Vector Spaces
MANUSCRIPT

William F. Barnes
[1]

July 21, 2024

Contents

Chapter 1

Vector Spaces

1 Foundations

1.1 Ket Notation

After a tour through calculus, vectors, and their holy marriage in vector calculus, one is well aware that a vector $\vec{a}$ is defined as a list of numbers or variables. We begin by replacing the 'arrow' notation with *ket* notation popularized by Paul Dirac:

$$\vec{a} = |a\rangle$$

To write out the vector explicitly, list the components inside the $|\rangle$ symbol:

$$\vec{a} = \langle a_1, a_2, a_3\rangle = |a_1, a_2, a_3\rangle$$

Of course, the number of components need not be three, and the coordinate system implied need not be Cartesian.

1.2 Complex Components

All vector components are assumed to be complex numbers unless restricted by circumstance. A complex number z has two components, real and imaginary, such that

$$z = \alpha + i\beta \, ,$$

where α, β are real numbers, and i is the imaginary unit:

$$i = \sqrt{-1}$$

The same complex number z can be expressed in polar form

$$z = r\, e^{i\phi} \, ,$$

where

$$r = |z| = \sqrt{\alpha^2 + \beta^2}$$

is the magnitude, and

$$\phi = \arctan\left(\beta/\alpha\right)$$

is the complex phase of z. Every complex number z a complex conjugate $\bar{z} = z^*$ that inverts the imaginary component:

$$\bar{z} = z^* = \alpha - i\beta = r\, e^{-i\phi}$$

Complex numbers obey special operations for addition, multiplication, and division. For two complex numbers z_j with $j = 1, 2$, we have

$$\begin{aligned} z_1 \pm z_2 &= (\alpha_1 + \alpha_2) \pm i\,(\beta_1 + \beta_2) \\ z_1 \cdot z_2 &= (\alpha_1\alpha_2 - \beta_1\beta_2) + i\,(\alpha_1\beta_2 + \alpha_2\beta_1) \\ z_1/z_2 &= z_1 \cdot z_2^* / \left|z_2\right|^2 \end{aligned}$$

1.3 Sets

A *set* is generally defined as a collection of distinct, well-defined objects. Perhaps the most common set is the real numbers, denoted $\mathbb{R}$. Distinguishing the set of integers $\mathbb{Z}$ from the irrational numbers $\mathbb{Q}'$, we can relate each set using the *union* operator:

$$\mathbb{R} = \mathbb{Z} \cup \mathbb{Q}'$$

An individual member of a set is called an *element*. For example, the set $\mathbb{C}$ of complex numbers is comprised of all elements z. This is formally denoted using the *in* symbol $\in$ as

$$z \in \mathbb{C} \, .$$

Mathematical statements can be shortened further by introducing the *for all* symbol $\forall$, along with the *there exists* symbol $\exists$. For instance, the idea that 'for all complex numbers z there exits a complex conjugate z^*' can be written as:

$$\forall z \in \mathbb{C} \, , \, \exists\, z^* \in \mathbb{C}$$

1.4 Spaces

A *space* is a set with some kind of ordered structure. For instance, the space of all ordered pairs of real numbers, i.e., all two-dimensional vectors with real components, is denoted $\mathbb{R}_2$.

For a less trivial example, we may define a space $\mathbb{L}_2\,[a,b]$ of all functions $\{f\,(x)\}$ obeying

$$\int_a^b \left|f\,(x)\right|^2\, dx < \infty \, .$$

2 Vector Space

A *vector space*, for a given vector $|A\rangle$, contains the set of all allowed vectors that $|A\rangle$ could have been. Formally, we say that a vector space $\mathcal{V}$ is comprised of complex elements $\{|a\rangle\}$ that obeys the *vector space axioms*.

2.1 Vector Space Axioms

In the axioms that follow, consider any three vectors $|a\rangle$, $|b\rangle$, $|c\rangle$ in the vector space $\mathcal{V}$. Let α, β be two nonzero complex scalars.

Addition

Vectors still obey the familiar rules for addition. Embedded in the definition are the notions of commutativity and associativity:

$$|a\rangle + |b\rangle = |b\rangle + |a\rangle$$
$$|a\rangle + (|b\rangle + |c\rangle) = (|a\rangle + |b\rangle) + |c\rangle$$

Scalar Multiplication

Multiplying a vector by a complex number results in a new vector that is parallel to the original. In particular, this means:

$$\forall\,|a\rangle \in \mathcal{V}\,,\ \forall\alpha \in \mathcal{C}\ :\ \exists\,\alpha\,|a\rangle \in \mathcal{V}$$

As expected, scalar multiplication follows the rules of commutativity and associativity:

$$\alpha\,(\beta\,|a\rangle) = (\alpha\beta)\,|a\rangle$$
$$\alpha\,(|a\rangle + |b\rangle) = \alpha\,|a\rangle + \alpha\,|b\rangle$$
$$(\alpha + \beta)\,|a\rangle = \alpha\,|a\rangle + \beta\,|a\rangle$$

Zero Vector

There exists a *zero vector* in the vector space that does not contribute to any sum. In the language of symbols, this precisely means

$$\exists\,|0\rangle \in \mathcal{V}\ :\ \forall\,|a\rangle \in \mathcal{V}\,,$$

or in practice, for addition:

$$|a\rangle + |0\rangle = |a\rangle$$

The zero vector plays an expected role in scalar multiplication:

$$0\,|a\rangle = |0\rangle$$

Additive Inverse

Every vector has a 'negative' version of itself called the *additive inverse*. That is:

$$\forall\,|a\rangle \in \mathcal{V}\ :\ \exists\,|-a\rangle \in \mathcal{V}\,,$$
$$|a\rangle + |-a\rangle = |0\rangle$$

2.2 Uniqueness

Uniqueness of Zero Vector

The first non-axiomatic issue to address is whether there exist multiple zero vectors in a given vector space. To capture this concern, take two vectors $|a\rangle$, $|b\rangle$ and add a unique zero vector to each:

$$|a\rangle + |0\rangle = |a\rangle$$
$$|b\rangle + |0\rangle' = |b\rangle$$

Using the shorthand $|a\rangle + |b\rangle = |c\rangle$, add the two equations to get

$$|c\rangle + |0\rangle + |0\rangle' = |c\rangle\ .$$

As it appears, the combination $|0\rangle + |0\rangle'$ can only be the zero vector itself:

$$|0\rangle + |0\rangle' = |0\rangle$$

Evidently, $|0\rangle'$ plays an indistinguishable role from $|0\rangle$. In conclusion, we find there exists exactly one (abbreviated $\exists!$) zero vector per vector space:

$$\exists!\,|0\rangle \in \mathcal{V}$$

Uniqueness of Additive Inverse

In a similar spirit, we can show that the additive inverse of a vector is unique. Take two copies of a vector $|a\rangle$ and add a unique additive inverse vector to each:

$$|a\rangle + |-a\rangle = |0\rangle$$
$$|a\rangle + |-a\rangle' = |0\rangle$$

Adding the two equations, we have

$$2\,|a\rangle + |-a\rangle + |-a\rangle' = |0\rangle\ ,$$

which is only true if

$$|-a\rangle + |-a\rangle' = 2\,|-a\rangle$$
$$|-a\rangle' = |-a\rangle\ ,$$

telling us the additive inverse is unique. Our declaration of the additive inverse becomes more specific:

$$\forall\,|a\rangle \in \mathcal{V}\ :\ \exists!\,|-a\rangle \in \mathcal{V}$$

Subtraction

The notion of subtraction can be formally introduced after establishing uniqueness of the additive inverse:

$$|a\rangle - |b\rangle = |a\rangle + |-b\rangle$$

2.3 Identities

Multiplication by One

From the above axioms, it's easy to show that the only scalar that multiplies into a vector to return the same vector is one:

$$1\,|a\rangle = |a\rangle$$

Multiplication by Zero

Similarly we can can ask which scalar multiplies into a vector to return the zero vector:

$$\alpha\,|a\rangle = |0\rangle$$

The answer is obviously zero, but can we prove it? To do so, take the above statement as a proposition, and then let $\alpha \to -\alpha$:

$$-\alpha\,|a\rangle = |0\rangle$$

If both statements are to be true, then it can only be that $\alpha = -\alpha$, which is only satisfied by $\alpha = 0$.

Condensed Notation

For a vector $|a\rangle$ and a scalar α, scalar multiplication is often represented as:

$$\alpha\,|a\rangle = |\alpha a\rangle$$

Likewise, the addition of two vectors $|a\rangle$, $|b\rangle$ is written:

$$|a\rangle + |b\rangle = |a+b\rangle$$

2.4 Applications

Cartesian Plane

The set of vectors based at the origin in the Cartesian plane qualifies as a vector space, provided that the 'usual' rules (from trigonometry) of vector addition and scalar multiplication are allowed.

Real n-tuples

Another vector space is the set of real-valued n-tuples

$$a = (a_1, a_2, \cdots, a_n)$$

obeying the addition rule

$$a + b = (a_1 + b_1, a_2 + b_2, \cdots, a_n + b_n)$$

and the multiplication rule

$$\lambda a = (\lambda a_1, \lambda a_2, \cdots, \lambda a_n)\ .$$

3 Inner Product

3.1 Bra Notation

There is another way to write vectors as they pertain to vector spaces using the so-called *bra notation*:

$$\langle a|$$

The so-called bra-vector is related to $\vec{a} = |a\rangle$, but is not identical.

Particularly, the bra-vector is called the *dual vector* of $|a\rangle$, also called a *linear functional*. The vector space occupied by $\langle a|$ is called the *dual space* to that occupied by $|a\rangle$. On their own, bra-vectors obey the same axioms as ket-vectors.

3.2 Inner Product

As an operator, a bra-vector $\langle b|$ can 'act on' a ket-vector $|a\rangle$ to produce a scalar:

$$\langle b|a\rangle = \alpha$$

Assuming $|a\rangle$, $|b\rangle$ are of the same vector space, the quantity $\langle b|a\rangle$ is called the *inner product* of vectors $|b\rangle$ and $|a\rangle$.

Conjugate

As an axiom, let us use up some available freedom to require that swapping $a \leftrightarrow b$ results in the complex conjugate of the original product:

$$\langle b|a\rangle = (\langle a|b\rangle)^* = \overline{\langle a|b\rangle}$$

Negative Product

It's possible for the inner product to yield a negative result. For two nonzero vectors $|a\rangle$, $|b\rangle$ satisfying

$$\langle b|a\rangle = \alpha\ ,$$

calculate the inner product $\langle b|-a\rangle$ to find:

$$\begin{aligned}\langle b|-a\rangle &= \langle b|0-a\rangle \\ &= \cancel{\langle b|0\rangle} - \langle b|a\rangle \\ &= -\alpha\end{aligned}$$

3.3 Linearity

The inner product obeys linearity rules much as an ordinary operator would. For the vectors $|a\rangle$, $|u\rangle$, $|v\rangle$, along with scalars α, β, we first have:

$$\langle a|\alpha u + \beta v\rangle = \alpha\,\langle a|u\rangle + \beta\,\langle a|v\rangle$$

Furthermore:

$$\begin{aligned}\langle \alpha u + \beta v | a\rangle &= \langle \alpha u | a\rangle + \langle \beta v | a\rangle \\ &= \overline{\langle a|\alpha u\rangle} + \overline{\langle a|\beta v\rangle} \\ &= \alpha^* \overline{\langle a|u\rangle} + \beta^* \overline{\langle a|v\rangle}\end{aligned}$$

Comparing each result, we have:

$$\overline{\langle a|\alpha u + \beta v\rangle} = \langle \alpha u + \beta v | a\rangle$$

3.4 Norm

By calculating $\langle a|a\rangle = \overline{\langle a|a\rangle}$, we readily find $\alpha = \alpha^* = \overline{\alpha}$, telling us the self-inner product always yields a real number:

$$\langle a|a\rangle \in \mathcal{R}$$

The square root the self-inner product is called the *norm* of the vector, axiomatically a positive number:

$$\|a\| = \sqrt{\langle a|a\rangle} > 0$$

Zero Norm

It immediately follows that the norm of any non-zero vector cannot itself be zero, with the only excepting being the zero vector:

$$\langle a|a\rangle = 0 \iff |a\rangle = |0\rangle$$

3.5 Applications

Complex Vector Space

Consider the vector space $\mathcal{C}_n$ whose elements are vectors containing n individual complex numbers, i.e.

$$\exists\, |x\rangle \in \mathcal{C}_n \;:\; |x\rangle = |x_1, x_2, ..., x_n\rangle$$

For two vectors $|a\rangle$ and $|b\rangle$ in $\mathcal{C}_n$, the inner product can be defined as

$$\langle a|b\rangle = \sum_{j=1}^{n} a_j^* b_j \ ,$$

or more generally, the definition can include weighting coefficients

$$\langle a|b\rangle = \sum_{j=1}^{n} a_j^* b_j w_j$$

for $w_j > 0 \in \mathcal{R}$.

Complex Function Space

By analogy to the inner product for vectors, a similar equation can be written for two complex functions $f(z)$, $g(z)$ defined in the interval $z \in [a, b]$ as

$$\langle f|g\rangle = \int_a^b f^*(z)\, g(z)\, dz \ .$$

Of course, the above can be generalized with a weighting function $w(z) > 0 \in \mathcal{R}$ such that

$$\langle f|g\rangle = \int_a^b f^*(z)\, g(z)\, w(z)\, dz \ .$$

4 Linear Combinations

Consider a vector space $\mathcal{V}$ admitting a set of n vectors $\{|\phi_j\rangle\}$ with $j = 1, 2, \ldots, n$. Introducing a set of n complex coefficients $\{c_j\}$, we construct a *linear combination*:

$$|a\rangle = \sum_{j=1}^{n} c_j\, |\phi_j\rangle$$

4.1 Span

The linear combination vector $|a\rangle$, along with all other linear combinations of $\{|\phi_j\rangle\}$, occupy a subspace $\mathcal{V}' \in \mathcal{V}$. In tighter terms, we say the vectors $\{|\phi_j\rangle\}$ *span* the vector space $\mathcal{V}'$.

4.2 Basis

If it turns out that $\mathcal{V}' = \mathcal{V}$, any vector allowed in $\mathcal{V}$ can be expressed as some linear combination of its elements. In this case, vectors $\{|\phi_j\rangle\}$ are called a *basis*, and the number n is a positive non-infinite integer called the *dimension* of the space.

4.3 Linear Independence

While the notion of 'span' makes sure there are 'not too few' basis vectors, we introduce *linear independence* to assure there aren't too many. That is, any basis vector $|\phi_k\rangle$ that can be expressed as a linear combination is *not* really a basis vector, and the dimension of the space may shrink by one.

Equivalently, we may argue that a set of linearly independent basis vectors only satisfies

$$\sum_{j=1}^{n} c_j\, |\phi_j\rangle = |0\rangle$$

when all coefficients $c_j = 0$. To show this we choose any two nonzero c_k and $c_{k'}$ (with the rest zero), reducing the above to

$$c_k |\phi_k\rangle = -c_{k'} |\phi_{k'}\rangle \ .$$

Clearly, the vector $|\phi_{k'}\rangle$ is not independent from $|\phi_k\rangle$ and either can be excluded from the basis.

4.4 Uniqueness of Coefficients

We can show that the coefficients c_j are unique for a given linear combination. Supposing we have a resultant vector $|a\rangle$ that that is 'arrived at' by two different sets of coefficients

$$|a\rangle = \sum_{j=1}^{n} c_j |\phi_j\rangle$$
$$|a\rangle = \sum_{j=1}^{n} c'_j |\phi_j\rangle \ .$$

Adding each equation and dividing by 2, we quickly find

$$|a\rangle = \sum_{j=1}^{n} \left(\frac{c_j + c'_j}{2} \right) |\phi_j\rangle \ ,$$

which only holds if every c_j is equal to c'_j.

5 Orthonormal Basis

5.1 Orthogonal Vectors

Two vectors $|\phi_j\rangle$, $|\phi_k\rangle$ are *orthogonal vectors* if their inner product is zero:

$$\langle \phi_j | \phi_k \rangle = 0$$

If *all* basis vectors $\{|\phi_j\rangle\}$ are mutually orthogonal, they constitute an *orthogonal basis*.

5.2 Normalized Basis

A basis vector is *normalized* if its self-inner product resolves to one:

$$\langle \phi_j | \phi_j \rangle = 1 \ ,$$

in which case the change of notation

$$|\phi_j\rangle \ \rightarrow \ |e_j\rangle$$

is made. Of course, one can always normalize each vector in an orthogonal basis by dividing out the norm:

$$|e_j\rangle = \frac{1}{\sqrt{\langle \phi_j | \phi_j \rangle}} |\phi_j\rangle$$

Orthonormal Basis

If all basis vectors are mutually orthogonal and have a norm of one, the set $\{|e_j\rangle\}$ is called an *orthonormal basis*. We summarize this by writing

$$\langle e_j | e_k \rangle = \delta_{jk} \ ,$$

where δ_{jk} is the Kronecker delta symbol:

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

5.3 Vector Components

Equipped with the notion of the orthonormal basis, let us reconsider the linear combination

$$|a\rangle = \sum_{j=1}^{n} a_j |e_j\rangle \ ,$$

and solve for the coefficients a_j.

Using what's sometimes called Fourier's trick, notice that projecting any bra-vector $\langle e_k|$ will trigger one inner product on the left, and n inner products on the right. However $n-1$ of these will be *zero*, and this plucks out the kth coefficient from the sum:

$$\langle e_k | a \rangle = \sum_{j=1}^{n} a_j \langle e_k | e_j \rangle = a_j \, \delta_{kj} = a_k$$

Evidently, any coefficient c_j can be reverse-engineered from a linear combination by the relation

$$a_j = \langle e_j | a \rangle \ .$$

The coefficients a_j are synonymous with the *components* of a vector. In pure bra-ket notation, a linear combination reads

$$|a\rangle = \sum_{j=1}^{n} \langle e_j | a \rangle |e_j\rangle \ ,$$

reminding us that the components of a vector are strictly related to the choice of basis.

5.4 Isomorphism

Consider a vector space $\mathcal{V}$ admitting a basis $\{|e_j\rangle\}$. A linear combination vector $|a\rangle$, in component form, can be written

$$|a\rangle = |a_1, a_2, \ldots, , a_n\rangle \ .$$

On the right side, we see that the (complex) components form an n-dimensional space of their own, namely $\mathcal{C}_n$.

To capture the 'one-to-oneness' between the original vector space and that occupied by its components, we say the n-dimensional inner product space is *isomorphic* with $\mathcal{C}_n$, or

$$\mathcal{V}_{(n)} \cong \mathcal{C}_n \; .$$

5.5 Gram-Schmidt Procedure

An arbitrary basis $\{|\phi_j\rangle\}$ can always be transformed into an orthonormal basis by the *Gram-Schmidt procedure*.

Denote $\{|e_j\rangle\}$ as the desired set of unit-normalized orthogonal vectors, and $\left\{|e'_j\rangle\right\}$ as a non-normalized version (a notational convenience). Starting with the $j=1$ vector, we write the easy result

$$\begin{aligned} |e'_1\rangle &= |\phi_1\rangle \\ |e_1\rangle &= |e'_1\rangle / \sqrt{\langle e'_1|e'_1\rangle} \; . \end{aligned}$$

Next, we need a new vector $|e'_2\rangle$ that involves $|\phi_2\rangle$ and is orthogonal to $|e_1\rangle$. This is achieved by writing

$$\begin{aligned} |e'_2\rangle &= |\phi_2\rangle - \langle e_1|\phi_2\rangle |e_1\rangle \\ |e_2\rangle &= |e'_2\rangle / \sqrt{\langle e'_2|e'_2\rangle} \; . \end{aligned}$$

Continuing for $j=3$, we need a vector $|e'_3\rangle$ that involves $|\phi_3\rangle$ and is orthogonal to $|e_1\rangle$, $|e_2\rangle$, satisfied by

$$|e'_3\rangle = |\phi_3\rangle - \langle e_1|\phi_3\rangle |e_1\rangle - \langle e_2|\phi_3\rangle |e_2\rangle \; ,$$

subject to the same normalization rule.

In the general $j=n$ case, this pattern extends to

$$\begin{aligned} |e'_n\rangle = |\phi_n\rangle &- \langle e_1|\phi_n\rangle |e_1\rangle \\ &- \langle e_2|\phi_n\rangle |e_2\rangle - \cdots - \langle e_{n-1}|\phi_n\rangle |e_{n-1}\rangle \; , \end{aligned}$$

normalized by

$$|e_n\rangle = \frac{1}{\sqrt{\langle e'_n|e'_n\rangle}} |e'_n\rangle \; .$$

Arbitrary Basis

Let us use the Gram-Schmidt procedure to produce an orthonormal basis from:

$$|\phi_1\rangle = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} , \; |\phi_2\rangle = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} , \; |\phi_3\rangle = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

Starting with $|e_1\rangle$, we have

$$|e_1\rangle = \frac{1}{\sqrt{\langle\phi_1|\phi_1\rangle}} |\phi_1\rangle = \frac{1}{\sqrt{2}} |\phi_1\rangle = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix} \; .$$

Next, $|e'_2\rangle$ is given by:

$$|e'_2\rangle = |\phi_2\rangle - \langle e_1|\phi_2\rangle |e_1\rangle = \begin{bmatrix} -1/2 \\ 1/2 \\ 1 \end{bmatrix}$$

$$|e_2\rangle = \frac{1}{\sqrt{6}} \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix}$$

Finally, for $|e'_3\rangle$, we have

$$|e'_3\rangle = |\phi_3\rangle - \langle e_1|\phi_3\rangle |e_1\rangle - \langle e_2|\phi_3\rangle |e_2\rangle = \begin{bmatrix} 2/3 \\ -2/3 \\ 2/3 \end{bmatrix} ,$$

normalizing to

$$|e_3\rangle = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \; .$$

Legendre Polynomials

Consider the set of real-valued polynomial functions of order no greater than four

$$P_4(x) = c_0 + c_1 x + c_2 x^2 + c_3 x^3 + c_4 x^4 \; ,$$

known as *Legendre polynomials*. Confining the x-domain to the window $-1 \leq x \leq 1$, we may introduce the inner product of two such polynomials $f(x)$, $g(x)$ as

$$\langle f|g\rangle = \int_{-1}^{1} f(x)\, g(x)\, dx \; .$$

Given $P_4(x)$, there are five vectors $|\phi_j\rangle = x^j$ with $j = 0, 1, 2, 3, 4$ form the basis of a five-dimensional vector space.

By the Gram-Schmidt procedure, we can normalize the basis $\{|\phi_j\rangle\}$ starting with

$$|e_0\rangle = \frac{1}{\sqrt{\int_{-1}^{1} dx}} |\phi_0\rangle = \frac{1}{\sqrt{2}} |\phi_0\rangle = \frac{1}{\sqrt{2}} \; .$$

Proceeding for $|e_1\rangle$, we have

$$\begin{aligned} |e'_1\rangle &= |\phi_1\rangle - \langle e_0|\phi_1\rangle |e_0\rangle \\ |e_1\rangle &= \frac{1}{\sqrt{\langle e'_1|e'_1\rangle}} |e'_1\rangle \; , \end{aligned}$$

reducing to

$$\begin{aligned} |e'_1\rangle &= |\phi_1\rangle - \cancel{\langle e_0|\phi_1\rangle |e_0\rangle} \\ |e_1\rangle &= \sqrt{\frac{3}{2}} |\phi_1\rangle = \sqrt{\frac{3}{2}}\, x \; . \end{aligned}$$

Continuing for $|e_2\rangle$, begin with

$$\begin{aligned} |e_2'\rangle &= |\phi_2\rangle - \langle e_0|\phi_2\rangle |e_0\rangle \\ &\quad - \cancel{\langle e_1||\phi_2\rangle |e_1\rangle} = |\phi_2\rangle - \frac{\sqrt{2}}{3} |e_0\rangle \ , \end{aligned}$$

where normalization requires calculating

$$\langle e_2'|e_2'\rangle = \int_{-1}^{1} \left(x^2 - \frac{1}{3}\right)^2 dx = \frac{8}{45} \ ,$$

landing us at

$$|e_2\rangle = \sqrt{\frac{5}{8}} \left(3x^2 - 1\right) \ .$$

Turning the same crank, it's straightforwardly shown that the remaining normalized basis vectors resolve to

$$\begin{aligned} |e_3\rangle &= \sqrt{\frac{7}{8}} \left(5x^3 - 3x\right) \\ |e_4\rangle &= \frac{3}{8\sqrt{2}} \left(35x^4 - 30x^2 + 3\right) \ . \end{aligned}$$

With an orthonormal basis on hand, we can expand an arbitrary function, such as $h(x) = x^4$, in terms of the basis as a linear combination:

$$|h\rangle = \sum_{j=0}^{4} h_j |e_j\rangle \ ,$$

where the components h_j are given by

$$h_j = \langle e_j|h\rangle \ .$$

By symmetry of $h(x) = x^4$, all odd h_j are zero, leaving three calculations to perform:

$$\begin{aligned} h_0 &= \frac{1}{\sqrt{2}} \int_{-1}^{1} x^4 \, dx = \frac{\sqrt{2}}{5} \\ h_2 &= \sqrt{\frac{5}{8}} \int_{-1}^{1} \left(3x^6 - x^4\right) dx = \sqrt{\frac{8}{5}} \frac{2}{7} \\ h_4 &= \frac{3}{8\sqrt{2}} \int_{-1}^{1} \left(35x^8 - 30x^6 + 3x^4\right) dx = \frac{1}{35} \frac{8\sqrt{2}}{3} \end{aligned}$$

As a reality check, we can readily verify that $|h\rangle$ still corresponds to x^4, as all other x^n terms cancel out:

$$\begin{aligned} |h\rangle &= \sum_{j=0}^{4} h_j |e_j\rangle \\ &= h_0 |e_0\rangle + h_2 |e_2\rangle + h_4 |e_0 4\rangle = |\phi_4\rangle \end{aligned}$$

6 Normed Vector Space

6.1 Normed Vector Space

The self-inner product of a vector $|a\rangle$, namely

$$\|a\| = \sqrt{\langle a|a\rangle} \in \mathcal{R} \geq 0$$

with

$$\|a\| = 0 \iff |a\rangle = 0$$

is the norm of the vector $|a\rangle$. It turns out that the notion of 'norm' extends to vector spaces.

A vector space $\mathcal{V}$ is said to be *normed* if two of its elements $|a\rangle$, $|b\rangle$, obey the triangle inequality:

$$\|a + b\| \leq \|a\| + \|b\|$$

A normed vector space must also contain the linearity relation

$$\|\alpha a\| = \sqrt{\langle \alpha a|\alpha a\rangle} = |\alpha| \, \|a\|$$

for a complex scalar α.

6.2 Two-Dimensional Systems

Maximum as Norm

Consider the vector space $\mathcal{R}_2$, i.e. pairs of real numbers (x, y). Let us show that the 'maximum' function

$$\|(x, y)\|_m = \max\{|x|, |y|\}$$

is a norm on $\mathcal{R}_2$.

Taking two vectors

$$\begin{aligned} |a\rangle &= (x_1, y_1) \\ |b\rangle &= (x_2, y_2) \ , \end{aligned}$$

the 'max' function tells us

$$\|a + b\|_m = \max\left(|x_1 + x_2|, |y_1 + y_2|\right) \ .$$

Note from the triangle inequality that the arguments sent to max {} function obey

$$\begin{aligned} |x_1 + x_2| &\leq |x_1| + |x_2| \\ |y_1 + y_2| &\leq |y_1| + |y_2| \ . \end{aligned}$$

Also observe that $|a\rangle$, $|b\rangle$ are subject to

$$\begin{aligned} |x_1| &\leq \|a\|_m \\ |y_1| &\leq \|a\|_m \\ |x_2| &\leq \|b\|_m \\ |y_2| &\leq \|b\|_m \ . \end{aligned}$$

Summing the x-equations and the y-equations, we find

$$\begin{aligned} |x_1| + |x_2| &\leq \|a\|_m + \|b\|_m \\ |y_1| + |y_2| &\leq \|a\|_m + \|b\|_m \ . \end{aligned}$$

Tracing back the inequality symbols, we may finally write

$$\begin{aligned}\|a+b\|_m &= \max\left(|x_1+x_2|\,,|y_1+y_2|\right)\\ &\leq \max\left(|x_1|+|x_2|\,,|y_1|+|y_2|\right)\\ &\leq \max\left(\|a\|_m+\|b\|_m, \|a\|_m+\|b\|_m\right)\\ &\leq \|a\|_m+\|b\|_m\ ,\end{aligned}$$

satisfying a requirement of a norm.

To complete the job we also establish a linearity relation:

$$\begin{aligned}\|\alpha a\|_m &= \max\left(|\alpha x_1|\,,|\alpha y_1|\right)\\ &= |\alpha|\max\left(|x_1|\,,|y_1|\right)\\ &= |\alpha|\,\|a\|_m\end{aligned}$$

Sum as Norm

Consider the (same) vector space $\mathcal{R}_2$, i.e. pairs of real numbers (x, y). Let us show that the 'sum' function

$$\|(x,y)\|_s = |x|+|y|$$

is a norm on $\mathcal{R}_2$.

Taking the two vectors

$$\begin{aligned}|a\rangle &= (x_1, y_1)\\ |b\rangle &= (x_2, y_2)\ ,\end{aligned}$$

the 'sum' function gives, using the same identities as above,

$$\begin{aligned}\|a+b\|_s &= |a|+|b|\\ &= |x_1+y_1|+|x_2+y_2|\\ &\leq |x_1|+|y_1|+|x_2|+|y_2|\\ &\leq \|a\|_m+\|b\|_m\ .\end{aligned}$$

To check for linearity, we write

$$\begin{aligned}\|\alpha a\|_s &= |\alpha x_1|+|\alpha y_1|\\ &= |\alpha|\left(|x_1|+|y_1|\right)\\ &= |\alpha|\,\|a\|_s\ .\end{aligned}$$

Unit Ball

In any vector space $\mathcal{V}$, the unit 'ball' $\mathcal{B}_1$ is defined as

$$\mathcal{B} = \{|a\rangle \in \mathcal{V}\ :\ \|a\| \leq 1\}\ .$$

Plotting the the 'max' function in the xy-plane, the unit ball resolves to a square frame of side 1, as $\max(a) = 1$ in the unit ball. In terms of the 'sum' function, the ball resolves to a filled diamond with points at $(0, \pm 1)$ and $(\pm 1, 0)$, generated by $|x|+|y| \leq 1$.

6.3 Identities

Cauchy-Bunyakovsky-Schwarz Inequality

There is a important fact called the *Cauchy-Bunyakovsky-Schwarz Inequality* that must be established.

For two vectors $|a\rangle$, $|b\rangle$, consider the nonzero sum $|a\rangle - \lambda\,|b\rangle$, or for short, $|a-\lambda b\rangle \neq |0\rangle$. Now expand the norm of this vector:

$$\begin{aligned}\|a-\lambda b\|^2 &= \langle a-\lambda b|a-\lambda b\rangle\\ &= \|a\|^2 - \lambda\,\langle a|b\rangle - \lambda^*\,\langle b|a\rangle + \lambda^2\|b\|^2\end{aligned}$$

Next choose

$$\lambda = \frac{\langle b|a\rangle}{\|b\|^2}\ ,$$

and the above becomes

$$\begin{aligned}\|a-\lambda b\|^2 = \|a\|^2 &- \frac{\langle b|a\rangle}{\|b\|^2}\,\langle a|b\rangle\\ &- \frac{\langle a|b\rangle}{\|b\|^2}\,\langle b|a\rangle + \frac{\langle a|b\rangle\,\langle b|a\rangle}{\|b\|^2\cancel{\|b\|^2}}\cancel{\|b\|^2}\ .\end{aligned}$$

The last two terms cancel exactly, and the norm on the left is defined to be positive and real. What's left is

$$\|a\|^2 - \frac{\langle a|b\rangle\,\langle b|a\rangle}{\|b\|^2} > 0\ ,$$

which simplifies to a very powerful result:

$$|\langle a|b\rangle| \leq \|a\|\|b\|$$

Triangle Inequality

For two vectors $|a\rangle$, $|b\rangle$, consider the nonzero sum $|a+\lambda b\rangle \neq |0\rangle$. Expand the norm of this vector,

$$\begin{aligned}\|a+\lambda b\|^2 &= \langle a-\lambda b|a+\lambda b\rangle\\ &= \|a\|^2 + \lambda\,\langle a|b\rangle + \lambda^*\,\langle b|a\rangle + \lambda^2\|b\|^2\\ &= \|a\|^2 + 2\mathrm{Re}\left(\lambda\,\langle a|b\rangle\right) + \lambda^2\|b\|^2\ ,\end{aligned}$$

and let $\lambda = 1$:

$$\|a+b\|^2 = \|a\|^2 + 2\mathrm{Re}\left(\langle a|b\rangle\right) + |b\|^2$$

By the Cauchy-Bunyakovsky-Schwarz inequality, the middle term is less than $\|a\|\|b\|$:

$$\begin{aligned}\|a+b\|^2 &\leq \|a\|^2 + 2\|a\|\|b\| + |b\|^2\\ &\leq \left(\|a\|+|b\|\right)^2\end{aligned}$$

Take a final square root of both sides and the proof is done.

Pythagorean Theorem

In the special case $\langle a|b\rangle = 0$, the triangle inequality reduces to the Pythagorean theorem:

$$\|a+b\|^2 = \|a\|^2 + \|b\|^2$$

Parallelogram Law

For two vectors $|a\rangle$, $|b\rangle$ and the linear combinations $|a+b\rangle$, $|a-b\rangle$, there exists a *parallelogram law* concerning vector addition:

$$\|x+y\|^2 + \|x-y\|^2 = 2\left(\|x\|^2 + \|y\|^2\right)$$

To prove this, write out the norm of $|a+b\rangle$ and $|a-b\rangle$

$$\begin{aligned} \|a+b\|^2 &= \|a\|^2 + \|b\|^2 + \langle a|b\rangle + \langle b|a\rangle \\ \|a-b\|^2 &= \|a\|^2 + \|b\|^2 - \langle a|b\rangle - \langle b|a\rangle \ , \end{aligned}$$

and add the resulting equations to finish the proof.

A complementary result comes from subtracting the equations:

$$\|a+b\|^2 - \|a-b\|^2 = 4\,\mathrm{Re}\left(\langle a|b\rangle\right)$$

One may imagine what combination on the left yields $\mathrm{Im}\left(\langle a|b\rangle\right)$ on the right. It's straightforward to show that the job is done by:

$$i\|a+ib\|^2 - i\|a-ib\|^2 = 4\,\mathrm{Im}\left(\langle a|b\rangle\right)$$

Polar Identity

The sum of $4\,\mathrm{Re}\left(\langle a|b\rangle\right)$ and $4\,\mathrm{Im}\left(\langle a|b\rangle\right)$, each given above, leads to the *polar identity*. First note that

$$4\,\mathrm{Re}\left(\langle a|b\rangle\right) + 4\,\mathrm{Im}\left(\langle a|b\rangle\right) = 4\left(\langle a|b\rangle\right) \ ,$$

which yields the identity

$$\begin{aligned} 4\left(\langle a|b\rangle\right) = \|a+b\|^2 - \|a-b\|^2 \\ + i\|a+ib\|^2 - i\|a-ib\|^2 \end{aligned}$$

Problem 1

Use the Cauchy-Bunyakovsky-Schwarz inequality to show that

$$1 < \int_0^{\pi/2} \sqrt{\sin(x)}\,dx < \sqrt{\frac{\pi}{2}} \ .$$

Hint: For the left, establish

$$\sin(x) \leq \sqrt{\sin(x)} \leq \sqrt{x}$$

and integrate, remembering

$$\int_0^{\pi/2} \sin(x)\,dx = 1 \ .$$

For the right, associate $a = \sqrt{\sin(x)}$ and $b = 1$ so that $\langle a|a\rangle = 1$ and $\langle b|b\rangle = \pi/2$.

7 Countably Finite System

7.1 Convergent Sequence

Suppose $\mathcal{V}$ is a normed vector space. The sequence of vectors

$$\{|a_j\rangle \in \mathcal{V}\} \ : \ j = (1, 2, 3, \dots)$$

is said to be *convergent* to the vector $|a\rangle$ if

$$\forall\, k_\epsilon \ : \ \exists\, \epsilon > 0$$

such that if $k > k_\epsilon$, then

$$\|a - a_k\| < \epsilon \ .$$

This statement inspires a definition of a convergent vector:

$$|a\rangle = \lim_{k\to\infty} |a_k\rangle$$

7.2 Cauchy Sequence

The (same) sequence of vectors $\{|a_j\rangle \in \mathcal{V}\}$ qualifies as a *Cauchy sequence* if

$$\forall\, k_\epsilon \ : \ \exists\, \epsilon > 0 \ ,$$

then

$$\|a_m - a_n\| < \epsilon$$

provided that $m, n > k_\epsilon$.

It's easy to show that any convergent sequence qualifies as a Cauchy sequence. For two vectors $|a_m\rangle$, $|a_n\rangle$, we know

$$\begin{aligned} \|a - a_m\| &< \alpha\epsilon \\ \|a - a_n\| &< \beta\epsilon \end{aligned}$$

for two parameters $\alpha, \beta > 0 \in \mathcal{R}$. Adding each, we have

$$\|a - a_m\| + \|a_n - a\| < (\alpha + \beta)\,\epsilon \ ,$$

which becomes, by the triangle inequality,

$$\|a - a_m + a_n - a\| < \epsilon \ ,$$

reducing to the fingerprint of a Cauchy sequence:

$$\|a_m - a_n\| < \epsilon$$

7.3 Complete Space

A normed vector spaces in which all Cauchy sequences converge is called a *complete space*, also known as a Banach space. In terms of an orthonormal basis, an arbitrary vector is given by

$$|a\rangle = \sum_{j=1}^{n} a_j |e_j\rangle \ ,$$

where clearly

$$\|a\| = \sqrt{|a|_1^2 + |a|_2^2 + \cdots + |a|_n^2} \ .$$

We may further consider a vector $|b\rangle$ that is itself a Cauchy sequence of vectors $|a^{(k \le m)}\rangle$ such that

$$|b\rangle = \sum_{k=1}^{m} |a^{(k)}\rangle = \sum_{j=1}^{n} \sum_{k=1}^{m} a_j^{(k)} |e_j\rangle = \sum_{j=1}^{n} b_j |e_j\rangle \ ,$$

where each coefficient b_j is itself a Cauchy sequence of the complex coefficients a_j:

$$b_j = \sum_{k=1}^{m} a_j^{(k)}$$

7.4 Supremum

Consider the complete infinite-dimensional space $\mathcal{C}_{ab}$ of complex-valued functions $f(x) \ : \ x \in [a, b]$. Here we define the 'supremum' function

$$\|f\|_{sup} = \max \{ |f(x)| \ : \ x \in [a, b] \} \ .$$

The 'sup' norm guarantees homogeneous convergence of a Cauchy sequence of functions $f^{(k)}(x)$ to a single function $f(x)$.

Unlike other norms we've encountered, the $\|f\|_{sup}$ does not bear a notion of inner product. Choosing a trivial example $f(x) = \cos x$, $g(x) = x$ in the interval $x \in [a, b]$, we have

$$\begin{aligned} \|f\|_{sup} &= 1 \\ \|g\|_{sup} &= \pi \\ \|f+g\|_{sup} &= 1 + \pi \\ \|f-g\|_{sup} &= -1 - \pi \ , \end{aligned}$$

which violates the parallelogram law:

$$\begin{aligned} \|f+g\|_{sup}^2 + \|f+g\|_{sup}^2 &\neq 2\left(\|f\|_{sup}^2 + \|g\|_{sup}^2\right) \\ (1+\pi)^2 + (-1-\pi)^2 &\neq 2\left(1+\pi^2\right) \\ 4\pi &\neq 0 \end{aligned}$$

7.5 Hilbert Space

A complete inner product space is called a *Hilbert space*, and we have shown that all finite-dimensional vector spaces are Hilbert spaces. The 'supremum' norm is a unique example of a complete space that is *not* a Hilbert space.

8 Countably Infinite System

The results of the previous section readily generalize to handle a *countably infinite* basis.

8.1 Fourier Series

Consider an orthonormal basis $\{|e_j\rangle\}$ containing an *infinite* number of basis vectors. An infinite linear combination

$$|x\rangle = \sum_{j=1}^{\infty} \langle e_j | x\rangle |e_j\rangle$$

is the *Fourier series* of the vector $|x\rangle$ in the basis $\{|e_j\rangle\}$, where $\langle x|e_j\rangle$ are *Fourier coefficients*.

8.2 Bessel Inequality

Let us show that any partial sum of a Fourier series is a Cauchy sequence.

Truncating the series at the nth term gives

$$|x^{(n)}\rangle = \sum_{j=1}^{n} \langle e_j | x\rangle |e_j\rangle$$

as a partial sum. Another truncation of the series with $m > n$ can be written

$$|x^{(m)}\rangle = |x^{(n)}\rangle + \sum_{j=n+1}^{m} \langle e_j | x\rangle |e_j\rangle \ ,$$

where the norm of the difference of the two vectors reads

$$\|x^{(m)} - x^{(n)}\|^2 = \sum_{j=n+1}^{m} |\langle e_j | x\rangle|^2 \ .$$

The above series is assured to be a positive real number, reducing the problem to showing that

$$\sum_{j=1}^{\infty} |\langle e_j | x\rangle|^2$$

does not diverge.

Proceed by writing out the $m \to \infty$ case, giving

$$\langle x - x^{(n)} | x - x^{(n)}\rangle = \|x\|^2 - \sum_{j=1}^{n} |\langle e_j | x\rangle|^2 \geq 0 \, .$$

Reading the equation form the right, we arrive at the *Bessel inequality*

$$\sum_{j=1}^{n} |\langle e_j|x\rangle|^2 \leq \|x\|^2 \ ,$$

proving that a partial sum of a Fourier series is a Cauchy sequence. Reading the above from left to right, we also have

$$\|x - x^{(n)}\| = \sqrt{\|x\|^2 - \sum_{j=1}^{n} |\langle e_j|x\rangle|^2} \ ,$$

which in the case of convergence ($n \to \infty$), we get the *Parseval relation*:

$$\|x\|^2 = \sum_{j=1}^{\infty} |\langle e_j|x\rangle|^2$$

8.3 Inner Product Space of Functions

The space $\mathcal{C}_2[a,b]$ of the inner product of two complex functions was written as

$$\langle f|g\rangle = \int_a^b f^*(z)\, g(z)\, dz \ .$$

For certain cases of the function $f(z)$, for instance a discontinuous function, the space $\mathcal{C}_2[a,b]$ is easily shown to not be complete with respect to the usual notion of norm,

$$\|f\| = \sqrt{\int_a^b |f|^2 \ dz} \ ,$$

implying that a Hilbert space of functions must be carefully discerned.

According to the Riesz-Fisher theorem, we may define a Hilbert of functions with a countably infinite basis, denoted $\mathcal{L}_2[a,b]$. Furthermore, the Stone-Weierstrass theorem states that the set of polynomials $\{|x_j\rangle\}$ with $j = 1, 2, \ldots$ forms a basis in $\mathcal{L}_2[a,b]$. Of course, we found such vectors to be non-orthogonal, corrected by the Gram-Schmidt process to produce the Legendre polynomials.

9 Operators

9.1 Definition

An *operator* L is a function that 'acts on' a vector $|x\rangle \in \mathcal{V}$ to create a new vector

$$L|x\rangle = |Lx\rangle = |y\rangle$$

that may or may not live in $\mathcal{V}$.

9.2 Linear Operator

An operator that maps a vector to its own vector space $\mathcal{V}$ is said to be *linear* if the relation

$$L|\alpha u + \beta v\rangle = \alpha L|u\rangle + \beta L|v\rangle$$

is satisfied, where α, β are complex scalars. Needless to mention, scalar multiplication is a special case of a linear operator.

Linearity Check

Interpreting vectors as functions, we can check whether certain operations for a function $f(x)$ qualify as linear operators. For example, the transformation

$$L(f(x)) = \sin(f(x))$$

fails when tested for linearity:

$$\begin{aligned} L(\alpha f(x)) = \sin(\alpha f(x)) &= \sin\alpha\cos(f(x)) \\ &\quad + \cos\alpha\sin(f(x)) \\ &\neq \alpha L(f(x)) \end{aligned}$$

The less trivial example

$$L(f(x)) = \int_0^1 \sin(xy)\, f(y)\, dy$$

does qualify as a linear operator, as the function f enters the integral linearly. Explicitly, we have

$$\begin{aligned} &L(\alpha f(x) + \beta g(x)) \\ &\quad = \int_0^1 \sin(xy)\,(\alpha f(y) + \beta g(y))\, dy \\ &\quad = \alpha\int_0^1 \sin(xy)\, f(y)\, dy + \beta\int_0^1 \sin(xy)\, g(y)\, dy \\ &\quad = \alpha L(f(x)) + \beta L(g(x)) \ . \end{aligned}$$

9.3 Adjoint Operator

Given an operator L, the *adjoint* operator $L^\dagger$, is defined such that

$$\langle b|L|a\rangle = \overline{\langle a|L^\dagger|b\rangle}$$

readily implying:

$$\langle b|L|a\rangle = \langle L^\dagger b|a\rangle$$

$$\left(L^\dagger\right)^\dagger = L$$

Two linear operators A and B always obey the relation

$$(AB)^\dagger = B^\dagger A^\dagger \ ,$$

proven by writing $\langle u|AB|v\rangle$ two different ways and comparing each right-hand result:

$$\langle u|AB|v\rangle = \langle (AB)^\dagger u|v\rangle$$

$$\langle u|AB|v\rangle = \langle A^\dagger u|Bv\rangle = \langle B^\dagger A^\dagger u|v\rangle$$

9.4 Hermitian Operator

An operator that is its own adjoint operator is called *self-adjoint*, also known as *Hermitian*:

$$L = L^\dagger \ \rightarrow \ \langle b|La\rangle = \langle Lb|a\rangle$$

If L is Hermitian, the operator $L^\dagger$ is called the *Hermitian conjugate* to L.

For self-adjoint operators, the quantity

$$\lambda = \langle b|\, L\, |a\rangle$$

is always real-valued.

We may also inquire whether the product AB of two Hermitian operators is itself Hermitian. Starting with $(AB)^\dagger = B^\dagger A^\dagger$, let $A = A^\dagger$ and $B = B^\dagger$ to find

$$AB = \left(B^\dagger A^\dagger\right)^\dagger = (BA)^\dagger \ ,$$

telling us AB is Hermitian only if $AB = BA$.

9.5 Anti-Hermitian Operator

An *Anti-Hermitian* operator is one that obeys

$$L = -L^\dagger \ .$$

For two Hermitian operators A, B, it turns out that $AB - BA$ is anti-Hermitian:

$$\begin{aligned}(AB - BA)^\dagger &= (AB)^\dagger - (BA)^\dagger \\ &= B^\dagger A^\dagger - A^\dagger B^\dagger = BA - AB\end{aligned}$$

Partial Derivative Operator

The partial derivative operator

$$A = \frac{\partial}{\partial x}$$

is an anti-Hermitian operator for certain boundary conditions. Consider two arbitrary function $f(x)$, $g(x)$ in the domain Ω where each function is zero on the boundary $\partial\Omega$.

Then, writing out $\langle f|\, A\, |g\rangle$ two different ways gives

$$\begin{aligned}\langle f|Ag\rangle &= \langle A^\dagger f|g\rangle \\ \int_\Omega f^* \partial_x g \, dx &= \int_\Omega (\partial_x)^\dagger f^* g \, dx \ .\end{aligned}$$

Integrating the left side by parts, we write

$$f^* g\big|_{\partial\Omega} - \int_\Omega \partial_x f^* g \, dx = \int_\Omega (\partial_x)^\dagger f^* g \, dx \ ,$$

where the boundary term equals zero by construction, indicating A to be anti-Hermitian.

9.6 Projector

For any fixed vector $|a\rangle$, the combination

$$P_a = |a\rangle \langle a|$$

is called the *projector* of $|a\rangle$. The projector does nothing on its own, but waits for a bra- or ket-vector to interact with the left or the right side, respectively. Acting on a vector $|x\rangle$, we have

$$P_a \,|x\rangle = |a\rangle \langle a|x\rangle = \langle a|x\rangle \,|a\rangle \ ,$$

which is $|a\rangle$ multiplied by a scalar.

Properties

The projector is a linear operator, easily verified by

$$\begin{aligned}P_a \,|\alpha u + \beta v\rangle &= |a\rangle \langle a| \left(|\alpha u\rangle + |\beta v\rangle\right) \\ &= |a\rangle \left(\alpha \langle a|u\rangle + \beta \langle a|v\rangle\right) \\ &= \alpha \langle a|u\rangle \,|a\rangle + \beta \langle a|v\rangle \,|a\rangle \\ &= \alpha P_a \,|u\rangle + \beta P_a \,|v\rangle \ ,\end{aligned}$$

and is also Hermitian:

$$\begin{aligned}\langle u|P_a v\rangle &= \langle u| \left(\langle a|v\rangle \,|a\rangle\right) \\ &= \langle a|v\rangle \langle u|a\rangle \\ &= \left(\langle u|a\rangle \langle a|\right) |v\rangle \\ &= \langle P_a u|v\rangle\end{aligned}$$

9.7 Identity Operator

Consider any vector $|x\rangle$ in the vector space $\mathcal{V}$ spanned by the basis vectors $\{|e_k\rangle\}$. Choosing any basis vector $|e_j\rangle$, apply a projector

$$P_{e_j} = |e_j\rangle \langle e_j|$$

to $|x\rangle$ to get

$$P_{e_j} \,|x\rangle = \langle e_j|x\rangle \,|e_j\rangle = x_j \,|e_j\rangle \ .$$

Summing over the index j, we find

$$\left(\sum_j P_{e_j}\right) |x\rangle = \sum_j x_j \,|e_j\rangle = |x\rangle \ .$$

The parenthesized quantity that leaves the vector unchanged is called the *identity* operator I. That is,

$$I\,|x\rangle = |x\rangle \ ,$$

where

$$I = \sum_j |e_j\rangle \langle e_j|$$

is also called the *completeness relation* for the basis. (It is possible for a basis to be *incomplete*, in which case the above sum is not equivalent to an identity operator.)

9.8 Commutator

For any two operators A and B, the difference

$$[AB] = AB - BA$$

defines their *commutator*, also known as the *commutation relation*. The result of $[AB]$ tells us which 'extra terms' emerge when swapping two operators. When the commutator evaluates to zero, the operators are said to *commute*.

For example, consider two operators

$$\begin{aligned} A &= \frac{\partial}{\partial x} \\ B &= x \end{aligned}$$

that can act on a function $f(x)$. Allowing the commutator to act on $f(x)$, we write

$$\begin{aligned} (AB - BA) \left| f \right\rangle &= \partial_x \left(x \left| f \right\rangle \right) - x \left(\partial_x \left| f \right\rangle \right) \\ &= \left| f \right\rangle + x\partial_x \left| f \right\rangle - x\partial_x \left| f \right\rangle \\ &= \left| f \right\rangle \ , \end{aligned}$$

telling us that that operators on hand do not commute, but instead obey

$$AB - BA = I \ .$$

10 Eigen-Calculations

If an operator L applied to a vector $\left| x \right\rangle$ results in a parallel vector $\lambda \left| x \right\rangle$, then $\left| x \right\rangle$ is called an *eigenvector* of L, and λ is the corresponding *eigenvalue*:

$$L \left| x \right\rangle = \lambda \left| x \right\rangle$$

10.1 Real Eigenvalues

It's straightforward to show that eigenvalues are always real if L is self-adjoint (Hermitian):

$$\left\langle x \right| L \left| x \right\rangle = \lambda \left\langle x | x \right\rangle \ \rightarrow \ \lambda = \frac{\left\langle x \right| L \left| x \right\rangle}{\left\langle x | x \right\rangle}$$

To establish this, write the eigenvalue problem

$$L \left| x \right\rangle = \lambda \left| x \right\rangle \ ,$$

and project an arbitrary vector $\left\langle y \right|$ to write

$$\left\langle y \right| L \left| x \right\rangle = \lambda \left\langle y | x \right\rangle \ ,$$

equivalent to

$$\left\langle L^\dagger y | x \right\rangle = \lambda \left\langle y | x \right\rangle \ .$$

Take the complex conjugate of each side to get

$$\overline{\left\langle L^\dagger y | x \right\rangle} = \lambda^* \overline{\left\langle y | x \right\rangle} \ ,$$

or

$$\left\langle x | L^\dagger y \right\rangle = \left\langle x \right| L^\dagger \left| y \right\rangle = \lambda^* \left\langle x | y \right\rangle$$

Finally, let us set $\left| y \right\rangle = \left| x \right\rangle$, and without loss of generality assume $\left| x \right\rangle$ is normalized, so we may take $\left\langle y | x \right\rangle = \left\langle x | y \right\rangle = 1$. After simplifying, we are lift with

$$\begin{aligned} \left\langle x \right| L \left| x \right\rangle &= \lambda \\ \left\langle x \right| L^\dagger \left| x \right\rangle &= \lambda^* \end{aligned}$$

If L is self-adjoint, we automatically have $L = L^\dagger$. This can only mean $\lambda = \lambda^*$, thus all λ are real.

10.2 Orthogonal Eigenvectors

For a linear self-adjoint operator L, we can show that two distinct eigenvalues λ_1, λ_2 correspond to two orthogonal eigenvectors.

Start with

$$\begin{aligned} \left\langle x_2 \right| L \left| x_1 \right\rangle &= \lambda_1 \left\langle x_2 | x_1 \right\rangle \\ \left\langle x_1 \right| L \left| x_2 \right\rangle &= \lambda_2 \left\langle x_1 | x_2 \right\rangle \ , \end{aligned}$$

and complex-conjugate the second equation to eliminate the $\left\langle x_2 | x_1 \right\rangle$-term:

$$\left\langle x_2 \right| L \left| x_1 \right\rangle = \left(\frac{\lambda_1}{\lambda_2} \right) \overline{\left\langle x_1 \right| L \left| x_2 \right\rangle} = \left(\frac{\lambda_1}{\lambda_2} \right) \left\langle x_2 \right| L \left| x_1 \right\rangle$$

For $\lambda_1 \neq \lambda_2$, the only reasonable conclusion is

$$\left\langle x_2 \right| L \left| x_1 \right\rangle = 0 \ \rightarrow \ \left\langle x_1 | x_2 \right\rangle = 0 \, ,$$

meaning $\left| x_1 \right\rangle$, $\left| x_2 \right\rangle$ must be orthogonal.

10.3 Equal Eigenvalues

If n eigenvalues are equal, one speaks of *n-fold degeneracy*. In this case, the corresponding eigenvectors are not necessarily orthogonal, in which case the vectors form a vector subspace of the original vector space that might admit its own orthonormal basis.

10.4 Calculating Eigenvectors

When the form of L is given, it's usually possible to solve for all eigenvalues λ_j. With n of these these established, the next move is to calculate the eigenvectors directly using

$$\begin{aligned} L \left| x^{(j)} \right\rangle &= \lambda_j \left| x^{(j)} \right\rangle \\ j &= 1, 2, \ldots, n \end{aligned}$$

Whether or not the eigenvectors form an orthonormal basis, we may express an arbitrary vector $|u\rangle$ as a linear combination:

$$|u\rangle = \sum_j C_j \left|x^{(j)}\right\rangle$$

Recall that the eigenvectors $\left|x^{(j)}\right\rangle$ are only orthogonal if L is self-adjoint, i.e. Hermitian.

10.5 Time Derivative Operators

Consider a vector space $\mathcal{V}$ of dimension n admitting a fixed orthonormal basis $\{|e_j\rangle\}$ where $j = 1, \ldots, n$. A time-varying vector $|u(t)\rangle$ is a linear combination of time-varying coefficients such that

$$|u(t)\rangle = \sum_j u_j(t) |e_j\rangle \ .$$

Single Time Derivative Operator

Now we introduce the single time-derivative operator

$$L = \frac{\partial}{\partial t} = \partial_t \ .$$

If we let L act on an eigenvector $\left|x^{(j)}(t)\right\rangle$, the result is equal to $|x(t)\rangle$ multiplied by its corresponding eigenvalue λ. That is,

$$L |x(t)\rangle = \lambda |x(t)\rangle \ ,$$

or

$$\sum_j \partial_t x_j(t) |e_j\rangle = \sum_j \lambda x_j(t) |e_j\rangle \ ,$$

implying n copies of the same separable differential equation

$$\partial_t x_j(t) = \lambda x_j(t)$$

for each index j.

Elementary methods give the solution for each equation

$$x_j(t) = x_j(t=0)\, e^{\lambda t} = x_{0j}\, e^{\lambda t} \ ,$$

telling us

$$\begin{aligned} |x(t)\rangle &= \sum_j x_{0j}\, e^{\lambda t} |e_j\rangle \\ &= e^{\lambda t} \sum_j x_{0j} |e_j\rangle = e^{\lambda t} |x_0\rangle \ . \end{aligned}$$

Perhaps not surprisingly, the time dependence evolves exponentially in time.

As a matter of technicality, a vector $|x(t)\rangle$ is best described as an *eigenfunction*, as the operator $L = \partial_t$ is trivial for time-independent vectors.

Double Time Derivative Operator

We also consider the double time-derivative operator $L = \partial_{tt}$. Using the same setup, it follows that each x_j is governed by the differential equation

$$\partial_{tt} x_j(t) = \lambda x_j(t) \ ,$$

whose solution is governed by λ.

For $\lambda = 0$, the coefficients evolve linearly in time:

$$\begin{aligned} &\lambda = 0 \\ &x_j(t) = x_{0j} + x_{1j}\, t \end{aligned}$$

For $\lambda \neq 0$, we have a linear combination of exponential terms:

$$\begin{aligned} &\lambda \neq 0 \\ &x_j(t) = x_{0j}\, e^{\sqrt{\lambda} t} + x_{1j}\, e^{-\sqrt{\lambda} t} \end{aligned}$$

11 Operator as Matrix

Consider a vector $|x\rangle$ living in vector space $\mathcal{V}$ that admits an orthonormal basis $\{|e_j\rangle\}$. As a linear combination of coefficients $\{x_j\}$, such a vector is written

$$|x\rangle = \sum_j x_j |e_j\rangle \ .$$

Suppose another vector $|y\rangle$, which is itself a linear combination of coefficients $\{y_j\}$ in the same basis, arises by applying a linear operator A onto $|x\rangle$:

$$|y\rangle = A|x\rangle = \sum_j y_j |e_j\rangle$$

The question now is, what can we discern about the operator A?

11.1 Matrix Elements

We proceed by 'solving for' any component y_j, which entails taking the inner product with a basis vector $\langle e_{k\neq j}|$ to get

$$\langle e_k| A |x\rangle = \sum_j \langle e_k| y_j |e_j\rangle = y_j \delta_{jk} = y_k \ ,$$

implying

$$\begin{aligned} y_j = \langle e_j| A |x\rangle &= \langle e_j| A \sum_k x_k |e_k\rangle \\ &= \sum_k \langle e_j| A |e_k\rangle\, x_k \ . \end{aligned}$$

That is, any y_j depends on each member of $\{x_j\}$ multiplied by a number

$$A_{jk} = \langle e_j | A | e_k \rangle$$

called a *matrix element*.

The set of matrix elements $\{A_{jk}\}$ is the *matrix* represented by the operator A. To restore the operator A in terms of its elements, begin with the identity $A = IAI$ and write out each identity operator explicitly to get

$$\begin{aligned} A = IAI &= \sum_j \sum_k |e_j\rangle \langle e_j| A |e_k\rangle \langle e_k| \\ &= \sum_j \sum_k |e_j\rangle A_{jk} \langle e_k| \ . \end{aligned}$$

Here we emphasize that the choice of basis vectors has direct impact on the components A_{jk}.

11.2 Matrix Algebra

Matrix Addition

Two operators A and B readily add to form a new operator C such that

$$\begin{aligned} A|x\rangle + B|x\rangle &= C|x\rangle \\ A_{jk} + B_{jk} &= C_{jk} \ , \end{aligned}$$

which of course requires A, B to be of equal dimension.

Scalar Multiplication

A scalar λ can be 'multiplied into' an operator A by scaling each component to create another operator B:

$$\begin{aligned} B &= \lambda A \\ B_{jk} &= \lambda A_{jk} \end{aligned}$$

Matrix Multiplication

Two operators A and B can 'multiply' to form a new operator C such that

$$\begin{aligned} A(B|x\rangle) &= C|x\rangle \\ AB &= C \ , \end{aligned}$$

which is generally an associative operation, but not commutative:

$$\begin{aligned} (AB)C &= A(BC) \\ AB &\neq BA \end{aligned}$$

The formula for matrix multiplication is calculated by brute force. $C = AB$ expands to

$$C = \sum_j \sum_k \sum_{j'} \sum_m |e_j\rangle A_{jk} \langle e_k | e_{j'}\rangle B_{j'm} \langle e_m|$$

$$C = \sum_j \sum_m |e_j\rangle \left(\sum_k A_{jk} B_{km} \right) \langle e_m| \ ,$$

telling us

$$C_{jm} = \sum_k A_{jk} \, B_{km} \ .$$

11.3 Matrix Transpose

For a given operator A with components A_{jk}, the *transpose* of A, denoted A^T, has components A_{kj}. That is, the transpose swaps rows $\leftrightarrow$ columns. Using the matrix multiplication rule, it's straightforward to show that the transpose of the product of two matrices equals the product of the individually transposed matrices in reversed order:

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

It's also straightforward to show the following determinant identity:

$$\det\left(A^\dagger\right) = \det\left((A^*)^T\right) = (\det(A))^*$$

11.4 Matrix Trace

For operators A represented by a square $(n \times n)$ matrix, a special quantity exists called the *trace* of the matrix. The trace is defined as the sum of the components along the diagonal:

$$\mathrm{tr}A = A_{11} + A_{22} + \cdots + A_{nn} = \sum_{k=1}^{n} A_{kk}$$

12 Hermitian Matrix

Recall that an operator A that is it's own adjoint operator, meaning $A = A^\dagger$ as appearing in the definition

$$\langle y| A |x\rangle = \overline{\langle x| A^\dagger |y\rangle} \ ,$$

where

$$|x\rangle , |y\rangle \ \in \mathcal{V} \ ,$$

is a Hermitian operator, synonymous with *Hermitian matrix*, where $A^\dagger$ is the *Hermitian conjugate*.

In component form, we note that

$$\langle y | Ax \rangle = \sum_{ij} A_{ij} \, y_i^* x_j \ ,$$

where meanwhile, an equivalent statement is

$$\overline{\langle x|A^\dagger y\rangle} = \sum_{ij} \left(A^\dagger_{ji}\right)^* x_j y_i^* \ ,$$

telling us that the components of a Hermitian operator obey

$$A^*_{ij} = A^\dagger_{ji} \ .$$

Anti-Hermitian Matrix

The properties of anti-Hermitian operators also extend to matrices. An anti-Hermitian matrix satisfies

$$\begin{aligned} A^\dagger &= -A \\ A^*_{ij} &= -A_{ij} \ . \end{aligned}$$

As a corollary, we note that if a matrix A is Hermitian, then iA is anti-Hermitian, and vise-versa.

Symmetric Matrix

When the components of a Hermitian matrix are all real-valued, the matrix is symmetric. Work this from the identity

$$A^*_{ij} = A^\dagger_{ji} \ ,$$

and notice that the complex conjugate A^* is just A again:

$$A_{ij} = A^\dagger_{ji}$$

Since $A = A^\dagger$ by definition, we further have

$$A_{ij} = A_{ji} \ ,$$

telling us A is symmetric.

As a special case of the Hermitian operator, we automatically have that the eigenvectors corresponding to non-equal eigenvalues of a symmetric matrix are orthogonal.

12.1 Commuting Operators

Now we derive the important fact that commuting Hermitian operators share an orthonormal basis. To begin, consider an operator A admitting an eigenvector $|\psi_n\rangle$ with corresponding eigenvalue a_n, and also a second operator B admitting an eigenvector $|\phi_n\rangle$ with corresponding eigenvalue b_n:

$$\begin{aligned} A\,|\psi_n\rangle &= a_n\,|\psi_n\rangle \\ B\,|\phi_n\rangle &= b_n\,|\phi_n\rangle \end{aligned}$$

Our key assumption is that any $|\psi_n\rangle$ can be written as a linear combination of $\{|\phi_n\rangle\}$, and vice-versa, which assumes each vector is a member of the same basis:

$$\begin{aligned} |\psi_n\rangle &= \sum_m \gamma_{mn}\,|\phi_m\rangle \\ |\phi_n\rangle &= \sum_m \tilde{\gamma}_{mn}\,|\psi_m\rangle \ . \end{aligned}$$

In the above, γ, $\tilde{\gamma}$ are matrix coefficients. We gain a restriction on γ, $\tilde{\gamma}$ by the substitution

$$\begin{aligned} |\psi_n\rangle &= \sum_m \gamma_{mn} \sum_{m'} \tilde{\gamma}_{m'm}\,|\phi_{m'}\rangle \\ &= \sum_{m'} \left(\sum_m \gamma_{mn}\tilde{\gamma}_{m'm} \right) |\phi_{m'}\rangle \ , \end{aligned}$$

implying the delta relation

$$\delta_{m'n} = \sum_m \gamma_{mn}\tilde{\gamma}_{m'm} \ .$$

To gain two more delta relations, compute $A\,|\psi_n\rangle$ two different ways to write

$$\begin{aligned} A\,|\psi_n\rangle &= A \sum_m \gamma_{mn}\,|\phi_m\rangle \\ &= A \sum_m \sum_{m'} \gamma_{mn}\tilde{\gamma}_{m'm}\,|\psi_{m'}\rangle \end{aligned}$$

and

$$a_n\,|\psi_n\rangle = \sum_{m'} \sum_m \gamma_{mn}\tilde{\gamma}_{m'm} a_{m'}\,|\psi_{m'}\rangle \ .$$

Comparing each ride side we find

$$\frac{1}{a_n} \sum_m \gamma_{mn}\tilde{\gamma}_{m'm} a_{m'} = \delta_{m'n} \ ,$$

and a similar exercise for for $B\,|\phi_n\rangle$ yields

$$\frac{1}{b_n} \sum_m \gamma_{mn}\tilde{\gamma}_{m'm} b_{m'} = \delta_{m'n} \ .$$

Anticipating a commutation calculation, let the operator B act on $A\,|\psi_n\rangle$

$$\begin{aligned} BA\,|\psi_n\rangle &= B \sum_{m'} \sum_m \gamma_{mn}\tilde{\gamma}_{m'm} a_{m'}\,|\psi_{m'}\rangle \\ &= \sum_{m'} \sum_\alpha \sum_m \gamma_{mn}\tilde{\gamma}_{m'm} a_{m'} \gamma_{\alpha m'} b_\alpha\,|\phi_\alpha\rangle \\ &= \sum_{m'} \sum_\rho \sum_\alpha \tilde{\gamma}_{\rho\alpha}\gamma_{\alpha m'} b_\alpha \delta_{nm'} a_n\,|\psi_\rho\rangle \\ &= \sum_{m'} \sum_\rho \delta_{\rho m'}\delta_{nm'} a_n b_\rho\,|\psi_\rho\rangle \\ &= a_n b_n\,|\psi_n\rangle \ , \end{aligned}$$

which simplifies nicely.

Similarly, we must let A act on $B\left|\psi_n\right\rangle$. Begin by calculating $B\left|\psi_n\right\rangle$ to get

$$\begin{aligned}B\left|\psi_n\right\rangle &= B\sum_m \gamma_{mn}\left|\phi_m\right\rangle \\ &= \sum_m \gamma_{mn} b_m \left|\phi_m\right\rangle \\ &= \sum_m \sum_{m'} \gamma_{mn} b_m \tilde{\gamma}_{m'm}\left|\psi_{m'}\right\rangle \ .\end{aligned}$$

The rest goes as

$$\begin{aligned}AB\left|\psi_n\right\rangle &= A\sum_m \sum_{m'} \gamma_{mn} b_m \tilde{\gamma}_{m'm}\left|\psi_{m'}\right\rangle \\ &= \sum_m \sum_{m'} \gamma_{mn} b_m \tilde{\gamma}_{m'm} a_{m'}\left|\psi_{m'}\right\rangle \\ &= \sum_{m'} \delta_{nm'} b_n a_{m'}\left|\psi_{m'}\right\rangle \\ &= a_n b_n \left|\psi_n\right\rangle \ .\end{aligned}$$

To finish, let us write the commutator of A and B to conclude

$$\begin{aligned}[AB]\left|\psi_n\right\rangle &= (AB - BA)\left|\psi_n\right\rangle \\ &= (a_n b_n - a_n b_n)\left|\psi_n\right\rangle = 0\,,\end{aligned}$$

which evaluates to *zero*. That is, we get a zero commutator of two operators whose eigenvectors share an orthonormal basis.

13 Matrix in Hilbert Subspace

13.1 Laplacian Operator

In the Hilbert space of functions $\mathcal{L}_2[-1,1]$, one can determine the components of the Laplacian operator

$$B = \partial_{xx}$$

of a subspace spanned by an orthonormal basis.

For instance, taking

$$\begin{aligned}\left|e_1\right\rangle &= \frac{1}{\sqrt{2}} \\ \left|e_2\right\rangle &= \frac{1}{\sqrt{2}}\left(\sin\left(\frac{\pi}{2}x\right) + \cos(\pi x)\right) \\ \left|e_3\right\rangle &= \frac{1}{\sqrt{2}}\left(\sin\left(\frac{\pi}{2}x\right) - \cos(\pi x)\right) \ ,\end{aligned}$$

the corresponding matrix B is calculated from

$$B = \begin{bmatrix}\left\langle e_1\right|B\left|e_1\right\rangle & \left\langle e_1\right|B\left|e_2\right\rangle & \left\langle e_1\right|B\left|e_3\right\rangle \\ \left\langle e_2\right|B\left|e_1\right\rangle & \left\langle e_2\right|B\left|e_2\right\rangle & \left\langle e_2\right|B\left|e_3\right\rangle \\ \left\langle e_3\right|B\left|e_1\right\rangle & \left\langle e_3\right|B\left|e_2\right\rangle & \left\langle e_3\right|B\left|e_3\right\rangle\end{bmatrix} \ ,$$

where

$$\left\langle e_j\right|B\left|e_k\right\rangle = \int_{-1}^{1} e_j^*(x)\,\partial_{xx} e_k(x)\,dx \ .$$

Carrying out each integral, find

$$B = \frac{-\pi^2}{8}\begin{bmatrix}0 & 0 & 0 \\ 0 & 5 & -3 \\ 0 & -3 & 5\end{bmatrix} \ ,$$

which is a Hermitian matrix by inspection.

13.2 Two Operators

In the (same) Hilbert space of functions $\mathcal{L}_2[-1,1]$, one can determine the components of a derivative operator $A = \partial_x$ and a Laplacian operator $B = \partial_{xx}$ of a subspace spanned by an orthonormal basis.

Using an example orthonormal basis

$$\begin{aligned}\left|e_1\right\rangle &= \frac{1}{\sqrt{2}} \\ \left|e_2\right\rangle &= \frac{1}{\sqrt{2}}\sin(\pi x) \\ \left|e_3\right\rangle &= \frac{1}{\sqrt{2}}\cos(\pi x) \ ,\end{aligned}$$

we first check that the subset is closed under operations A:

$$\begin{aligned}A\left|e_1\right\rangle &= \partial_x\left(\frac{1}{\sqrt{2}}\right) = 0 \\ A\left|e_2\right\rangle &= \partial_x\left(\frac{1}{\sqrt{2}}\sin(\pi x)\right) \\ &= \frac{\pi}{\sqrt{2}}\cos(\pi x) = \pi\left|e_3\right\rangle \\ A\left|e_3\right\rangle &= \partial_x\left(\frac{1}{\sqrt{2}}\cos(\pi x)\right) \\ &= -\frac{\pi}{\sqrt{2}}\sin(\pi x) = -\pi\left|e_2\right\rangle\end{aligned}$$

Each nontrivial result can be written in terms of the original basis vectors. Therefore $A\left|x\right\rangle$, where $\left|x\right\rangle$ is an arbitrary linear combination of the basis vectors, will result in a vector in the same subspace. Expressing A as a matrix requires calculating $A_{jk} = \left\langle e_j\right|A\left|e_k\right\rangle$, resulting in

$$A = \pi\begin{bmatrix}0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0\end{bmatrix} \ ,$$

which is a *not* a Hermitian matrix.

Repeating the same exercise using $B = \partial_{xx}$ results (of course) in a Hermitian matrix

$$B = -\pi^2 \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} .$$

Interestingly, since the operation ∂_{xx} is two instances of ∂_x, it should follow that $AA = B$, easily checked by matrix multiplication:

$$\pi^2 \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} = -\pi^2 \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

13.3 Pauli Matrices

Consider the set of three 2×2 Hermitian matrices

$$\sigma_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$
$$\sigma_2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}$$
$$\sigma_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

called the *Pauli matrices*, where $i = \sqrt{-1}$. Interestingly, each matrix $(k = 1, 2, 3)$ has the property

$$\sigma_k \sigma_k = I ,$$

where I is the two-dimensional identity matrix. As a consequence, it follows that

$$\sigma_k^{2m} = I$$
$$\sigma_k^{2m+1} = \sigma_k$$

for integer m. Furthermore, we have

$$\sigma_2 \sigma_3 = i\sigma_1$$
$$\sigma_3 \sigma_1 = i\sigma_2$$
$$\sigma_1 \sigma_2 = i\sigma_3 .$$

Eigenvectors and Eigenvalues

Eigenvectors and eigenvalues of the Pauli matrices are determined by

$$\sigma_k \left|x\right\rangle = \lambda_k \left|x\right\rangle .$$

Working with σ_1 as an example, we write

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \lambda_1 \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} ,$$

giving two equations

$$x_2 = \lambda_1 x_1$$
$$x_1 = \lambda_1 x_2 ,$$

having two nontrivial branches $\lambda_1 = 1$ and $\lambda_1 = -1$, implying either $x_1 = x_2$, or respectively, $x_1 = -x_2$.

By standard means, find one normalized eigenvector per eigenvalue:

$$\lambda_1 = +1 \rightarrow \left|x^{(+)}\right\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$
$$\lambda_1 = -1 \rightarrow \left|x^{(-)}\right\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

Repeating the exercise on $\sigma_2 \left|y\right\rangle = \lambda_2 \left|y\right\rangle$, and a third time on $\sigma_3 \left|z\right\rangle = \lambda_3 \left|z\right\rangle$, we find the eigenvalues are always $\lambda_k = \pm 1$. The corresponding normalized eigenvectors turn out to be

$$\left|y^{(+)}\right\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} i \\ 1 \end{bmatrix}$$
$$\left|y^{(-)}\right\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ i \end{bmatrix}$$
$$\left|z^{(+)}\right\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$
$$\left|z^{(-)}\right\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ 1 \end{bmatrix} .$$

14 Functions of Operators

14.1 Motivation

Recall that the single time derivative operator $L = \partial_t$ as it appears in the problem $L\left|x\right\rangle = \lambda\left|x\right\rangle$ has exponentially-evolving eigenvectors $\left|x(t)\right\rangle = e^{\lambda t}\left|x_0\right\rangle$, where $\left|x_0\right\rangle$ is the initial state in a fixed orthonormal basis. The more general statement

$$\left|x(t)\right\rangle = e^{Lt}\left|x_0\right\rangle$$

leads to the same solution, easily shown by taking a time derivative:

$$\partial_t \left|x(t)\right\rangle = L\, e^{Lt}\left|x_0\right\rangle = L\left|x(t)\right\rangle$$

14.2 Time Evolution Operator

The combination e^{Lt} is generally known as a *time evolution* operator. The exponential function, despite having an operator in its argument, readily expands as

$$\begin{aligned} e^{Lt} &= \sum_{k=0}^{\infty} \frac{1}{k!} (Lt)^k \\ &= I + (Lt) + \frac{1}{2!}(Lt)^2 + \frac{1}{3!}(Lt)^3 + \cdots . \end{aligned}$$

Grouping even terms and odd terms separately, the above reads

$$e^{Lt} = \left(I + \frac{1}{2!}(Lt)^2 + \frac{1}{4!}(Lt)^4 + \cdots \right) + \left(Lt + \frac{1}{3!}(Lt)^3 + \frac{1}{5!}(Lt)^5 + \cdots \right) .$$

Making the substitution $L = -i\tilde{H}$, we further have

$$e^{-i\tilde{H}t} = \left(I - \frac{1}{2!}\left(\tilde{H}t\right)^2 + \frac{1}{4!}\left(\tilde{H}t\right)^4 - \cdots \right) - i\left(\tilde{H}t - \frac{1}{3!}\left(\tilde{H}t\right)^3 + \frac{1}{5!}\left(\tilde{H}t\right)^5 - \cdots \right) ,$$

simplifying nicely to

$$e^{-i\tilde{H}t} = \cos\left(\tilde{H}t\right) - i\sin\left(\tilde{H}t\right) ,$$

giving away two more functions where an operator may naturally embed.

14.3 Pauli Matrix Operators

In the special case that $\tilde{H}$ is equal to any of the Pauli matrices $\{\sigma_k\}$ with $k = 1, 2, 3$ up to a proportionality constant μ such that $\tilde{H} = \mu\sigma_k$, the above reduces to

$$e^{-i\sigma_k \mu t} = I\cos(\mu t) - i\sigma_k \sin(\mu t) ,$$

which removes the operator from any infinite series. Explicitly:

$$e^{-i\sigma_1 \mu t} = \begin{bmatrix} \cos(\mu t) & -i\sin(\mu t) \\ -i\sin(\mu t) & \cos(\mu t) \end{bmatrix}$$

$$e^{-i\sigma_2 \mu t} = \begin{bmatrix} \cos(\mu t) & -\sin(\mu t) \\ -\sin(\mu t) & \cos(\mu t) \end{bmatrix}$$

$$e^{-i\sigma_3 \mu t} = \begin{bmatrix} e^{-i\mu t} & 0 \\ 0 & e^{i\mu t} \end{bmatrix}$$

15 Unitary Operators

Consider a vector space $\mathcal{V}$ admitting two different sets of basis vectors $\{|e_j\rangle\}$ and $\{|\tilde{e}_j\rangle\}$. In terms of coefficients x_j and $\tilde{x}_j$, a given vector $|x\rangle$ is a linear combination in each basis:

$$|x\rangle = \sum_j x_j |e_j\rangle = \sum_j \tilde{x}_j |\tilde{e}_j\rangle$$

Any coefficient(s) x_j, which exist in the vector space $\tilde{\mathcal{V}}$, can be isolated by exploiting the orthogonality between each $|e_j\rangle$ such that

$$x_j = \langle e_j|x\rangle = \sum_k \langle e_j| x_k |\tilde{e}_k\rangle = \sum_k \tilde{x}_k \langle e_j|\tilde{e}_k\rangle ,$$

which applies similarly to $\tilde{x}_j$:

$$\tilde{x}_j = \langle \tilde{e}_j|x\rangle = \sum_k \langle \tilde{e}_j| x_k |e_k\rangle = \sum_k x_k \langle \tilde{e}_j|e_k\rangle$$

15.1 Unitary Matrix

In component notation, the above has implicated two matrices

$$U_{jk} = \langle \tilde{e}_j|e_k\rangle$$
$$\tilde{U}_{jk} = \langle e_j|\tilde{e}_k\rangle ,$$

which can be interpreted as a set of coordinate transformation matrices that carry $\{x_j\} \in \tilde{\mathcal{V}} \to \{\tilde{x}_j\} \in \tilde{\mathcal{V}}$, and vice versa.

These matrices are intricately related, which we first notice by the observation

$$\tilde{U}_{jk} = \langle e_j|\tilde{e}_k\rangle = \overline{\langle \tilde{e}_k|e_j\rangle} = U^*_{kj} = U^\dagger_{jk} ,$$

telling us each matrix is the other's Hermitian conjugate (dropping component notation):

$$\tilde{U} = U^\dagger$$
$$U = \tilde{U}^\dagger$$

Now, the double transformation $\{x_j\} \to \{\tilde{x}_j\} \to \{x_j\}$ (and vice versa) tells us that each combination $\tilde{U}U$ and $U\tilde{U}$ is an identity matrix:

$$\tilde{U}U = I$$
$$U\tilde{U} = I$$

Combining the two previous results yields four identities:

$$U^\dagger U = I$$
$$UU^\dagger = I$$
$$\tilde{U}\tilde{U}^\dagger = I$$
$$\tilde{U}^\dagger\tilde{U} = I$$

Any operator satisfying the above equations is called *unitary*.

15.2 Effect on Inner Product

An important consequence of unitary operators arises when dealing with the inner product of two vectors $|x\rangle, |y\rangle \in \mathcal{V}$. Calculating the inner product of $U|x\rangle$ and $U|y\rangle$, we find

$$\langle Ux|Uy\rangle = \langle U^\dagger Ux|y\rangle = \langle Ix|y\rangle = \langle x|y\rangle ,$$

indicating that the inner product is unaffected by the operations.

In most applications (namely in two- and three-dimensional space), unitary operations correspond to rotations and reflections of the coordinates $\{x_j\} \in \tilde{\mathcal{V}}$.

15.3 Effect on Basis Vectors

Next, we examine how operator U acts directly on the basis vectors $\{|e_j\rangle\} \in \mathcal{V}$. Supposing we set $|x\rangle = |e_j\rangle$ and $|y\rangle = |e_k\rangle$, we find

$$\langle Ue_j|Ue_k\rangle = \langle e_j|e_k\rangle = \delta_{jk} \ ,$$

telling us that the combinations $|Ue_j\rangle$ and $|Ue_k\rangle$ are members of a second orthogonal basis $\{|g_j\rangle\}$ that can be generated from the original $\{|e_j\rangle\}$ via

$$\begin{aligned} |g_j\rangle &= U\,|e_j\rangle \\ j &= 1, 2, 3, \ldots, N \ . \end{aligned}$$

As a matrix, recall that the operator U can be written as

$$U = \sum_{jk} |e_j\rangle\, U_{jk}\, \langle e_k| = \sum_{jk} |e_j\rangle\, \langle \tilde{e}_j|e_k\rangle\, \langle e_k| \ ,$$

where the completeness relation

$$I = \sum_k |e_k\rangle\, \langle e_k|$$

reduces the above to

$$U = \sum_j |e_j\rangle\, \langle \tilde{e}_j| \ .$$

Right away, we find

$$U\,|\tilde{e}_j\rangle = \sum_j |e_j\rangle\, \langle \tilde{e}_j|\tilde{e}_j\rangle = |e_j\rangle \ .$$

Similarly, we deduce

$$\tilde{U} = \sum_j |\tilde{e}_j\rangle\, \langle e_j| \ ,$$

which leads to

$$\tilde{U}\,|e_j\rangle = |\tilde{e}_j\rangle \ .$$

In other words, the matrix U takes the jth vector from the primed basis and churns out the jth vector from the unprimed basis. Or, the $\tilde{U}$ matrix takes the jth vector from the unprimed basis and computes the corresponding primed vector.

In component form, these results read:

$$|e_k\rangle = U\,|\tilde{e}_k\rangle = \sum_{ij} |\tilde{e}_i\rangle\, U_{ij}\, \langle \tilde{e}_j|\tilde{e}_k\rangle = \sum_j U_{jk}\,|\tilde{e}_j\rangle$$

$$|\tilde{e}_k\rangle = \tilde{U}\,|e_k\rangle = \sum_{ij} |e_i\rangle\, \tilde{U}_{ij}\, \langle e_j|e_k\rangle = \sum_j \tilde{U}_{jk}\,|e_j\rangle$$

Project $\langle \tilde{e}_j|$ into the first equation and $\langle e_j|$ into the second to recover the component form of each matrix:

$$\begin{aligned} U_{jk} &= \langle \tilde{e}_j|e_k\rangle = \langle \tilde{e}_j|\, U\, |\tilde{e}_k\rangle \\ \tilde{U}_{jk} &= \langle e_j|\tilde{e}_k\rangle = \langle e_j|\, \tilde{U}\, |e_k\rangle \end{aligned}$$

Sanity Check

For a sanity check, let us apply U to a vector $|x\rangle \in \mathcal{V}$. Calculating this, we have

$$\begin{aligned} U\,|x\rangle &= \sum_{jk} x_k\,|e_j\rangle\, \langle \tilde{e}_j|e_k\rangle \\ &= \sum_j \left(\sum_k U_{jk} x_k \right) |e_j\rangle = |x'\rangle \ , \end{aligned}$$

where

$$x'_j = \sum_k U_{jk} x_k$$

is the rotated vector component in the same basis $\{|e_j\rangle\}$.

15.4 Rotations

Two Dimensions

The simplest nontrivial case involving unitary operators addresses rotations in the two-dimensional plane. Consider a Cartesian space spanned by the orthonormal basis $|e_1\rangle = \hat{x}$, $|e_2\rangle = \hat{y}$. A second orthonormal basis $\{|\tilde{e}_j\rangle\}$ is oriented at angle ϕ with respect with respect to the original. In particular:

$$\begin{aligned} |\tilde{e}_1\rangle &= \cos(\phi)\,|e_1\rangle + \sin(\phi)\,|e_2\rangle \\ |\tilde{e}_2\rangle &= -\sin(\phi)\,|e_1\rangle + \cos(\phi)\,|e_2\rangle \end{aligned}$$

Using the above formula for U_{jk} applied to standard two-dimensional geometry, we quickly find the components of U to be

$$\begin{aligned} U_{xx} &= \langle \tilde{e}_1|e_1\rangle = \cos\phi \\ U_{xy} &= \langle \tilde{e}_1|e_2\rangle = -\sin\phi \\ U_{yx} &= \langle \tilde{e}_2|e_1\rangle = \sin\phi \\ U_{yy} &= \langle \tilde{e}_2|e_2\rangle = \cos\phi \ , \end{aligned}$$

and similarly for $U^\dagger = \tilde{U}$:

$$\begin{aligned} \tilde{U}_{xx} &= \langle e_1|\tilde{e}_1\rangle = \cos\phi \\ \tilde{U}_{xy} &= \langle e_1|\tilde{e}_2\rangle = \sin\phi \\ \tilde{U}_{yx} &= \langle e_2|\tilde{e}_1\rangle = -\sin\phi \\ \tilde{U}_{yy} &= \langle e_2|\tilde{e}_2\rangle = \cos\phi \end{aligned}$$

There is an important different between the two operators. Matrix U calculates the components of a rotated vector in a fixed basis. Matrix $U^\dagger = \tilde{U}$ transforms the coordinates of a fixed vector when the basis is rotated.

Three Dimensions

In three dimensions, we extend the two-dimensional case to write three matrices (presumably named after aviation terms)

$$R_z(\alpha) = \begin{bmatrix} \cos(\alpha) & -\sin(\alpha) & 0 \\ \sin(\alpha) & \cos(\alpha) & 0 \\ 0 & 0 & 1 \end{bmatrix} = \text{‘yaw’}$$

$$R_y(\beta) = \begin{bmatrix} \cos(\beta) & 0 & \sin(\beta) \\ 0 & 1 & 0 \\ -\sin(\beta) & 0 & \cos(\beta) \end{bmatrix} = \text{‘pitch’}$$

$$R_x(\gamma) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\gamma) & -\sin(\gamma) \\ 0 & \sin(\gamma) & \cos(\gamma) \end{bmatrix} = \text{‘roll’} ,$$

where a general rotation of a vector in three dimensions is the product

$$R = R_z(\alpha) R_y(\beta) R_x(\gamma) \ .$$

Due to commutivity rules, the order in which the matrices are applied *does* affect the result. Formally, the above matrices correspond to an *intrinsic* rotation, having *Tait–Bryan* angles α, β, γ. The domain restriction on each angle is as follows:

$$\begin{aligned} 0 &\leq \alpha < 2\pi \\ 0 &\leq \beta < \pi \\ 0 &\leq \gamma < 2\pi \end{aligned}$$

15.5 Effect on Operator

Now we examine what happens to the components of an operator A when undergoing a change of basis vectors summarized by

$$|x\rangle = \sum_j x_j |e_j\rangle = \sum_j \tilde{x}_j |\tilde{e}_j\rangle \ .$$

For some vector $|x\rangle$ in the vector space $\mathcal{V}$, the operation $A|x\rangle$ yields a vector $|y\rangle$, also in $\mathcal{V}$. Expressing this calculation in two different orthonormal bases, we have

$$\begin{aligned} \sum_j A_{ij} x_j &= y_i \\ \sum_j \tilde{A}_{ij} \tilde{x}_j &= \tilde{y}_i \ . \end{aligned}$$

Substituting

$$\begin{aligned} x_j &= \sum_k \tilde{x}_k \langle e_j | \tilde{e}_k \rangle \\ y_i &= \sum_k \tilde{y}_k \langle e_i | \tilde{e}_k \rangle \end{aligned}$$

into the first equation, we end up with

$$A\tilde{U} |\tilde{x}\rangle = \tilde{U} |\tilde{y}\rangle \ ,$$

where multiplying both sides by U gives

$$UA\tilde{U} |\tilde{x}\rangle = U\tilde{U} |\tilde{y}\rangle = I |\tilde{y}\rangle = |\tilde{y}\rangle \ .$$

Meanwhile, we already already know $|\tilde{y}\rangle = \tilde{A} |\tilde{x}\rangle$ by construction, and we conclude

$$\tilde{A} = UA\tilde{U} \ .$$

Of course, this result can be attained more directly by substituting

$$A = \sum_{i'} \sum_{j'} |e_{i'}\rangle A_{i'j'} \langle e_{j'}|$$

into $\tilde{A}_{ij} = \langle \tilde{e}_i | A | \tilde{e}_j \rangle$.

16 Differential Equations

16.1 Schrodinger Equation

The chief equation of quantum mechanics is conveniently framed as an eigenvalue problem. As such, the famous Schrodinger equation reads

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

where $i = \sqrt{-1}$, $\hbar$ is Planck's constant, and H is a Hermitian operator called the *Hamiltonian* of size $n \times n$. The symbol $|\psi(t)\rangle$ is the *quantum state vector*, which like any other vector, resolves to components provided an orthonormal basis exists:

$$|\psi(t)\rangle = (\psi_1(t), \psi_2(t), \psi_3(t), \dots)$$

Naming the supporting orthornormal basis $|e_j\rangle$, we explicitly have

$$|\psi(t)\rangle = \sum_j \psi_j(t) |e_j\rangle$$

and

$$H = \sum_j \sum_k |e_j\rangle H_{jk} \langle e_k| \ ,$$

letting us state the problem in component form:

$$\begin{aligned} i\hbar \sum_j \frac{d}{dt} \psi_j(t) |e_j\rangle & \\ &= \sum_j \sum_k \sum_{j'} |e_j\rangle H_{jk} \psi_{j'}(t) \langle e_k | e_{j'} \rangle \\ &= \sum_j \sum_k H_{jk} \psi_k(t) |e_j\rangle \end{aligned}$$

With the j-sum present on each side, the above reduces to:

$$i\hbar \frac{d}{dt} \psi_j(t) = \sum_k H_{jk} \psi_k(t)$$

TISE

Assuming that H admits n eigenvalues E_j and n eigenvectors represented by $|\phi^{(j)}\rangle$, we may also write the *time-independent Schrodinger equation*, or *TISE*:

$$H\,|\phi^{(j)}\rangle = E_j\,|\phi^{(j)}\rangle$$

Next, since H is a Hermitian operator, its eigenvectors $|\phi^{(j)}\rangle$ form an orthonormal basis for which the quantum state vector can be expressed as a linear combination:

$$|\psi\,(t)\rangle = \sum_j C_j\,(t)\,|\phi^{(j)}\rangle$$

Applying the H-operator to both sides of the above, thereby writing the Schrodinger equation, gives for any given index j,

$$i\hbar\frac{d}{dt}C_j\,(t) = E_j C_j\,(t)\ ,$$

solved by

$$C_j\,(t) = C_j\,(t=0)\,e^{-iE_j t/\hbar}\ .$$

Note that the initial value of each $C_n\,(0)$ is calculated from the initial condition $|\psi\,(0)\rangle$ according to

$$\langle\phi^{(k)}|\psi\,(0)\rangle = \sum_j C_j\,(t)\,\langle\phi^{(k)}|\phi^{(j)}\rangle = C_k\ ,$$

where

$$C_k = \langle\phi^{(k)}|\psi\,(0)\rangle = \sum_j \left(\phi_j^{(k)}\right)^*\psi_j\,(0)\ .$$

The evolution of the quantum state vector can thus be written

$$|\psi\,(t)\rangle = \sum_j C_j\,(0)\,e^{-iE_j t/\hbar}\,|\phi^{(j)}\rangle\ .$$

16.2 Hamiltonian Matrix

Consider a two-component vector that presumes the existence of an orthonormal basis

$$|u\,(t)\rangle = \begin{bmatrix} u_1\,(t) \\ u_2\,(t) \end{bmatrix}$$

that relates to the time derivative operator by

$$i\partial_t\,|u\,(t)\rangle = \hat{H}\,|u\,(t)\rangle\ ,$$

where $\hat{H}$ is the dimensionalized *Hamiltonian matrix*, having form

$$\hat{H} = \begin{bmatrix} 0 & \delta \\ \delta & 0 \end{bmatrix}\ ,$$

with δ being constant.

Solving the eigenvalue problem

$$\hat{H}\,|x^{(j)}\rangle = \lambda_j\,|x^{(j)}\rangle$$

for this case, we quickly find two eigenvalues

$$\begin{aligned} \lambda_+ &= \delta \\ \lambda_- &= -\delta\ , \end{aligned}$$

and two corresponding eigenvectors

$$\begin{aligned} |x^{(+)}\rangle &= \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix} \\ |x^{(-)}\rangle &= \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \end{bmatrix}\ . \end{aligned}$$

With the eigenvectors for the operator on hand, the vector $|u\,(t)\rangle$ can be expressed as a linear combination

$$\begin{aligned} |u\,(t)\rangle &= C_+\,(t)\,|x^{(+)}\rangle + C_-\,(t)\,|x^{(-)}\rangle \\ &= \sum_j C_j\,(t)\,|x^{(j)}\rangle\ . \end{aligned}$$

Apply $\hat{H}$ to both sides to distill a differential equation

$$i\frac{\partial}{\partial t}C_j\,(t) = C_j\,(t)\,\lambda_j\ ,$$

solved by

$$C_j\,(t) = C_j\,(t=0)\,e^{-i\lambda_j t}\ .$$

The updated general solution now reads

$$|u\,(t)\rangle = C_+\,(0)\,e^{-i\delta t}\,|x^{(+)}\rangle + C_-\,(0)\,e^{i\delta t}\,|x^{(-)}\rangle\ ,$$

where the coefficients $C_\pm\,(0)$ are determined by the initial conditions of the system. Since the eigenvectors $|x^{(\pm)}\rangle$ form an orthonromal basis, the coefficients are easily isolated from $|u\,(t=0)\rangle$:

$$C_j\,(0) = \langle x^{(j)}|u\,(0)\rangle$$

Note finally that the exponential terms can be traded for trigonometric terms by Euler's formula to give

$$\begin{aligned} |u\,(t)\rangle = &\frac{1}{\sqrt{2}}\cos\left(\delta t\right)\begin{bmatrix} C_+ + C_- \\ C_+ - C_- \end{bmatrix} \\ &+ \frac{i}{\sqrt{2}}\sin\left(\delta t\right)\begin{bmatrix} -C_+ + C_- \\ -C_+ - C_- \end{bmatrix}\ . \end{aligned}$$

16.3 Damped Harmonic Oscillator

The differential equation that governs the damped harmonic oscillator reads

$$\frac{d^2}{dt^2}x(t) - b\frac{d}{dt}x(t) + \omega_0^2 x(t) = 0 ,$$

which is a second-order equation. However, this problem can be turned into a system of first-order equations by defining a vector

$$|u(t)\rangle = \begin{bmatrix} x(t) \\ dx(t)/dt \end{bmatrix}$$

along with a matrix

$$A = \begin{bmatrix} 0 & 1 \\ -\omega_0^2 & -b \end{bmatrix} ,$$

so the problem may be rewritten

$$\frac{d}{dt}|u(t)\rangle = A|u(t)\rangle .$$

Solving the eigenvalue problem

$$A|q^{(j)}\rangle = \lambda_j |q^{(j)}\rangle ,$$

we quickly find

$$\lambda = -\frac{b}{2} \pm \sqrt{\frac{b^2}{4} - \omega_0^2} ,$$

with corresponding eigenvectors

$$|q^{(+)}\rangle = \frac{1}{\sqrt{\lambda_+^2 + 1}} \begin{bmatrix} 1 \\ \lambda_+ \end{bmatrix}$$
$$|q^{(-)}\rangle = \frac{1}{\sqrt{\lambda_-^2 + 1}} \begin{bmatrix} 1 \\ \lambda_- \end{bmatrix} .$$

Note that the eigenvectors $|q^{(\pm)}\rangle$ are linearly independent but not orthogonal. Regardless of this, the general solution can be written as a linear combination

$$\begin{aligned} |u(t)\rangle &= C_+(t)|q^{(+)}\rangle + C_-(t)|q^{(-)}\rangle \\ &= \sum_j C_j(t)|q^{(j)}\rangle . \end{aligned}$$

Applying the A-operator to both sides of the above produces

$$\begin{aligned} \sum_j \frac{d}{dt}(C_j(t))|q^{(j)}\rangle &= A\sum_j C_j(t)|q^{(j)}\rangle \\ &= \sum_j C_j(t)\lambda_j|q^{(j)}\rangle , \end{aligned}$$

immediately implying

$$\frac{d}{dt}C_j(t) = C_j(t)\lambda_j ,$$

solved by

$$C_j(t) = C_j(t=0)e^{\lambda_j t} .$$

The general solution now reads

$$|u(t)\rangle = \sum_j C_j(0)e^{\lambda_j t}|q^{(j)}\rangle .$$

Setting $t = 0$ in the above gives

$$|u(0)\rangle = \sum_j C_j(0)|q^{(j)}\rangle .$$

Note however that the basis vectors $|q^{(k)}\rangle$ are not orthogonal, thus the coefficients $C_j(0)$ cannot be isolated by taking the inner product with $\langle q^{(k)}|$. To proceed, write the above in component form to get

$$u_k(0) = \sum_j C_j(0)q_k^{(j)} .$$

This is a linear system of the form $A|x\rangle = |b\rangle$ with $|u(0)\rangle$ playing the role of $|b\rangle$, the vector components $q_k^{(j)}$ serving as matrix components, and the components $C_j(0)$ corresponding to $|x\rangle$.