Consider a separable Hilbert space Ω and a complete, normalized (though not orthogonal) discrete set $\{ | \alpha> \}$ in it. Naively, one may hope that the completeness of the set implies the existence of a measure I_α that would allow an expansion of the unity operator in the form:

$$\hat{I} = \sum_\alpha | \alpha> I_\alpha <\alpha | \ . \tag{1}$$

It is easy to see that, if that were the case then we would have

$$I_\alpha = \sum_\beta S_{\alpha\beta} \ , \tag{2}$$

where $S = Q^{-1}$ is the inverse to the matrix $Q_{\alpha\beta} = | <\alpha | \beta> |^2$:

$$\sum_\mu S_{\alpha\mu} | <\mu | \beta> |^2 = \delta_{\alpha\beta} \ .$$

(inverse to a symmetric matrix is symmetric). Unfortunately, the expansion of the form (1) is not possible, in general. There are examples where it is possible, notably the representation of $\hat{I}$ as an integral $(1/\pi) \int | \alpha> <\alpha | \ d^2\alpha$ over the complex plane of coherent states; however this is the only example I know and the set $\{ | \alpha> \}$ in this case is (a) overcomplete and (b) continuous.

Let us define a conjugate set $\{ | \tilde{\alpha}> \}$ by the condition

$$<\tilde{\alpha} | \beta> = \delta_{\alpha\beta} = <\tilde{\beta} | \alpha> \ .$$

Since, obviously, for any state $| \psi> \in \Omega$ we can write

$$| \psi> = \sum_\alpha | \alpha> <\tilde{\alpha} | \psi> \ ,$$

it follows that the bi-orthogonal set *does* allow expansions of the unity:

$$\hat{I} = \sum_\alpha | \alpha> <\tilde{\alpha}| \ = \sum_\alpha | \tilde{\alpha}> <\alpha | \ .$$

The members of the conjugate set are expressed in terms of the original set as follows:

$$| \tilde{\alpha}> = \sum_\beta | \beta> <\tilde{\beta} | \tilde{\alpha}> \ ,$$

and the matrix $\tilde{\Phi}_{\alpha\beta} = <\tilde{\alpha} | \tilde{\beta}>$ is inverse to $\Phi_{\alpha\beta} = <\alpha | \beta>$:

$$\sum_\mu <\tilde{\alpha} | \tilde{\mu}> <\mu | \beta> = \delta_{\alpha\beta} \ .$$

The assumption that expansion (1) exists, leads to a contradiction. Indeed, assuming (1) we should be able to write

$$| \beta> = \sum_\alpha | \alpha> I_\alpha <\alpha | \beta> \ ,$$

whence

$$<\tilde{\mu} | \beta> = \delta_{\mu\beta} = \sum_\alpha <\tilde{\mu} | \alpha> I_\alpha <\alpha | \beta> = I_\mu <\mu | \beta> \neq \delta_{\mu\beta} \ .$$

The bi-orthogonal set is convenient in problems of finding the eigenvalues of operators whose matrices are given in the $\{\,|\,\alpha\rangle\}$ representation. Let H be a hermitean operator in Ω with the eigenvalues E_m: $H\,|\,m\rangle = E_m\,|\,m\rangle$. Determination of these eigenvalues reduces to the diagonalization of the "ugly" matrix

$$H_{\alpha\beta} \equiv \langle\tilde{\alpha}\,|\,H\,|\,\beta\rangle\ ,$$

and not of the hermitean matrix

$$\tilde{H}_{\alpha\beta} = \langle\alpha\,|\,H\,|\,\beta\rangle = \sum_{\mu}\langle\alpha\,|\,\mu\rangle\,H_{\mu\beta}\ .$$

as one could naively think. Indeed, any eigenstate $|\,E\rangle$ can be expanded in $\{\,|\,\alpha\rangle\}$ as follows:

$$|\,E\rangle = \sum_{\beta}\,|\,\beta\rangle\langle\tilde{\beta}\,|\,E\rangle\ ,$$

and determination of the coefficients $\langle\tilde{\beta}\,|\,E\rangle \equiv C_\beta$ leads to a secular equation of the form

$$\sum_{\beta}\left[H_{\alpha\beta} - E\,\delta_{\alpha\beta}\right]C_\beta = 0\ .$$

Even though the matrix $H_{\alpha\beta}$ is not a hermitean matrix (e.g., its diagonal elements are, in general, complex), its eigenvalues are real and equal to E_m. It is easy to prove the following useful property: $\sum_{\alpha}H_{\alpha\alpha} = \mathrm{tr}\,H$. Indeed,

$$\langle\tilde{\alpha}\,|\,H\,|\,\alpha\rangle = \sum_{m}\langle\tilde{\alpha}\,|\,m\rangle E_m\langle m\,|\,\alpha\rangle = \sum_{m}\langle m\,|\,\alpha\rangle\langle\tilde{\alpha}\,|\,m\rangle E_m\ ,$$

whence the trace is obtained by summing over α and using expansion of the unity operator.

Both matrices $H_{\alpha\beta}$ and $\tilde{H}_{\alpha\beta}$ give an unambiguous representation of the operator H:

$$H = \sum_{\alpha\beta}|\,\alpha\rangle\langle\tilde{\alpha}\,|\,H\,|\,\beta\rangle\langle\tilde{\beta}\,| \equiv \sum_{\alpha\beta}|\,\alpha\rangle H_{\alpha\beta}\langle\tilde{\beta}\,|\ ;$$

$$H = \sum_{\alpha\beta}|\,\tilde{\alpha}\rangle\langle\alpha\,|\,H\,|\,\beta\rangle\langle\tilde{\beta}\,| \equiv \sum_{\alpha\beta}|\,\tilde{\alpha}\rangle\tilde{H}_{\alpha\beta}\langle\tilde{\beta}\,|\ ,$$

and both possess a set of real eigenvalues and a set of eigenstates that is complete in Ω, but whereas the eigenvalues of the non-hermitean matrix $H_{\alpha\beta}$ coincide with those of the operator H, the eigenvalues of the hermitean $\tilde{H}_{\alpha\beta}$ are those of a different operator

$$H_{\mathrm{mod}} = \sum_{\alpha\beta}|\,\alpha\rangle\langle\alpha\,|\,H\,|\,\beta\rangle\langle\tilde{\beta}\,| \quad \neq H\ ,$$

which coincides with H only when $\{\,|\,\alpha\rangle\}$ is an orthonormal set. If we diagonalize $\langle\alpha\,|\,H\,|\,\beta\rangle$, that is find the eigenstates of H_{mod}, and then evaluate the average of H in these states, we will, generally, find a set of upper bounds to the eigenvalues of $H_{\alpha\beta}$ – as follows from the variational principle. To find the eigenvalues of H we must first calculate $\langle\tilde{\alpha}\,|\,\tilde{\beta}\rangle$ and then $H_{\alpha\beta} = \sum_{\mu}\langle\tilde{\alpha}\,|\,\tilde{\mu}\rangle H_{\mu\beta}$.

Example: coupled well problem.

Let ϕ_i $(i=1, 2)$ be the single-well functions (neglecting the presence of the other well). Define $\Phi_{ij} = <\phi_i \mid \phi_j>$ or

$$\Phi = \begin{bmatrix} 1 & t \\ t^* & 1 \end{bmatrix} .$$

The matrix $\tilde{\Phi}$ defined by $\tilde{\Phi}_{\alpha\beta} = <\tilde{\phi}_i \mid \tilde{\phi}_j>$ is of the form

$$\tilde{\Phi} = \Phi^{-1} = \frac{1}{1 - |t|^2} \begin{bmatrix} 1 & -t \\ -t^* & 1 \end{bmatrix} .$$

The matrix S_{ij} which is inverse to $Q_{ij} = |<\phi_i \mid \phi_j>|^2$ is given by

$$S = \frac{1}{1 - |t|^4} \begin{bmatrix} 1 & -|t|^2 \\ -|t|^2 & 1 \end{bmatrix} ,$$

and the "measures" I_i by

$$I_i = \frac{1}{1 + |t|^2} , \quad (i=1, 2) .$$

We see explicitly that $\sum_i |i> I_i <i| \neq \hat{I}$. Let us construct the conjugate set:

$$\tilde{\phi}_1 = \phi_1 <\tilde{\phi}_1 \mid \tilde{\phi}_1> + \phi_2 <\tilde{\phi}_2 \mid \tilde{\phi}_1> = \frac{\phi_1 - t^* \phi_2}{1 - |t|^2} ;$$

$$\tilde{\phi}_2 = \phi_1 <\tilde{\phi}_1 \mid \tilde{\phi}_2> + \phi_2 <\tilde{\phi}_2 \mid \tilde{\phi}_2> = \frac{\phi_2 - t \phi_1}{1 - |t|^2} ;$$

Suppose we know the matrix elements of an operator H in the basis ϕ_i:

$$<\phi_i \mid H \mid \phi_j> \equiv \tilde{H}_{ij} = \begin{bmatrix} \varepsilon_1 & V \\ V^* & \varepsilon_2 \end{bmatrix} ,$$

where, without a loss in generality we assume $\varepsilon_1 + \varepsilon_2 = 0$ and $\varepsilon_2 - \varepsilon_1 \equiv \Delta$. The eigenvalues of $\tilde{H}_{ij}$ are $\tilde{\lambda} = \pm\sqrt{|V|^2 + \Delta^2/4}$. Let us now construct H_{ij}:

$$H_{ij} \equiv <\tilde{\phi}_i \mid H \mid \phi_j> = \frac{1}{1 - |t|^2} \begin{bmatrix} \varepsilon_1 - tV^* & V - t\varepsilon_2 \\ V^* - t^*\varepsilon_1 & \varepsilon_2 - t^*V \end{bmatrix} ,$$

and determine its eigenvalues

$$(1 - |t|^2)\, \lambda = -\mathrm{Re}\,(tV^*) \pm \sqrt{[\mathrm{Re}\,(tV^*)]^2 + (|V|^2 + \Delta^2/4)(1 - |t|^2)} .$$

It can be shown that if the eigenvalues are arranged in the order of ascending magnitude, then $\lambda_i < \tilde{\lambda}_i$. Let us explicitly demonstrate this, taking $V = t$ (energy overlap integral with a unit uniform barrier) and $\Delta \ll |V|$ (near the resonance). In this example, we find

$$\lambda_1 \approx \frac{-|t|}{1 - |t|} - \frac{\Delta^2}{8|t|} \approx \tilde{\lambda}_1 - \frac{|t|^2}{1 - |t|} ;$$

$$\lambda_2 \approx \frac{|t|}{1 + |t|} + \frac{\Delta^2}{8|t|} \approx \tilde{\lambda}_2 - \frac{|t|^2}{1 + |t|} .$$

notes added in Aug 2003

Nonorthogonal states and sensorics

A common problem in sensorics is to determine the contribution of each of several known functions into an experimentally measured quantity,

$$F = \sum_i c_i f_i \,,$$

where F is an experimentally measured function of some variable or group of variables *x* and $\{f_i\}$ is a set of distinguishable functions of same variables *x*.

Typical examples of the set $\{f_i\}$ are distinguishable fluorescent spectra $\{f_i(\lambda)\}$ or distinguishable temporal signals $\{f_i(t)\}$. In either case, the problem is to determine the weights $\{c_i\}$ from an experimental measurement *F*. In this case, the word *distinguishable* has a simple mathematical meaning: the functions $\{f_i\}$ must be *linearly independent*. This means they form a basis in the vector space

$$\mathrm{H} = \mathrm{span}\,\{f_i\}.$$

Evidently, the basis needs not be (and usually is not) orthogonal with respect to a scalar product defined in H,

$$M_{ij} = \left\langle f_i \middle| f_j \right\rangle$$

The matrix M_{ij} in general is not a diagonal matrix. Of course, without any loss in generality, we can consider the basis functions normalized to unity, $M_{ii} = 1$, for all *i*.

The general solution to the problem can be written in terms of the conjugate basis $\{\widetilde{f}_i\}$, defined by $\left\langle \widetilde{f}_i \middle| f_j \right\rangle = \delta_{ij}$, i.e.,

$$\widetilde{f}_i = \sum_j \left[M^{-1}\right]_{ij} f_j \,.$$

Whence we have

$$c_i = \left\langle \widetilde{f}_i \middle| F \right\rangle \,.$$

The scalar product in H can be defined in a variety of ways and in practice it should be defined to suit the problem at hand. For bell-shaped distributions the most common definition is probably in terms of an integral over the appropriate set of variables:

$$\left\langle f_i \middle| f_j \right\rangle \equiv \int f_i^*(x) f_j(x)\, dx \,,$$

The asterisk has been placed on the first function merely to emphasize that the procedure works equally well for complex functions and their linear combinations with complex coefficients c_i, i.e. in a Hilbert space H. The dimensionality of this space, i.e. the number of basis states $\{f_i\}$ in principle does not have to be finite, but the procedure of inverting the matrix of scalar products may be tricky in the case of infinite dimensionality.

In what follows I assume a finite-dimensional set $\{f_i\}$ comprising *n* functions. It is worth noting that, besides the useful functions we are seeking (representing the fluorescent markers $\{f_i(\lambda)\}$ or fiducial signals $\{f_i(t)\}$), the set may include parasitic or noise functions (e.g. represented by their

spectral decomposition). The procedure will automatically filter out all such functions so long as we have included them in the set.

As an example consider the problem of *least-square fit*, which is a common procedure to minimize an integrated deviation of the assumed curve from the actual data points. I will show that the least square fit is a particular example of the described above procedure.

To be specific, suppose we are dealing with an experimentally measured waveform $F(t)$ that we wish to fit to a linear combination of n known functions $\{f_i(t)\}$,

$$F(t) = \sum_{i=1}^{n} c_i f_i(t) .$$

We assume that $F(t)$ was measured at m successive points in time t_α ($\alpha = 1,2,3...,m$), and denote the measured values as $F(t_\alpha) = u_\alpha$. The least square procedure corresponds to finding the minimum with respect to the coefficients $\{c_i\}$ of the deviation function $D\{c_i\}$:

$$D\{c_i\} = \sum_{\alpha=1}^{m} \left(\sum_{i=1}^{n} c_i f_i(t_\alpha) - u_\alpha \right)^2 .$$

As is well known, the necessary condition for the minimum is that it is at least a local minimum, i.e. the coefficients c_i must satisfy the set of n equations $\partial D / \partial c_i = 0$. This yields a matrix equation of the form $\mathbf{Ac} = \mathbf{b}$, where

$$A_{ij} = \sum_{\alpha=1}^{m} f_i(t_\alpha) f_j(t_\alpha)$$

$$b_j = \sum_{\alpha=1}^{m} f_j(t_\alpha) u_\alpha$$

Exactly the same result is obtained in the procedure described above if the scalar product is defined by an integral over t approximated as a sum over α of functions evaluated at t_α, viz.

$$\left\langle f_i \middle| f_j \right\rangle \equiv \frac{m}{T} \int_0^T f_i(t) f_j(t)\, dt \quad \approx \quad \sum_{\alpha=1}^{m} f_i(t_\alpha) f_j(t_\alpha)$$

In this case, evidently, $M_{ij} = A_{ij}$ and we have

$$c_i = \left\langle \tilde{f}_i \middle| F \right\rangle = \sum_{\alpha=1}^{m} \tilde{f}_i(t_\alpha) u_\alpha = \sum_{\alpha=1}^{m} \sum_{j=1}^{n} M_{ij}^{-1} f_j(t_\alpha) u_\alpha$$

which is the same as $\mathbf{c} = \mathbf{M^{-1}}\, \mathbf{b}$.

Clearly, if points t_α are given, e.g. uniformly distributed over the interval of time variation, the least square fit is probably the best procedure and it is gratifying that our procedure produces the least square fir automatically. However, we have an insight as to how to choose points t_α better! Indeed, we can follow the Gauss integration procedure. For any desired basis set $\{f_i\}$ we can choose the points t_α (and the corresponding weights w_α) so as to make exact the evaluation of all n^2 integrals as a "Gaussian" sum,

$$\int f_i(t) f_j(t)\, dt = \sum_{\alpha} w_\alpha\, f_i(t_\alpha) f_j(t_\alpha) .$$