Here is a summary of concepts involved with vector spaces.

For our purposes a **field**, k, is a subset of the complex numbers that is closed under addition, multiplication, and additive and multiplicative inverses. Examples are the rationals, \mathbf{Q} , the reals, \Re , and the complex numbers, \mathbf{C} . In general one can assume one is working with $k = \Re$ although we will need \mathbf{C} at some point.

- 1. Vector Space. A vector space over a field k is a non empty set V together with a well defined addition '+' and a scalar multiplication '.' by elements of the field k, satisfying the following axioms.
 - **Operations.** For all $u, v \in V$ and $c \in k$ we have $u + v \in V$ and $cu \in V$. Sometimes, we write $c \cdot u$ but this \cdot is often omitted.
 - Addition properties. The addition is commutative (u + v = v + u) and associative (u + (v + w) = (u + v) + w).
 - Additive identity and inverse. There is a zero vector, denoted 0 such that u + 0 = 0 + u = u for all u. Moreover, for each u there is a u^* such that $u + u^* = u^* + u = 0$. The u^* can be shown to be uniquely defined by u and is denoted as -u. The zero vector is also shown to be unique!
 - Distributivity and unitariness. The two operations interact naturally:

$$c(u+v) = cu + cv$$
, $(c+d)u = cu + du$, $(cd)u = c(du)$.

Moreover, 1(u) = u for all u.

As noted above this course, the field k is usually \Re the field of real numbers. In that case, we drop the phrase "over the field \Re ".

- 2. Subspace If V is a vector space over k and W is a non empty subset, then we say that W is a subspace of V if we have:
 - For all $w_1, w_2 \in W$ we have $w_1 + w_2 \in W$.
 - For all $c \in k$ and $w \in W$ we have $cw \in W$.

We note that the vector 0 will always belong to a subspace as soon as it is non empty, since $w \in W$ implies $0w = 0 \in W$ by the second subspace condition above. Hence, you may replace the condition of W being non empty by the simpler condition $0 \in W$, as done in the book. 3. (optional) A challenging example. Here is an exotic example of a vector space which should be studied to verify your understanding of the above definition.

Let $V = \Re$ be made into a vector space over the usual field of real numbers \Re as follows:

• We define a new addition \oplus on V by the formula:

$$v \oplus w = v + w - 1$$

where the operations on the right hand side are the usual operations in real numbers.

• We define a new scalar multiplication \odot by \Re on V by the formula:

$$c \odot v = cv + 1 - c$$

where, as before, the operations on the right are the usual operations in real numbers.

It is instructive to verify all the axioms from these definitions. You should also identify what -v means. This example should be kept in mind while analyzing all the following concepts.

You should also make an enhanced version of this example by taking V to be \Re^n as the set, but using the above definitions of addition and scalar multiplication, suitably generalized.¹

4. A Universal example. Let S be any non empty set and consider

 $F_S^k = \{f: S \to k \mid \text{ where } f(s) = 0 \text{ for all except finitely many } s \in S. \}$

It can be shown that every vector space can be described in this manner, but finding such an explicit S can be tedious and it is better to use the basic definition. If $k = \Re$ we may drop it from the notation.

It is easy to verify how F_S^k is a vector space by defining (f+g)(s) = f(s) + g(s)and cf(s) = cf(s). The extra condition on f is not necessary, but it is essential if you want to claim that every vector space has a standard structure!

5. Basic examples. Here are some of the standard examples.

 $w_1 \oplus w_2 = \psi^{-1}(\psi(w_1) + \psi(w_2))$ and $c \odot w = \psi^{-1}(c\psi(w))$.

¹Can you make other examples of such weird operations? Here is a general hint and secret of all such constructions.

Let V be any vector space over a field k. Let ψ be any bijective (ii.e. injective and surjective) function from V to itself.

Define a new vector space W which uses the same V as an underlying set but defines operations as follows.

It can be shown that W is a vector space "isomorphic" to V which means essentially the same as V. See below for explanation of "isomorphic".

Can you guess the ψ for the example given above? Hint: try $x \to x + \alpha$ for some α .

• Euclidean spaces. The space k^n consisting of all *n*-tuples of elements of k, usually written as a column.

This can be described as F_S^k where S is the set $\{1, 2, 3, \dots, n\}$. A typical

function f in the vector space may be displayed as $\begin{pmatrix} f(1) \\ f(2) \\ \cdots \\ f(n) \end{pmatrix}$. This leads to

the usual notation for k^n . For our purposes k will almost always be \Re in which case we are talking about the familiar \Re^n .

• The case of an infinite S. If we take $S = \{1, 2, \dots, n, \dots\}$, the set of natural numbers, then we find it convenient to display $f \in F_S$ as

$$f(1) + f(2)x + f(3)x^2 + \dots + f(n+1)x^n + \dots$$

Note that the description of F_S implies that after some large enough exponent N, the coefficients are all zero and we have a set of polynomials.

The book denotes F_S^{\Re} by the symbol \mathbb{P} . A general notation for F_S^k is also k[x] which is the ring of polynomials in x with coefficients in k, where, we have chosen to ignore the usual multiplication of polynomials!

We now note that if we keep the same set $S = \{1, 2, \dots, n, \dots\}$ but drop the special condition on functions we get a much bigger set, namely

$$H = \{f : S \to k\}.$$

As before, any such $f \in H$ be displayed as $f(1) + f(2)x + f(3)x^2 + \dots + f(n + 1)x^n + \dots$

Since there is no special condition on the function, we now get power series! The general notation for this set is k[[x]], the ring of power series in x with coefficients in k, where, as before, we ignore the usual product of power series. It can be shown that $H = F_T^k$ for some convenient set T, but finding such a Tis a daunting task!

There is also a well known subset of H when $k = \Re$, namely the set of convergent power series. To write it as F_T^{\Re} is an even more daunting task!

- spaces of matrices: $Mat_{m,n}$, the set of m by n matrices with matrix addition and scalar multiplication.
- function spaces: If X is any set then F^ℜ_X = {f : X → ℜ} is a vector space with the usual addition of functions and multiplication of functions by a scalar.
 (i.e. (f + g)(x) = f(x) + g(x), (α f)(x) = α f(x))

- direct sums: If U and V are vector spaces and $U \oplus V = \{(u, v) \mid v \in U, v \in V\}$ then $U \oplus V$ is a vector space with addition $(u_1, v_1) + (u_2, v_2) = (u_1 + v_1, u_2 + v_2)$ and scalar multiplication $\lambda(u, v) = (\lambda u, \lambda v)$. This allows us to define vector spaces tailored to specific problems and to view more complex spaces as constructed of simpler ones. For example we can, properly interpreted, view \Re^7 as $\Re^2 \oplus \Re^5$.
- 6. Important Remark: A subspace of a vector space V is simply a subset of V that is closed under linear combinations. Except for the following exotic example, all of the vector spaces we will encounter in this course either fall in the above list of examples or are subspaces of members of this list. This means that once a prospective vector space has been identified as a subset of a known vector space (e.g. the above examples) then all that is required to verify that it is a vector space is to check that it is closed under linear combinations.

Example: Let $C^1(\Re)$ be the set of all differentiable functions $f : \Re \to \Re$. This is a vector space since:

- it is a subset of the vector space F_{\Re}^{\Re}
- If $f, g \in C^1(\Re)$ and $\alpha, \beta \in \Re$ then $\frac{d}{dx}(\alpha f + \beta g) = \alpha \frac{d}{dx}f + \beta \frac{d}{dx}g$ so $(\alpha f + \beta g) \in C^1(\Re)$
- 7. Basic structures in a vector space. Now let V be a k-vector space (i.e. vector space over a field k).

Foe any subset $A \subset V$, we define its **span**:

Span $A = \{c_1v_1, \dots + c_mv_m \mid \text{ where } c_i \in k, v_i \in A \text{ and } m \text{ is some non negative integer.} \}.$

Note that Span A can be described as the set of all possible **linear combinations** of elements of A. Note that even when A is infinite, we only allow finitely many elements of it at a time! Also note that m is allowed to be zero and it gives the combination 0, by a standard convention.

We say that a set A spans V or is a spanning set for V if Span A = V.

A subset $A \subset V$ is said to be **linearly dependent** if there are elements $v_1, \dots, v_m \in A$ such that $c_1v_1 + \dots + c_mv_m = 0$ for some $c_1, \dots, c_m \in k$ with at least one non zero element c_i among them.

In application, other convenient forms of this condition are used. One such version is:

A subset $A \subset V$ is said to be linearly dependent if there is some $v \in A$ such that v is a linear combination of some elements $w_1, \dots, w_r \in A$ which are distinct from v. A compact way of saying this is to write that $v \in Span A \setminus \{v\}$.

A set $A \subset V$ is said to be **linearly independent**, if it is not linearly dependent. We often drop the word "linearly" from these terms.

A subset $A \subset V$ is said to be **a basis** of V if

Span A = V and A is linearly independent.

Convention: If V is any vector space then $\{O\}$, the set consisting only of O, the zero vector of V is a subspace of V. It is logically true that the empty set, Φ is independent and by convention it is a basis for $\{O\}$.

We say that V is **finite dimensional** if it has a finite basis.

Important Observation: If V is a vector space and B_1 and B_2 are bases for V with $B_1 \subset B_2$ then $B_1 = B_2$.

proof: If $B_1 \neq B_2$ then there would have to be $b \in B_2$ that is not in B_1 . However B_1 is a spanning set for V which means that b_2 must be a linear combination of elements in B_1 . But $B_1 \subset B_2$ so b_2 is a linear combination of the other elements of B_2 which says that B_2 is dependent which is a contradiction.

Definition: The number of elements in a basis for V, is said to be **the dimension** of V. We write dim V or dim_k V if we wish to identify k.

We will soon argue that the dimension is a well defined number for any vector space, i.e. every basis of a vector space has the same number of elements.

Examples:

- The standard basis $\{e_1, e_2, \dots, e_n\}$ is linearly independent and a spanning set for \Re^n , thus the dimension of \Re^n is n.
- If V is a vector space, recall that by convention the empty set Φ is a basis for the zero subspace {O}. Thus the zero subspace has dimension 0 since Φ has 0 elements.

Infinite Dimensional Spaces A vector space that is not finite dimensional (i.e. does not have a finite basis is said to be infinite dimensional. It is true that even infinite dimensional spaces do have bases (linearly independent spanning sets). However we don't say that the dimension is the "number of elements in a basis". For infinite dimensional spaces we need to use the finer notion of "number of "

called "cardinality²" which distinguishes between different infinite sets. Properly dealing with "cardinality" would take too much time so, for our purposes, we won't try to define the dimension of a vector space that is not finite dimensional but will rather simply say that all such spaces are infinite dimensional.

- 8. Homomorphisms, Isomorphisms and Automorphisms. Given k-vector spaces V, W a map $T : V \to W$ is said to be a linear transformation if it satisfies these two conditions:
 - $T(v_1 + v_2) = T(v_1) + T(v_2)$ for all $v_1, v_2 \in V$.
 - T(cv) = cT(v) for all $c \in k$ and $v \in V$.

We may also use the term "homomorphism" (meaning similar structure) to denote such a map.

There are two concepts associated with the notion of a linear transformation (homomorphism).

First is "the image of T" which can be formally denoted as

$$T(V) = \{T(v) \mid v \in V\}.$$

Second is "the Kernel of T" which can be defined as:

$$Ker \ T = \{ v \in V \mid T(v) = 0 \}.$$

It is easy to verify that both the Kernel and the Image are respectively subspaces of V and W.

The homomorphism T is **injective (or "one to one ")** iff Ker T = 0 where we have used a slightly abused notation 0 in place of $\{0\}$. This abuse is routinely done!

The homomorphism T surjective (or "onto") if T(V) = W i.e. W is its total image.

The homomorphism T is said to be an **isomorphism** if it is both injective and surjective i.e. bijective. The word "iso" denotes sameness and the concept says that the two vector spaces with an isomorphism mapping one to the other are essentially the same. They, can be treated as replacements for each other in analyzing their properties.

An isomorphism of V to itself is called an automorphism.

²Two sets A, B are said to have the same cardinality if there is a bijective map from A to B. As the example of the sets $\{1, 2, 3, \dots, n \dots\}$ and $\{2, 4, 6, \dots, 2n, \dots\}$ shows, a set can have the same cardinality as a proper subset. This suggests that one has to be very careful in dealing with infinite cardinality.

9. A Fundamental Theorem (existence of bases). This is the fundamental theorem in the theory of vector spaces.

Let V be a vector space with a spanning set A. Then there is a subset B of A such that B is a basis of V. That is B is also a spanning set for V and B is independent.

We shall give a complete proof in case A is finite and a partial proof in case A is infinite.

Proof.

If A is independent, then it is a basis and we are done. If A is dependent, then there is some vector $v \in A$ which is a linear combination of vectors in $A_1 = A \setminus \{v\}$.

We now claim that Span $A = Span A_1$ by the following argument.

Proof of claim. Note that any $w \in Span A$ can be written as: $w = cv + w_1$ where $c \in k$ and $w_1 \in Span A_1$.

By assumption, $v \in Span A_1$, so cv and hence w belongs to $Span A_1$. Thus $Span A \subset Span A_1$. Clearly $Span A_1 \subset Span A$ since $A_1 \subset A$.

This shows $Span A = Span A_1$.

Thus, if V = Span A and A is dependent, then we get a proper subset A_1 of A such that $V = Span A_1$. We can now apply the same argument to A_1 and either get a basis or a smaller spanning set A_2 .

In case A is a finite set to begin with, this cannot continue indefinitely and we must get a basis at some stage.

We remark that we may run into an empty subset of A, in case the vector space is the zero space $\{0\}$. However, in this case the whole set A can only be $\{0\}$ or empty and we have nothing to argue!

It is also possible to do the above proof in a reversed process. We can start with independent subsets of A and enlarge them as much as possible. We argue that eventually we should get a basis.

In case the set A is infinite requires a more general inductive principle called Zorn's Lemma. This lemma allows us (we omit the details here) to prove that there are maximal independent subsets in any vector space. Such sets are easily seen to be spanning sets.

Theorem A: Suppose V is a vector space and $F = \{F_1, F_2, \dots, F_s\}$ is a finite subset of V. Suppose further that the set of vectors $B = \{B_1, B_2, \dots, B_s, B_{s+1}\}$ is contained in the linear span of F. Then B is linearly dependent set.

Proof: For each $i = 1 \cdots s + 1$ there is a linear combination

$$a_{i,1}F1 + a_{i,2}F_2 + \dots + a_{i,s}F_s = B_i$$

Let A be the $s + 1 \times s$ matrix

$$\begin{array}{c} \begin{array}{c} a_{1,1} & a_{1,2} & \cdots & a_{1,s} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,s} \\ \vdots & \vdots & \vdots & \vdots \\ a_{i,1} & a_{i,2} & \cdots & a_{i,s} \\ \vdots & \vdots & \vdots & \vdots \\ a_{s,1} & a_{s,2} & \cdots & a_{s,s} \\ a_{s+1,1} & a_{s+1,2} & \cdots & a_{s+1,s} \end{array} \end{array}$$
Then letting $F = \begin{pmatrix} F_1 \\ F_2 \\ \vdots \\ F_i \\ \vdots \\ F_s \end{pmatrix}$ and $B = \begin{pmatrix} B_1 \\ B_2 \\ \vdots \\ B_i \\ \vdots \\ B_s \\ B_{s+1} \end{pmatrix}$ we write $\begin{array}{c} B_1 \\ B_2 \\ \vdots \\ B_i \\ B_s \\ B_{s+1} \end{array}$
 $AF = B$

Now we augment A by the $s + 1 \times s + 1$ identity matrix to get $\langle A|I \rangle$ which we then row reduce to get $\langle R|Q \rangle$ where R is an REF of A and Q is an invertible matrix such that QA = R.

Since R has more rows than pivots its last row must be all zeros and therefore the last row of Q is a row of the consistency matrix which cannot be all zero since it is a row of an invertible matrix! Let this last row be $C = [c_1, c_2, \dots, c_{s+1}]$ then $CA = [0, 0, \dots, 0]$ the row s zeros.

But recall that we had AF = B so

$$C(AF) = CB$$

Since $C(AF) = (CA)F = [[0, 0, ..., 0] \begin{pmatrix} F_1 \\ F_2 \\ \vdots \\ F_i \\ \vdots \\ F_s \end{pmatrix} = O$ and

 $CB = c_1F_1 + c_2F_2 + \cdots + c_sF_s + c_{s+1}F_{s+1}$ and not all of the c_i are 0 we have that B is linearly dependent.

10. **Example:** Let P_3 be the vector space of polynomials of degree at most 3 and let $F_1 = x, F_2 = x^2, F_3 = x^3$ and $B_1 = x + x^2, B_2 = x - 4x^3, B_3 = x + 3x^2 + x^3,$ $B_4 = 2x - x^2 + x^3$ then according to the theorem the four elements of B = $\{B_1, B_2, B_3, B_4\}$ are linearly dependent since they are linear combinations of the three members of $F = \{F_1, F_2, F_3\}$. Moreover, we can calculate a dependence relation. Here

$$1 F_1 + 1 F_2 + 0 F_3 = B_1$$

$$1 F_1 + 0 F_2 + -4 F_3 = B_2$$

$$1 F_1 + 3 F_2 + 1 F_3 = B_3$$

$$2 F_1 - 1 F_2 + 1 F_3 = B_4$$

so setting

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & -4 \\ 1 & 3 & 1 \\ 2 & -1 & 1 \end{pmatrix}$$

We have $A \begin{pmatrix} F_1 \\ F_2 \\ F_3 \end{pmatrix} = \begin{pmatrix} B_1 \\ B_2 \\ B_3 \\ B_4 \end{pmatrix}$
and $REF(\langle A|I \rangle) = \begin{pmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & -1 & -4 & -1 & 1 & 0 & 0 \\ 0 & 0 & -7 & -3 & 2 & 1 & 0 \\ 0 & 0 & 0 & -\frac{32}{7} & \frac{5}{7} & \frac{13}{7} & 1 \end{pmatrix}$
According to the theorem it must be true that

$$\left(\begin{array}{ccc} -\frac{32}{7} & \frac{5}{7} & \frac{13}{7} & 1 \end{array}\right) \begin{pmatrix} B_1 \\ B_2 \\ B_3 \\ B_4 \end{pmatrix} = 0$$

In fact $\frac{-32}{7}(x+x^2) + \frac{5}{7}(x-4x^3) + \frac{13}{7}(x+3x^2+x^3) + (1)(2x-x^2+x^3) = 0$

11. Theorem B: Any two bases of a vector space have the same number of elements. **Proof:** We prove the theorem for finite dimensional vector spaces. The theorem is true for infinite dimensional spaces but requires the idea of *cardinality* to replace

"number of" and that would take too much time for an introductory course.

Suppose the vector space has two linearly independent spanning sets, B and C and that B has m elements and C has n elements. We can assume $m \leq n$. If m < n then $c_1, c_2, \dots, c_m, c_{m+1}$ are all in the linear span of the m elements of B and are therefore linearly dependent, a contradiction. Thus m = n.

12. **Definition:** If V is a vector space then the **dimension** of V is the number of elements in any basis of V.

The above theorem says that dimension is well-defined since all bases have the same number of elements.

Theorem: Any n + 1 elements in a vector space of dimension n is linearly dependent.

Proof: This follows immediately from Theorem A.

Corollary: If V is a vector space of dimension n and W is a subspace of V then $dim(W) \leq n$.

Corollary: If V is a vector space of dimension n then any n linearly independent elements of V is a basis of V.

Proof: Let $B = \{b_1, b_2, \dots, b_n\}$ be a set of *n* linearly independent elements of *V*. We need only show that *V* is in the linear span of *B* since we know it is linearly independent. Let *v* be any element of *V*. Then $\{b_1, b_2, \dots, b_n, v\}$ is a set with n + 1 elements in an *n*-dimensional vector space so it is dependent. This means that there are scalars α_i , not all zero, such that

$$\alpha_1 b_1 + \alpha_2 b_2 + \dots + \alpha_n b_n + \alpha_{n+1} v = O$$

It must be true that $\alpha_{n+1} \neq 0$ since otherwise

$$\alpha_1 b_1 + \alpha_2 b_2 + \dots + \alpha_n b_n = O$$

but that is not possible since If $\alpha_{n+1} \neq 0$ then $v = -\frac{\alpha_1}{\alpha_{n+1}}b_1 + -\frac{\alpha_2}{\alpha_{n+1}}b_2 + \cdots + -\frac{\alpha_n}{\alpha_{n+1}}b_n$ and we are done.

Corollary: If V is a vector space of dimension n and W is a subspace of V of dimension n then W = V.

Proof: The basis for W has n independent elements so it must be a basis for V.

13. Extension of independent sets to Bases

Theorem: Suppose V is a vector space and A is a linearly independent subset of V then there is a basis B of V which contains A.

Proof: Although the theorem is true in general, the proof for infinite dimensional spaces requires extra tools and we will restrict our attention to the case when V is a subspace of some finite dimensional space W. Let W have dimension N.

Suppose L is the linear span of A and $A = \{a_1, a_2, \dots, a_s\}$. A is a basis for L so if L = V then V has a basis containing A.

If $L \neq V$ then there is $a_{s+1} \in V$ that is not in L. It follows then that $\{a_1, a_2, \dots, a_s, a_{s+1}\}$ is linearly independent. If this were not the case then there would be

$$\alpha_1 a_1 + \alpha_2 a_2 + \dots + \alpha_s a_s + \alpha_{s+1} a_{s+1} = O$$

with not all of the $\alpha_i = 0$. But α_{s+1} must be 0 or a_{s+1} is in the span of A. However if this is the case then A is linearly dependent.

If $\{a_1, a_2, \dots, a_s, a_{s+1}\}$ does not span V then the process can be repeated to find a_{s+2}, \dots, a_{s+k} . Since $s+k \leq N$ and N is the maximum number of elements in an independent subset of V the process must stop at which time $\{a_1, a_2, \dots, a_s, a_{s+k}\}$ is a spanning set for V. But since it is independent it must be a bais and s = k = N.

14. A matrix approach to extension of independent sets to bases

The following is a method for implementing the previous theorem in the case where V is \Re^n and A is a linearly independent subset of \Re^n .

Suppose $A_1, A_2, ..., A_r$ are independent vectors in \Re^n . Then the above theorem says that one can find vectors $A_{r+1}, ..., A_n$ such that the vectors $A_1, ..., A_n$ are linearly independent and are therefore a basis for \Re^n . If A_1 is the matrix with columns $A_1, A_2, ..., A_r$ then this is equivalent to saying that we can add n - rcolumns to A_1 to get a matrix A with n linearly independent columns. The following construction does this.

Form $(A_1|I)$ by augmenting A_1 by the *n* by *n* identity and let *R* be an REF of $(A_1|I)$. Let *C* be the consistency matrix of A_1 . Since $rank(A_1) = r$ and has *n* rows, the REF of A_1 has exactly n-r zero rows and hence the consistency matrix *C* has n-r rows. If C^t is the transpose of *C* then its columns are independent since they are columns of an invertible matrix. Then $A = A_1|C^t$, is *n* by *n* and we claim its columns are linearly independent. We will prove this later as an application of orthogonality.

Example: Let
$$A_1 = \begin{pmatrix} 1 & 3 \\ 0 & 1 \\ 2 & 4 \\ 3 & -1 \end{pmatrix}$$
 then the REF of $\langle A|I \rangle$ is $\begin{pmatrix} 1 & 3 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & -2 & 2 & 1 & 0 \\ 0 & 0 & 0 & 7 & \frac{-3}{2} & 1 \end{pmatrix}$.
The consistency matrix is $C = \begin{pmatrix} -2 & 2 & 1 & 0 \\ 0 & 7 & \frac{-3}{2} & 1 \end{pmatrix}$. So $A = \langle A_1|C^t \rangle = \begin{pmatrix} 1 & 3 & -2 & 0 \\ 0 & 1 & 2 & 7 \\ 2 & 4 & 1 & \frac{-3}{2} \\ 3 & -1 & 0 & 1 \end{pmatrix}$

To see that A has independent rows we can, for instance calculate its REF and

	$\begin{pmatrix} 1 \end{pmatrix}$	3	-2	0 \
get	0	1	2	7
	0	0	9	$\frac{25}{2}$
	$\left(0 \right)$	0	0	$\left(\frac{3\overline{14}}{9}\right)$