IV. Matrix Mechanics

We now turn to the a pragmatic aspect of QM: given a particular problem, how can we translate the Dirac notation into a form that might be interpretable by a computer? As hinted at previously, we do this by mapping Dirac notation onto a *complex vector space*. The operations in Hilbert space then reduce to linear algebra that can easily be done on a computer. This formalism is completely equivalent to the Dirac notation we've already covered; in different contexts, one will prove more useful than the other.

1. States can be represented by vectors

First, we will begin with an arbitrary complete orthonormal basis of states $\{|\phi_i\rangle\}$. Then, we know that we can write any other state as:

$$|\psi\rangle = c_1|\phi_1\rangle + c_2|\phi_2\rangle + c_3|\phi_3\rangle + ... = \sum_i c_i|\phi_i\rangle$$

How are these coefficient determined? Here, we follow a common trick and take the inner product with the j^{th} state:

$$\left(\langle\phi_j|\right)|\psi\rangle = \left(\langle\phi_j|\right)\sum_i c_i|\phi_i\rangle = \sum_i c_i\left(\langle\phi_j|\right)|\phi_i\rangle = \sum_i c_i\delta_{ij}$$

Since the Kronecker delta is only non-zero when i=j, the sum collapses to one term:

$$\langle\phi_j|\psi\rangle = c_j$$

The simple conclusion of these equations is that *knowing the coefficients is equivalent to knowing the wavefunction*. If we know $|\psi\rangle$, we can determine the coefficients through the second relation. Vice versa, If we know the coefficients, we can reconstruct $|\psi\rangle$ by performing the sum $\sum_i c_i|\phi_i\rangle$. Thus, if we fix this arbitrary basis, we can throw away all the basis state and just keep track of the *coefficients* of the ket state:

$$|\psi\rangle \rightarrow \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ \dots \end{pmatrix}_\phi$$

In harmony with the intuitive arguments made previously, here we associate the ket states with column vectors. Notice the small

subscript "ϕ", which reminds us that this vector of coefficients represents $|\psi\rangle$ in the $\{|\phi_i\rangle\}$ basis. If we were really careful, we would keep this subscript at all times; however, in practice we will typically know what basis we are working in, and the subscript will be dropped.

How to we represent the corresponding bra state $\langle\psi|$ as a vector? Well, we know that

$$\langle\psi| = (|\psi\rangle)^\dagger = \left(\sum_i c_i |\phi_i\rangle\right)^\dagger = \sum_i \langle\phi_i| c_i^* .$$

Now, as noted before, we expect to associate bra states with row vectors, and the above relation shows us that the elements of this row vector should be the *complex conjugates* of the column vector:

$$\boxed{\langle\psi| \rightarrow \begin{pmatrix} c_1^* & c_2^* & c_3^* & \ldots \end{pmatrix}_\phi}$$

Noting that bra states and ket states were defined to be Hermitian conjugates of one another, we see that *Hermitian conjugation in state space corresponds to taking the complex conjugate transpose of the coefficient vector.*

Now, the vector notation is totally equivalent to Dirac notation; thus, anything we compute in one representation should be exactly the same if computed in the other. As one illustration of this point, it is useful to check that this association of states with vectors preserves the inner product:

$$\langle\psi|\psi'\rangle = \left(\sum_i \langle\phi_i| c_i^*\right)\left(\sum_j c_j' |\phi_j\rangle\right) = \sum_{ij} c_i^* c_j' \langle\phi_i|\phi_j\rangle = \sum_{ij} c_i^* c_j' \delta_{ij} = \sum_i c_i^* c_i'$$

$$\begin{pmatrix} c_1^* & c_2^* & c_3^* & \ldots \end{pmatrix}_\phi \cdot \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ \ldots \end{pmatrix}_\phi = \sum_i c_i^* c_i'$$

So, the two ways of writing the inner product give identical results, and overlaps between the vectors have the same quantitative (and qualitative) meaning as their bra-ket counterparts.

2. Operators Become Matrices

In order to complete our transition to linear algebra we need to determine how operators act on the coefficient vectors described in

the previous section. Before we do this, we need to consider the operator:

$$\hat{O} = \sum_i |\phi_i\rangle\langle\phi_i|$$

Acting this operator on an arbitrary state:

$$\hat{O}|\psi\rangle = \sum_i |\phi_i\rangle\langle\phi_i|\psi\rangle$$

However, we showed above that $\langle\phi_j|\psi\rangle = c_j$, the coefficients of the state $|\psi\rangle$. Thus,

$$\hat{O}|\psi\rangle = \sum_i |\phi_i\rangle\langle\phi_i|\psi\rangle = \sum_i |\phi_i\rangle c_i = |\psi\rangle$$

Thus, $\hat{O}$ acting on any state gives the same state back. The operator that accomplishes this is the identity operator, $\hat{1}$, and so we write:

$$\hat{1} = \sum_i |\phi_i\rangle\langle\phi_i|$$

and say that this is a *resolution of the identity*.

With this in hand, we manipulate the expression for the bra-ket sandwich of and arbitrary operator $\hat{A}$:

$$\langle\psi|\ \hat{A}\ |\psi'\rangle$$

$$\hat{1} = \sum_i |\phi_i\rangle\langle\phi_i| \qquad \hat{1} = \sum_j |\phi_j\rangle\langle\phi_j|$$

$$\Rightarrow \sum_{ij} \langle\psi|\phi_i\rangle\langle\phi_i|\hat{A}|\phi_j\rangle\langle\phi_j|\psi'\rangle$$

$$\Rightarrow \sum_{ij} c_i^* \langle\phi_i|\hat{A}|\phi_j\rangle c_j'$$

We can re-write this last line as a standard (row)x(matrix)x(column) multiplication if we define the matrix $\mathbf{A}$, whose elements are $A_{ij} = \langle\phi_i|\hat{A}|\phi_j\rangle$. This association is so common, in fact, that even in Dirac notation, these objects are typically referred to as “matrix elements”. So, to summarize, we can write:

$$\hat{A} \rightarrow \begin{pmatrix} \langle \phi_1 | \hat{A} | \phi_1 \rangle & \langle \phi_1 | \hat{A} | \phi_2 \rangle & \langle \phi_1 | \hat{A} | \phi_3 \rangle & \ldots \\ \langle \phi_2 | \hat{A} | \phi_1 \rangle & \langle \phi_2 | \hat{A} | \phi_2 \rangle & \ldots & \\ \langle \phi_3 | \hat{A} | \phi_1 \rangle & \ldots & & \\ \ldots & & & \end{pmatrix}_{\phi} \equiv \mathbf{A}$$

This matrix has all the same action on row and column vectors that the operator $\hat{A}$ has on bras and kets:

$$\hat{A}|\psi\rangle \leftrightarrow \mathbf{A} \cdot \mathbf{c} \qquad \langle \psi | \hat{A} \leftrightarrow \mathbf{c}^{\dagger} \cdot \mathbf{A}$$

It is also easy to show that the product of two operators is correctly represented by the product of their matrix representations. Further, similar to the case of vectors, the adjoint of $\hat{A}$ is the complex conjugate transpose of $\mathbf{A}$. Using the above associations, we can write every operation in Hilbert space in terms of matrix-vector multiplications, which are easily handled by a computer.

3. Some Interesting Matrix Properties

We are now going to prove a number of interesting things about particular classes of matrices which will prove useful later. These same identities can be proven for the raw operators, but the results are somewhat more familiar when one has the matrix formulation in hand.

Now, since each Hermitian operator defines its own orthonormal basis, we often be interested in making a change of basis from the eigenbasis of one Hermitian operator, $\hat{A}$, to that of another, $\hat{B}$. Denote the eigenbasis of $\hat{A}$ by $\{|\varphi_\alpha\rangle\}$ and the eigenbasis of $\hat{B}$ by $\{|\chi_\alpha\rangle\}$. Then according to what we know about Hermitian operators $\{|\varphi_\alpha\rangle\}$ and $\{|\chi_\alpha\rangle\}$ are both orthonormal bases. Thus, we can write any state as

$$|\psi\rangle = \sum_\alpha a_\alpha |\varphi_\alpha\rangle \quad \text{or} \quad |\psi\rangle = \sum_\alpha b_\alpha |\chi_\alpha\rangle$$

Our task is to get $\{a_\alpha\}$ from $\{b_\alpha\}$ (or vice versa). This is accomplished using our favorite trick; take the inner product of each equation with $|\chi_\beta\rangle$:

$$\langle \chi_\beta | \psi \rangle = \sum_\alpha a_\alpha \langle \chi_\beta | \varphi_\alpha \rangle \quad \text{or} \quad \langle \chi_\beta | \psi \rangle = \sum_\alpha b_\alpha \langle \chi_\beta | \chi_\alpha \rangle$$

Equating the r.h.s. and making use of the orthonormality of the $\{|\chi_\alpha\rangle\}$ gives:

$$\sum_\alpha a_\alpha \langle \chi_\beta | \varphi_\alpha \rangle = \sum_\alpha b_\alpha \langle \chi_\beta | \chi_\alpha \rangle = \sum_\alpha b_\alpha \delta_{\alpha\beta} = b_\beta$$

$$\Rightarrow \sum_\alpha a_\alpha \langle \chi_\beta | \varphi_\alpha \rangle = b_\beta$$

This leads to the definition of the **Transformation matrix**:

$$\mathbf{T} \equiv \begin{pmatrix} \langle \chi_1 | \varphi_1 \rangle & \langle \chi_1 | \varphi_2 \rangle & \langle \chi_1 | \varphi_3 \rangle & \cdots \\ \langle \chi_2 | \varphi_1 \rangle & \langle \chi_2 | \varphi_2 \rangle & \cdots & \cdots \\ \langle \chi_3 | \varphi_1 \rangle & \cdots & \cdots & \\ \cdots & \cdots & & \end{pmatrix}$$

The columns of this matrix are the coefficients of the "new" basis states ($\{|\chi_\alpha\rangle\}$) in terms of "old" ones ($\{|\varphi_\alpha\rangle\}$) and it allows us to transform from one basis to another using simple matrix algebra:

$$\boxed{\mathbf{b} = \mathbf{T}\mathbf{a}}$$

Now, what about the reverse transformation (i.e. $\mathbf{b}$ to $\mathbf{a}$)? Well, our designation of the "old" ($\{|\varphi_\alpha\rangle\}$) and "new" ($\{|\chi_\alpha\rangle\}$) bases was completely arbitrary; we can change the direction of the transformation by simply switching the roles of $\{|\chi_\alpha\rangle\}$ and $\{|\varphi_\alpha\rangle\}$. For example, we can obtain the $\mathbf{b}$ to $\mathbf{a}$ transformation matrix by simply swapping the letters φ and χ in our definition of $\mathbf{T}$

$$\mathbf{S} \equiv \begin{pmatrix} \langle \varphi_1 | \chi_1 \rangle & \langle \varphi_1 | \chi_2 \rangle & \langle \varphi_1 | \chi_3 \rangle & \cdots \\ \langle \varphi_2 | \chi_1 \rangle & \langle \varphi_2 | \chi_2 \rangle & \cdots & \cdots \\ \langle \varphi_3 | \chi_1 \rangle & \cdots & \cdots & \\ \cdots & \cdots & & \end{pmatrix}$$

This matrix satisfies:

$$\mathbf{a} = \mathbf{S}\mathbf{b}.$$

However, looking at our definition of $\mathbf{S}$, we see that it is just the Hermitian conjugate of $\mathbf{T}$! This leads to an important result:

$$\mathbf{a} = \mathbf{S}\mathbf{b} = \mathbf{T}^\dagger \mathbf{b} = \mathbf{T}^\dagger (\mathbf{T}\mathbf{a}) = \mathbf{T}^\dagger \mathbf{T}\mathbf{a}$$

Reading from right to left, this shows that $\mathbf{T}^\dagger\mathbf{T}$ acting on any vector $\mathbf{a}$ *gives back the same vector*. Thus, we conclude that:

$$\mathbf{T}^\dagger \mathbf{T} = \mathbf{1}$$

Matrices that satisfy this special property are called **unitary** matrices. Any property we are interested in will be *invariant to unitary*

transformations. From a physical perspective, this is because unitary transforms correspond to a change of basis and we know that the basis we choose to represent things is arbitrary and should not matter. From a mathematical point of view, this results from the fact that unitary matrices will always occur in Hermitian conjugate pairs in our results (because of the bra-ket structure of the inner product) and $\mathbf{T}^{\dagger}\mathbf{T}$ is the identity.

The transformation matrix also allows us to change the basis for an operator. Denote the matrix representation of $\hat{H}$ in the eigenbasis of $\hat{A}$ by:

$$\begin{pmatrix} \langle\varphi_1|\hat{H}|\varphi_1\rangle & \langle\varphi_1|\hat{H}|\varphi_2\rangle & \langle\varphi_1|\hat{H}|\varphi_3\rangle & \ldots \\ \langle\varphi_2|\hat{H}|\varphi_1\rangle & \langle\varphi_2|\hat{H}|\varphi_2\rangle & \ldots & \\ \langle\varphi_3|\hat{H}|\varphi_1\rangle & \ldots & & \\ \ldots & & & \end{pmatrix} \equiv \mathbf{H}_A$$

Then we have that a matrix element of $\hat{H}$ can be represented in the $\hat{A}$ basis by:

$$\langle\psi|\hat{H}|\psi'\rangle = \mathbf{a}^{\dagger}\mathbf{H}_A\mathbf{a}'$$

and in the $\hat{B}$ basis by:

$$\langle\psi|\hat{H}|\psi'\rangle = \mathbf{b}^{\dagger}\mathbf{H}_B\mathbf{b}'$$

Now, using the fact that $\mathbf{a}' = \mathbf{T}^{\dagger}\mathbf{b}'$ (and the Hermitian conjugate relation $\mathbf{a}^{\dagger} = \mathbf{b}^{\dagger}\mathbf{T}$),

$$\Rightarrow \langle\psi|\hat{H}|\psi'\rangle = \mathbf{a}^{\dagger}\mathbf{H}_A\mathbf{a}' = \mathbf{b}^{\dagger}\mathbf{T}\mathbf{H}_A\mathbf{T}^{\dagger}\mathbf{b}'$$

Comparing this last equation with the definition of $\mathbf{H}_B$ leads to the conclusion that under a change of basis from $\hat{A}$ to $\hat{B}$, an arbitrary matrix transforms as:

$$\mathbf{H}_B = \mathbf{T}\mathbf{H}_A\mathbf{T}^{\dagger}$$

One important special case of this relation is when $\hat{H} = \hat{A}$. Then, we know the matrix elements:

$$\begin{pmatrix} \langle\varphi_1|\hat{A}|\varphi_1\rangle & \langle\varphi_1|\hat{A}|\varphi_2\rangle & \langle\varphi_1|\hat{A}|\varphi_3\rangle & \ldots \\ \langle\varphi_2|\hat{A}|\varphi_1\rangle & \langle\varphi_2|\hat{A}|\varphi_2\rangle & \ldots & \\ \langle\varphi_{31}|\hat{A}|\varphi_1\rangle & \ldots & & \\ \ldots & & & \end{pmatrix} = \begin{pmatrix} a_1 & 0 & 0 & 0 \\ 0 & a_2 & 0 & \ldots \\ 0 & 0 & a_3 & \ldots \\ 0 & \ldots & \ldots & \ldots \end{pmatrix} = \mathbf{A}_A$$

That is, the matrix that represents $\hat{A}$ *is diagonal in the eigenbasis of* $\hat{A}$. This allows us to very easily represent $\hat{A}$ in any other eigenbasis:

$$\mathbf{A}_B = \mathbf{T}\mathbf{A}_A\mathbf{T}^\dagger$$

Where we recall that $\mathbf{A}_A$ is just a diagonal matrix. In practice, we will very often want to work with a matrix in its eigenbasis and only use the transformation rules to move to other bases when necessary.

As an example, consider a **function of a matrix**. This is defined by the power series expansion of the function:

$$f(\mathbf{A}) \equiv \mathbf{1} + f'(0)\mathbf{A} + \frac{f''(0)}{2!}\mathbf{A}\mathbf{A} + \frac{f'''(0)}{3!}\mathbf{A}\mathbf{A}\mathbf{A} + \ldots$$

One important example of this is the **exponential** of a matrix, which we will use quite frequently:

$$e^{\mathbf{A}} \equiv \mathbf{1} + \mathbf{A} + \frac{\mathbf{A}\mathbf{A}}{2!} + \frac{\mathbf{A}\mathbf{A}\mathbf{A}}{3!} + \ldots$$

If we transform from whatever arbitrary basis we are in into the eigenbasis, we can write $\mathbf{A}$ in diagonal form ($\mathbf{A} = \mathbf{T}\mathbf{A}_A\mathbf{T}^\dagger$) and the function becomes:

$$f(\mathbf{A}) = \mathbf{1} + f'(0)\mathbf{T}\mathbf{A}_A\mathbf{T}^\dagger + \frac{f''(0)}{2!}\mathbf{T}\mathbf{A}_A\cancel{\mathbf{T}^\dagger}\mathbf{T}\mathbf{A}_A\mathbf{T}^\dagger + \frac{f'''(0)}{3!}\mathbf{T}\mathbf{A}_A\cancel{\mathbf{T}^\dagger}\mathbf{T}\mathbf{A}_A\cancel{\mathbf{T}^\dagger}\mathbf{T}\mathbf{A}_A\mathbf{T}^\dagger + \ldots$$

$$\Rightarrow f(\mathbf{A}) = \mathbf{1} + f'(0)\mathbf{T}\mathbf{A}_A\mathbf{T}^\dagger + \frac{f''(0)}{2!}\mathbf{T}\mathbf{A}_A\overset{1}{\mathbf{A}_A}\mathbf{T}^\dagger + \frac{f'''(0)}{3!}\mathbf{T}\mathbf{A}_A\overset{1}{\mathbf{A}_A}\overset{1}{\mathbf{A}_A}\mathbf{T}^\dagger + \ldots$$

and noting that $\mathbf{1} = \mathbf{T}\mathbf{T}^\dagger$ and that the coefficients (which are *numbers*) commute with the transformation matrix:

$$\Rightarrow f(\mathbf{A}) = \mathbf{T}\mathbf{T}^\dagger + \mathbf{T}f'(0)\mathbf{A}_A\mathbf{T}^\dagger + \mathbf{T}\frac{f''(0)}{2!}\mathbf{A}_A\mathbf{A}_A\mathbf{T}^\dagger + \mathbf{T}\frac{f'''(0)}{3!}\mathbf{A}_A\mathbf{A}_A\mathbf{A}_A\mathbf{T}^\dagger + \ldots$$

$$\Rightarrow f(\mathbf{A}) = \mathbf{T}\left(\mathbf{1} + f'(0)\mathbf{A}_A + \frac{f''(0)}{2!}\mathbf{A}_A\mathbf{A}_A + \frac{f'''(0)}{3!}\mathbf{A}_A\mathbf{A}_A\mathbf{A}_A + \ldots\right)\mathbf{T}^\dagger$$

$$\Rightarrow f(\mathbf{A}) = \mathbf{T}f(\mathbf{A}_A)\mathbf{T}^\dagger$$

Thus, *functions of matrices transform just like matrices* when we change basis. Why is this important? In its eigenbasis, we know that $\hat{A}$ is represented by a diagonal matrix, and it is trivial to apply a function to a diagonal matrix; the result is a diagonal matrix, with the diagonal elements given by $f(x)$ evaluated at each of the eigenvalues:

$$f(\mathbf{A}_A) = f\begin{pmatrix} e_1 & 0 & 0 & 0 \\ 0 & e_2 & 0 & 0 \\ 0 & 0 & e_3 & 0 \\ 0 & 0 & 0 & \ldots \end{pmatrix} = \begin{pmatrix} f(e_1) & 0 & 0 & 0 \\ 0 & f(e_2) & 0 & 0 \\ 0 & 0 & f(e_3) & 0 \\ 0 & 0 & 0 & \ldots \end{pmatrix}$$

Thus, we can apply a function to a matrix in three steps: 1) Change basis to the eigenbasis of $\hat{A}$ ($\mathbf{A} = \mathbf{T}\mathbf{A}_A\mathbf{T}^\dagger$) 2) Apply $f(x)$ to the diagonal elements of $\mathbf{A}_A$ to obtain $f(\mathbf{A}_A)$ 3) Transform back to the original basis ($f(\mathbf{A}) = \mathbf{T}f(\mathbf{A}_A)\mathbf{T}^\dagger$)

B. Discrete Variable Representation (DVR)

For one dimensional systems, we are left with a puzzle: on the one hand, it is natural to work in the (continuous) position representation (as in wave mechanics) but on the other hand, we need a discrete representation to translate things on to a computer. The discrete variable representation (DVR) is one solution to this problem. The DVR works in the harmonic oscillator basis. To refresh our memory, recall that the matrix elements of $\hat{p}^2$ are:

$$\langle m|\hat{p}^2|n\rangle = -\tfrac{1}{2}\langle m|\left(\hat{a}^\dagger - \hat{a}\right)^2|n\rangle = \tfrac{1}{2}\langle m|\left(\hat{a}^\dagger\hat{a}^\dagger - \hat{a}\hat{a}^\dagger - \hat{a}^\dagger\hat{a} + \hat{a}\hat{a}\right)|n\rangle$$
$$= -\tfrac{1}{2}\left(\sqrt{n+2}\sqrt{n+1}\delta_{m,n+2} - (2n+1)\delta_{m,n} + \sqrt{n-1}\sqrt{n}\delta_{m,n-2}\right)$$

while the matrix elements of $\hat{q}$ are:

$$\langle m|\hat{q}|n\rangle = \tfrac{1}{\sqrt{2}}\langle m|\left(\hat{a}^\dagger + \hat{a}\right)|n\rangle = \tfrac{1}{\sqrt{2}}\left(\sqrt{n+1}\delta_{m,n+1} + \sqrt{n}\delta_{m,n-1}\right)$$

As a result, we can write both of these operators as **matrices** in the harmonic oscillator basis:

$$\mathbf{Q} \equiv \tfrac{1}{\sqrt{2}}\begin{pmatrix} 0 & \sqrt{1} & 0 & \ldots \\ \sqrt{1} & 0 & \sqrt{2} & \ldots \\ 0 & \sqrt{2} & 0 & \sqrt{3} \\ \ldots & \ldots & \sqrt{3} & \ldots \end{pmatrix} \qquad \mathbf{P}^2 \equiv \tfrac{1}{2}\begin{pmatrix} 1 & 0 & -\sqrt{2} & \ldots \\ 0 & 3 & 0 & -\sqrt{6} \\ -\sqrt{2} & 0 & 5 & 0 \\ \ldots & -\sqrt{6} & 0 & \ldots \end{pmatrix}$$

Thus far, we have just been applying the rules of matrix mechanics. Now, if our Hamiltonian takes the standard form $\hat{H} = \frac{\hat{p}^2}{2m} + V(\hat{q})$ then we can easily translate this into matrix mechanics: $\mathbf{H} = \frac{\mathbf{P}^2}{2m} + V(\mathbf{Q})$. The only tricky part is how we evaluate $V(\mathbf{Q})$; in DVR we follow the

prescription above and diagonalize $\mathbf{Q}$, evaluate $V(\mathbf{Q})$ in the eigenbasis and then transform back to the harmonic oscillator basis: $V(\mathbf{Q}) = \mathbf{T} V(\mathbf{Q}_q) \mathbf{T}^\dagger$. As mentioned above, evaluating $V(\mathbf{Q}_q)$ is simple; it is just a diagonal matrix with diagonal elements given by the function V evaluated at the eigenvalues of $\mathbf{Q}$.

Therefore, we know how to represent $\hat{H} = \frac{\hat{p}^2}{2m} + V(\hat{q})$ as a matrix in the harmonic oscillator basis. We have made no approximations (yet). However, we have a bit of a problem, because the matrices $\mathbf{Q}$ and $\mathbf{P}^2$ are infinite-dimensional. This is a bit of a problem on a (finite) computer. The key approximation we make is that of a **truncated basis**: we select some large but finite number of basis functions, N. For example, we might choose the N lowest energy states of the harmonic oscillator if we are interested in the low energy states of $\hat{H} = \frac{\hat{p}^2}{2m} + V(\hat{q})$. Once we have selected our truncated basis, we perform the entire problem as if these functions form a *complete* basis. As a result, $\mathbf{Q}$ and $\mathbf{P}^2$ go from being infinite dimensional matrices to being $N \times N$ matrices that we can store on a computer. The key observation is that, as I make N larger and larger, my truncated basis becomes closer to the complete basis and all the results I compute in the truncated basis (e.g. energy eigenvalues, expectation values, etc.) must approach the corresponding complete basis results. Since one can easily deal with N up to several thousand on a computer, this makes it possible to obtain very precise answers for one-dimensional problems.

So, to summarize, here are the major steps in DVR:

1) Choose a harmonic oscillator frequency, ω, and number of basis functions, N. In principle any ω will work if N is large enough, because every harmonic oscillator defines its own complete basis. However, a well-chosen ω will approach the exact result more rapidly with N.
2) Build $\mathbf{Q}$ and $\mathbf{P}^2$ in the harmonic oscillator basis. These will be $N \times N$ matrices.

3) Diagonalize $\mathbf{Q}$ to obtain the diagonal matrix $\mathbf{Q}_q$ and the unitary matrix $\mathbf{T}$.
4) Compute the potential using $V(\mathbf{Q}) = \mathbf{T} V(\mathbf{Q}_q) \mathbf{T}^\dagger$.
5) Construct $\mathbf{H} = \frac{\mathbf{P}^2}{2m} + V(\mathbf{Q})$ and compute whatever you need (e.g. eigenvalues, eigenstates, average values) using matrix mechanics.

C. Variational method

One important point is that the DVR representation is **not unique**. The ambiguity comes because we have chosen to compute the matrix representation of certain "basic" operators ($\hat{p}^2$ and $\hat{q}$) directly, while other operators ($V(\hat{q})$) have been represented as functions of the "basic" matrices. Toe see where the ambiguity arises, let's pretend we've chosen a very small basis: N=3. There are two ways we could compute $\mathbf{P}^2$. The first is to compute its matrix elements directly:

$$\mathbf{P}^2 \equiv \tfrac{1}{2}\begin{pmatrix} 1 & 0 & -\sqrt{2} \\ 0 & 3 & 0 \\ -\sqrt{2} & 0 & 5 \end{pmatrix}$$

The second way is to compute the matrix representation of $\hat{p}$ and then square the matrix:

$$\mathbf{P}\cdot\mathbf{P} = \tfrac{1}{\sqrt{2}}\begin{pmatrix} 0 & i & 0 \\ -i & 0 & i\sqrt{2} \\ 0 & -i\sqrt{2} & 0 \end{pmatrix}\begin{pmatrix} 0 & i & 0 \\ -i & 0 & i\sqrt{2} \\ 0 & -i\sqrt{2} & 0 \end{pmatrix} = \tfrac{1}{2}\begin{pmatrix} 1 & 0 & -\sqrt{2} \\ 0 & 3 & 0 \\ \sqrt{2} & 0 & 2 \end{pmatrix}$$

And the results are clearly different. Note that if we had chosen a larger N, these two routes would have given (essentially) the same answer. However, it is clear that, for finite N the way we choose to represent things in terms of matrices **matters**.

The foundation of variational approaches is the Raleigh-Ritz **variational principle**: Given any state $|\psi_T\rangle$ (called the trial state) and any Hamiltonian $\hat{H}$, the average energy for the trial state is always greater than or equal to the ground state energy of the Hamiltonian:

$$E_T \equiv \frac{\langle \psi_T | \hat{H} | \psi_T \rangle}{\langle \psi_T | \psi_T \rangle} \geq E_0 .$$

To prove that this is the case, let us expand $|\psi_T\rangle$ in terms of the eigenstates of $\hat{H}$:

$$|\psi_T\rangle = \sum_i \alpha_i |\phi_i\rangle$$

where each $|\phi_i\rangle$ satisfies the eigenvalue equation:

$$\hat{H}|\phi_i\rangle = E_i |\phi_i\rangle .$$

Now, we plug this form for $|\psi_T\rangle$ into the expression for the average energy:

$$E_T \equiv \frac{\langle \psi_T | \hat{H} | \psi_T \rangle}{\langle \psi_T | \psi_T \rangle} = \frac{\sum_j \alpha_j^* \langle \phi_j | \hat{H} \sum_i \alpha_i |\phi_i\rangle}{\sum_j \alpha_j^* \langle \phi_j | \sum_i \alpha_i |\phi_i\rangle} = \frac{\sum_j \alpha_j^* \langle \phi_j | \sum_i \alpha_i \hat{H} |\phi_i\rangle}{\sum_j \alpha_j^* \langle \phi_j | \sum_i \alpha_i |\phi_i\rangle}$$

$$= \frac{\sum_j \alpha_j^* \langle \phi_j | \sum_i \alpha_i E_i |\phi_i\rangle}{\sum_j \alpha_j^* \langle \phi_j | \sum_i \alpha_i |\phi_i\rangle} = \frac{\sum_{ij} \alpha_j^* \alpha_i E_i \cancel{\langle \phi_j | \phi_i \rangle}}{\sum_{ij} \alpha_j^* \alpha_i \cancel{\langle \phi_j | \phi_i \rangle}} \quad \delta_{ij}$$

$$= \frac{\sum_i \alpha_i^* \alpha_i E_i}{\sum_i \alpha_i^* \alpha_i}$$

Now, we note that $\alpha_i^* \alpha_i \equiv w_i$ is a real, positive number. Thus the trial energy is just a weighted sum of the eigenvalues of $\hat{H}$:

$$E_T = \frac{\sum_i w_i E_i}{\sum_i w_i} .$$

Now, we make use of the fact that every eigenvalue, E_i, of $\hat{H}$ is by definition greater than or equal to the lowest eigenvalue, E_0, of $\hat{H}$. Thus,

$$\sum_i w_i E_i \geq \sum_i w_i E_0$$

with equality holding only if w_i is zero for every state except the ground state (or ground ***states*** if E_0 is degenerate). Making use of this for the trial energy:

$$E_T \geq \frac{\sum_i w_i E_0}{\sum_i w_i} = \frac{E_0 \sum_i w_i}{\sum_i w_i} = E_0$$

Thus we see that the average energy for any trial state must be greater than or equal to the ground state energy. This gives us a clear-cut way to determine the "best" approximation to the ground state out of a set of trial functions: whichever $|\psi_T\rangle$ has the lowest average energy is physically the best approximation to the ground state.

To formulate this in practical terms, assume that $|\psi_T\rangle$ depends on a set of *adjustable parameters*, $\mathbf{c}$. Then, consider the possible trial energies that can be obtained by varying $\mathbf{c}$. We write

$$E_T(\mathbf{c}) = \frac{\langle \psi_T(\mathbf{c}) | \hat{H} | \psi_T(\mathbf{c}) \rangle}{\langle \psi_T(\mathbf{c}) | \psi_T(\mathbf{c}) \rangle}$$

And we see that the trial energy becomes a (many dimensional) function of the variable parameters. According to the variational theorem, the "best" choice of the parameters for describing the ground state are the ones that minimize $E_T(\mathbf{c})$. This implies that the gradient of $E_T(\mathbf{c})$ is zero at the variationally optimal choice of $\mathbf{c}$. If we call the optimal set of parameters $\mathbf{c}_0$, then we can find these parameters as the solution of the equation:

$$\nabla_{\mathbf{c}} E_T(\mathbf{c})\big|_{\mathbf{c}_0} = 0$$

To make the connection with DVR, consider a large (but not complete) set of orthonormal states $\{|\phi_i\rangle\}$. Most commonly, this set is obtained by truncation of some known complete orthonormal set (e.g. the harmonic oscillator energy eigenstates) to a set containing only N states. The $\{|\phi_i\rangle\}$ are not eigenstates of the Hamiltonian we are interested in; the eigenstates are the objects we want to approximate, so it would be foolish to assume they were known from the outset. We will be interested in trial states that are linear combinations of the $\{|\phi_i\rangle\}$:

$$|\psi_T(\mathbf{c})\rangle = \sum_{i=1}^{N} c_i |\phi_i\rangle$$

Clearly, as the number of basis functions becomes large ($N \to \infty$), the basis becomes complete and we can write any state (including the

eigenstates of $\hat{H}$ as a linear combination of the $\{|\phi_i\rangle\}$. Thus, in the $N \to \infty$ limit, *the choice of basis does not matter*. We will get exactly the same answer (and in particular the same answer as obtained with DVR) for any choice of the $\{|\phi_i\rangle\}$. However, different choices of $\{|\phi_i\rangle\}$ will give different answers for finite N. In this respect, the variational principle also gives one an idea of which basis is "best": the one that gives the lowest average energy rigorously gives a better approximation to the ground state. This is somewhat magical, since we **do not know** the ground state, but we can tell when we are close to it.

Given the definition of the ket trial state, the bra state is just the hermitian conjugate:

$$\langle\psi_T(\mathbf{c})| = \sum_{i=1}^{N} \langle\phi_i| c_i^*$$

And the trial energy is given by

$$E_T(\mathbf{c},\mathbf{c}^*) = \frac{\langle\psi_T(\mathbf{c}^*)|\hat{H}|\psi_T(\mathbf{c})\rangle}{\langle\psi_T(\mathbf{c}^*)|\psi_T(\mathbf{c})\rangle} = \frac{\sum_{i=1}^{N}\langle\phi_i|c_i^*\hat{H}\sum_{j=1}^{N}c_j|\phi_j\rangle}{\sum_{i=1}^{N}\langle\phi_i|c_i^*\sum_{j=1}^{N}c_j|\phi_j\rangle} = \frac{\sum_{i,j=1}^{N}c_i^*\langle\phi_i|\hat{H}|\phi_j\rangle c_j}{\sum_{i,j=1}^{N}\langle\phi_i|\phi_j\rangle c_i^* c_j}$$

$$= \frac{\sum_{i,j=1}^{N}c_i^*\langle\phi_i|\hat{H}|\phi_j\rangle c_j}{\sum_{i=1}^{N}c_i^* c_i}$$

Thus, the energy is a function of both $\mathbf{c}$ and $\mathbf{c}^*$. From complex variable calculus, we know that these are to be treated as *independent* variables, so the variational theorem gives us two conditions for the optimal choices of $\mathbf{c}$ and $\mathbf{c}^*$:

$$0 = \frac{\partial E_T(\mathbf{c},\mathbf{c}^*)}{\partial c_k^*} = \frac{\partial}{\partial c_k^*}\frac{\sum_{i,j=1}^{N}c_i^*\langle\phi_i|\hat{H}|\phi_j\rangle c_j}{\sum_{i=1}^{N}c_i^* c_i} = \frac{\sum_{j=1}^{N}\langle\phi_k|\hat{H}|\phi_j\rangle c_j}{\sum_{i=1}^{N}c_i^* c_i} - \frac{\sum_{i,j=1}^{N}c_i^*\langle\phi_i|\hat{H}|\phi_j\rangle c_j}{\left(\sum_{i=1}^{N}c_i^* c_i\right)^2}c_k$$

$$= \frac{\sum_{j=1}^{N}\langle\phi_k|\hat{H}|\phi_j\rangle c_j}{\sum_{i=1}^{N}c_i^* c_i} - \frac{E_T(\mathbf{c},\mathbf{c}^*)c_k}{\sum_{i=1}^{N}c_i^* c_i}$$

$$0 = \frac{\partial E_T(\mathbf{c},\mathbf{c}^*)}{\partial c_k} = \frac{\partial}{\partial c_k}\frac{\sum_{i,j=1}^{N} c_i^* \langle \phi_i | \hat{H} | \phi_j \rangle c_j}{\sum_{i=1}^{N} c_i^* c_i} = \frac{\sum_{i=1}^{N} c_i^* \langle \phi_i | \hat{H} | \phi_k \rangle}{\sum_{i=1}^{N} c_i^* c_i} - \frac{\sum_{i,j=1}^{N} c_i^* \langle \phi_i | \hat{H} | \phi_j \rangle c_j}{\left(\sum_{i=1}^{N} c_i^* c_i\right)^2} c_k^*$$

$$= \frac{\sum_{i=1}^{N} c_i^* \langle \phi_i | \hat{H} | \phi_k \rangle}{\sum_{i=1}^{N} c_i^* c_i} - \frac{E_T(\mathbf{c},\mathbf{c}^*) c_k^*}{\sum_{i=1}^{N} c_i^* c_i}$$

Collapsing each equation

$$\Rightarrow 0 = \sum_{j=1}^{N} \langle \phi_k | \hat{H} | \phi_j \rangle c_j - E_T(\mathbf{c},\mathbf{c}^*) c_k$$

and

$$\Rightarrow 0 = \sum_{i=1}^{N} c_i^* \langle \phi_i | \hat{H} | \phi_k \rangle - E_T(\mathbf{c},\mathbf{c}^*) c_k^* .$$

We can write these two equations in matrix notation:

$$\Rightarrow 0 = \mathbf{H}\mathbf{c} - E_T(\mathbf{c},\mathbf{c}^*)\mathbf{c}$$

$$\Rightarrow \mathbf{H}\mathbf{c} = E_T(\mathbf{c},\mathbf{c}^*)\mathbf{c}$$

and

$$\Rightarrow 0 = \mathbf{H}\mathbf{c}^\dagger - E_T(\mathbf{c},\mathbf{c}^*)\mathbf{c}^\dagger$$

$$\Rightarrow \mathbf{c}^\dagger\mathbf{H} = \mathbf{c}^\dagger E_T(\mathbf{c},\mathbf{c}^*)$$

Clearly, however, these two equations are hermitian conjugates of one another, and so we will typically just write the ket equation. These equations are just the eigenvalue equations for the matrix $\mathbf{H}$; thus finding the variationally linear combination of an orthonormal set simply corresponds to finding the eigenvalues and eigenvectors of $\mathbf{H}$. The lowest eigenvalue approximates the ground state energy, the second lowest approximates the first excited state, etc. One can actually prove a modified variational principle for the excited states defined in this manner: for any finite N, the mth eigenvalue of $\mathbf{H}$ is always greater than or equal to the mth eigenvalue of $\hat{H}$.

In practice, we can apply the variational method the same way DVR is used: First, we need to choose an appropriate orthonormal basis $\{|\phi_i\rangle\}$. Second, we need to build the matrix representation of the Hamiltonian in the HO basis:

$$H_{mn} \equiv \langle m | \frac{p^2}{2m} + V(x) | n \rangle$$

This is usually the most tedious step of the calculation, as it involves a good deal of algebra deriving the general matrix elements and a nontrivial amount of computer work to make sure the matrix is built properly. Finally, we diagonalize the matrix $\mathbf{H}$ to obtain the approximate eigenvalues. As an additional check, we should increase N and re-run the calculation to see that the result is converged.

Thus, the only difference between DVR and variational mechanics is the way one builds $\mathbf{H}$: in the former case, one builds it from the constituent matrices $\mathbf{Q}$ and $\mathbf{P}^2$ while in the latter, one builds $\mathbf{H}$ directly. The variational principle is unique among matrix representation methods because it guarantees that the approximate energy is an **upper bound** to the true energy. This gives us a rigorous way of evaluating "how close" we are to the correct answer and which basis is "best" for a given problem: one always seeks the lowest energy. Because of its unique properties, the variational method is the technique of choice for many applications.