Suppose we have a Hamiltonian $\hat{H} = \hat{H}_0 + \hat{H}_I$ and that we know how to diagonalize (find the eigenvalues and eigenstates of $\hat{H}_0$. If we then write things in the basis of the eigenstates of $\hat{H}_0$, we can write any operator, $\hat{\mathcal{O}}$ in that basis by just multiplying the operator on both sides by $\hat{\mathbf{1}}$:

$$\hat{\mathcal{O}} = \hat{\mathbf{1}}\hat{\mathcal{O}}\hat{\mathbf{1}} = \left(\sum_n |n\rangle\langle n|\right)\hat{\mathcal{O}}\left(\sum_m |m\rangle\langle m|\right) = \sum_{m,n} |n\rangle \mathcal{O}_{nm} \langle m|,$$

where $\mathcal{O}_{nm} \equiv \langle n|\hat{\mathcal{O}}|m\rangle$. We can represent these quantities in matrix notation as

$$\vec{\mathcal{O}} = \begin{pmatrix} \mathcal{O}_{11} & \cdots & \mathcal{O}_{1n} \\ \vdots & \ddots & \vdots \\ \mathcal{O}_{n1} & \cdots & \mathcal{O}_{nn} \end{pmatrix}, \; |m\rangle = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \text{ and } \langle m| = \begin{pmatrix} 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \end{pmatrix},$$

where the column and row vectors have components that are all zero except in the m$^{th}$ position (index). Our unperturbed Hamiltonian can then be represented by a diagonal matrix with the element $H_{0,mm} = E_m$.

The unperturbed Hamiltonian is a bit more complicated. In principle, all of its elements might be non-zero. At the very least, though, we can always make the diagonal elements zero by an appropriate change of Hamiltonians. To see this, it is easier to work with just a two-dimensional space. Then our unperturbed Hamiltonian is just $\hat{H}_0 = E_1|1\rangle\langle 1| + E_2|2\rangle\langle 2|$, which in a matrix representation becomes

$$\vec{H}_0 = \begin{pmatrix} E_1 & 0 \\ 0 & E_2 \end{pmatrix},$$

and in general, the interaction Hamiltonian is

$$\hat{H}_I = \sum_{m,n} H_{I,mn} |n\rangle\langle m|,$$

which when represented as a matrix is just

$$\vec{H}_I = \begin{pmatrix} H_{I,11} & H_{I,12} \\ H_{I,21} & H_{I,22} \end{pmatrix}.$$

Suppose that we were to now write the total Hamiltonian in the basis of eigenstates of the unperturbed Hamiltonian. That would just be

$$\ddot{H}_{tot} = \ddot{H}_0 + \ddot{H}_I = \begin{pmatrix} E_1 + H_{I,11} & H_{I,12} \\ H_{I,21} & E_2 + H_{I,22} \end{pmatrix} = \begin{pmatrix} E_1 + H_{I,11} & 0 \\ 0 & E_2 + H_{I,22} \end{pmatrix} + \begin{pmatrix} 0 & H_{I,12} \\ H_{I,21} & 0 \end{pmatrix}.$$

(The representation in terms of dyads of kets and bras should be clear, so I don't write it here.) We see that we have just written $\hat{H}_{tot} = \hat{H}'_0 + \hat{H}'_I$, where the new Hamiltonian $\hat{H}'_0$ has the same eigenstates as does $\hat{H}_0$, just with different eigenvalues. That is,

$$\hat{H}'_0 |1\rangle \to \begin{pmatrix} E_1 + H_{I,11} & 0 \\ 0 & E_2 + H_{I,22} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} E_1 + H_{I,11} \\ 0 \end{pmatrix} = \left(E_1 + H_{I,11}\right) \begin{pmatrix} 1 \\ 0 \end{pmatrix} \to \left(E_1 + H_{I,11}\right) |1\rangle$$

And similar for $\hat{H}'_0 |2\rangle$.

So we can always absorb the diagonal elements of our interaction Hamiltonian into the unperturbed Hamiltonian without having to do any new calculations of eigenstates, while the eigenvalues change in a trivial manner. We are then left with an interaction Hamiltonian that can only produce transitions between different eigenstates of the unperturbed Hamiltonian.