# PHYSICS 204A: MATH METHODS

IAN LIM
LAST UPDATED OCTOBER 2, 2019

These notes were taken for Physics 204A, *Math Methods*, as taught by Nemanja Kaloper at the University of California, Davis in fall quarter 2019. I live-TEXed them using Overleaf, and as such there may be typos; please send questions, comments, complaints, and corrections to itlim@ucdavis.edu.

Many thanks to Arun Debray for the LATEX template for these lecture notes: as of the time of writing, you can find him at https://web.ma.utexas.edu/users/a.debray/.

## Contents



Lecture 1.

## Wednesday, September 25, 2019

Reading assignment: read Ch. 2 and 3 of the course text (Arfken/Weber). This is basic linear algebra and vector analysis.

The purpose of this course is to learn formal aspects of quantum mechanics. We'll focus on doing analysis in Hilbert space. It's a remarkable fact about the natural world that most of our physical world is well-approximated by linear systems.

In the simplest form, we may think of vectors as arrays of numbers,

$$(v_1, v_2, \ldots). \tag{1.1}$$

But we can also think of some real function $f(x) : \mathbb{R} \to \mathbb{R}$ as a collection of numbers too, just by taking its values at arbitrarily close points.

**Definition 1.2.** A *linear vector space* over a field $F$, denoted $L(F)$, is a set $\{|v\rangle\}$ with an addition operation $+$ such that

- for $|v\rangle, |u\rangle \in L, |u\rangle + |v\rangle = |w\rangle \in L$ (closure)
- for $c \in F$, $c|v\rangle \in L$ (scalar multiplication).

These axioms directly imply that any linear combination of vectors in $L$ is also in the vector space:

$$c_v|v\rangle + c_u|u\rangle \in L. \tag{1.3}$$

This leads us naturally to the notion of *linear (in)dependence*. Suppose we take some vectors $|v_k\rangle \in L$ and make a linear combo,

$$\sum_k c_k|v_k\rangle. \tag{1.4}$$

**Definition 1.5.** A set of vectors $\{|v_k\rangle\}$ is *linearly dependent* if there exists some $\{c_k\}$ not all zero such that

$$\sum_k c_k|v_k\rangle = |0\rangle, \tag{1.6}$$

and such that $|0\rangle \notin \{|v_k\rangle\}$.

We need this last condition because otherwise we could simply take the coefficient of the $|0\rangle$ vector to be 1 and then arrive at a trivial solution.

**Definition 1.7.** If a set of vectors is not linearly dependent, it is *linearly independent*.

The next question we might ask is as follows: what is the size of the biggest set of linearly independent vectors we can construct for a given vector space?

**Definition 1.8.** The maximum number of linearly independent vectors associated to a given vector space is called the *dimension*.

**Example 1.9.** We may consider an infinitely differentiable ($C^\infty$) function. It has a Taylor expansion

$$f(x) = \sum \frac{f^{(n)}(0}{n!} x^n, \tag{1.10}$$

which we may think of as an expansion in the basis $(1, x, x^2, \ldots)$.

So this is a vector space with countably infinite dimension. But we can have uncountably infinite-dimensional spaces too, e.g. the space of Fourier-transformable functions in a basis $e^{ikx}, k \in \mathbb{R}$. These factors are not just linearly independent; introducing an appropriate inner product, they are orthogonal.

It follows that for a vector space $L_D$ of dimension $D$, any set with more than $D$ vectors must be linearly dependent, i.e. $\exists \bar{c}_v, \bar{c}_k$ not all zero such that

$$\bar{c}_v |v\rangle + \sum_k \bar{c}_k |v_k\rangle = 0. \tag{1.11}$$

Moreover $\bar{c}_v \neq 0$ or else the original set $|v_k\rangle$ would be linearly dependent. Hence we can divide through, define $\hat{c}_k = \bar{c}_k / \bar{c}_v$, and write

$$|v\rangle = \sum_k \hat{c}_k |v_k\rangle. \tag{1.12}$$

That is, we have *decomposed* a general vector $|v\rangle$ in terms of its components $\hat{c}_k$ with respect to a basis $|v_k\rangle$.

We might now be interested in adding more structure to our vector space. Consider $L_D$ with $|v\rangle, |u\rangle$.

**Definition 1.13.** We define an *inner product* by

$$\langle v|u\rangle : (|v\rangle, |u\rangle) \to F \tag{1.14}$$

as a map from the input vectors to the field over which the vector space is defined, with the following properties:

- $\langle u|(\lambda|v_1\rangle + \mu|v_2\rangle) = \lambda\langle u|v_1\rangle + \mu\langle u|v_2\rangle$ (linearity)
- $\langle v|u\rangle = \langle u|v\rangle^*$
- $\langle v|v\rangle \geq 0$, with equality only for the zero vector (positive semi-definite).

If we choose a basis $\{|v_k\rangle\}$, then if our vectors $|u\rangle, |v\rangle$ have some expansion in this basis then by linearity that

$$\langle u|v\rangle = \sum_k \hat{c}_k^v \langle u|v_k\rangle, \tag{1.15}$$

and we can expand each of these inner products as

$$\langle u|v_k\rangle = (\langle v_k|u\rangle)^* = \sum_n c_n^{u*} \langle v_n|v_k\rangle. \tag{1.16}$$

It follows that we can write a general inner product as

$$\langle u|v\rangle = \sum_{n,k} = c_k^v c_n^{u*} \langle v_n|v_k\rangle. \tag{1.17}$$

Moreover if we could choose a nice basis which had a special property of *orthogonality* or better yet *orthonormality*, we could reduce this to a single sum

$$\langle u|v\rangle = \sum_k c_k^v c_k^* \tag{1.18}$$

in terms of the components alone.

Lecture 2.

## Monday, September 30, 2019

Last time, we wrote a general form for the dot (inner) product,

$$\langle u|v\rangle = \sum_{n,m} c_n^{*u} c_m^v \langle \phi_n|\phi_m\rangle, \tag{2.1}$$

where the $|\phi\rangle$s are basis vectors and $u, v$ have expansions

$$|u\rangle = \sum_n c_n^u |\phi_n\rangle. \tag{2.2}$$

This is some quadratic form (it depends only quadratically on the components). And indeed it would be very nice if we could define $\langle \phi_N|\phi_m\rangle = \delta_{nm}$, so that our double-sum collapses to a single sum.

**Orthonormality** Let us suppose we start with a basis $\{|\phi_n\rangle\}$ for a vector space $L_D$. We shall show that we can construct a new basis $\{|\chi_n\rangle\}^D$ such that $\langle \chi_m|\chi_n\rangle = \delta_{mn}$ has the desired property.

WLOG let us number the basis vectors $|\phi_1\rangle, |\phi_2\rangle, \ldots$ and consider some inner products. The inner product

$$\langle \phi_1|\phi_1\rangle = N_1 \tag{2.3}$$

is some value $N_1$. If $N_1 = 1$ then we are done; otherwise, define

$$|\chi_1\rangle \equiv \frac{1}{\sqrt{N_1}} |\phi_1\rangle \tag{2.4}$$

so that

$$\langle \chi_1|\chi_1\rangle = \frac{\langle \phi_1|\phi_1\rangle}{N_1} = 1. \tag{2.5}$$

Hence $|\chi_1\rangle$ is a unit vector.

Consider the next vector $|\phi_2\rangle$. If

$$\langle \chi_1|\phi_2\rangle = 0, \tag{2.6}$$

then we can normalize and get

$$|\chi_2\rangle = \frac{|\phi_2\rangle}{\sqrt{N_2}}, \tag{2.7}$$

where $N_2 = \langle \phi_2|\phi_2\rangle$. Otherwise, we first subtract off the projection of the first normalized vector,

$$|\hat{\chi}_2\rangle = |\phi_2\rangle - \langle \chi_1|\phi_2\rangle |\chi_1\rangle, \tag{2.8}$$

so that

$$\langle \chi_1|\hat{\chi}_2\rangle = \langle \chi_1|\phi_2\rangle - \langle \chi_1|\phi_2\rangle \underbrace{\langle \chi_1|\chi_1\rangle}_{=1} = 0. \tag{2.9}$$

Hence by our definition, $|\chi_1\rangle$ and $|\hat{\chi}_2\rangle$ are orthogonal and we can just normalize. Defining

$$\hat{N}_2 = \langle \hat{\chi}_2|\hat{\chi}_2\rangle, \tag{2.10}$$

we have

$$|\chi_2\rangle = \frac{|\hat{\chi}_2\rangle}{\sqrt{\hat{N}_2}}, \tag{2.11}$$

which is a unit vector and normal to $|\chi_1\rangle$.

We continue by induction, subtracting off projections and normalizing. This is the *Gram-Schmidt procedure*.[1] Notice also that because $\langle \phi_n|\phi_m\rangle = \langle \phi_m|\phi_n\rangle^*$, we can consider values of $m, n$ to be entries in a Hermitian matrix. Recalling that $\langle u|u\rangle \geq 0$, the norm is perfectly well-defined and indeed we can see that orthonormalization is equivalent to diagonalizing a Hermitian matrix.

[1] Note that the procedure is a little more subtle in the infinite-dimensional case.

**More inner products** On a function space, we can define an inner product

$$\langle f|g\rangle = \int_{-\infty}^{\infty} dx\, f^*(x)g(x). \tag{2.12}$$

These inner products come with strings attached; our functions usually have to satisfy some integrability properties in order for the inner products to be well-defined. Often the functions we're interested in come from differential equations. Most of the ones we encounter in physics are second-order so these functions ought to be twice-differentiable.[2] And we should also require that our functions are square-integrable so that the integral is well-defined.

We can then define

$$\begin{aligned}\langle u|u\rangle &= \sum_{n,m} c_n^* c_m \langle \phi_n|\phi_m\rangle \\ &= \sum_n |c_n|^2. \end{aligned}\tag{2.13}$$

And this is none other than the generalization of Pythagoras's theorem.

**Schwarz inequality** From the axioms, we can prove the following inequality.

$$|\langle f|g\rangle|^2 \le \langle f|f\rangle\langle g|g\rangle. \tag{2.14}$$

We may define the linear combination

$$|f-\lambda g\rangle \tag{2.15}$$

and consider its norm (squared)

$$\langle f-\lambda g|f-\lambda g\rangle = \langle f|f\rangle - \lambda^*\langle g|f\rangle - \lambda\langle f|g\rangle + \lambda\lambda^*\langle g|g\rangle. \tag{2.16}$$

This is obviously non-negative, given the axioms. We can extremize this by taking derivatives with respect to $\lambda, \lambda^*$ (which we may treat as linearly independent, since they are complex) and find that

$$\lambda^* = \frac{\langle g|f\rangle}{\langle g|g\rangle}, \quad \lambda = \frac{\langle f|g\rangle}{\langle g|g\rangle}. \tag{2.17}$$

A bit of manipulation yields the Schwarz inequality.

Let us also note that we can in general translate between the ket notation and vector notation. For a ket vector $|u\rangle$ we can associate the bra (row) vector $\langle u| = (c_1^*, c_2^* \ldots, c_D^*)$. Then the vector inner product is the same as old-fashioned row-column multiplication.

**Bessel inequality** Suppose we rewrite the vector $|u\rangle$ in a weird way, as

$$|u\rangle = \sum_n{}' c_n^u|\phi_n\rangle + |\Delta u\rangle, \tag{2.18}$$

where we take some terms and separate them out (so the sum $\sum'$ omits some indices). We know that

$$|\Delta u\rangle = u - \sum_n{}' c_n^8|\phi_n\rangle \neq 0, \tag{2.19}$$

so that

$$0 < \langle \Delta u|\Delta u\rangle \tag{2.20}$$

$$= \langle u - \sum{}' c_n^u\phi_n | u - \sum{}' c_n^u \phi_n\rangle \tag{2.21}$$

$$= \langle u|u\rangle - \sum{}' |c_n^u|^2. \tag{2.22}$$

Check the cross-terms with the definition of $u$ to get this final term. Rearranging, we get the Bessel inequality, which says that

$$\langle u|u\rangle > \sum{}' |c_n^u|^2, \tag{2.23}$$

i.e. the norm of a vector is greater than the partial sums of the squares of the components.

[2] This is a little too strong, actually. They can have finitely many discontinuities and this is still okay.

**Linear operators** We are primarily interested in linear operators, i.e. linear maps from the vector space to itself obeying

$$A(\lambda|\phi\rangle + \mu|\chi\rangle) = \lambda A|\phi\rangle + \mu A|\chi\rangle. \tag{2.24}$$

We can define two operators to be equal if they have the same action on all vectors, i.e.

$$A|\phi\rangle = B|\phi\rangle \tag{2.25}$$

for all $\phi \in L$.

In particular there's a nice way that we can rewrite the identity operator, as

$$\mathbb{I} = \sum_n |\phi_n\rangle\langle\phi_n|. \tag{2.26}$$

Let's prove this: by definition, $\mathbb{I}|u\rangle = |u\rangle$. On the other side, we see that

$$\sum_n |\phi_n\rangle\langle u| =\rangle \sum_n c_n^u |\phi_n\rangle \equiv |u\rangle. \tag{2.27}$$

Provided that $\{|\phi_n\rangle\}$ is a complete basis, this operator is indeed the identity.

**The delta function** The Dirac delta function is defined in such a way that

$$\int_a^b f(t)\delta(x-t)dt = f(x), \tag{2.28}$$

provided that $x$ is in the interval $(a, b)$. We could think of this as an inner product, however. We have a function and its shadow on the delta function picks out a value. The delta function isn't properly square-integrable, but we may consider it as having a good inner product with functions in our function space (as the limit of some sequence of square-integrable functions, if you like).

Now we'll do something strange. Let us express the delta function in a function basis,

$$\delta(x-t) = \sum_n c_n(t)\phi_n(x). \tag{2.29}$$

The $t$ dependence must be in the coefficients since the functions themselves are just given. How do we find the coefficients? Just take the integral

$$\int dx\, \phi_m^*(x)\delta(x-t) = \int dx\, \sum_n c_n(t)\phi_n(x)\phi_n^*. \tag{2.30}$$

This is super easy to evaluate. On the RHS we have a Kronecker delta $\delta_{nm}$ by the orthonormality of the basis, and on the left side we have the evaluation of the basis vector $\phi_m^*$ at $t$, i.e.

$$c_m(t) = \phi_m^*(t). \tag{2.31}$$

Hence

$$\delta(x-t) = \sum_n \phi_n^*(t)\phi_n(x). \tag{2.32}$$

We can se that this had to be the case by substituting our expression for the delta function into an integral:

$$f(x) = \int dt\, \delta(x-t)f(t) = \sum_n \phi_n(x) \int dt\, \phi_n^*(t)f(t) \tag{2.33}$$

$$= \sum \langle\phi_n|f\rangle\phi_n(x), \tag{2.34}$$

which is none other than the components of $f$ in the basis $\phi_n$.

To make our discussion more concrete, let us consider analytic functions which have Taylor expansion

$$f(x) = \sum \frac{f^{(n)}(0)}{n!}x^n, \tag{2.35}$$

defined over the interval $[-1, 1]$. Hence $\{1, x, x^2, x^3, \ldots\}$ form a complete basis set for arbitrarily differentiable functions. They are certainly not orthogonal in general, e.g. $\int_{-1}^1 dx\, 1 \cdot x^2 \neq 0$. But we can make them orthonormal with Gram-Schmidt.

Under this inner product, we have

$$\int_{-1}^1 dx\, 1 \cdot 1 = 2, \tag{2.36}$$

so our first normalized vector is $1/\sqrt{2}$. We can check $x$:

$$\int_{-1}^{1} dx\, x \cdot x = 2/3, \tag{2.37}$$

so the next normalized vector is $\sqrt{3/2}x$. Continuing this way, we see that $x^2$ and $x$ are already orthogonal but

$$\int_{-1}^{1} dx\, \frac{1}{\sqrt{2}} x^2 = \frac{2}{3\sqrt{2}}, \tag{2.38}$$

so our first unit vector is not orthogonal to $x^2$. We can instead define

$$\hat{x}^2 = x^2 - \sqrt{\frac{2}{3}}. \tag{2.39}$$

which is now orthogonal to the first unit vector $1/\sqrt{2}$ and to the second unit vector $\sqrt{3/2}x$. We can normalize $\hat{x}^2$ and determine the third unit vector in this set, which is $\frac{3x^2}{2} - 1$ (we think).

Let us remark that the most general operators that can be diagonalized are *normal* operators, i.e. those satisfying

$$[A, A^\dagger] = 0. \tag{2.40}$$

Clearly, one set of operators that are not normal are the raising and lowering operators, whose commutator is $[a, a^\dagger] = 1$.

Lecture 3.

## Wednesday, October 2, 2019

Today we'll continue discussing operators. We've discussed the identity operator,

$$\mathbb{I} = \sum |\phi_i\rangle\langle\phi_i|, \tag{3.1}$$

which maps any vector into itself. More generally, we can define an operator as follows:

**Definition 3.2.** An *operator* is a map $A : |\psi\rangle \to |\bar{\psi}\rangle$ where $|\psi\rangle, |\bar{\psi}\rangle \in L_D$ are in the same vector space $L_D$.[3] A *linear operator* is an operator obeying the linearity property

$$A(\mu|\psi\rangle + \nu|\chi\rangle) = \mu A|\psi\rangle + \nu A|\chi\rangle. \tag{3.3}$$

If we have a set of linear operators, we may define an addition operation on operators as

$$(A + B)|\psi\rangle = A|\psi\rangle + B|\psi\rangle \tag{3.4}$$

and scalar multiplication as

$$(kA)|\psi\rangle = k(A|\psi\rangle). \tag{3.5}$$

Hence we can take linear combinations of linear operators as

$$kA + lB, \tag{3.6}$$

and under this definition we see that linear operators form a vector space.

But there's another way to combine operators, namely by *composition*. That is, given operators $A$ and $B$ we can define a new operator $AB$ defined by the composition

$$AB|\psi\rangle = A(B|\psi\rangle). \tag{3.7}$$

Composition must satisfy certain properties with respect to the other operations we've defined, namely distributivity with respect to addition:

$$A(B + C) = AB + AC. \tag{3.8}$$

Let us note that the product of operators (composition) is generally not commutative,

$$AB \neq BA \tag{3.9}$$

in general. We know this from matrix multiplication.

[3] The math-inclined among us may talk about the space of inner automorphisms on the vector space.

Alternately, we could define a composition rule using a commutator (bracket),

$$[A, B] = AB - BA. \tag{3.10}$$

It's not to hard to check that this rule also satisfies distributivity over addition. If we wished, we could also prove (by crunching through the commutators) the Jacobi identity,

$$[[A, B], C] + [[B, C], A] + [[C, A], B] = 0. \tag{3.11}$$

This is related to the Bianchi identity in differential geometry.

Under our standard composition rule, we can define *inverses*. That is, if an operator $A$ acts as

$$|\psi'\rangle = A|\psi\rangle, \tag{3.12}$$

then the inverse $A^{-1}$ (if it exists) is the operator such that

$$A^{-1}|\psi'\rangle = |0\rangle. \tag{3.13}$$

That is, $A^{-1}A|\psi\rangle = |\psi\rangle$.

As we know, not every operator is invertible. Consider the operator which just sends some vector to zero (i.e. it has a null eigenvector). If

$$A|\psi\rangle = 0 \tag{3.14}$$

then the inverse is not well-defined: $A^{-1}|0\rangle = ?$.

**Definition 3.15.** The adjoint $A^\dagger$ of an operator $A$ is defined by

$$\langle\chi|A^\dagger|\psi\rangle = \langle\psi|A|\chi\rangle^*. \tag{3.16}$$

**Definition 3.17.** If an operator is self-adjoint, $H^\dagger = H$, then we call it Hermitian.

Notice that

$$\langle\psi|H|\psi\rangle = \langle\psi|H^\dagger|\psi\rangle = \langle\psi|H|\psi\rangle^* \tag{3.18}$$

by the definition of the adjoint and hermiticity. Therefore the diagonal matrix elements of $H$ are *real numbers*, i.e. their eigenvalues are real.

We could have also taken an operator which was *anti-Hermitian*, $H^\dagger = -H$, which implies that the diagonal elements are instead purely imaginary by the same argument.

**Definition 3.19.** A *unitary* operator is an operator obeying the property

$$U^{-1} = U^\dagger. \tag{3.20}$$

It's clear that we can restrict to the real case, in which case Hermitian matrices become symmetric matrices and anti-Hermitian matrices become antisymmetric (sometimes called skew-symmetric). Our unitary matrices reduce to orthogonal matrices.

**Example 3.21.** Consider the Hilbert space of smooth square-integrable functions over the real line, $f \in L$. Our inner product is the integral

$$\int dx\, f^*(x)g(x). \tag{3.22}$$

Define the operator $D = -\frac{d}{dx}$. What is the adjoint $D^\dagger$? We have

$$\int f^*(-\frac{d}{dx})g = f^*g|_a^b + \int (\frac{d}{dx}f^*)g. \tag{3.23}$$

The boundary term vanishes based on the boundary conditions, i.e. given that $f, g$ vanish at infinity. Taking the complex conjugate to get the adjoint, what is left is

$$-\left[\int (\frac{d}{dx}f^*)g\right]^* = -\int g^* Df. \tag{3.24}$$

So $D$ is not Hermitian but $iD$ is (it adds a minus sign to fix the sign in the integration by parts).

Note also that when the integration region is finite, the boundary conditions become nontrivial. However, if $f$ and $g$ vanish at the boundary (e.g. $[-1, 1]$) then we restore hermiticity.

Recall we said that we could assign matrix elements to an operator with respect to some set of vectors,

$$\langle\chi|A|\psi\rangle. \tag{3.25}$$

In fact, it's a Sisyphean task to do this for all sets of vectors, but fortunately (thanks to linearity) it suffices to compute the matrix elements in some (complete) basis. With a basis $|\psi_i\rangle$ we can define

$$\langle\psi_n|A|\psi_k\rangle = A_{nk}. \tag{3.26}$$

For recall that the identity can be written as $\mathbb{I} = \sum_k |\psi_k\rangle\langle\psi_k|$, and suppose $|\psi\rangle$ has some decomposition in the basis

$$|\psi\rangle = \sum_k \langle\psi_k|\psi\rangle|\psi_k\rangle \tag{3.27}$$

Then

$$\begin{aligned} A|\psi\rangle &= \mathbb{I}\mathbb{A}\mathbb{I}|\psi\rangle \\ &= \sum_{k,n} |\psi_k\rangle\langle\psi_k|A|\psi_n\rangle\langle\psi_n|\psi\rangle. \end{aligned}$$

Thus we can recognize the components of $|\psi\rangle$ in our basis, which are given by $\langle\psi_n|\psi\rangle = C_n$. Hence

$$\begin{aligned} A\left(\sum_k \langle\psi_k|\psi\rangle|\psi_k\rangle\right) &= \sum_k |\psi_k\rangle\left(\sum_n \langle\psi_k|A|\psi_n\rangle\langle\psi_n|\psi\rangle\right) \\ &= \sum_k |\psi_k\rangle\left(\sum_n A_{kn}C_n\right) \\ &= \sum |\psi_k\rangle C_k^{\bar{\psi}} \end{aligned}$$

in terms of the components of some new vector $|\bar{\psi}\rangle$. That is, if we know the matrix elements of $A$ in some basis and we know the components of the vector in that basis, we can uniquely determine the components of its image under $A$ in the same basis.

In the end, this is just abstract matrix multiplication. That is,

$$C_n^{\bar{\psi}} = \sum_k A_{nk}C_k^{\psi}. \tag{3.28}$$

We can also write the operator $A$ in terms of its matrix elements:

$$\begin{aligned} A &= \mathbb{I}A\mathbb{I} \\ &= \sum_{n,k} |\psi_k\rangle\langle\psi_k|A|\psi_n\rangle\langle\psi_n| \\ &= \sum_{n,k} A_{kn}|\psi_k\rangle\langle\psi_n|. \end{aligned}$$

This also tells us immediately that the matrix elements of the identity in any orthonormal basis are as we could have guessed– $\mathbb{I}_{nk} = \delta_{nk}$, the Kronecker delta.

It also follows that the matrix elements of the adjoint of an operator obey

$$(A^\dagger)_{nk} = A^*_{kn}. \tag{3.29}$$

This gives us another statement of hermiticity– equivalently, a hermitian operator is one whose matrix elements obey

$$A^*_{kn} = A_{nk}. \tag{3.30}$$

And thus

$$A^\dagger = \sum_{k,n} A^*_{nk}|\psi_k\rangle\langle\psi_n|. \tag{3.31}$$

Let's check that for hermitian operators, the expectation value is non-negative,

$$\langle\psi|A|\psi\rangle \geq 0. \tag{3.32}$$

Writing $A$ in terms of its matrix elements, we have

$$\begin{aligned}\langle\psi|A|\psi\rangle &= \sum_{n,k}\langle\psi|\psi_k\rangle A_{kn}\langle\psi_n|\psi\rangle \\ &= C_k^{*\psi} A_{kn} C_n^{\psi}.\end{aligned}$$

Suppose we have two orthonormal bases for the same space, $\{|\psi_k\rangle\}, \{|\psi_k'\rangle\}$. It follows that the new basis has some decomposition in the old basis. That is, the set $\{|\psi_k\rangle, |\psi_1'\rangle\}$ is linearly dependent and so

$$|\psi_n'\rangle = \sum c_k^{n\prime}|\psi_k\rangle, \tag{3.33}$$

in terms of some coefficients $c_k^{n\prime}$. It's also true that we can go back,

$$|\psi_n\rangle = \sum c_k^n|\psi_k'\rangle. \tag{3.34}$$

Certainly we can write this decomposition as the action of an operator $U$:

$$U|\psi_n\rangle = \sum c_k^{n\prime}|\psi_k\rangle, \tag{3.35}$$

and moreover $U$ must be invertible.