polynomials of degree at most 2 with coefficients prove that $\{1, \theta, \theta^2\}$ is a basis for P_2 and thus P_2 has dimension 3. This means it is isomorphic to $\mathbb{F}^3$. Similarly, prove that P_n, the space of all polynomials in the variable θ with degree at most n has $\{1, \theta, \cdots, \theta^n\}$ as a basis and is thus isomorphic to $\mathbb{F}^n$.

Example 0.2. Consider $X = \sum_{i=1}^{\infty} X_i$, each $X_i = \mathbb{F}$. This is the space $\mathbb{F} \oplus \mathbb{F} \cdots$ *(countably many copies)*. This can be regarded as the space of column vectors $x = \begin{pmatrix} x^1 \\ . \\ . \\ x^n \\ . \\ . \end{pmatrix}$ of infinite depth; this amounts to saying that each $x \in X$ is an infinite sequence. $\mathbb{N} \xrightarrow{x} \mathbb{F}$ with $x(i) = x^i \in \mathbb{F}$. It is much more convenient to think of this as the vector space of all polynomials in the variable θ, $X := \{a = a_0 + a_1\theta + \cdots + a_n\theta^n \mid 0 \leq n \in \mathbb{N}\}$; then show that $\{1, \theta, \theta^2, \cdots\}$ is a basis of X and thus this space is infinite-dimensional.*(The polynomial $0 \in X$ is by agreement of degree $-\infty$; for this $a_i = 0$ for each i, recall that $a_0 + a_1\theta + \cdots + a_d\theta^d$ has degree d if $a_d \neq 0$)*. We denote this space by $\mathbb{F}[\theta]$.

Remark 0.2. In illustration 1,2 above, it should be noted that a polynomial over $\mathbb{F}$ i.e. an element $a = a(\theta) \in \mathbb{F}[\theta]$ is not a 'polynomial function' $\mathbb{F} \xrightarrow{\tilde{a}} \mathbb{F}$ given by $\tilde{a}(\lambda) := a_0 + a_1\lambda + \cdot s + a_n\lambda^n$. Indeed, suppose $\mathbb{F}$ is the field with only tow elements 0 and 1 i.e $\mathbb{F} = \mathbb{Z}_2$. The polynomial $a(\theta) = \theta + \theta^2 \in \mathbb{F}[\theta]$ is not the zero polynomial but *(since $1+1 = 0$ in $\mathbb{Z}_2$)* we have $\lambda + \lambda^2 = \tilde{a}(\lambda) = 0$ for each $\lambda \in \mathbb{F} = \mathbb{Z}_2$. Thus if $\tilde{\mathbb{F}}$ is the vector space of all polynomial functions $\mathbb{F} \to \mathbb{F}$ *(prove that $\tilde{\mathbb{F}}$ is a vector space)* and we consider $\mathbb{F}[\theta] \xrightarrow{f} \mathbb{F}$ given by $f(a(\theta)) = \tilde{a}$, this is not an isomorphism in general. Prove however, that if $\mathbb{F}$ has an infinite number of elements *(in particular if characteristic of* $\mathbb{F}$ *is* 0*)* the function $f(a) := \tilde{a}$ defined by $a(\theta) = a_0 + a_1\theta + \cdots + a_n\theta^n \to \tilde{a}(\lambda) := a_0 + a_1\lambda + \cdots + a_n\lambda^n$ is an isomorphism.

Hint:

A nonzero polynomial $a(\theta)$ of degree n has at most n zeros, so $\tilde{a}(\lambda) = 0$ is not possible for all λ. Thus no nonzero polynomial goes to the zero polynomial function. Prove that $f(a) := \tilde{a}$ is a linear transformation. After that? If you can't proceed, read the next subsection (1.5).

Example 0.3. Assuming that $\mathbb{F}$ has $n+1$ distinct elements $\lambda_0, \cdots, \lambda_n$, consider the polynomials $L_i(\theta) := \Pi_{j\neq i}\dfrac{\theta - \lambda_j}{\lambda_i - \lambda_j}$ so that

$$L_i(\lambda_j) = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

, $0 \leq i,\ j \leq n$. Given $a(\theta) = \displaystyle\sum_{i=0}^{n} \mu_i L_i(\theta)$, we have $a(\lambda_j) = \mu_j$ and thus a is the zero polynomial only when each $\mu_j = 0$. This means the set $\{L_0(\theta), \cdots, L_n(\theta)\}$ is linearly independent in the space P_{n+1} of polynomials in θ with degree at most n, and as such, spans a vector space of dimension $n+1$ which must be then all of P_{n+1} since $dim(P_{n+1}) = n+1$. In other words, each polynomial $a(\theta) = a_0 + \cdots + a_n\theta^n$ can be expressed uniquely as $a(\theta) = \displaystyle\sum_{i=0}^{n} a(\lambda_i)L_i(\theta)$; this formula is called the Lagrange interpolation formula.

(d) If S spans X then any linearly independent subset T os S can be completed to a basis $T \cup S_0$ for $<< S >>= X$ where $S_0 \subseteq S$ and $S_0 \cap T = \phi$

(e) In general, if P is a subspace of X, there are many subspaces Q such that $P \oplus Q = X$.

(f) If P and Q are finite dimensional subspaces of X, $dim\,(P + Q) + dim\,(P \cap Q) = dim\,P + dim\,Q$.

Where $P + Q :=<< P \cup Q >>$; in particular, $dim\,(P \oplus Q) = dim\,P + dim\,Q$.

For instance, $dim\,(\mathbb{F}^n) = dim\,\mathbb{F} + \cdots + \dim\,\mathbb{F}$ (n times)

$= 1 + \cdots + 1 = n$

(we have $dim\,\mathbb{F} = 1$ since $\{1\}$ clearly a basis because we can write any $\lambda \in \mathbb{F}$ as $\lambda.1$) because $\mathbb{F}^n = \mathbb{F} \oplus \cdots \oplus \mathbb{F}$ *(n times).*

1.5 Suppose $X \xrightarrow{A} Y$ is a linear transformation. Prove the following:

(i) The image of A *(also called 'the image of X under A' and the 'range of A')* $A(X) := \{A(x) \mid x \in X\}$ is a subspace of Y,

(ii) The kernel of A, $ker\, A := \{x \in X \mid A(x) = 0\}$, also denoted by $A^{-1}\{0\}$, is a subspace of X, and

(iii) A is injective iff $ker\, A = \{0\}$

Hint: *(If $ker\, A = \{0\}$ and $Ax = Au$, we have $a(x-u) = A(x) - A(u) = 0$ so that $x - u = 0$; if A is injective, we have $A(x) = 0 = A(0)$ forcing $x = 0$).*

(iv) If $dim\, X = n < \infty$ with $\{e_1, \cdots, e_n\}$ as a basis and W is a vector space, an assignment $Ae_i = w_i \in W$ where $w_i \in W$ are arbitrarily chosen, will supply a unique linear transformation *(also denoted by A)* $X \xrightarrow{A} W$ with the formula $A(x) := A(\sum_{i=1}^{n} x^i e_i) = \sum_{i=1}^{n} x^i A(c_i) = \sum_{i=1}^{n} x^i w_i$ for $x = \sum_{i=1}^{n} x^i e_i$.

(i) This sentence includes a claim with its proof. Extend this to an infinite-dimensional space

Hint:*(The same proof works).*

(ii) Use this to show that if $dim\, X = n < \infty$, and <u>rank</u> of $A = \rho(A) := dim\,(A(X))$, nullity of $A = \nu(A) := dim\,(ker\, A)$, we have $\rho(A) + \nu(A) = n = dim\, X$.

Hint:

Show that if $ker\, A \oplus Q = X$ then $Q \xrightarrow{B} Im\, A$ defined by $Bq = Aq$ is an isomorphism; this is obvious since $Im\, B = Im\, A$ and thus A is surjective and $ker\, B = Q \cap ker\, A = \{0\}$ i.e. B is injective. Now if $\{e_1, \cdots, e_q\}$ generate $ker\, A$, this set is linearly independent in X (prove this) and ($\because << X >>= X$ so that any linearly independent set in X can be extended to a basis of X by (d) on page 7) can be extended to a basis $\{e_1, \cdots, e_q, e_{q+1}, \cdots, e_n\}$; clearly $<< e_{q+1}, \cdots, e_n >> \oplus ker\, A = X$. As illustration, take $\mathbb{R}^2 \xrightarrow{A} \mathbb{R}^3$ defined by $A\begin{pmatrix} x^1 \\ x^2 \end{pmatrix} := \begin{pmatrix} x^1 \\ x^1 + x^2 \\ x^3 \end{pmatrix}$ Prove that it is a linear transformation and $Im\, A$ has $\left\{\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}\right\}$ as a basis concluding that rank $A = 2$. Verify that nullity *of A is 0. As another illustration, take* $\mathbb{F}^3 \xrightarrow{A} \mathbb{F}^3$ *given by* $A\begin{pmatrix} x^1 \\ x^2 \\ x^3 \end{pmatrix} = \begin{pmatrix} x^1 - x^2 + 2x^3 \\ 2x^1 + x^2 \\ -x - 2x^2 + 2x^3 \end{pmatrix}$,

observe that $\begin{pmatrix} a \\ b \\ c \end{pmatrix} \in ker\, A$ *iff* $\frac{a}{-2} = \frac{b}{4} = \frac{c}{3}$ *holds and prove that* $\left\{\begin{pmatrix} -2 \\ 4 \\ 3 \end{pmatrix}\right\}$ *is a basis for ker A* *so that A has nullity 1 and rank 2.*

(a) As an exercise, show that $T^2 = 0$ ensures $Im\, T \subseteq ker\, T$ but consider $\mathbb{F}^2 \xrightarrow{T} \mathbb{F}^2$ given by $T\begin{bmatrix} x^1 \\ x^2 \end{bmatrix} = \begin{bmatrix} x^2 \\ 0 \end{bmatrix}$ to conclude that $T^2 = 0$ need not mean $T = 0$.

(b) Moreover, prove that if T^2 and T have the same *rank*, for $X \xrightarrow{T} X$, then $ker\, T = ker\, T^2$ *(they have the same dimension; we consider $dim\, X < \infty$)* and $ker\, T = \{0\}$.

(c) show that $ker\, T = \{0\}$ iff each linearly independent subset of X is carried to a linearly independent subset by T. Also, prove that $ker\, T = \{0\}$ iff $X \xrightarrow{T} W$ is an isomorphism *(Now revisit the task at the end of the remark on page 6 where $\mathbb{F}[\theta] \xrightarrow{T} \tilde{\mathbb{F}}$ given by $T(a(\theta)) = \tilde{a}$)* onto $T(X)$.

(d) For $dim\, X = dim\, Y < \infty$, $X \xrightarrow{A} Y$ being linear, A is surjective $\leftrightarrow$ A is an isomorphism.

Hint: *(rank of $A = dim\, Y$ iff nullity of $A = 0$ iff rank of $A = dim\, X$)*

(e) Given $X_i \xrightarrow{B_i} Y_i \xrightarrow{A_i} Z_i$, $i = i, 2$ one can define $X_1 \oplus X_2 \xrightarrow{B_1 \oplus B_2} Y_1 \oplus Y_2 \xrightarrow{A_1 \oplus A_2} Z_1 \oplus Z_2$ with $(A_1 \oplus A_2).(B_1 \oplus B_2) = A_1B_1 \oplus A_2B_2$ and $(A_1 \oplus A_2) + (B_1 \oplus B_2) = (A_1 + B_1) \oplus (A_2 + B_2)$ *(Take $(A_1 \oplus A_2)(y_1 \oplus Y_2) = A_1(y_1) \oplus A_2(y_2)$ etc. Recall (f) form page 7 saying $P + Q := << P \cup Q >>$. Take everything finite dimensional for comfort).*

1.6 Let X be a vector space and choose an indexed basis $e = \{e_i \mid i \in I\}$. If Y is another vector space and $A(e_i) \in Y$ are assigned i.e. if $e \xrightarrow{A} Y$ is a function, it extends uniquely to a linear transformation $X \xrightarrow{A} Y$ *(again denoted by A)* via the formu8la $A(x) := \sum A(e_i)x^i$ for $x = \sum e_i x^i$. choose $e \xrightarrow{e^i} \mathbb{F}$ by taking

$$e^i(e_j) = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if} i \neq j \end{cases}$$

to produce linear form $X \xrightarrow{e^i} \mathbb{F}$; this acts on $x = \sum_k e_k x^k$ to produce $e^i(x) = x^i$. If $x \neq 0$ so that some $x^i \neq 0$, $e^i \in X'$ is non zero*($\because e^i(x) = x^i$)* and thus for $X \xrightarrow{\Gamma} X''$ defined on page *(3)* via the formula $< \Gamma(x) \mid \varphi >:=< \varphi \mid x >$, we have $< \Gamma(x) \mid e^i >=< e^i \mid x >= x^i \neq 0$ i.e. $\Gamma(x)$ is nonzero. This means that $ker\, \Gamma = \{0\}$. Since Γ was proved to be linear on page 3, we conclude that Γ is injective.

1.6.1 If $dim\, X = n < \infty$, we find $< \varphi \mid x >=< \varphi \mid \sum_{i=1}^{n} e_i x^i >= \sum_{i=1}^{n} < \varphi \mid e_i > x^i = \sum_{i=1}^{n} < \varphi \mid e_i >< e^i \mid x >$. This holds at each $x \in X$ so we have $\varphi = \sum_{i=1}^{n} \varphi_i e^i$ where $\varphi_i =< \varphi \mid e_i >\in \mathbb{F}$.

In other words, each $\varphi \in X'$ can be written uniquely as a linear combination of the forms $e^i \in X'$, $1 \leq i \leq n$. This means $\{e^i \mid 1 \leq i \leq n\}$ is a basis for X' **and we conclude** $X =<< e_1, \cdots, e_n >> \Longrightarrow X' =<< e^1, \cdots, e^n >>$.$\circledast$

But this means $dim\, X = dim\, X' = n$ and then $dim\, X'' = n$. Thus the injective linear $X \xrightarrow{\Gamma} X''$ must be an isomorphism ((d)on page 9). We identify therefore $x \in X$ with $\Gamma(x) \in X''$ which is given by $< \Gamma x \mid \varphi >=< \varphi \mid x >$, in particular, $\Gamma(e_i)$ with e_i. This permits us to write $< \varphi \mid x >$ also as $< x \mid \varphi >$ (i.e. $\varphi(x) \in \mathbb{F}$ as $x(\varphi)$) for $x \in X$,$\varphi \in X'$ and we do it beginning now. The implication $\circledast$ now exhibits an isomorphism $X \cong X'$ via $e_i \to e^i$ which is involutive in the same sense that $X'' = X'$ and *(prove this)* $A'' = A$ for $X \xrightarrow{A} Y$.

Exercise

Prove that for $X \xrightarrow{A} Y \xrightarrow{B} Z$, we have $(BA)' = A'B'$ and that $(id_X)' = id_X$

Hint:

Just use $< A'(\psi) \mid x >: - < \psi \mid A(x) >$, the defining formula for A' given on page (2) in (1.3.2).

Example 0.4. Essentially we know that there is only one finite-dimensional vector space of dimension n; this is $\mathbb{F}^n$ with basis $e_i = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \to i$-th place*(see page (5))*. Then while $(\mathbb{F}^n)'$ is *(isomorphic to)* $\mathbb{F}^n$, we write its elements as row vectors $x = [x_1, \cdots, x_n]$. At present, it may be convenient to take $\varphi \in (\mathbb{F}^n)'$ as $\varphi = [\varphi_1, \cdots, \varphi_n]$, $\varphi_i \in \mathbb{F}$, $\varphi = \sum_{i=1}^{n} \varphi_i e^i$ *(where of course $e^i = [0, \cdots, 1, 0, \cdots]$)*

for notational consistency. Then, for $x = \begin{pmatrix} x^1 \\ . \\ . \\ . \\ x^n \end{pmatrix} \in X \cong \mathbb{F}^n$ and $\varphi = [\varphi_1, \cdots, \varphi_n] \in X' \cong \mathbb{F}^n$, we have $< \varphi \mid x >= [\varphi_1, \cdots, \varphi_n] \begin{pmatrix} x^1 \\ . \\ . \\ . \\ x^n \end{pmatrix} = \sum_{i=1}^{n} \varphi_i x^i \in \mathbb{F}$. *(but don't say 'dot product'; that is still in distant future).*

That is, the action of the row vector φ which is a linear form on $\mathbb{F}^n$ on the column vector x which is a vector in $\mathbb{F}^n$ is the matrix product of φ with x.

Exercise

For $X \xrightarrow{A} Y$, prove that

(a) $ker\, A' = (Im\, A)^0$, (b) $Im\, A' = (ker\, A)^0$ and $rank\, A' = rank\, A'$, $nullity\, A = nullity\, A'$. *(of course, for $dim\, X < \infty$, $dim\, Y < \infty$)*, where we define the annihilator P^0 of a subset $P \subseteq X$ by $P^0 := \{\varphi \in X' \mid < \varphi \mid p >= 0 \text{ at each } p \in P\}$.

(Then prove that P^0 is a subspace of X' regardless of whether P is or is not a subspace of X *(Just note that $< \lambda\varphi + \varsigma \mid p >= \lambda < \varphi \mid p > + < \varsigma \mid p >= 0$ for $\varphi, \varsigma \in X'$).*

Now if $<< X >>=<< e_1, \cdots, c_n >>$ and $W =<< e_1, \cdots, e_r >>$ is a subspace of X we clearly have $W^0 =<< e^{r+1}, \cdots, e^n >>$ ($< \varphi \mid p >=< \sum \varphi_i e^i \mid \sum_{i=1}^{r} e_i p^i >= \sum_{j=1}^{n} \sum_{i=1}^{r} \varphi_j < e^j \mid e_i > p^i = 0$ demands that $\varphi \in << e^{r+1}, \cdots, e^n >>$). Since $<< e^{r+1}, \cdots, e^n >>$, we conclude that $dim\, W + dim\, W^0 = dim\, X$. It should be also clear that $W^{00} = W$.

Then $< A'(\psi) \mid x >=< \psi A x >$ says that $\psi \in ker\, A'$ iff $< \varphi \mid Ax >= 0$ at each $x \in X$ i.e. iff $\varphi \in (Im\, A)^0$. Recalling that $W^{00} = W$, if we use A' for A, we have $ker\, A = ker\, A'' = (Im\, A')^0$ which means $(ker\, A)^0 = (Im\, A')^{00} = Im\, A'$. Since nullity of A' is $dim\,(ker\, A') = dim\,(Im\, A)^0$ and since $dim\,Y' = rank\,(A') + nullity\,A' = dim\,(Im\,A') + dim\,(Im\,A)^0 = dim\,(Im\,A) + dim\,(Im\,A)^0$

we have $rank\,A = rank\,A'$, similarly, $nullity\,A = nullity\,A'$)

1.7 Matrices and linear transformations

We work with a linear transformation $X \xrightarrow{A} Y$ and remind ourselves that for $\varphi(x) \in \mathbb{F}$ is $< \varphi \mid x >$.

Then $\mid Y >< \varphi$ stands for the function $X \to Y$ which acts on $x \in X$ by returning $\mid y >< \varphi \mid x >$ where we write

(i) $\mid x >, \mid y >$ etc. for vectors x, y in X and Y etc. calling them 'ket's and $< \varphi \mid, < \psi \mid$, etc. for forms φ, ψ in X' and Y' etc. calling them 'bra's so that $< \varphi \mid x >$ is a $bar(c)ket \in \mathbb{F}$ and $\mid x >< \varphi \mid$ is a ket-bra which is a linear transformation $X \to X$ returning $\mid x >< \varphi \mid w > \in X$ on input $\mid w > \in X$.

This is Dirac notion and terminology for linear algebra.

1.7.1 If $X \xrightarrow{|y><\varphi|} Y$ is not zero then

(i) $Im(\|y >< \varphi\|) =<< y >>$ ($\because$ $v \in Im(\|y >< \varphi\|)$ iff $v = \|y >< \varphi\|x >$ for some $x \in X$ i.e. a scalar multiple of $\|y >$; note that since we have here $< \varphi\| \neq 0$, we have $0 \neq \varphi_i =< \varphi\|e_i >$. for some basis vector e_i and then any $\lambda \in \mathbb{F}$ is $< \varphi \mid e_i >$ for the choice $x = \frac{e_i\lambda}{\varphi_i}$ since then $< \varphi \mid x >=< \varphi \mid \frac{e_i\lambda}{\varphi_i} >= \frac{<\varphi|e_i>}{\varphi_i}\lambda = \lambda$ so that $y\lambda =\mid y >< \varphi \mid \frac{e_i\lambda}{\varphi_i} >$ for each $y\lambda$ in $<< y >$.).

(ii) $ker\,(\mid y >< \varphi \mid) =<< \varphi >>^0$ *(recall $X'' = X$)* ($x \in ker\,(\mid y >< \varphi \mid)$ iff $\mid y >< \varphi \mid x >= 0$ iff $< \varphi \mid x >= 0$ since $\mid y >\neq 0$. This means, since $0 =< \varphi \mid x >=< x \mid \varphi >$, that $x \in << \varphi >>^0$)

(iii) Bilinearity: $\mid \lambda y + v >< \varphi \mid x >=\mid \lambda y >< \varphi \mid x > + \mid v >< \varphi \mid x >= [\lambda \mid y >< \varphi \mid + \mid v >< \varphi](x)$ at each $x \in X$ so that $\mid \lambda y + v >< \varphi \mid= \lambda \mid y >< \varphi \mid + \mid v >< \varphi \mid$, and $\mid>< \varphi + \varsigma\mu \mid x >=\mid y >< \varphi \mid x > + \mid y >< \varsigma \mid x > \mu = [\mid y >< \varphi \mid +\mu \mid y >< \varsigma \mid](x)$ at each x, so $\mid y >< \varphi + \varsigma\mu \mid=\mid y >< \varphi \mid + \mid>< \varsigma \mid \mu$ for each $y, v \in Y$, $\varphi, \varsigma \in X'$, $\lambda, \mu \in \mathbb{F}$

(iv) For $W \xrightarrow{A} X \xrightarrow{|><\varphi|} Y \xrightarrow{B} Z$ we have $B \circ (\mid y >< \varphi \mid x >=\mid By >< \varphi \mid x >$ at each $x \in X$ so that $B\circ \mid y >< \varphi \mid=\mid B(y) >< \varphi \mid$ $(X \to X)$ and $(\because < \varphi \mid Aw >= ((\mid y >< \varphi \mid) \circ A)w$ at each $w \in X$ so that $(\mid y >< \varphi \mid) \circ A =\mid y >< A'\varphi' \mid$ $(W \to Y)$

(v) Since $< A'\psi \mid x >=< \psi \mid Ax >$ for $X \xrightarrow{A} Y$, we have $< (\mid y >< \varphi \mid)'\psi \mid x >=< \psi \mid |y >< \varphi \mid x >$ at each $x \in X$ which means $(\mid y >< \varphi \mid)'(\psi) =< \psi \mid |y >< \varphi \mid>$

$=\mid \psi y >< \varphi \mid$

$= (\mid \varphi >< y \mid)(\psi) =\mid \varphi >< y \mid \psi >$

(we have $X \xrightarrow{|y><\varphi|} X \xrightarrow{\psi} \mathbb{F}$, then by (iv), $\psi\circ \mid y >< \varphi \mid = \mid \psi(y) >< \varphi \mid$ but $\psi(y) =< \psi \mid y >=< y \mid \psi >$ since $y \in Y = Y''$ and thus $\psi(y)(\varphi) \in X'$ is simply $(< y \mid \psi >)\varphi$ which we are writing as $\mid \varphi >< y \mid \psi >$; note that $< y \mid \psi >=< \psi \mid y >\in \mathbb{F}$). at each $\psi\imath Y'$.

Thus we conclude:

$(|>< \varphi \mid)' = \mid \varphi >< y \mid$ *where* $\mid y >\in Y$ *on LHS has been written as* $< y \mid\in Y''$ *while* $< \varphi \mid\in X'$ *on LHS has been written as* $\mid \varphi >\in X'$ *as a ket since we have*

$Y' \xrightarrow{(|y><\varphi|)'=|><y|} X'$ *just as we have* $Z \xrightarrow{|w><\alpha|} W$ *for* $\mid w >\in Y$, $< \alpha \mid\in Z'$

(vi) For $X \xrightarrow{|y><\varphi|} Y \xrightarrow{|z><\psi|} Z$

$(\mid Z >< \psi \mid) \circ (\mid y >< \varphi \mid) =< \psi \mid y >\mid z >< \varphi \mid$ where $y \in Y$, $\psi \in Y'$, $\varphi \in X'$ and $x \in Z$ *(use* $B \circ (\mid y >< \varphi \mid) = \mid By >< \varphi \mid$ *established in (iv) with* $B = \mid z >< \psi \mid$*, noting that* $< \psi \mid y \in \mathbb{F}$*)*

(vii) For $X =<< e_1, \cdots, e_n >>$ so that $X' =<< e', \cdots, e' >>$ *(see ⊛ on page 10)*, we have $\sum_{i=1}^{n} \mid e_i >< e^i \mid e_j >= \mid e_j >$ which means $id_X = \sum_{i=1}^{n} \mid e_i >< e^i \mid$ (because $\sum_{i=1}^{n} \mid e_i >< e^i \mid$ is seem to be the function $e \to X$ given by $e_j \to e_j$) and $\sum_{i=1}^{n} \mid e_i >< e_i \mid e^j >= \mid e^j >$

$(\because < e_i \mid e^j >=< e^j \mid e_i >=$

$$\delta_{ij} == \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

)

so that $id'_X = \sum_{i=1}^{n} \mid e^i >< e_i \mid$.

(viii) Suppose $X =<< e_1, \cdots, e_n >>$, $Y =<< d_1, \cdots, d_m >>$. Given $X \xrightarrow{A} Y$, we have $A(e_i) \in Y$ and can write $A(e_i) = \sum_{j=1}^{m} \mid d_j > a_i^j$ for scalar $a_i^j \in \mathbb{F}$ uniquely; $a_i^j =< d^j \mid Ae_i >$. Then $A = A \circ id_X = A \circ (\sum_{i=1}^{n} \mid e_i >< e^i \mid$ *(see (vii) above)* $= \sum_{i=1}^{n} \mid Ae_i >< e^i \mid = \sum_{i=1}^{n}\sum_{j=1}^{m} \mid d_j > a_i^j < e^i \mid = \sum_{i=1}^{n}\sum_{j=1}^{m} a_i^j \mid d_j >< e^i \mid$ which means that the $m.n$ linear transformation $X \xrightarrow{|d_j><e^i|} Y$ which act on $x \in X$ to return the vector $\mid d_j >< e^i \mid x >= \mid d_j > X^i$ form a basis for the space $L(X, Y)$ since we just saw that any $X \xrightarrow{A} Y$ can be written uniquely as a linear combination $A = \sum_{i=1}^{n}\sum_{j=1}^{n} a_i^j \mid d_j >< e^i \mid$ in terms of these $\mid d_j >< e^i \mid$. The $m.n$ scalars a_j^j are written as an $m \times n$ matrix $a = [a_i^j]$ with i indexing the columns and j indexing the rows and this is called the matrix associated with the linear transformation $X \xrightarrow{A} Y$ with reference to the given indexed bases $e = \{e_i, \cdots, e_n\}$ of X and $d = \{d_1 \cdots, d_m\}$ of Y.

Just how do we write this matrix? To begin, recall that when we say $y \in Y$ is uniquely written as $\sum_{j=1}^{m} d_j y^j$, the scalars y^j are $< d^j \mid y > \in \mathbb{F}$ where $\{d^1, \cdots, d^m\}$ is the basis of Y' associated with the basis $d = \{d_1, \cdots, d_m\}$ defined by $< d^j \mid d_k > :=$

$$\delta_{jk} = \begin{cases} 1 & \text{if } j = k \\ \\ 0 & \text{if } j \neq k \end{cases}$$

. Here $A(e_i) \in Y$ and so there scalars $(A(e_i))^j$ are $< d^j \mid Ae_i >$ which we have written as a_i^j; they depend on both $\{e_1, \cdots, e_n\}$ and $\{d_1, \cdots, d_m\}$ *(and of course on $\{e^1, \cdots, e^n\}$ and $\{d^1, \cdots, d^n\}$).*

The matrix $a = [a_i^j]$ is obtained by writing the m-deep columns $A(e_i)$ as the i-the column and thus

$$a = \begin{pmatrix} a_1^1 & a_2^1 & . \; . \; . & . \; . \; . & a_n^1 \\ a_1^2 & a_2^2 & . \; . \; . & . \; . \; . & a_n^2 \\ \\ . & . & . & . & \\ . & . & . & . & \\ . & . & . & . & \\ \\ a_1^m & a_2^m & . \; . \; . & . \; . \; . & a_n^m \end{pmatrix}$$

with a_i^j at the intersection of the j-th row and the i-the column.

(ix) The discussion in *(viii)* will be summarized by writing $(X, \underline{\underline{e}}) \xrightarrow[A]{a=[a_i^j]} (Y, \underline{\underline{\alpha}})$ and we keep writing $a_i^j = < d^j \mid Ae_i >$ so that we never forget that the matrix representation is in terms of the indexed bases $\underline{\underline{e}}$ and $\underline{\underline{d}}$ *(and of course their 'dual' or 'reciprocal' bases $\{e^1, \cdots, e^n\}$, $\{d^1, \cdots, d^m\}$)*; this is important because if the indexing is changed, the location of a_i^j is charged in the matrix and if either of the bases id changed, the entries will be different.

If $(X, \underline{\underline{e}} \xrightarrow[A]{a=[a_i^j]} (Y, \underline{\underline{d}}) \xrightarrow[B]{b=[b_j^k]} (Z, \underline{\underline{c}})$, $1 \leq i \leq n$, $1 \leq j \leq m$, $1 \leq k \leq p$, we have $(b \circ a)_i^k = < c^k \mid (B \circ A)e_i > = < C^k \mid B(\sum_{j=1}^{m} a_i^j \mid d_j >) > = \sum_{j=1}^{m} < C^k B d_j > a_i^j = \sum_{j=1}^{m} b_j^k a_i^j$ so that the composition of the linear transformation $B \circ A$ corresponds exactly to the composition of their matrix representation in the given bases.

Further, since $Y' = << d^1, \cdots, d^m >>$ and $X' = << e^1, \cdots, e^n >>$, we have the matrix representation of A' given by $Y' \xrightarrow[a']{A'} X'$ where $(a')_j^i = < e_i \mid A'(d^j) > = < A(e_i) \mid d^j > = < d^j \mid A(e_i) > = a_i^j$ *(note that we use $Y'' = Y$, $X'' = X$, and the fact that $< \varphi \mid x > = < x \mid \varphi >$ for kets $\mid x >$ and bras $< \varphi \mid$)*

Also, $(X,\underline{\underline{e}}) \xrightarrow[I]{id_X} (X,\underline{\underline{e}})$ has $I_i^e =< e^l \mid id_X(e_i) >=< e^l \mid e_i >$

$$= \begin{cases} 1 & \text{if } l = i \\ 0 & \text{if } l \neq i \end{cases}$$

, $1 \leq i \leq n$, $1 \leq l \leq n$; thus the identity operator with respect to the same basis has the representation given by the identity *(square)* matrix which has on-diagonal entries 1 and off-diagonal entries 0.

We shall now supply some examples to illustrate the discussion. We use the standard terminology: Choice of a basis $\underline{\underline{e}}$ for X is called a coordinatization of X, in the representation $x = \sum e_i x^i = \sum \mid e_i > x^i$. the scalars x^i are called the coordinates or the components of the vector x with respect to the given coordinatization, and a linear transformation $X \xrightarrow{A} X$ will be frequently called a 'linear operator on X'.

1.8 Some illustrations.

1.8.1 Many authors do not wish to emphasize the need for indexing a basis, at least in the beginning. There is a reason for it. Consider the basis $\left\{ e_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, e_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$ for $\mathbb{F}^2$; the indexing would mean that $\{e_2, e_1\}$ is not a basis ($\{e_1, e_2\}$ but a different basis *(for this reason, ordered basis is usually the term for what we are calling indexed basis)*. To describe the action of a linear operator $\mathbb{F}^2 \xrightarrow{A} \mathbb{F}^2$ on basis vectors, one could say either (i) $Ae_1 = \alpha e_1$, $Ae_2 = \beta e_1 + \gamma e_2$ *(in which ordering the basis is irrelevant)*, or (ii) A has matrix $\begin{bmatrix} \alpha & \beta \\ 0 & \gamma \end{bmatrix}$ with respect to the ordered basis $\{e_1, e_2\}$ *(in which case ordering the basis is crucial, since one could easily but of course mistakenly, take this to mean $Ae_1 = \alpha e_1 + \beta e_2$, $Ae_2 = \gamma e_2$)*. The best thin to do seems to be: keep writing "$a_i^j =< d^j \mid Ae_i >, 1 \leq j \leq m \; 1 \leq i \leq n$ and $a = [a_i^j]$ is an $m \times n$ matrix for $(X,\underline{\underline{e}}) \xrightarrow{A} (Y,\underline{\underline{d}})$: $Ae_i = \sum \mid d_j > a_i^j$". Linear algebra is better understood without matrices but the subject is so computational in application that de-emphasizing matrices is almost scandalous.

1.8.2 (i) Find the matrix of $\mathbb{F}^3 \xrightarrow{A} \mathbb{F}^2$ given by $A \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2x + 3y - z \\ 4x - y + 2z \end{pmatrix}$ relative to the basis

$$\left\{\underbrace{\begin{pmatrix}1\\1\\0\end{pmatrix}}_{=b_1},\underbrace{\begin{pmatrix}1\\2\\3\end{pmatrix}}_{=b_2},\underbrace{\begin{pmatrix}1\\3\\5\end{pmatrix}}_{=b_3}\right\}\text{ of }\mathbb{F}^3\text{ and }\left\{\underbrace{\begin{pmatrix}1\\2\end{pmatrix}}_{=c_1},\underbrace{\begin{pmatrix}2\\3\end{pmatrix}}_{=c_2}\right\}\text{ of }\mathbb{F}^2.$$

Hint:

That the two sets given are indeed bases can be verified by checking them for linear independence. For instance, we know that $\mathbb{F}^3$ *has three elements in any basis and the system* $\lambda_1 b_1 + \lambda_2 b_2 + \lambda_3 = 0$ *has no nontrivial solutions(verify) so that* $\{b_1, b_2, b_3\}$ *are linearly independent and must form a basis of* $\mathbb{F}^3$*. Next, if* $\begin{pmatrix}a\\b\end{pmatrix} \in \mathbb{R}^2$ *, solving* $x\begin{pmatrix}1\\2\end{pmatrix} + y\begin{pmatrix}2\\3\end{pmatrix} = \begin{pmatrix}a\\b\end{pmatrix}$ *supplies* $x = -3a + 2b$, $y = 2a - b$ *so that in the given basis* $\{c_1, c_2\}$ *of* $\mathbb{F}^2$*, we must have this vector as* $\begin{pmatrix}-3a+2b\\2a=b\end{pmatrix}$ *and directly calculating therefore, we have* $Ab_1 = \begin{pmatrix}5\\3\end{pmatrix} = -9c_1 + 7c_2$ $Ab_2 = c_1 + 2c_2$, $Ab_3 = 4c_1 + c_2$*. The desired matrix will be obtained by these column vectors* Ab_1, Ab_2, Ab_3 *so that it is* $a = \begin{pmatrix}-9 & 1 & 4\\7 & 2 & 1\end{pmatrix}$*. The vector* $\begin{pmatrix}2\\5\\-3\end{pmatrix}$ *of* $\mathbb{F}^3$ *is calculated to be* $\begin{pmatrix}11\\-21\\12\end{pmatrix}$ *in the basis* $\{b_1, b_2, b_3\}$ *(verify that any* $\begin{pmatrix}a\\b\\c\end{pmatrix} \in \mathbb{F}^3$ *in the basis* $\{b_1, b_2, b_3\}$ *will be* $\begin{pmatrix}-a+2b-c\\5a-5b+2c\\-3a+3b-3c\end{pmatrix}$*) and calculating* $A\begin{pmatrix}2\\5\\-3\end{pmatrix} = \begin{pmatrix}4+15+3\\8-5-6\end{pmatrix} = \begin{pmatrix}22\\-3\end{pmatrix}$ *by the supplied formula, writing this as* $\begin{pmatrix}-66-6\\44+3\end{pmatrix}$ *is the basis* $\{c_1, c_2\}$*, we get* $\begin{pmatrix}-72\\47\end{pmatrix}$ *which is precisely* $\begin{pmatrix}-9 & 1 & 4\\7 & 2 & 1\end{pmatrix}\begin{pmatrix}11\\-21\\12\end{pmatrix}$ *as it should be and displays the action of* A *as matrix action.*

(ii) For $\mathbb{F}^3 \xrightarrow{A} \mathbb{F}^2$ given by $A\begin{pmatrix}x\\y\\z\end{pmatrix} = \begin{pmatrix}3x+2y-4z\\x-5y+3z\end{pmatrix}$ with

$B_1 = \left\{ \underbrace{\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}}_{=u_1}, \underbrace{\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}}_{=u_2}, \underbrace{\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}}_{=u_3} \right\}$ and $B_2 = \left\{ \underbrace{\begin{pmatrix} 1 \\ 3 \end{pmatrix}}_{=v_1}, \underbrace{\begin{pmatrix} 2 \\ 5 \end{pmatrix}}_{=v_2} \right\}$ as bases, the representation of A is $\begin{pmatrix} -7 & -33 & -13 \\ 4 & 19 & 8 \end{pmatrix}$. Verify this and the action displayed on $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ to be $\begin{pmatrix} -13x - 20y + 26z \\ 8x + 11y - 15z \end{pmatrix}$ in the coordinatizon B_2 for $A \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ as given.

1.8.3 Prove that if $X \xrightarrow{A} Y$ is a linear transformation, there exists a basis $\underline{\underline{e}}$ for X and a basis $\underline{\underline{d}}$ for Y for which A has the matrix representation. $a = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$ where I_r is the identity matrix of order r and r is the rank of A; take $dim\, X = n < \infty$, $dim\, Y = m < \infty$.

Hint:

we know that $r = dim\,(A(\alpha))$ hence $dim(ker\, A) = n - r$ ((ii) on page 8). Let $\{e_{r+1}, \cdots, e_n\}$ be a basis for $ker\, A$ then as a linearly independent subset of X, this can be extended to a basis $\underline{\underline{e}} = \{e_1, \cdots, e_r, e_{r+1}, \cdots, e_n\}$ of X. Write $d_j = A(e_j)$ for $1 \le j \le r$. Since $X \xrightarrow{A} A(X)$ is surjective, it must be an isomorphism $X \xrightarrow{A} A(X)$ ((d) on page 9) and thus these r vectors $d_j \in A(X)$ will form a basis for the r-dimensional space $A(X)$; then they are linearly independent in Y and we can extend it to a basis $\underline{\underline{d}} = \{d_1, \cdots, d_m\}$ of Y (read the HINT for (ii) on page 8 if you are uncertain about the process of getting $\underline{\underline{e}}$ and $\underline{\underline{d}}$). Then

$d_1 = A(e_1) = 0.d_1 + 0.d_2 + \cdots + 0.d_m$

$d_2 = A(e_2) = 0.d_1 + 1.d_2 + \cdots + 0.d_m$

$\cdots \quad \cdots \quad \cdots$

$d_r = A(e_r) = 0.d_1 + \cdots + 1.d_r + 0.d_{r+1} \cdots + 0.d_m$

$0 = A(e_{r+1}) = 0.d_1 + \cdots + 0.d_m$ ($\because e_{r+1} \in ker\, A$)

$\cdots \quad \cdots \quad \cdots$

$0 = a(en) = 0.d_1 + \cdots + 0.d_m$ ($\because e_n \in ker\, A$)

and therefore $(X, \underline{\underline{e}}) \xrightarrow[a]{A} (Y, \underline{\underline{d}})$ *must be* $a = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$ *as advertized.*

1.8.4 What is $dim\,(L(X,Y))$ if $dim\,X = n < \infty$, $dim\,Y = m < \infty$?

Hint:

Fix a basis $\underline{\underline{e}} = \{e_1, \cdots, e_n\}$ *for* X *and a basis* $\underline{\underline{d}} = \{d_1, \cdots, d_m\}$ *for* Y. *Then the corresponding* $(X, \underline{\underline{e}}) \xrightarrow[a]{A} (Y, \underline{\underline{d}})$, $A \leftrightarrow a$, *is bijective. Since the vector space of all* $m \times n$*matrices(over* $\mathbb{F}$*) is of dimension* $m.n$ *(with basis* $\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \end{pmatrix}$, $\begin{pmatrix} 0 & 1 & 0 & \cdots \\ 0 & . & . & \cdots \end{pmatrix}$, $\begin{pmatrix} 0 & \cdots & 0 \\ 1 & \cdots & . \\ 0 & . & \cdots \end{pmatrix}$ *etc; prove this),* $dim\,(L(X,Y)) = m.n$.

$\|$ *But it is far better to simply appeal to the fact that* $\{\mid d_j >< e_i \mid |1 \leq i \leq n\, 1 \leq j \leq m\}$ *forms a basis for* $L(X,Y)$ *(see page 15)*

1.8.5 If $X \times X \xrightarrow{B} \mathbb{F}$ is such that $B(\lambda x + u, w) = \lambda B(x.w) + B(u,w)$ and $B(x, u + w\lambda) = B(x,u) + B(x,w)\lambda$, we say it is a bilinear form on X *(compare 1.3 page 2)*. Show that if B and C are bilinear forms on X, $(B + C)(x,u) := B(x,u) + C(x,u)$ makes the collection of all bilinear forms on X a vector space*(with the obvious scalar multiplication)*. Call this $\underline{\underline{\text{Bilim}(X)}}$.

(i) If $\underline{\underline{e'}} = \{e'_1, \cdots, e'_n\}$ is a basis of X' show that $\beta^j_i(x,u) := \beta'_j(x)\beta'_i(u)$ forms a basis of $\underline{\underline{Bilim\,(X)}}$.

Hint:

If $B \in \underline{\underline{Bilim\,(X)}}$ *and* $b^j_i := B(e_j, e_i) \in \mathbb{F}$ *where* $\underline{\underline{e}} = \{e_1, \cdots, e_n\}$ *is the corresponding basis for* $X = X''$, *we get* $(\sum_{i,j} b^j_i \beta^j_i)(e_s, e_t) = \sum_{i,j} b^j_i \beta^j_i(e_s, e_t) = \sum b^i_i e'_j(e_s)e'_i(e_t) = \sum b^j_i \delta_{js}\delta_{it} = b^s_t = B(e_s, e_t)$; *since* $\{(e_s, e_t) \mid 1 \leq s \leq n,\, 1 \leq t \leq n\}$ *forms a basis of* $X \times X$, *we get all of* B *from this; thus the* β^j_i *span* $\underline{\underline{Bilim\,X}}$*(verify of course that* β^j_i *are bilinear). If* $B = \sum b^j_i \beta^j_i = 0$, *we have* $B(e_s, e_t) = b^t_s = 0$ *for each* s, t *which means this set is linearly independent and is thus a basis for* $\underline{\underline{Bilin\,(X)}}$. *In particular dim* $(\underline{\underline{Bilin\,(X)}} = (dim\,X)^2$.

(ii) Entering $b^j_i = B(e_j, e_j)$ into n^2 matrix at the intersection of the j-th row and i-th column will get us a matrix b which is called the matrix of the bilinear form B with respect to the basis $\underline{\underline{e}}$ of X. For example, the bilinear form on $\mathbb{F}^2$ given by $B\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right) := 2x_1y_1 - 3x_1y_2 + x_2y_2$, with respect to the basis $\left\{ e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}$ has the matrix $\begin{pmatrix} 2 & -1 \\ 2 & 0 \end{pmatrix}$ *(Verify this).*

(iii) Given a Bilinear form $X \times X \xrightarrow{B} \mathbb{F}$ on X, the function $X \xrightarrow{Q} \mathbb{F}$ supplied by $Q(x) := B(x, x)$ is called the quadratic form associated to B.

Show that the correspondence established by *(ii)* above, $B \leftrightarrow b$ is bijective in the sense that $B(x, u) = x^t bu$ *(x^t is the transpose of the column vector x)* and that any n^2-matrix b will raise a bilinear form $B(x, u) := x^t bu$. Further, show that $Q \leftrightarrow q$ establishes a similar correspondence between quadratic forms Q and symmetric n^2-matrices q via $Q(x, x) := x^t qx$.

(iv) Show that if B raises the quadratic form Q, we can get the bilinear form B from

$2B(u, v) = B(u, u) + B(u, v) + B(v, u) + B(v, v) - B(u, u) - B(v, v)$

$= B(u + v, u + v) - B(u, u) - f(v, v)$

$= Q(u + v) - Q(u) - Q(v)$

that is, $B(u, v)$ can be defined from the quadratic form Q via $B(u, v) := \frac{1}{2}[Q(u+v) - Q(u) - Q(v)]$ (provided of course that $2 \neq 0$ in $\mathbb{F}$ i.e. $\mathbb{F} \neq \mathbb{Z}_2$)

Note:

Frequently, to save space, $\begin{pmatrix} x^1 \\ x^2 \end{pmatrix} \in \mathbb{F}^n$ will be written $[x^1. \cdots x^n]^t$ or $(x^1, \cdots, x^n)$.

I am faithfully following linear algebra and Group Representation *(Volume I)* by Ronald Shaw *(Academic Press 1982).* This handout covers the selection from the first *(25)* pages from the first chapter. There is a copy in our central library but it is perhaps easier to work through the handout compared to the some what terse presentation of the book.