Week 4&6

Linear algebra

F. Sorin (MX)

Ecole Polytechnique Fédérale de Lausanne

Overview

- General concept of Matrix analysis as applied to change of Basis problems
- Review of important linear algebra notions

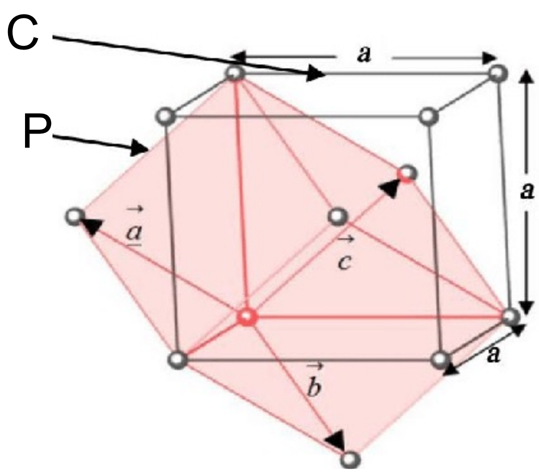
- The postulate of Quantum Mechanics and the need for advanced linear algebra

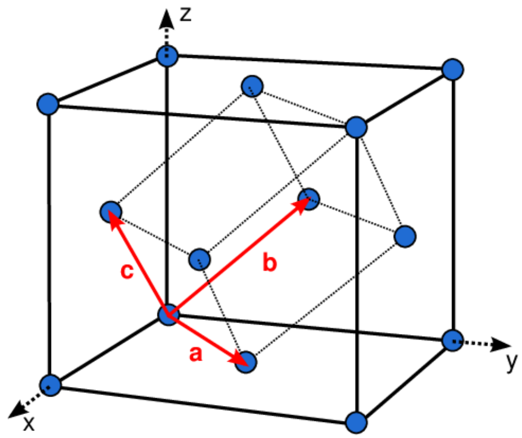
- Advanced Linear algebra: Vector spaces and Hilbert spaces, self-adjoint operators, unitary operators, spectral theorem.

- Example of applications:
  - Bloch theorem
  - Brillouin zone

Towards linear algebra

- The matrix formalism in vector spaces to change basis and manipulate vectors and operators in different basis is very important and useful.
- Changing unit cells and using the reciprocal space entails to change coordinates. The matrix formalism is a powerful tool to do that.





BCC Bravais lattice: $\mathcal{B}' = \{O, \vec{a}', \vec{b}', \vec{c}'\}$

$$\vec{a}' = \frac{1}{2}(-\vec{a} + \vec{b} + \vec{c})$$

$$\vec{b}' = \frac{1}{2}(\vec{a} - \vec{b} + \vec{c})$$

$$\vec{c}' = \frac{1}{2}(\vec{a} + \vec{b} - \vec{c})$$

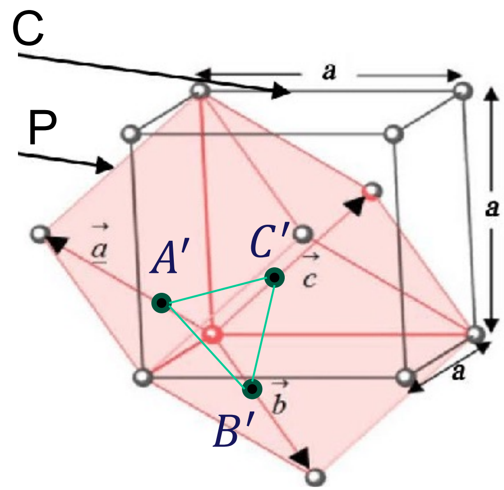
FCC Bravais lattice: $\mathcal{B}'' = \{O, \vec{a}'', \vec{b}'', \vec{c}''\}$

$$\vec{a}'' = \frac{1}{2}(\vec{b} + \vec{c})$$

$$\vec{b}'' = \frac{1}{2}(\vec{a} + \vec{c})$$

$$\vec{c}'' = \frac{1}{2}(\vec{a} + \vec{b})$$

Towards Linear Algebra



- An interesting exercise is to verify that indeed the normal to a plan (hkl) is the direction in the reciprocal space [hkl].

- Let's consider a plan P'(h'k'l') that intercepts the axis $\vec{a}', \vec{b}', \vec{c}'$

at points $A' = \begin{pmatrix} 1/h' \\ 0 \\ 0 \end{pmatrix}$, $B' = \begin{pmatrix} 0 \\ 1/k' \\ 0 \end{pmatrix}$ and $C' = \begin{pmatrix} 0 \\ 0 \\ 1/l' \end{pmatrix}$.

- This plan is indeed by construction a (*h'k'l'*) crystal plan. And still by construction of the reciprocal space, the direction $\vec{N}' = h' \, \vec{a}'' + k'\vec{b}'' + l'\vec{c}''$ should be perpendicular to the plan P'(h'k'l'), since the reciprocal space of a BCC is the FCC Bravais lattice.
  In other words, $\vec{N}'$ should be perpendicular to the vectors $\overrightarrow{A'B'}$ and $\overrightarrow{A'C'}$ that generate P'.

- To verify this, we can express all vectors in the orthonormal $\mathcal{B} = \{O, \vec{x}, y, \vec{z}\}$ and apply the dot product in this convenient basis.

- The first question is then:

Is there a matrix that enables to express the coordinates of a vector in the orthonormal basis $\mathcal{B} = \{O, \vec{a}, \vec{b}, \vec{c}\}$ from its coordinate in the $\mathcal{B}' = \{O, \vec{a}', \vec{b}', \vec{c}'\}$ basis ?

Towards Linear Algebra

- In the $\mathcal{B}' = \{O, \vec{a}', \vec{b}', \vec{c}'\}$ basis, we have: $\vec{a}' = \begin{pmatrix}1\\0\\0\end{pmatrix}$; $\vec{b}' = \begin{pmatrix}0\\1\\0\end{pmatrix}$; $\vec{c}' = \begin{pmatrix}0\\0\\1\end{pmatrix}$

- We are looking for a matrix $M' = \begin{pmatrix}m_{11} & m_{12} & m_{13}\\ m_{21} & m_{22} & m_{23}\\ m_{31} & m_{32} & m_{33}\end{pmatrix}$ such that : $M'\vec{a}'_{B'} = \vec{a}'_{B}$

Coordinates of $\vec{a}'$ in $\mathcal{B}'$

Coordinates of $\vec{a}'$ in $\mathcal{B}$

And similarly for $\vec{b}'$ *and* $\vec{c}'$

- We find that $M' = \frac{1}{2}\begin{pmatrix}-1 & 1 & 1\\ 1 & -1 & 1\\ 1 & 1 & -1\end{pmatrix}$ and that M' is reversible.
- In the same way we can find a matrix M'' that transforms the coordinates in $\mathcal{B}'' = \{O, \vec{a}'', \vec{b}'', \vec{c}''\}$ into coordinates in $\mathcal{B}$:

$$M'' = \frac{1}{2}\begin{pmatrix}0 & 1 & 1\\ 1 & 0 & 1\\ 1 & 1 & 0\end{pmatrix}$$

- From the linearity of matrix operations, we can express the coordinates of the vectors $\overrightarrow{A'B'}$ and $\overrightarrow{A'C'}$ in the $\mathcal{B}$ basis also by multiplying them by M'.
- To prove the orthogonality of P' and $\vec{N}'$, we finally need to take the dot product of $\mathrm{M}'\overrightarrow{A'B'}$ and $M''\vec{N}'$, as well as of $\mathrm{M}'\overrightarrow{A'C'}$ and $M''\vec{N}'$.

Important concepts in Linear Algebra

- Linear algebra for engineers typically cover:
  - Linear system of equations;
  - Matrix formalism, their manipulation, multiplication.
  - Rank of a matrix:

  The dimension of the vector space generated by its columns (or rows).

  i.e. maximal number of linearly independent columns.

  - Addition and multiplication by a scalar have all the nice regular properties: commutativity, associativity…
  - Multiplication:

  For two matrices A (kxp) and B (pxn):

A *kxn* matrix

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1j} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2j} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{i1} & a_{i2} & \dots & a_{ij} & \dots & a_{1n} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{k1} & a_{k2} & \dots & a_{kj} & \dots & a_{kn} \end{bmatrix}$$

$$AB = [(AB)_{ij}]_{\substack{i=1,\dots,k \\ j=1,\dots,n}}$$

$$= \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1l} & \dots & a_{1p} \\ a_{21} & a_{22} & \dots & a_{2l} & \dots & a_{2p} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{i1} & a_{i2} & \dots & a_{il} & \dots & a_{ip} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{k1} & a_{k2} & \dots & a_{kl} & \dots & a_{kp} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1j} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2j} & \dots & b_{2n} \\ \vdots & \vdots & & \vdots & & \vdots \\ b_{l1} & b_{l2} & \dots & b_{lj} & \dots & b_{ln} \\ \vdots & \vdots & & \vdots & & \vdots \\ b_{p1} & b_{p2} & \dots & b_{pj} & \dots & b_{pn} \end{bmatrix}$$

$$(AB)_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \dots + a_{il}b_{lj} + \dots + a_{ip}b_{pj} = \sum_{l=1}^{p} a_{il}b_{lj}$$

Important concepts in Linear Algebra

- Matrix multiplication:
  - Associative: A(BC) = (AB)C
  - Not commutative!

  - Function of matrices:

  In the same way that a function of a variable f(x) can be constructed through its Taylor series, functions f(M) of a squared matrix M can be defined through the corresponding Taylor series. Hence for the exponential:

$$\exp(M) = \mathbf{1} + \mathbf{M} + \frac{\mathbf{M}^2}{2} + \cdots = \sum_{n=0}^{\infty} \frac{\mathbf{M}^n}{n!}$$

  - Note that for two matrices (operators) that do not commute, ie $H_1 H_2 \neq H_2 H_1$, then the exponents also don't commute.
  - The non-commutativity of the product of matrices, and the facts that some do commute, has profound meaning in physics / engineering.
  - Example:

  If we consider the differential equation $\frac{d\vec{x}}{dt} = \mathbf{M}\vec{x} \quad \Longrightarrow \quad \vec{x}(t) = \exp(t\mathbf{M})\,\vec{x}(0)$

  If the matrix, or the operator as we will define later, varies over time: $M = M_1$ until t_1, then. $M = M_2$, we have:

$$\vec{x}(t_1+t_2) = \exp(t_2\mathbf{M}_2)x(t_1) = \exp(t_2\mathbf{M}_2)\exp(t_1\mathbf{M}_1)x(0) \neq \exp(t_1\mathbf{M}_1+t_2\mathbf{M}_2)x(0)$$

  The order at which we apply a matrix operator must be respected.

Important concepts in Linear Algebra

- Transpose Matrix:
  - For $A = (a_{ij})$ a kxn matrix, the transpose matrix of A is the nxk matrix: $A^T = (a_{ji})$.
  - For two matrices A and B with the proper size:

$$(A^T)^T = A \qquad (AB)^T = B^T A^T$$

$$(A+B)^T = A^T + B^T \qquad (A_1 A_2 \ldots A_m)^T = A_m^T \ldots A_2^T A_1^T$$

$$(\alpha A)^T = \alpha A^T$$

- A matrix A is symmetric if $A = A^T$ and anti-symmetric if $A = -A^T$.
  Every matrix is the sum of a symmetric and an anti-symmetric matrix.

- Trace of **square mattrices**:
  - The trace of a matrix $A = (a_{ij})$ is the sum of the diagonal terms: $\operatorname{tr} A = \sum_{i=1}^{n} a_{ii}$
  - For two square matrices A and B, and α a scalar:

$$\operatorname{tr}(A+B) = \operatorname{tr} A + \operatorname{tr} B$$

$$\operatorname{tr}(\alpha A) = \alpha \operatorname{tr} A;$$

$$\operatorname{tr} A = \operatorname{tr} A^T;$$

$$\operatorname{tr}(AB) = \operatorname{tr}(BA).$$

  - The trace is independent of the basis onto which the operator is defined !

Important concepts in Linear Algebra

- Inverse Matrix:

  - If A is a square matrix (real or complex), B is the inverse of A if AB = BA = I

  I is the identity matrix nxn with diagonal coefficients equal to 1, and off-diagonal coefficients equal to 0.

  B is unique ! It is noted A^{-1}. In finite dimensions, it is equivalent to say:

  - A is invertible
  - The equation Ax = b has a unique solution.
  - If A is a square matrix of order n, *rank(A) = n*.
  - The linear application $x \to Ax$ is injective
  - The linear application $x \to Ax$ is surjective

$$(A^{-1})^{-1} = A;$$

$$(AB)^{-1} = B^{-1}A^{-1}$$

$$(A^n)^{-1} = (A^{-1})^n;$$

$$(\alpha A)^{-1} = \frac{1}{\alpha}A^{-1};$$

$$(A^T)^{-1} = (A^{-1})^T.$$

Important concepts in Linear Algebra

- The determinant of a matrix was invented to evaluate if a nxn matrix is of rank n with a single number, and not lines / columns operations.

- For $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, $det(A) = ad - bc$. If $det(A) = 0$, it gives a relation between the two lines that indicates if they are linearly dependent

- Key result: A nxn matrix is inversible if an only if $det(A) \neq 0$

- There are many ways to derive the determinant. A practical one is the Laplace formula:

Let $A = (a_{ij})$ be a square matrix of order n. Let [A_{ij}] be the submatrix of A obtained by deleting row i and column j. The minor-ij M_{ij} and the cofactor-ij C_{ij} are defined by

$$M_{ij} = \det[A_{ij}], \qquad C_{ij} = (-1)^{i+j} M_{ij}$$

$$det(A) = \sum_{j=1}^{n} a_{ij} C_{ij}$$

- One can also show that for two square matrices: $det(AB) = det(A)\, de\,t(B) = det(BA)$

- This is very important as it ensures that the determinant is independent of the basis, so the inversible property is a function of the linear transformation associated to A.

Algebraic structures – Vector Space

- A vectorial space adds one more opportunity for operations now between elements of different nature. It indeed combines elements from different sets (scalars and vectors) where operations are possible within each set and across sets.

If (K,+,.) is a commutative Field, we call K-vector space a set E with an internal law $+_E$, (noted +) and an external law from K x E into E

$$(\lambda, x) \longrightarrow \lambda x$$

Such that:

- (E,+) is an abelian group;

- $\forall(\lambda,\mu) \in K^2, \forall x \in E, \ (\lambda+\mu)x = \lambda x + \mu x$

  $\forall \lambda \in K, \forall(x,y) \in E^2, \ \lambda(x+y) = \lambda x + \lambda y$

  $\forall(\lambda,\mu) \in K^2, \forall x \in E, \ \lambda(\mu x) = (\lambda\mu)x$

  $\forall x \in E, \ 1x = x.$

- Examples:
  - The set of vectors is a $\mathbb{R}$-vector space;
  - Complex numbers can be defined as a $\mathbb{R}$-vector space for usual laws;
  - Polynomial can be built as a vector space, even an algebra

- **An algebra** is a K-vector space E with an additional internal law $*$ distributive on +, and that can be associative, commutative, and can have an identity.

  Example: the cross product of two vectors in $\mathbb{R}^3$ is an internal law that is not associative.

Vector Spaces – basic definitions

- Let's consider the general case of $\mathbb{C}$ $-$vector spaces, and an arbitrary nature of the vector (vectors $\mathbb{C}^n$,functions, matrices etc...)

- A subspace of a vector space V is a subset of V that is also a vector space. To verify that a subset U of V is a subspace you must check that U contains the vector 0 (neutral for addition), and that U is closed under addition and scalar multiplication.

- A list $(v_1, \dots, v_n)$ of vectors in V is a **finite number** of vectors of length n.

The **span** of a list of vectors $(v_1, \dots, v_n)$ in V, denoted as $\text{span}(v_1, \dots, v_n)$, is the set of all linear combinations of these vectors:

$$\text{span}(v_1, \dots, v_n) = \{u \in V, \exists (a_1, \dots, a) \in \mathbb{C}^n, u = a_1 v_1 + \dots + a_n v_n\}$$

- A vector space V is said to be **finite dimensional** if it is spanned by some list of vectors in V:

$$\exists\ (v_1, \dots, v_n) \in V, \forall u \in V, \exists (a_1, \dots, a_n) \in \mathbb{C}^n, u = a_1 v_1 + \dots + a_n v_n$$

  - If V is not finite dimensional, it is infinite dimensional. In such case, no list of vectors from V can span V.

  - Example: the space of $\mathbb{C}$-polynomilal is of infinite dimension.

Polynomials

- Polynomials are extremely important mathematical objects in all aspects of applied mathematics. They are used to extrapolate functions, can be found in the resolution of differential equations, appear in many models, etc…
- Deriving, integrating, manipulating and finding roots of polynomials is hence important skills to master.

- The construction of polynomials is very interesting but a full explanation is beyond the scope of this class. Here are the main steps:
  - We can define the set of sequences in $\mathbb{R}$ with a finite number of terms that are non-zero.

    Equivalently: $K^{(\mathbb{N})} = \{(a_n) \in \mathbb{R}, n \in \mathbb{N},$ such that $(\exists N, \forall k \geq N, a_k = 0)\}$.

  - To each polynomial *P* of degree n in $\mathbb{R}$ or $\mathbb{C}$, one can associate a unique sequence in $K^{(\mathbb{N})}$ such that:

    $$P(z) = \sum_{k=0}^{n} a_k z^k$$

  - The degree is defined as the greatest integer n for which $a_n \neq 0$.

This unicity can be easily demonstrated by induction on the degree of the polymer.

  - One then defines, on top of the usual addition operation, the multiplication as follow:
    - $For\ (a_n)$ & $(b_n)\ \in K^{(\mathbb{N})}, (c_n) = \sum_{k=0}^{n} a_k b_{n-k}$
  - One can then show that $K^{(\mathbb{N})}$ with this internal law, the regular addition law, and the external product is an algebra.
  - If we note $X \in K^{(\mathbb{N})}$ with $X = \{0,1,0,0, \dots 0, \dots\}$. It is easy to show that $X^k = \{0,0, \dots 0,1,0, \dots\}$
  - The algebra can be generated by the basis of the polynomials $X^k_{k \in \mathbb{N}}$
  - In other words, the set of polynomials of degree n is of dimension n+1 and is generated by the polynomials $X^k_{0 \leq k \leq n}$

Vector Spaces – Basis

- A **basis** of V is a list of vectors in V that both spans V and is linearly independent.
  - A list of vectors $(v_1, \dots, v_n)$ is said to be linearly independent if the equation:
    $a_1 v_1 + \cdots + a_n v_n = 0$
    has for only solution $\forall i, a_i = 0$
    ie: one cannot express one vector of the set as linear expression of the others.

- A basis of V is a list of vectors in V that both spans V and is linearly independent:
  - **The dimension** of a finite dimensional vector space V is the length of the shortest list of vectors that span V.
  - For a finite dimensional vector space, a list of vectors of length dimV is a basis if it is linearly independent or if it is a spanning list.
  - All basis of a finite dimensional vector space have the same length.
    - Cannot be of smaller length, otherwise dimV would be reduced by 1.
    - Cannot be of greater length: **in a vector space of dimension n, there cannot be a list of linearly independent vector of length n+1** (see Annexe)
  - All list of linearly independent vector of length the dimension of V form a basis for V !

Vector Spaces – Linear transformations

- Let U and V be vector spaces over K. A function T : V → U is called a linear transformation if, for all u, v ∈ V and α ∈ K:
  - T(u + v) = T(u) + T(v)
  - T(αu) = αT(u)
- If the image of T is in V, ie if T : V → V, T is called a **linear operator.**
- Example: Let V denote the space of real polynomials p(x) of a real variable x with real coefficients. Here are two linear operators:
  - Differentiation: Tp = p'. This operator is linear because $(p_1 + p_2)' = p_1' + p_2'$ and (ap)′ = ap'.
  - Multiplication by x: Sp = xp is also a linear operator.
- To every linear operator T, one can associate a matrix that acts on the vectors of V (finite dimension).
  - This matrix can be defined by the effect of T on the vectors of a given basis
  - So the matrix will depend on the basis chosen !
- If $(v_1, \dots, v_n)$ forms a basis of V, Tv_1 is a vector in V so it can be written as a linear combination of the basis vectors:

$$Tv_j = T_{1j} v_1 + T_{2j} v_2 + \dots + T_{nj} v_n$$

Vector Spaces – Linear transformations

- Which we can generalize as:

$$Tv_j \quad \longleftrightarrow \quad \begin{pmatrix} T_{11} & \cdots & T_{1j} & \cdots & T_{1n} \\ T_{21} & \cdots & T_{2j} & \cdots & T_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ T_{n1} & \cdots & T_{nj} & \cdots & T_{nn} \end{pmatrix} \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \; j\text{-th position}$$

Which forms the matrix, in the particular basis, of the operator.

- The notions discussed on matrices above apply to operators !
  - A linear operator T : V → V is said to be injective if Tu = Tv, with u,v ∈ V, implies u = v.
  - T is injective if and only if nullT = {0}, with the subspace: $\text{null } T \;=\; \{v \in V;\; Tv = 0\}$
  - The range of T is the subspace image of V under the map T: $\text{range } T \;=\; \{Tv;\; v \in V\}$
  - T is surjective if rangeT = V
  - In infinite dimension, T is bijective if it is injective **and** surjective.
  - In finite dimension, T is bijective and has an inverse if it is injective **or** surjective, just like its associated matrix !
  - dim(rangeT) = rank(T)

From Operators to Matrices

- **Commutativity**: the composition of two operators is associated with the product of matrices.
- Two operators will commute in terms of their composition, if their associated matrices commute with respect to the multiplication of matrices.

- The commutator [·,·] of two operators X,Y is defined as [X,Y]≡ XY − YX.
  - Two operators X, Y commute if [X,Y] = 0.

- The trace and determinant of operators are defined the same way as above, and do not depend on the basis chosen for the associated matrix.

- **Eigen values and eigen vectors** of operators:
  - An eigen vector u form a linear operator T is a vector that satisfies $Tu = \lambda u$.
  - λ is called an eigenvalue.
  - For a given eigenvalue λ, there maybe several linearly independent eigen vectors of T. We say that λ generates a sub-space of a given dimension ≥ 1.
  $$U_\lambda = \{u \in V, Tu = \lambda u\}$$
  The eigenvalue is then said to be degenerate if dim U_λ>1.

  - The set of eigenvalues of T is called the spectrum of T.
  - Set of eigenvectors of T corresponding to λ = null(T − λI).
  - The eigen values are found solving $\det(T - \lambda I) = 0$.
  - $P(\lambda) = \det(T - \lambda I)$ is called the characteristic polynomial of T.

Diagonalizable Matrices

- ***Theorem***: Let T be a linear operator, and assume $\lambda_1,...\lambda_n$ are distinct eigenvalues of T and $u_1,...u_n$ are corresponding nonzero eigenvectors. Then $(u_1,...u_n)$ are linearly independent.
- This is true for $\mathbb{R}$ and $\mathbb{C}$ vectorial spaces: they can form a basis for the space !!
- If A is a nxn matrice with n distinct and non zero eigenvalues, it is diagonalizable !
- A square matrix A is diagonalizable if it is similar to a diagonal matrix, ie there exists an invertible matrix of passage P, and a diagonal matrix D, such that $A = PDP^{-1}$
- P is the matrix of the eigen vectors of A !
- Equivalently, A is diagonalizable if there exists a basis of its eigen vectors.
- The associated linear operator T is diagonalizable if there is a basis of the vectorial space V formed by the eigenvectors of T.
- In a more general sense, if the dimension of the sub-spaces of the eigen values of A (nxn) add up to n, then it is diagonalizable.
  So if $\lambda_1,...\lambda_k$ are distinct eigen vectors each generating a subspace $U_{\lambda_i} = \{u \in V, Tu = \lambda_i u\}$,

  A is diagonalizable if $\sum_i \dim(U_{\lambda_i}) = n$

Diagonalizable Matrices

- Roots and polynomial algebra:
  - A polynomial of degree *n* in $\mathbb{R}$ can have a maximum of *n* roots, and the polynomials $(X-\alpha)^{\beta}$ are irreducible factors, very much like prime numbers for numbers.
  - In particular, $\alpha \in \mathbb{R}$ is the root of a polynomial P, of order $\beta \in \mathbb{R}$, if there is a polynomial Q such that

  $P = (X-\alpha)^{\beta}$ Q.

  A polynomial in $\mathbb{R}$ is said split, if $\exists \alpha_i \in \mathbb{R}, \beta_i \in \mathbb{N}$ such that $P(X) = \prod_i (X-\alpha_i)^{\beta_i}$

  If $\beta_i > 1$, the root is said degenerate. If *deg(P)* = *n*, then $n = \sum_i \beta_i$

  - **Fundamental result**: every polynomials in $\mathbb{C}$ has at least one root.
  - Corollary: every polynomial in $\mathbb{C}$ is split.
  - Example : factorize P(x) = x^4-1

- The characteristic polynomial of a matrix in $\mathbb{C}$ can be split.

- Algebraic multiplicity: the order of the eigen value as a polynomial root, β_i : $P(X) \propto (X-\lambda_i)^{\beta_i}$

- Geometric multiplicity: the dimension of the sub-space generated by λ_i: $\dim(U_{\lambda_i})$

- One can show that
  - $\dim(U_{\lambda_i}) \leq \beta_i$
  - A is diagonalizable if and only if $\forall \lambda_i, \dim\left(U_{\lambda_i}\right) = \beta_i$

Inner Product

- An inner product on a vector space V over $\mathbb{R}$ *or* $\mathbb{C}$ is a map from an ordered pair (u, v) of vectors in V to a number $\langle u|v\rangle$ in $\mathbb{R}$ *or* $\mathbb{C}$. The axioms for $\langle u|v\rangle$ are inspired by the axioms for the dot product of vectors:

1. $\langle v|v\rangle \geq 0$, for all vectors $v \in$ V.

2. $\langle v|v\rangle = 0$ if and only if $v = 0$.

3. $\langle u|v_1 + v_2\rangle = \langle u|v_1\rangle + \langle u|v_2\rangle$. Additivity in the second entry.

4. $\langle u|\alpha v\rangle = \alpha \langle u|v\rangle$, $\alpha \in \mathbb{C}$. Homogeneity in the second entry.

5. $\langle u|v\rangle = \langle v|u\rangle^*$. Conjugate exchange symmetry.

So: $\langle \alpha u|v\rangle = \langle v|\alpha u\rangle * = \alpha^* \langle v|u\rangle * = \alpha^* \langle u|v\rangle$

- The norm of a vector is also noted: $|v|^2 = \langle v|v\rangle \geq 0$

- Here we already take the Dirac notation:
  - ket $|v\rangle$ is a vector;
  - Bra $\langle v|$ is a linear operator acting on a vector via the dot product.

- The complex conjugate is noted with a*. If we consider a real vector space, the conjugate is just the real number unchanged and we find the symmetry of the dot product in $\mathbb{R}^3$

Inner Product

- Two vectors are orthogonal if $\langle u|v\rangle = 0$.

- Schwartz inequality: $|\langle u|v\rangle| \leq |u||v|$

- A list of vectors is said to be orthonormal if all vectors have norm one and are pairwise orthogonal.
  - Consider a list (e_1,...,e_n) of orthonormal vectors in V. Orthonormality means that:

$$\langle e_i, e_j\rangle = \delta_{ij}$$

  - We also have:

$$\begin{aligned} |a_1 e_1 + \ldots + a_n e_n|^2 &= \langle a_1 e_1 + \ldots + a_n e_n \,,\, a_1 e_1 + \ldots + a_n e_n\rangle \\ &= \langle a_1 e_1 \,, a_1 e_1\rangle + \ldots + \langle a_n e_n \,, a_n e_n\rangle \\ &= |a_1|^2 + \ldots + |a_n|^2 \,. \end{aligned}$$

  - So a set of orthonormal vectors are necessarily linearly independant.

- An orthonormal basis of V is a list of orthonormal vectors that is also a basis for V. Let (e_1,...,e_n) denote an orthonormal basis. Then any vector v can be written as

$$v = a_1 e_1 + \ldots + a_n e_n \qquad \text{with} \qquad \langle e_i, v\rangle = \langle e_i \,, a_i e_i\rangle = a_i$$

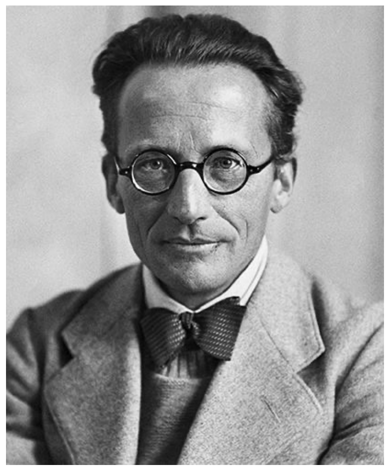
- This represents the projection of v on a vector e_i.

Linear Algebra as a profound mathematical formalism

- Matrix analysis is a great tools in Materials Science to handle anisotropic properties of materials.
- The tensor formalism in continuum and fluid mechanics is also greatly facilitated by a good understanding of matrix formalism.

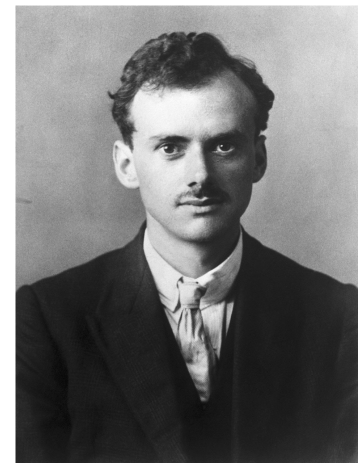
- In the 19th and 20th centuries, several physicists had the wonderful intuition to treat analysis problems of the wave mechanics with geometry, vectors, and linear algebra approaches.

David Hilbert
1862 - 1943



Erwin Schrödinger
1887 - 1961

John Von Neumann
1903 - 1957



Paul Dirac
1902 - 1984

- Notions missing:
  - Hilbert spaces
  - Adjunct functions;
  - Hermitian and Unitary operators
  - Spectral theorem

Quantum Mechanics: from waves to vectors

- In classical mechanics the motion of a particle is usually described using the time-dependent position $\overrightarrow{x(t)}$ as the dynamical variable.

- In wave mechanics the dynamical variable is a wave-function. This wavefunction depends on position and on time and it is a complex number.

- When all three dimensions of space are relevant we write the wavefunction as $\Psi(\vec{x}, t)$

- For one dimensional problems, the Schrödinger equation governs the evolution in space and time of the wave function for a non-relativistic particle:

$$i\hbar\frac{\partial\Psi}{\partial t}(x,t) = \left(-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x,t)\right)\Psi(x,t)$$

- Note that Ψ must be complex otherwise the LHS is complex but the RHS would be real.

- The equation is linear so any linear combination of waves solution of the Schrödinger equation is also a solution.

- The density probability is defined as:

$$P(x,t) = \rho(x,t) \equiv \Psi^*(x,t)\Psi(x,t) = |\Psi(x,t)|^2 \quad \text{with} \quad \int_{-\infty}^{\infty} dx\, |\Psi(x,t)|^2 = 1$$

- Note that the Schrödinger equation implies that $\frac{d}{dt}\int_{-\infty}^{\infty} dx\, |\Psi(x,t)|^2 = 0$

Quantum Mechanics: a linear algebra formalism

- A first hint that the formalism of wavefunction could be interpreted as a geometric / linear algebra problem, is that the set of complex functions that have the square of their norm integrable is a Hilbert vectorial space.

- Let's look at the time independent form of the equation, via a separation of variable:

$$\Psi(x,t) \;=\; e^{-iEt/\hbar}\,\psi(x)$$

we get

$$\left(-\frac{\hbar^2}{2m}\frac{d^2}{dx^2}+V(x)\right)\psi(x) \;=\; E\,\psi(x)$$

or

$$\hat{H}\,\psi(x) \;=\; E\,\psi(x)$$

- We recognize here an eigen vector problem where the wave function could be seen as a vector, and H an operator (a matrix) that acts on such vector. The matrix has unusual coefficients as it could host operations on functions such as derivatives.

- A solution ψ associated with an energy E is called an energy eigenstate of energy E. The set of all allowed values of E is called the spectrum of the Hamiltonian H. A degeneracy in the spectrum occurs when there is more than one solution ψ for a given value of the energy.

Quantum Mechanics: a linear algebra formalism

- To solve such an equation, we usually don't impose the normalize requirement to the wave function, but only that it is continuous and bounded, and that its derivatives are bounded.

- With discrete energy levels (note that the equation could also lead to a continuum of energy in some cases), one obtains a set of eigenvectors $\psi_n(x)$ with associated energy levels E_n.

- A good example is the case of a free particle in a box in one dimension.

$$\left(-\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + V(x)\right)\psi(x) = E\,\psi(x)$$

with $V(x) = \begin{cases} 0, & -\frac{L}{2} \leq x \leq \frac{L}{2} \\ +\infty, & x > \frac{L}{2} \ and \ x < -\frac{L}{2} \end{cases}$

- The solutions:

$$\psi_n(x) = \begin{cases} \sqrt{\frac{2}{L}}\sin(k_n x) & \text{for } n \text{ even} \\ \sqrt{\frac{2}{L}}\cos(k_n x) & \text{for } n \text{ odd.} \end{cases}$$

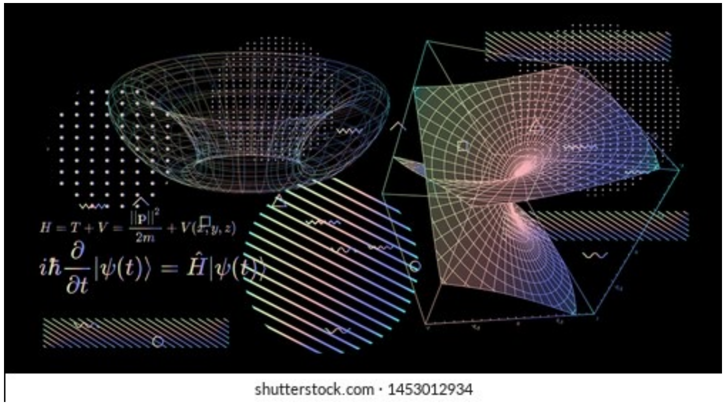
with $k_n = \frac{n\pi}{L}$ $\quad E_n = \hbar\omega_n = \frac{n^2\pi^2\hbar^2}{2mL^2}$

- Since the Schrödinger equation is linear, all linear combinations of the $\psi_n(x)$ are also solution of the equation for a proper E.
- If we define the product of wave function as $\langle\psi_m(x)|\psi_n(x)\rangle = \int_{-L/2}^{L/2}\psi^*{}_m(x)\,\psi_n(x)dx$, one can show that it is zero except for *m* = *n.*

Postulates of Quantum Mechanics

- In other words, we can create a function that will be a linear combination of a set of functions that are orthonormal for a certain inner product.

- This calls for a linear algebra formalism !



- Quantum Mechanics Postulates (there are various ways to organize them):

Postulate 1: Superposition principle

To each physical system is associated a **Hilbert space** $\mathcal{E}_{\boldsymbol{H}}$. The state of a system is entirely defined at each instant by a normalized vector of the system:

$$|\psi(t)\rangle = \sum c_i \; |\psi_i\rangle$$

where the $|\psi_i\rangle$ notation is the Dirac “ket” notation representing a vector.

The $|\psi_i\rangle$ are the state vectors that form an **orthonormal basis**.

Note that a state can be shifted by a phase factor without changing the physical meaning. However, the phase factor of the coefficients c_i cannot be ignored.

Postulates of Quantum Mechanics

- **Postulate 2: Operators**

For every physical quantity A, we can associate a linear **self-adjoint operator $\hat{A}$** (Hermitian operator in finite dimension) that acts on the Hilbert space $\mathcal{E}_{\boldsymbol{H}}$

$$\left\langle \varphi \middle| \hat{A}\psi \right\rangle = \left\langle \hat{A}\varphi \middle| \psi \right\rangle$$

- **Postulate 3: Experimental outcome**

For a given physical quantity A, whatever the state of the system $|\psi\rangle$ right before a measurement, the only possible outcome of a measurement are the **(real) eigen values** of the observable $\hat{\boldsymbol{A}}$.

- **Postulate 4: Projection principle**

  - The probability to find an eigen value a_α of an observable $\hat{\boldsymbol{A}}$ is given by:

$$P(a_\alpha) = \sum_{r_\alpha=1}^{n_\alpha} |\langle \alpha, r_\alpha | \psi \rangle|^2$$

Where n_α is the dimension of the sub-space generated by a_α, and the $|\alpha, r_\alpha\rangle$ the associated orthonormal eigenvectors.

  - The new state right after the measurement is **the projection of $|\psi\rangle$ on this sub-space generated by a_α**.

Postulates of Quantum Mechanics

- **Postulate 5: Time evolution**

For a state $|\psi(t)\rangle$ at time t, as long as the system is not subjected to any type of observation (no external intervention / interaction), its evolution over time is governed by the Schrödinger equation:

$$i\hbar\frac{d|\psi(t)\rangle}{dt} = \hat{H}|\psi(t)\rangle$$

Where $\hat{H}$ is the energy observable, or the Hamiltonian of the system.

- This postulate is not demonstrated but is a model that is confronted to experiment and that has done quite well thus far.

- It is hard to develop a physical intuition to this formalism, and to the world of infinitesimally small in general.

- A solid understanding of the underlying mathematical formalism is however of great help to handle quantum mechanics that is essential to understand many properties of materials.
- Missing notions:
  - Dot product and orthonormal basis for a set of functions;
  - Hilbert spaces;
  - Self-adjoint (Hermitian) operators;
  - Adjunct functions;
  - Spectral theorem.

Advanced Linear Algebra: Hilbert Spaces

- The postulate of QM stipulated that a system must be defined on a Hilbert Space.

- A Hilbert space H is a real or complex **inner product** space that is also a **complete metric** space with respect to the distance function induced by the inner product

- Inner product space is simply a vectorial space with an inner product.

- A complete metric is the property that every Cauchy sequence of H with respect to the metric converges in H.
  This means that there is no "missing point": for example, $\mathbb{Q}$ is not complete because you can have sequences in $\mathbb{Q}$ converging towards irrational numbers like $\sqrt{2}$.

Example: $\mathbb{R}$-polynomials is not a Hilbert space: $e^x = \sum \frac{x^k}{k!}$

- The completeness of the Hilbert space used in QM is particularly important for infinite dimension spaces.
  It expends the notion of finite-dimensional Euclidian spaces (which are Hilbert spaces) to infinite-dimensional ones.

If a state $|\psi\rangle = \sum c_i \, |\psi_i\rangle$, the infinite sum must converge in the vectorial space. It is enough that the sum of the norms converges (if it converges absolutely, then the series also converges to a vector) and this is exactly the Cauchy completeness condition.

Advanced Linear Algebra: Projection Operator

- Consider a subset U of vectors in V (not necessarily a subspace). Then we can define a subspace $U^\perp$, called the orthogonal complement of U as the set of all vectors orthogonal to the vectors in U:

$$U^\perp = \{v \in V \,|\, \langle v, u \rangle = 0, \text{ for all } u \in U\}$$

- It is a simple yet important notion in QM since after a measurement, the eigen state is projected on the subspace generated by the measured eigenvalue of the corresponding observable.

- If U is a subspace of V, then $V = U \oplus U^\perp$

- Given this decomposition any vector $v \in V$ can be written uniquely as $v = u + w$ where $u \in U$ and $w \in U^\perp$. One can define a linear operator P_U, called the orthogonal projection of V onto U, that acting on v gives the vector u.

- Since P_U is the identity on U, it is easy to show that $P_U^2 = P_U$ and so it only has 0 or 1 as eigenvalues.
- Then:

$$P_U = \text{diag}(\underbrace{1, \ldots 1}_{n \text{ entries}}, \underbrace{0, \ldots, 0}_{k \text{ entries}})$$

- With $n = Tr(P_U) = \dim U = \text{rank} P_U$

- Every vectorial space of finite dimension, has an orthonormal basis for a given inner product.

Advanced Linear Algebra: Orthonormal Basis

- Every vectorial space of finite dimension has an orthonormal basis for a given inner product.
  - By definition, a finite dimensional space has a list of vector that spans it. Hence it must have a basis.
  - The orthonormal basis can be constructed from an existing arbitrary basis via the Gram-Schmidt algorithm.

  From an arbitrary basis of a vectorial space V with $\dim(V) = n$, and $(v_1, \ldots., v_n)$ a basis of V, you can built the orthonormal basis $(e_1, \ldots., e_n)$ with:

$$u_1 = v_1 \quad ; \quad e_1 = \frac{u_1}{\|u_1\|}$$

$$u_2 = v_2 - \langle v_2, u_1 \rangle u_1, \qquad e_2 = \frac{u_2}{\|u_2\|}$$

$$\forall k \geq 2, u_k = v_k - \sum_{i=1}^{k-1} \langle v_k, u_i \rangle \, u_i \text{ and } e_k = \frac{u_k}{\|u_k\|}$$

  - The existence of an orthonormal basis can also be shown by induction over the dimension of the vectorial space.

Advanced Linear Algebra: Orthonormal Basis

- How about for infinite dimension ?
  - You need to use the axiom of choice to prove the existence of a basis for all vectorial spaces.
  - Some vector spaces can't have an inner product, so there is no possibility of a notion of orthogonality
    - Example: spaces defined on an finite field, or non-ordered field.
  - Every vector space that has a **Countable** basis, can have an orthonormal basis
    - A set is countable if either it is finite or there exists an injective function from it into the natural numbers; this means that each element in the set may be associated to a unique natural number.
    - The Gram-Schmidt construction can be applied !
  - A basis is countable if and only if the space is **Separable**.
    - A separable space is a space that contains a countable and dense subset.
  - Actually, you can show that every Hilbert space has an orthonormal basis ! The basis is not necessarily countable however.

Advanced Linear Algebra: Adjoint Operators

- We consider a linear operator T on a vector space V that has an inner product;

- The linear operator $T^\dagger$ on V called the adjoint of T, is constructed as follow: for u,v vectors of V:

$$\langle u\,, Tv\rangle \;=\; \langle T^\dagger u\,, v\rangle$$

  - $T^\dagger$ is a linear operator:

  - For T and S two linear operators: $(ST)^\dagger = T^\dagger S^\dagger$

  - The adjoint of the adjoint is the original operator: $(S^\dagger)^\dagger = S$

- If we apply adjoint formula to vectors of an orthonormal basis, we get:

$$\begin{aligned} \langle T^\dagger e_i, e_j\rangle &= \langle e_i, T e_j\rangle \\ \langle T^\dagger_{ki} e_k, e_j\rangle &= \langle e_i, T_{kj} e_k\rangle \\ (T^\dagger_{ki})^* \delta_{kj} &= T_{jk}\delta_{ik} \\ (T^\dagger)^*_{ji} &= T_{ij} \end{aligned}$$

or $(T^\dagger)_{ij} = (T_{ji})^*$

- Over an orthonormal basis, the adjoint matrix is the transpose and complex conjugate.

- This is a very useful operator and is typically different from T. When the adjoint $T^\dagger$ happens to be equal to T, the operator is said to be *self-adjoint or Hermitian*

Advanced Linear Algebra: self-adjoint Operators

- Self-adjoint (or Hermitian in finite dimension) operators are linear operators T for which. T = T†.

- One can show that: $T = T^\dagger$ if and only if $\langle v, Tv\rangle \in \mathbb{R}$ for all v

$$\langle v, Tv\rangle = \langle T^\dagger v, v\rangle \models \langle Tv, v\rangle = \langle v, Tv\rangle^* \qquad \langle v, Tv\rangle = \langle Tv, v\rangle = \langle v, T^\dagger v\rangle$$

So $\langle v, (T - T^\dagger)v\rangle = 0$

- One can show that this implies that T = T†.

- Two other very important results:
  - **The eigenvalues of Hermitian operators are real**: $Tv = \lambda v$

$$\langle v, Tv\rangle = \langle v, \lambda v\rangle = \lambda\langle v, v\rangle \quad \text{and} \quad \langle v, Tv\rangle = \langle Tv, v\rangle = \langle \lambda v, v\rangle = \lambda^* \langle v, v\rangle$$

  - Different eigenvalues of a Hermitian operator correspond to orthogonal eigenfunctions:

$$Tv_1 = \lambda_1 v_1, \qquad Tv_2 = \lambda_2 v_2 \qquad \text{with} \qquad \lambda_1 \neq \lambda_2$$

$$\langle v_2, Tv_1\rangle = \langle v_2, \lambda_1 v_1\rangle = \lambda_1\langle v_2, v_1\rangle \qquad \langle v_2, Tv_1\rangle = \langle Tv_2, v_1\rangle = \langle \lambda_2 v_2, v_1\rangle = \lambda_2\langle v_2, v_1\rangle$$

So we must have: $\langle v_1, v_2\rangle = 0$

Advanced Linear Algebra: Unitary Operators

- An operator U in a complex vector space V is said to be a unitary operator if it is surjective and does not change the magnitude of the vector it acts upon.

$$|Uu| = |u|\,, \ \text{for all } u \in V$$

- The definition is useful even for infinite dimensional spaces. Note that, nullU = 0, and U is injective. Since U is also assumed to be surjective, a unitary operator U is always invertible.

- A more common definition: $U^\dagger U \ = \ UU^\dagger \ = \ I$

$$\langle Uu, Uu\rangle = \langle u, u\rangle \quad \text{so} \quad \langle u, U^\dagger U\, u\rangle = \langle u, u\rangle \quad \rightarrow \quad \langle u\,,\, (U^\dagger U - I)u\rangle \ = \ 0 \ \ \text{for all} \ \ u$$

- Unitary operators preserve inner products in the following sense:

$$\langle Uu\,,\, Uv\rangle \ = \ \langle u\,,v\rangle$$

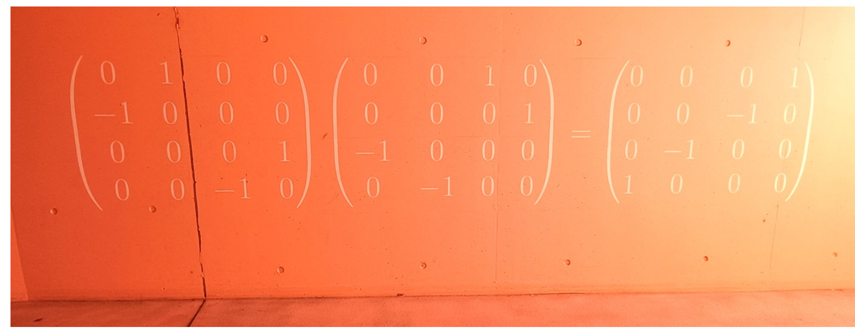
- It is actually an equivalent definition of a unitary operator:

A unitary operator is a bounded linear operator U : H → H on a Hilbert space H for which the following hold:

- U is surjective, and
- U preserves the inner product of the Hilbert space.

Advanced Linear Algebra: Spectral Theorem

- Hence for Hermitian operators, we can have a basis of orthonormal eigen vectors that form a basis, with physical, real, eigenvalues. An eigen states can be expressed in this basis!
- If A is Hermitian on V of finite dimension, then there exists an orthonormal basis of V consisting of eigenvectors of A. Each eigenvalue is real.



- Most observables in quantum mechanics will be Hermitian operators !
- Two commutative Hermitian operators can be diagonalized on a similar basis of eigen vectors.
- This also works with another kind of operator: Unitary operator

Linear Algebra: finite dimension vectorial spaces

- A vector space V is said to be finite dimensional if it is spanned by a finite list of vectors in V:
  - A basis of V is a list of vectors in V that both spans V and is linearly independent:
  - The dimension of a finite dimensional vector space V is the length of the shortest list of vectors that span V.
  - *There cannot be a list of n+1 linearly independent vectors in V (see anexe for proof).*
  - *Any list of linearly independent vectors of length n = dimV is a basis of V*

- A linear operator in V, and associated matrix, has the following equivalent properties:
  - The columns (lines) of the associated matrix are linearly independent;
  - The operator is injective;
  - The operator is surjective;
  - The matrix in invertible;
  - $\det(A) \neq 0$

- A matrix A is diagonalizable if it is similar to a diagonal matrix, i.e. there exist an invertible matrix P, and a diagonal matrix D, such that $P^{-1}AP = D$.
  - Equivalently, A is diagonalizable if there exist a basis of its eigen vectors.
  - The associated linear operator T is diagonalizable if there is a basis of the vectorial space V formed by the eigenvectors of T.
  - A matrix nxn with n distinct and non-zero eigenvalues is diagonalizable.
  - If the dimension of the sub-spaces of the eigen values of A (nxn) add up to n, then it is diagonalizable.

Advanced Linear Algebra: Adjoint Operators

- We consider a linear operator T on a vector space V that has an inner product, the linear operator $T^\dagger$ on V called the adjoint of T, is constructed as follow: for u,v vectors of V:

$$\langle u\,, Tv\rangle \;=\; \langle T^\dagger u\,, v\rangle$$ Which also implies that: $$(T^\dagger)_{ij} = (T_{ji})^*$$

- Self-adjoint (or Hermitian in finite dimension) operators are linear operators T for which $T = T^\dagger$.

- Important results::
    - **The eigenvalues of Hermitian operators are real.**
    - Different eigenvalues of a Hermitian operator correspond to **orthogonal eigenfunctions.**

- **Spectral theorem (finite dimension)**:

If A is Hermitian operator on a Hilbert space V of finite dimension, then there exists an orthonormal basis of V consisting of eigenvectors of A. Each eigenvalue is real.

- This is equivalent to say that A can be diagonalized;
- It is also equivalent to the fact that the sub-spaces of the eigenvalues of V are orthogonal, and the sum of their dimension is equal to dimV.
- The spectral theorem actually applies to Normal operators, defined as operators for which $[T, T^\dagger] = 0$. This includes self-adjoint and Unitary matrices.

Advanced Linear Algebra: Unitary Operators

- **Spectral theorem (infinite dimension)**:

In infinite dimension, the problem is more complex and the theorem holds only in certain conditions (that are almost always met in QM). It applies to certain types of operators:
  - Compact self-adjoint operators;
  - Bounded self-adjoint operators.

- Example: the set of square-integrable functions from I in $\mathbb{R}$ to $\mathbb{C}$ is a Hilbert space often defined with the inner product: $f, g \in H, \langle f, g \rangle = \int_0^L \overline{f(x)} g(x) dx.$

- An operator U in a complex vector space V is said to be **a unitary operator** if it is surjective and does not change the magnitude of the vector it acts upon: $|Uu| = |u|\,,\ \text{for all } u \in V$

- The definition is useful even for infinite dimensional spaces. Note that, nullU = 0, and U is injective. Since U is also assumed to be surjective, a unitary operator U is always invertible.

- A more common definition: $U^\dagger U = UU^\dagger = I$

$$\langle Uu, Uu \rangle = \langle u, u \rangle \quad \text{so} \quad \langle u, U^\dagger U\, u \rangle = \langle u, u \rangle \quad \rightarrow \quad \langle u\,, (U^\dagger U - I)u \rangle = 0 \quad \text{for all} \quad u$$

- Unitary operators preserve inner products in the following sense: $\langle Uu\,, Uv \rangle = \langle u\,, v \rangle$

- Normal operator is one for which $[T, T^\dagger] = 0, or\ T T^\dagger = T^\dagger \mathrm{T}$. It is immediate that Hermitian and Unitary operators are also normal.

Linear Algebra in Quantum Mechanics

- **Spectral decomposition**: in finite dimension, a self-adjoint operator can be diagonalized, hence possess a set of orthonormal eigenvectors that form a basis. If a_α are its eigenvalues, that can be degenerate, hence span a sub-space of dimension n_α and eigenvectors $|\alpha, r_\alpha\rangle$, one can write:

$$\hat{A} = \sum_{\alpha} \sum_{r_\alpha=1}^{n_\alpha} a_\alpha |\alpha, r_\alpha\rangle \langle \alpha, r_\alpha|$$

- This is based on the concept of outer product which is an operator $|\psi\rangle\langle\varphi|$. For an orthogonal basis, $\widehat{P_\alpha} = \sum_{r_\alpha=1}^{n_\alpha} |\alpha, r_\alpha\rangle\langle\alpha, r_\alpha|$ is a projector on the sub-space of a_α.

- For an object in state $|\psi\rangle$, the probability to find an eigen value a_α of an observable $\widehat{\boldsymbol{A}}$ is given by:

$$P(a_\alpha) = \langle\psi|\widehat{P_\alpha}|\psi\rangle = \sum_{r_\alpha=1}^{n_\alpha} |\langle\alpha, r_\alpha|\psi\rangle|^2$$

where n_α is the dimension of the sub-space generated by a_α, and the $|\alpha, r_\alpha\rangle$ the associated orthonormal eigenvectors.

- **Commuting observables**:
  - If two normal operators commute on a Hilbert space, there exists a basis of common eigenvectors.
  - This is quite powerful and is used for example in the quantum numbers of orbitals in the Hydrogen atom, or to prove the Bloch theorem.

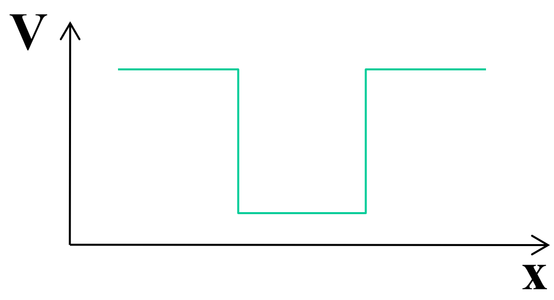
Common potentials in QM

- Free Particle: V = 0

$$\hat{H} = -\frac{\hbar^2}{2m}\vec{\nabla}^2 + V(\vec{r})$$

$$\Psi(r,t) = \frac{1}{V^{1/2}} e^{-i(\vec{k}.\vec{r}-\omega t)} \qquad \vec{p} = \hbar\vec{k}, \mathrm{E} = \hbar\omega \qquad \rho(\vec{r}) = |\Psi(\vec{r},t)|^2 = \frac{1}{V}$$

- Quantum well: quantization of energy states



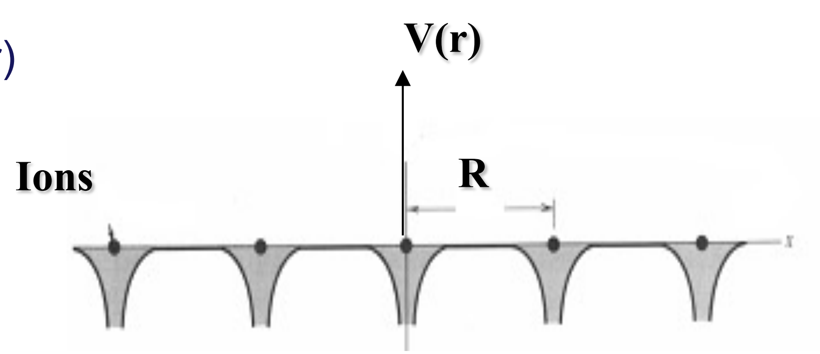
Infinite well: $$E_n = \frac{\hbar^2}{2m}\left(\frac{(n+1)\pi}{L}\right)^2$$

- In a crystal solid: Periodic Potential: V(r+R) = V(r)

- Bloch Theorem:

$$u_{k,n} = f_{k,n}(\vec{r})e^{i\vec{k}.\vec{r}}$$

$$f_{k,n}(\vec{r}+\vec{R}) = f_{k,n}(\vec{r})$$



Bloch Theorem

- **Bloch theorem**:

The solution of the Schrödinger equation in a periodic lattice takes the form of a plane wave modulated by a periodic function:

$$u_{k,n} = f_{k,n}(\vec{r})e^{i\vec{k}.\vec{r}}$$

$$f_{k,n}(\vec{r}+\vec{R}) = f_{k,n}(\vec{r})$$

- The translational symmetry of the problem reflects on the Hamiltonian:

$$\hat{H} = -\frac{\hbar^2}{2m}\vec{\nabla}^2 + V(\vec{r})$$ With $V(\vec{r}+\vec{a}) = V(\vec{r})$ for every vector $\vec{a}$ of the direct lattice.

- Noting $\widehat{T_{\boldsymbol{a}}}$ the translation operator, one sees immediately that the symmetry of the system implies that: $\left[\hat{H}, \widehat{T_{\boldsymbol{a}}}\right] = 0$

- $\hat{H}$ *and* $\widehat{T_{\boldsymbol{a}}}$ being normal operators, they can be diagonalized over a basis of common eigenvectors.

- Eigenvalues of $\widehat{T_{\boldsymbol{a}}}$ must be of norm 1, so they verify that the probability of finding the particle is periodic. Hence, in 1D, they are of the form e^{iqa} with $q \in \left]-\frac{\pi}{a}, \frac{\pi}{a}\right]$.
- For $\varphi(x)$ an eigen function of $\hat{H}$ *and* $\widehat{T_{\boldsymbol{a}}}$, $\varphi(x+a) = e^{iqa}\varphi(x)$

- One can write that $\varphi(x) = e^{iqx} \times e^{-iqx}\,\varphi(x)$. Calling $f(x) = e^{-iqx}\,\varphi(x)$, we see that:

$$f(x+a) = e^{-iq(x+a)}\,\varphi(x+a) = e^{-iqx}\,e^{-iqa}\,e^{iqa}\varphi(x) = e^{-iqx}\,\varphi(x) = f(x)$$

Electrons in a Periodic Potential

- Electrons in a periodic potential can be treated in two well-known approaches:

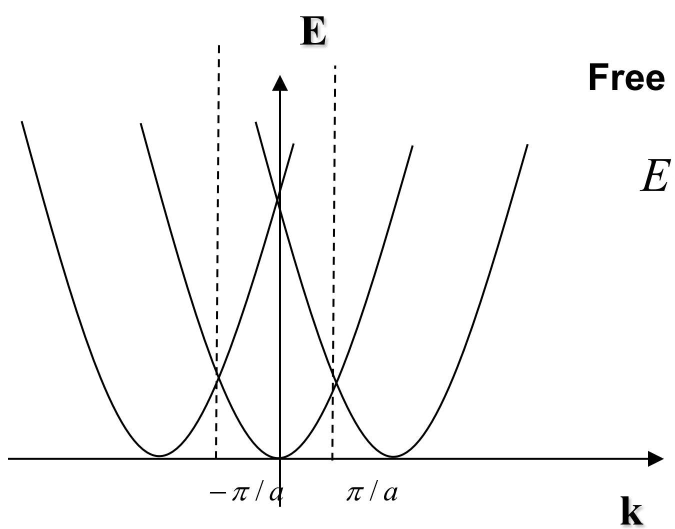
    - Tight bending: electrons strongly attached to their atom.

Wave functions are linear combination of atomic orbital with weak overlap from one site to the next.

Bands are formed as the atoms get closer together (ie overlap is increased)

- Nearly free electron model: a perturbation problem

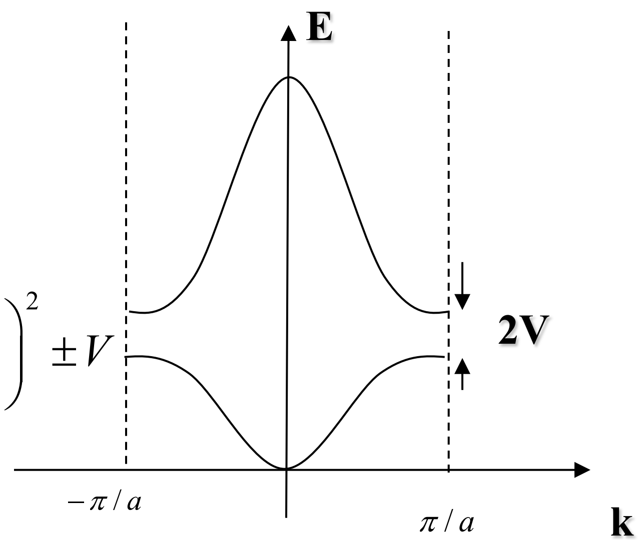
  The periodic potential is treated as a small perturbation, the effect of which is to split the degenerate energy level at the Brillouin zone edge.



Free Electron:

$$E(\vec{k}) = \frac{\hbar^2 k^2}{2m}$$

$$E_{\pm} = \frac{\hbar^2}{2m}\left(\frac{k}{2}\right)^2 \pm V$$



The Brillouin Zone

- From the Bloch theorem, we know that an electron in a periodic potential has a wave function of the form:

$$u_{k,n} = f_{k,n}(\vec{r})e^{i\vec{k}.\vec{r}}$$

$$f_{k,n}(\vec{r}+\vec{R}) = f_{k,n}(\vec{r})$$

- For a given k vector, several solutions could be possible, expressed by the label n.
- If we consider a vector **D** of the direct lattice, $u_{k,n}(\boldsymbol{r}+\boldsymbol{D}) = e^{i\boldsymbol{k}.\boldsymbol{D}}\, u_{k,n}(\boldsymbol{r})$. So the norm is unchanged by a translation along the crystal lattice, which is expected.
- For a large crystal of size ($N_1\mathbf{a_1}$, $N_2\mathbf{a_2}$, $N_3\mathbf{a_3}$) with ($\mathbf{a_1}$,$\mathbf{a_2}$, $\mathbf{a_3}$) being the primitive lattice basis, we can apply the periodic boundary conditions and obtain a quantification of the k number:

$$\boldsymbol{k} = \sum_{i=1}^{3} \frac{n_i}{N_i} \boldsymbol{a_i^*} \text{ , } n_i \in \mathbb{Z}$$

- For large N, the k states are very close together and form quasi-continuum of states.
- If **K** is in the reciprocal lattice, for **k** = $\mathbf{k_0}$ + **K**, we have:

$u_{k,n}(\boldsymbol{r}) = e^{i\boldsymbol{k}.\boldsymbol{r}}\, f_{k,n}(\boldsymbol{r}) = e^{i\boldsymbol{k_0}.\boldsymbol{r}}\, e^{i\boldsymbol{K}.\boldsymbol{r}}\, f_{k,n}(\boldsymbol{r})$. But the function $e^{i\boldsymbol{K}.\boldsymbol{r}}\, f_{k,n}(\boldsymbol{r})$ verifies:

$e^{i\boldsymbol{K}.(\boldsymbol{r}+\boldsymbol{D})}\, f_{k,n}(\boldsymbol{r}+\boldsymbol{D}) = e^{i\boldsymbol{K}.\boldsymbol{r}}\, f_{k,n}(\boldsymbol{r})$ by definition of the reciprocal lattice.

So $u_{k,n}(\boldsymbol{r})$ and $u_{k_0,n}(\boldsymbol{r})$ represents a similar solution with the same energy.

Quantum states are hence fully defined within an elementary cell of the reciprocal space called the Brillouin zone.

Splitting of energy levels at the Brillouin Zone

- Before applying a small periodic potential, the electrons are free in the crystal and are represented by plane waves:

$$\psi_k(r) = \frac{1}{\sqrt{\Omega}} e^{i\mathbf{k}.r} \text{ and } \quad E(\vec{k}) = \frac{\hbar^2 k^2}{2m}$$

- In 1D, the Brillouin zone is $\left]-\frac{\pi}{a}, \frac{\pi}{a}\right]$, and the energy level for $k_{\pm} = \pm\frac{\pi}{a}$ $(\mathcal{E}_0 = \frac{\hbar^2\pi^2}{2ma^2})$, is

degenerated with two states $\left|+\frac{\pi}{a}\right\rangle$ and $\left|-\frac{\pi}{a}\right\rangle$.

- A small periodic potential will create a perturbation that will lift the degeneracy.
- Since V is periodic, we can develop it as a Fourier series and to a first approximation, consider only the first harmonics:

$$V(x) = \nu(e^{i\frac{2\pi x}{a}} + e^{-i\frac{2\pi x}{a}}) = 2\nu\times\cos(\frac{2\pi x}{a}).$$

- This perturbation adds to the Hamiltonian at the first order, with a matrix in the basis of the degenerated states:

$$\hat{H} + \hat{V} = \begin{pmatrix} \mathcal{E}_0 & \nu \\ \nu & \mathcal{E}_0 \end{pmatrix}$$

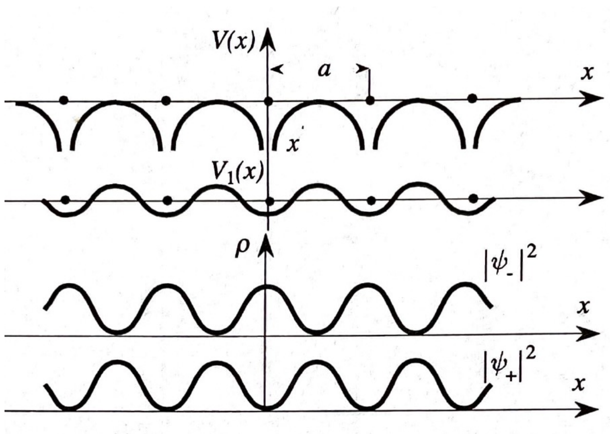
Splitting of energy levels at the Brillouin Zone

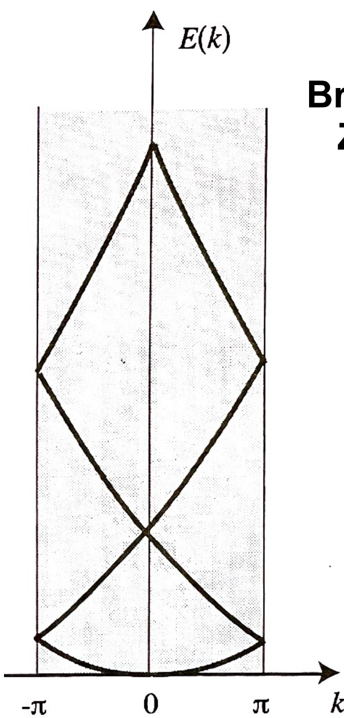
- Two new eigenstates emerge: $|\psi_{\pm}\rangle = \frac{1}{\sqrt{2}}\left(\left|+\frac{\pi}{a}\right\rangle \pm \left|-\frac{\pi}{a}\right\rangle\right)$

  with energies:

$$\mathcal{E}_{+} = \mathcal{E}_0 + |\nu| \text{ and } \mathcal{E}_{-} = \mathcal{E}_0 - |\nu|$$

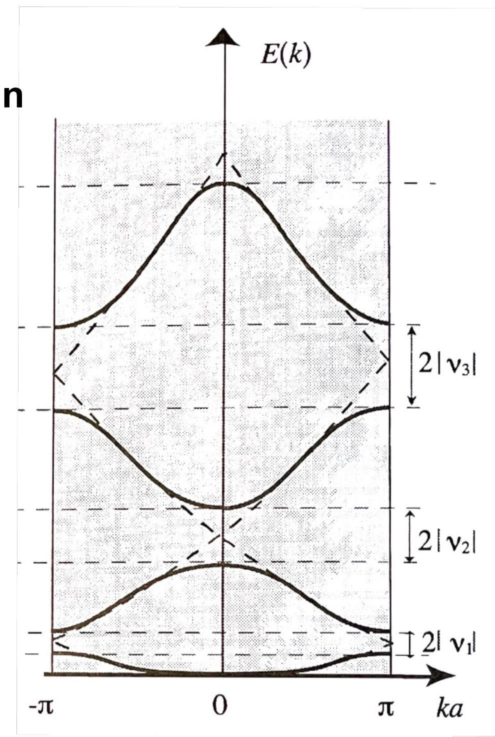
The energy level $\mathcal{E}_0$ is then split, and the region of energy $\mathcal{E}_0 - |\nu| < \mathcal{E} < \mathcal{E}_0 + |\nu|$ has no eigenstate: it is a gap of energy for which no stationary solution is found.



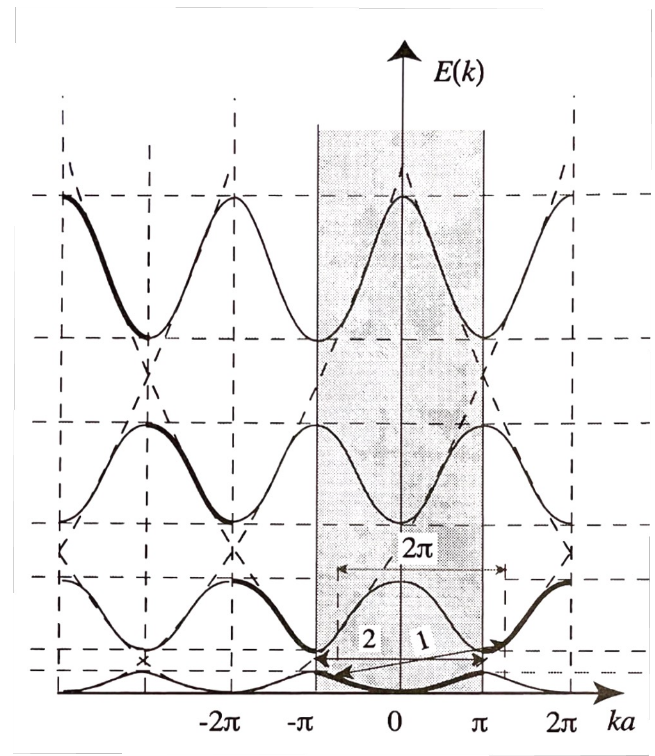


Free Electrons

Brillouin Zone



Periodic potential



SUMMARY

- We reviewed matrix manipulation concepts and how they can apply to a change of coordinates.
- We briefly reviewed wave mechanics concepts and found the linear algebra formalism hidden in their expressions
- We reminded the postulates of Quantum Mechanics
- We reviewed important linear algebra concepts;
- We introduced new concepts such as Hilbert space, self-adjoint operators, Hermitian and unitary operators
- We started to apply these notions to the understanding of fundamental concepts in the quantum treatment of materials properties

Annexe

- If V is a vector space with dimV=n, there cannot be a list of n+1 linearly independent vectors in V.
- Proof by induction (from notes of Prof. Isidora Milin @ Harvard University)

Base Case. When $n = 1$, let $\{v\}$ be a basis for V, and $v_1, v_2 \in V$ be any two vectors in V. Then, $\exists a_1, a_2 \in F$ such that $v_1 = a_1 v$ and $v_2 = a_2 v$. If $a_1 a_2 = 0$, then either $v_1 = \vec{0}$ or $v_2 = \vec{0}$, which, in both cases, makes $\{v_1, v_2\}$ into a linearly dependent set. Otherwise, we can write $\frac{1}{a_1} v_1 - \frac{1}{a_2} v_2 = \vec{0}$, so that $\{v_1, v_2\}$ is again linearly dependent.

Inductive Step. Suppose that in any $n-1$-dimensional vector space any n vectors are linearly dependent. We want to show this implies that any $n+1$ vectors $v_1, v_2, \ldots, v_n$ in any n-dimensional vector space V are linearly dependant.

Let $e_1, \ldots, e_n$ be a basis for V. Then, all vectors in V are expressible as linear combinations of the e_i's, so we can find scalars $a_{i,j}$, where $1 \leq i \leq n+1$, $1 \leq j \leq n$ such that:

$$v_1 = a_{1,1} e_1 + a_{1,2} e_2 + \ldots + a_{1,n} e_n$$

$$v_2 = a_{2,1} e_1 + a_{2,2} e_2 + \ldots + a_{2,n} e_n$$

$$\ldots$$

$$v_n = a_{n,1} e_1 + a_{n,2} e_2 + \ldots + a_{n,n} e_n$$

$$v_{n+1} = a_{n+1,1} e_1 + a_{n+1,2} e_2 + \ldots + a_{n+1,n} e_n$$

Now, consider the scalars "in the first column" i.e. $a_{1,1}, a_{2,1}, \ldots, a_{n,1}, a_{n+1,1}$. If all of them are 0, then we have that $v_1, v_2, \ldots, v_n, v_{n+1} \in V' = span(e_2, \ldots e_n)$. Now, $dim(V') = n - 1$ and $v_1, \ldots v_n \in V'$, so by the inductive hypothesis, $v_1 \ldots v_n$ are linearly dependent, which implies $v_1, \ldots, v_n, v_{n+1}$ are linearly dependent as well.

So we now assume that $\exists a_{i,1} \neq 0$. WLOG (i.e. by reordering the v_i's), assume $a_{1,1} \neq 0$. Now consider the following n vectors:

$$v_2' = v_2 - \frac{a_{2,1}}{a_{1,1}} v_1 = a_{2,2}' e_2 + \ldots + a_{2,n}' e_n$$

$$\ldots$$

$$v_n' = v_n - \frac{a_{n,1}}{a_{1,1}} v_1 = a_{n,2}' e_2 + \ldots + a_{n,n}' e_n$$

$$v_{n+1}' = v_{n+1} - \frac{a_{n+1,1}}{a_{1,1}} v_1 = a_{n+1,2}' e_2 + \ldots + a_{n+1,n}' e_n$$

Note that $v_2', \ldots, v_n', v_{n+1}'$ are n vectors in the $n-1$-dimensional vector space $V' = span(e_2, \ldots, e_n)$, so that by the inductive hypothesis, they are linearly dependent. That is, $\exists \lambda_2, \ldots, \lambda_{n+1} \in F$, not all zero, such that

$$\lambda_2 v_2' + \ldots + \lambda_n v_n' + \lambda_{n+1} v_{n+1}' = \vec{0}$$

This implies that:

$$\lambda_2 (v_2 - \frac{a_{2,1}}{a_{1,1}} v_1) + \ldots + \lambda_n (v_n - \frac{a_{n,1}}{a_{1,1}} v_1) + \lambda_{n+1} (v_{n+1} - \frac{a_{n+1,1}}{a_{1,1}} v_1) = \vec{0}$$

and we finnally get:

$$-(\lambda_2 \frac{a_{2,1}}{a_{1,1}} + \ldots + \lambda_n \frac{a_{n,1}}{a_{1,1}} + \lambda_{n+1} \frac{a_{n+1,1}}{a_{1,1}}) v_1 + \lambda_2 v_2 + \ldots + \lambda_n v_n + \lambda_{n+1} v_{n+1} = \vec{0}$$

Thus, $v_1, v_2, \ldots, v_n, v_{n+1}$ are linearly dependent, which ends the proof of the inductive step. $\square$