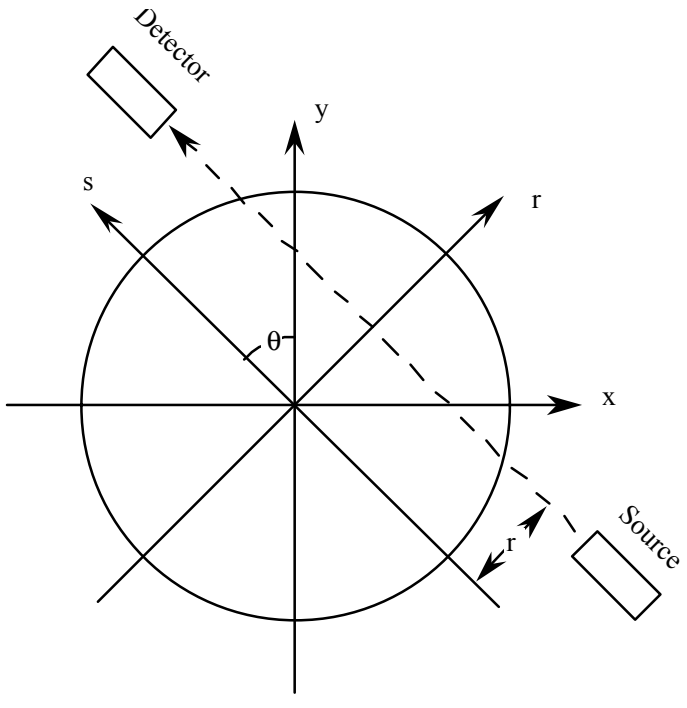
Analytic Tomographic Reconstruction

Much of medical imaging is performed by acquiring x-ray projection data onto a sheet of film or other planar detector. Chest radiography, mammography, angiography, and even dental x-rays are examples of "plain-film" radiography or "x-ray projection imaging". In these techniques, the three-dimensional structure in an object (i.e., human body) is projected as an x-ray shadowgram onto a two-dimensional detector. Often this provides sufficient information to satisfy the clinical requirements of the diagnostic task. One can use projection imaging to find lung tumors, breast cancer, bone fractures, and other disease states.

There are many cases, however, where using projection imaging techniques is inadequate for the diagnostic task. This occurs because the superposition of the complex anatomical structure of the human body onto a two-dimensional image can obscure subtle disease states such as soft-tissue tumors surrounded by soft-tissue organs (brain, liver, kidney). For this applications, the invention of tomography in the 1970's provided a revolution in the way that radiographic information was formed and utilized in medical practice. The term "tomography" means imaging ("graphy") a cut or slice ("tomo") through the body, and generating a tomographic image conceptually is equivalent to slicing up the human body like a loaf of bread, then pulling up a single slice and looking at its cut surface. This allows the diagnostician to see anatomical detail in one section of the body without superposition of overlying or underlying structure. Tomography is a central theme in many types of medical imaging used today, and can be formed using x-ray imaging, nuclear medicine, magnetic resonance imaging, and ultrasound.

The technique of forming a tomographic image using x-rays or radionuclides ("nuclear medicine"), and to a certain extent using magnetic resonance imaging (MRI) are equivalent. We will focus our discussion on x-ray tomography and radionuclide tomography since these are nearly identical. In x-ray tomography, data are acquired by scanning the patient with an x-ray tube and detector so that x-ray "projection data" are acquired across the body at multiple angles. The process is similar in radionuclide tomography in which we use a detector to acquire photons which are emitted by a radiopharmaceutical taken up by different structures (i.e., organs, tissues, lesions) in the body. We then use a mathematical formula (callen an "algorithm") to determine the distribution of linear attenuation coefficients (in the case of x-ray imaging) or radionuclide concentration ($\mu Ci/cm^3$) in the body (in the case of radionuclide tomography) from the x-ray or radionuclide projection data acquired with the imaging equipment. A gray-scale can be assigned to the possible values of the attenuation coefficient μ to generate an image that shows the anatomy of the body for that particular tomographic slice. Our goal in this chapter is to describe the mathematics with which we transform the experimental projection measurements to calculate the map of attenuation coefficients across the object.



We begin by describing an object which lies in the x-y plane by the "density function" μ(x,y). For x-ray imaging, the function fμx,y) describes the spatial distribution of linear attenuation coefficients in one plane of the object. (For radionuclide imaging, the function μ(x,y) describes the spatial distribution of radionuclide in the body. We will concentrate on x-ray tomography, although the mathematics for radionuclide tomography are essentially the same.) As mentioned above, we begin the tomographic process by acquiring "projection" data, obtained by scanning an x-ray source and detector across opposite sides of the object as shown in Figure 1. The source emits a beam of x-rays that is attenuated by the structure in the object while the x-ray beam passes through the object. The x-ray beam emerging from the x-ray beam then is recorded with a radiation detector. Mathematically, it is convenient to define a stationary coordinate system with x-y axes which are fixed in the stationary object. Correspondingly, we define a rotating coordinate system with rs axes such that the s-axis is parallel to the x-ray beam passing through the object (Fig 1). For a monoenergetic x-ray spectrum and a parallel-beam geometry, the photon fluence Φ recorded with the object in the beam is related to the photon fluence Φ_o measured with no object

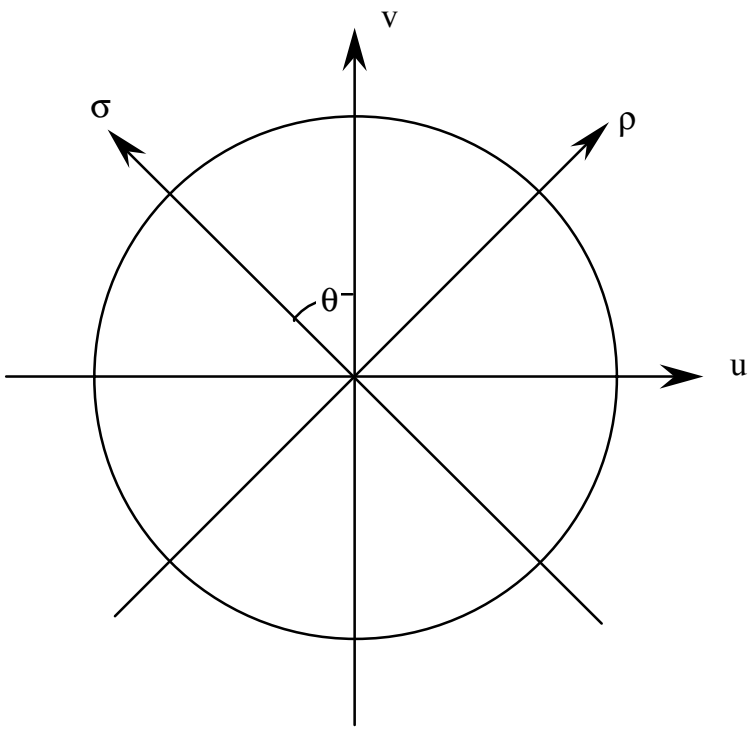


Fig 2: Rotated coordinate axis for tomographic data

$$\Phi = \Phi_o \exp\left[- \int_{-\infty}^{+\infty} \mu(s) ds \right] \qquad (17\text{-}1)$$

In a x-ray imaging measurement, we record the attenuated x-ray intensity Φ acquired at the angle θ with respect to the y-axis. Each measurement obtained with the source-detector pair is called a "projection" or "ray-sum", and can be defined geometrically by the angle θ between the projection axis (i.e.,the "s-axis") and the y axis (Fig 1),and by the perpendicular distance r between the projection line and the s axis.

$$p_\theta(r) = -\ln\left[\frac{\Phi}{\Phi_o}\right] = \int_{r\theta} \mu(s)\, \delta(x\cos\theta + y\sin\theta - s)\, ds = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \mu(x\cos\theta + y\sin\theta)\, dx\, dy \qquad (17\text{-}2)$$

where the argument of the delta function δ describes the linear path along which the x-ray beam travels in generating the projection image (Figure 2). Our goal is to determine the density function (i.e., the map of linear attenuation coefficients across the object) μ(x,y) from the set of projection data $p_\theta(r)$ acquired across different radial positions r and for different projection angles θ. This will be performed using one of two equivalent techniques called Fourier Reconstruction and Filtered Backprojection. These techniques will be described separately.

Fourier Reconstruction

The method of Fourier reconstruction relies on the following derivation, which shows that for a given rotational angle θ, the Fourier transform of the projection data $p_\theta(r)$ is equal to the Fourier transform of the density function μ(x,y) along the r-axis (Fig 2). This can be shown by calculating the Fourier transform of $p_\theta(r)$ as

$$P(\rho,\theta) = \Im[p_\theta(r)] = \int_{-\infty}^{+\infty} p_\theta(r)\ \exp(-2\pi i\rho r)\, dr \tag{17-3}$$

$$= \int_{-8}^{+\infty}\left[\int_{-\infty}^{+\infty}\mu(r,s)ds\right]\exp(-2\pi i\rho r)\, dr = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\mu(r,s)\exp(-2\pi i\rho r)\, dr\, ds$$

$$= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\mu(r,s)\exp[-2\pi i(\rho r+\sigma s)drds]\Big|_{s=0}$$

We therefore see that the Fourier transform P(ρ,θ) of the projection function $p_\theta(r)$ is equal to the Fourier transform of our density function μ(x,y) along the line s = 0 (i.e., along the rotated r-axis in the xy-plane). If this operation is repeated for different values of θ, we can obtain enough samples of the P(ρ,θ) and use interpolation to fill in the entire spatial-frequency plane which is expressed in terms of the coordinates (ρ,θ) or (u,v). The resulting function P(u,v) then can be inverse Fourier transformed to obtain the density function μ(x,y)

$$\mu(x,y) = \Im^{-1}[P(u,v)] = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} P(u,v)\exp[2\pi i(xu+yv)]\, du\, dv \tag{17-4}$$

where P(u,v) is obtained by multiple measurements of P(ρ,θ)

$$P(\rho,\theta) = \Im[p_\theta(r)] \tag{17-5}$$

Convolution Backprojection (or "Filtered Backprojection")

The Fourier Reconstruction method derived above is a general result which can be used for x-ray CT, radionuclide tomography, and magnetic resonance imaging. In fact, while it is generally used for magnetic resonance imaging, it is not often used for x-ray CT or radionuclide tomography. This occurs because the Fourier Reconstruction methods requires that the entire set of projection data must be acquired and Fourier transformed to determine P(u,v) ≡ P(ρ,θ), which we have shown is equivalent to the Fourier transform of the density function μ(x,y) according to Equation 17-4. After P(u,v) is determined, it is inverse transformed to derive the density function μ(x,y) which in the case of x-ray CT is a map (image) of the linear attenuation coefficients across the slice of the object, and which in the case of radionuclide tomography is a map (image) of the radionuclide density across the slice of the object. An alternative method called "Filtered Backprojection" or "Convolution Backprojection" provides an alternative method of achieving the same result. The advantage of "Filtered Backprojection" is that

much of the calculation needed to form the image can be performed on each set of projection data $p_\theta(r)$ for a given angle θ. Compared with Fourier Reconstruction, Filtered Backprojection allows a greater fraction of the calculation to be performed during the acquisition of the projection data. Thus, the image is available for viewing much sooner after the completion of the scan than in Fourier Reconstruction.

The method of Filtered Backprojection can be derived by starting with Equation 17-4, which shows that the density function $\mu(x,y)$ is obtained as the inverse Fourier Transform of P(u,v), expressed in cartesian coordinates in the spatial-frequency domain (u,v). According to Equation 17-5, P(u,v) can be obtained as the Fourier transform of the set of projection data $p_\theta(r)$ acquired with the imaging system. Thus, from Equation 17-4, we have

$$\mu(x,y) = \Im^{-1}\,[P(u,v)] = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} P(u,v)\,\exp[2\pi i(xu + yv)]\,du\,dv \qquad (17\text{-}6)$$

We now write the Fourier transform P(u,v) in polar coordinates $P(\rho,\theta) = P(u,v)$, using the relationships $u=\rho\cos\theta$ and $v=\rho\sin\theta$, and expressing the differential area element du dv as $\rho\,d\rho\,d\theta$, where (ρ,θ) are polar coordinates of (u,v) in the spatial frequency domain. Therefore, Equation 17-6 can be written as

$$\mu(x,y) = \int_{0}^{2\pi}\int_{0}^{+\infty} P(\rho,\theta)\,\exp[2\pi i(x\,\rho\cos\theta + y\,\rho\sin\theta)]\,\rho\,d\rho\,d\theta \qquad (17\text{-}7)$$

$$= \int_{0}^{\pi}\int_{-\infty}^{+\infty} |\rho|\,P(\rho,\theta)\,\exp[2\pi\rho i(x\cos\theta + y\sin\theta)]\,d\rho\,d\theta$$

We note from equation 17-7 that $\mu(x,y)$ can be obtained as the Fourier transform of the function $|\rho|\,P(\rho,q)$, where the Fourier transform is taken in terms of polar coordinates in the spatial frequency domain. While it appears difficult to perform this operation, we can circumvent this difficulty first by noting that $P(\rho,\theta)$ is obtained from $p_\theta(r)$, which represent the projection data measured from the object. In a real system, the projeciton data $p_\theta(r)$ are measured using detectors having finite width and using finite sampling intervals such that the spatial frequency content of $p_\theta(r)$ (and therefore of $P(\rho,\theta)$) is band-limited to some maximum spatial-frequency value M (i.e., we can assume safely that the Fourier transform P has no spatial frequencies above some maximum M). Thus, we can write Equation 17-7 as

$$\mu(x,y) = \int_{0}^{\pi}\int_{-M}^{M} |\rho|\,P(\rho,\theta)\,\exp[2\pi\rho i(x\cos\theta + y\sin\theta)]\,d\rho\,d\theta \qquad (17\text{–}8)$$

We also define a "kernel function", $\kappa(\rho) = M\{\Pi(\rho/2M)\text{-}\Lambda(\rho/M)$, where

$$\Pi(\rho/2M) = \begin{cases} 1 & \text{if } |\rho| < M \\ 0 & \text{if } |\rho| > M \end{cases} \qquad (17\text{-}9)$$

$$\Lambda(\rho/2M) = \begin{cases} 1-\frac{|\rho|}{M} & \text{if } |\rho| < M \\ 0 & \text{if } |\rho| > M \end{cases} \qquad (17\text{-}10)$$

We now can express the inner integral of equation 17-8 in terms of the filtered-projection data p^*_θ, where

$$\mu(x,y) = \int_0^\pi p^*_\theta(r)\ d\theta \qquad (17\text{–}11)$$

where

$$p^*_\theta(r) = \int_0^M |\rho|\ P(\rho,\theta)\ \exp[2\pi\rho i(x\cos\theta + y\sin\theta)]\ d\rho$$

$$= \int_{-\infty}^{+\infty} \kappa(\rho)\ P(\rho,\theta)\ \exp[2\pi\rho i(x\cos\theta + y\sin\theta)]\ d\rho$$

$$= \Im^{-1}[\kappa(\rho)\ P(\rho,\theta)] = \Im^{-1}[\kappa(\rho)] * \Im^{-1}[P(\rho,\theta)] \qquad (17\text{-}12)$$

Finally, we know that the inverse Fourier transform k(r) of the kernel $\kappa(\rho)$ can be expressed as

$$k(r) = \Im^{-1}[\kappa(\rho)] = \Im^{-1}[M\{II(\rho/2M)-\Lambda(\rho/M)] = 2M^2\ \text{sinc}(2Mr) - M^2\ \text{sinc}^2(Mr) \qquad (17\text{-}13)$$

while from Equation 17-3 we know that the inverse Fourier transform of $P(\rho,\theta)$ is the projection function $p_\theta(r)$

$$p_\theta(r) = \Im^{-1}[P(\rho,\theta)] \qquad (17\text{-}14)$$

Substituting Equation 17-13 and 17-14 into Equation 17-12, we have

$$p^*_\theta(r) = k(r) * p_\theta(r) \qquad (17\text{-}15)$$

allowing Equation 17-11 to be expressed as

$$\mu(x,y) = \int_0^\pi p^*_\theta(r)\ d\theta = \int_0^\pi [\,k(r) * p_\theta(r)\,]\ d\theta \qquad (17\text{–}16)$$

Equation 17-16 shows that we can acquire projection data $p_\theta(r)$ using the imaging system. This process is repeated at angular increments $\Delta\theta$ for n angular samples, across an entire half-rotation of π radians ($n = \pi/\Delta\theta$). After a single set of projection data $p_\theta(r)$ are acquired at a fixed angle θ, the projection data

are convolved or "filtered" with the kernel function k(r) then backprojected according to the integral in Equation 17-16. The backprojection of the entire set of filtered projection data forms the image represented by the density function μ(x,y) as shown in Equation 17-16.

Sampling Considerations

One important aspect of designing a tomographic imaging system is developing criteria for the sampling requirements of the projection data. In particular, we want to know how many detectors we want to use to acquire one projection data set (i.e., at one angle of the tomographic acquisition) and how many projection sets must be acquired for the entire tomographic projection set. Specifically, assume that we are reconstructing an object with a circular cross-section having a diameter L. Our imaging task is to resolve the object into volume elements (voxels) of width w along the x- and y-axes, and of thickness b along the z-axis. For simplicity, we also assume that we are using a parallel beam geometry in which the object is scanned in n equal spatial increments along the object. Therefore, if the width of each spatial increment is w, then the number of spatial samples n is equal to n = L/w. In acquiring the tomographic data, we must satisfy two sampling criteria.

1. According to Shannon's criterion, if the spatial resolution of the image system has a spatial resolution that corresponds to a spatial frequency M, then the sampling frequency for that object must be 2M. In other words, if the spatial resolution width of the imaging system is r, then the sampling interval must be r/2 (e.,g. the sampling interval must be 5 mm for a system with spatial resolution of 1 cm). Therefore, we will select the sampling increment $w \leq r/2$, so that the number of spatial samples n across the object of width L is N = L/w.

2. In addition, we need adequate numbers of samples so that the tomographic reconstruction algorithm is not underdetermined. Specifically, if the number of spatial samples across the object of width L is n = L/w, then the total number of cells N in the circular object is

$$N = \frac{\text{area of circular object}}{\text{area of individual cell}} = \frac{\pi L^2}{4w^2} = \frac{\pi n^2}{4} \qquad (17\text{-}17)$$

To determine $\pi n^2/4$ independent values of the attenuation coefficient μ with CT requires at least as many indpendent measurements. Therefore, $mn = \pi n^2/4$, showing us that the number of angular samples m is equal to

$$m = \frac{\pi n}{4} \qquad (17\text{-}18)$$

with m equally-spaced angular samples at angles 0 to π radians around the object.

Discrete Form of Filtered Backprojection

From Equations (17-3) and (17-15), we know that the tomogram can be reconstructed in the spatial domain by convolving the projection data ($p_\theta(r)$) with a convolution kernel (k(r)) having the form

$$k(r) = 2M^2\ \text{sinc}(2Mr) - M^2\ \text{sinc}^2(Mr) \qquad (17\text{-}19)$$

where

$$p^{*}_{\theta}(r) = k(r) * p_{\theta}(r) \tag{17-20}$$

Since modern tomographic reconstruction algorithms are implemented in digital form using a computer or other data processing system, we will evaluate Eq. (17-19) and Eq. (17-20) in discrete form. If we express Eq. (17-20) explicitedly as a convolution in integral form, from Eq. (17-19) we know that

$$p^{*}_{\theta}(r) = k(r) * p_{\theta}(r) = \int_{-\infty}^{+\infty} p_{\theta}(r')\left\{2M^2\,\mathrm{sinc}\left[2M(r-r')\right] - M^2\,\mathrm{sinc}^2\left[M(r-r')\right]\right\}dr \tag{17-21}$$

To evaluate the integral in discrete form, we note that the projection data $p_\theta(r)$ is band-limited with a maximum spatial frequency M. Therefore, we must sample the projection data with a sampling frequency 2M, in which the sampling interval is $\Delta r = (2M)^{-1}$. Therefore, if we select our data points so that r and r' are represented by the discrete variables r_n and r_m respectively, where

$$r \approx r_n = \frac{n}{2M} \text{ and } r' \approx r_m = \frac{m}{2M} \tag{17-22}$$

for n,m = 0, ±1, ±2, ..., then we can first evaluate the first half of the integral in discrete form, where

$$\int_{-\infty}^{+\infty} p_{\theta}(r')\left\{2M^2\,\mathrm{sinc}\left[2M(r-r')\right]\right\}dr \approx 2M^2 \sum_{m=-\infty}^{+\infty}\left\{p_{\theta}(r_m)\cdot\mathrm{sinc}\left[2M\left(\frac{n-m}{2M}\right)\right]\right\}\left[\frac{1}{2M}\right] \tag{17-23a}$$

$$= M\sum_{m=-\infty}^{+\infty} p_{\theta}(r_m)\,\mathrm{sinc}(n-m) = M p_{\theta}(r_n) \tag{17-23b}$$

since

$$\mathrm{sinc}(n-m) = \frac{\sin\pi(n-m)}{\pi(n-m)} = \begin{cases} 1 & \text{for n = m} \\ 0 & \text{for n} \neq \text{m} \end{cases} \tag{17-24}$$

Similarly,

$$\int_{-\infty}^{+\infty} p_{\theta}(r')\left\{M^2\,\mathrm{sinc}^2\left[M(r-r')\right]\right\}dr \approx M^2 \sum_{m=-\infty}^{+\infty}\left\{p_{\theta}(r_m)\cdot\mathrm{sinc}^2\left[M\left(\frac{n-m}{2M}\right)\right]\right\}\left[\frac{1}{2M}\right] \tag{17-25a}$$

$$= \frac{M}{2}\sum_{m=-\infty}^{+\infty} p_{\theta}(r_m)\,\mathrm{sinc}^2\left(\frac{n-m}{2}\right) \tag{17-25b}$$

However,

$$\mathrm{sinc}^2\left(\frac{n-m}{2}\right) = \left[\frac{\sin\frac{\pi}{2}(n-m)}{\frac{\pi}{2}(n-m)}\right]^2 = \begin{cases} 1 & \text{if n = m} \\ 0 & \text{if (n – m) is even} \\ \dfrac{4}{\pi^2(n-m)^2} & \text{if (n – m) isodd} \end{cases} \tag{17-26}$$

Substituting Eq. (17-26) into Eq. (17-25) yields

$$\int_{-\infty}^{+\infty} p_{\theta}(r')\left\{M^2\,\mathrm{sinc}^2\left[M(r-r')\right]\right\}dr \approx \frac{M}{2}\left\{p_{\theta}(r_n) + \frac{4}{\pi^2}\sum_{\mathrm{n-m\ odd}}\left[\frac{\mathrm{p}_{\theta}(r_m)}{(n-m)^2}\right]\right\} \tag{17-27}$$

Therefore, from Eqs. (17-21), (17-23), (17-25), and (17-27), we can express the discrete form of the convolution operation as

$$p_\theta^*(r) = k(r) * p_\theta(r) \approx M p_\theta(r_n) - \frac{M}{2}\left\{ p_\theta(r_n) + \frac{4}{\pi^2} \sum_{\text{n-m odd}} \left[\frac{\text{p}_\theta(r_m)}{(n-m)^2} \right] \right\} \tag{17-28}$$

$$p_\theta^*(r_n) = \frac{M}{2} p_\theta(r_n) - \frac{2M}{\pi^2} \sum_{\text{n-m odd}} \left[\frac{\text{p}_\theta(r_m)}{(n-m)^2} \right] \tag{17-29}$$

Finally, we note that the sampling interval w is

$$w = \Delta r = \frac{1}{2M} \quad \text{or} \quad \text{M} = \frac{1}{2\text{w}} \tag{17-30}$$

we know that

$$p_\theta^*(r_n) = \frac{1}{4w} p_\theta(r_n) - \frac{1}{\pi^2 w} \sum_{\text{(n-m) odd}} \left[\frac{\text{p}_\theta(r_m)}{(n-m)^2} \right] \tag{17-31}$$

where the sum is taken over all values of m where (n-m) has an odd value. We then use Eq. (17-16)

$$\mu(\text{x,y}) = \int_0^\pi \text{p}_\theta^*(\text{r})\ \text{d}\theta = \int_0^\pi [\ \text{k(r)} * \text{p}_\theta\text{(r)}\ ]\ \text{d}\theta \tag{17–32}$$

to reconstruct the image using backprojection, which we also can express in discrete form as

$$\mu(x,y) = \sum_{j=1}^{m} p_{\theta_j}^*(x\cos\theta_j + y\sin\theta_j) \cdot \Delta\theta \tag{17-33}$$

for projection data $p_\theta(r)$ taken at angular samples of $\theta = \theta_j$, and at an angular increment $\Delta\theta$.

Bioeng 230B: Spring 2002 **Hasegawa**

Iterative Reconstruction Algorithms

Concept of Detection Probabilities

The detection probability is used in iterative reconstruction techniques and equals the probability that a photon emitted from a source in the object is recorded by a detector element in the imaging system.

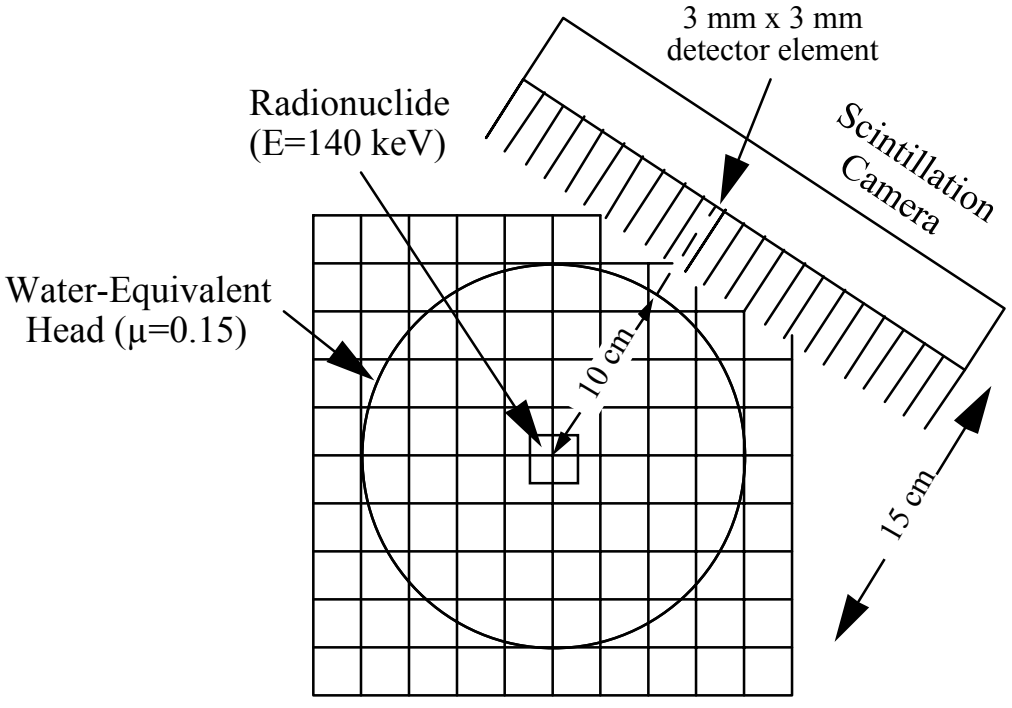
λ_b = concentration of a radionuclide at location b
n_d = number of counts measured at detector element d
P_{db} = probability that a photon emitted from location b is recorded in detector d

Detection probability is calculated from
(a) geometric efficiency of the collimator
(b) attenuation and scatter of photons in the object
(c) detection efficiency of colllimator

Example: A tumor in the center of the head has accumulated a radiopharmaceutical labeled with ^{99m}Tc (140 keV). Gamma rays emitted by the tumor are detected by a scintillation camera, as shown below. A photon emitted by the tumor is recorded by a 3 mm x 3 mm detector element on the scintillation camera where

(a) the head is circular with a diameter of 20 cm,
(b) the perpendicular distance from the tumor and detector is 15 cm,
(c) the head is water-equivalent with $\mu = 0.15\ cm^{-1}$ at 140 keV.
(d) the collimator is 4 cm long with 3 mm x 3 mm square holes aligned with the detector elements

Please calculate the probability that a photon emitted by the tumor is recorded by a 3 mm x 3 mm detector element for the following situations:



(a) the detector element located along the perpendicular between the tumor and the detector surface

(b) the detector element located at 30 degrees from the perpendicular between the tumor and the detector surface.

Solution:

We will assume that the probability of detecting a photon by the detector element is determined by (a) the geometric efficiency of the detector and (b) the attenuation of the photon by the tissue in the head, and that other effects are negligible in comparison.

(a) The geometrical efficiency η_{geom} of detecting a photon from a source 15 cm away from a 3 mm by 3 mm detector is

$$\eta_{geom} = \frac{A_{\text{det}}}{4\pi R^2} = \frac{(3mm)^2}{4\pi(15cm)^2} = 3.183 \text{ x } 10^{-7}$$

The transmission through the 10 cm thickness of the head (water-equivalent) is

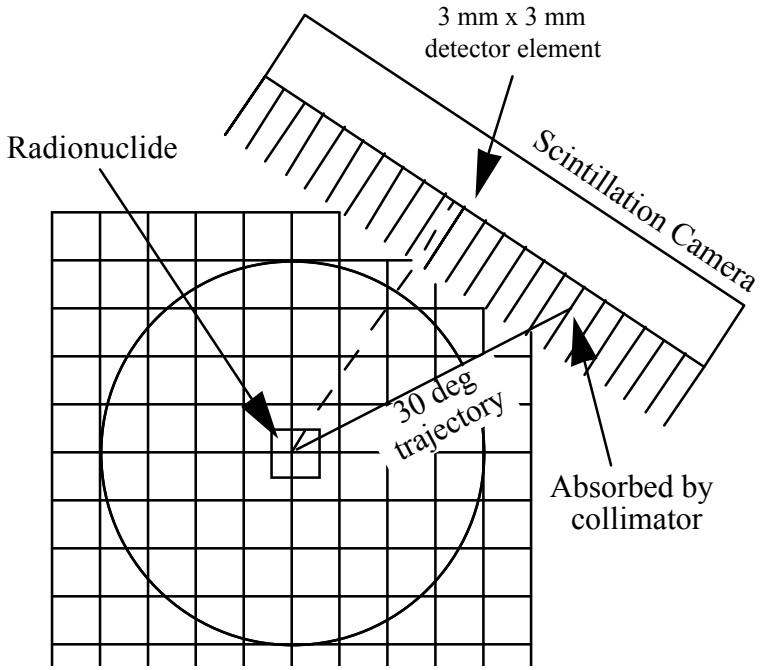
$$\eta_{\text{trans}} = \exp[-\mu_w x_w] = \exp[-(0.15 \text{ cm}^{-1})(10 \text{ cm})] = 2.23 \text{ x } 10^{-1}$$

Therefore, the total detection probability for this situation is

$$\eta_{\text{total}} = \eta_{\text{geom}} \ \eta_{\text{trans}} = (3.183 \text{ x } 10^{-7})(2.23 \text{ x } 10^{-1}) = 7.10 \text{ x } 10^{-8}$$

(b) As shown in the diagram (see next page), if the detector element is 30 degrees from the perpendicular between the tumor and detector surface, then any emiited photons will be absorbed by the collimator septa and will not reach the detector. Therefore, the total detection probability for this situation is

$$\eta_{\text{total}} = 0$$



Iterative Reconstruction Algorithm

In a radionuclide imaging device, we measure photons emitted by a radioactive object using an external detector array. We then can use the recorded projection data to calculate or mathematically "reconstruct" an image representing the radionuclide distribution at each point in the object. One

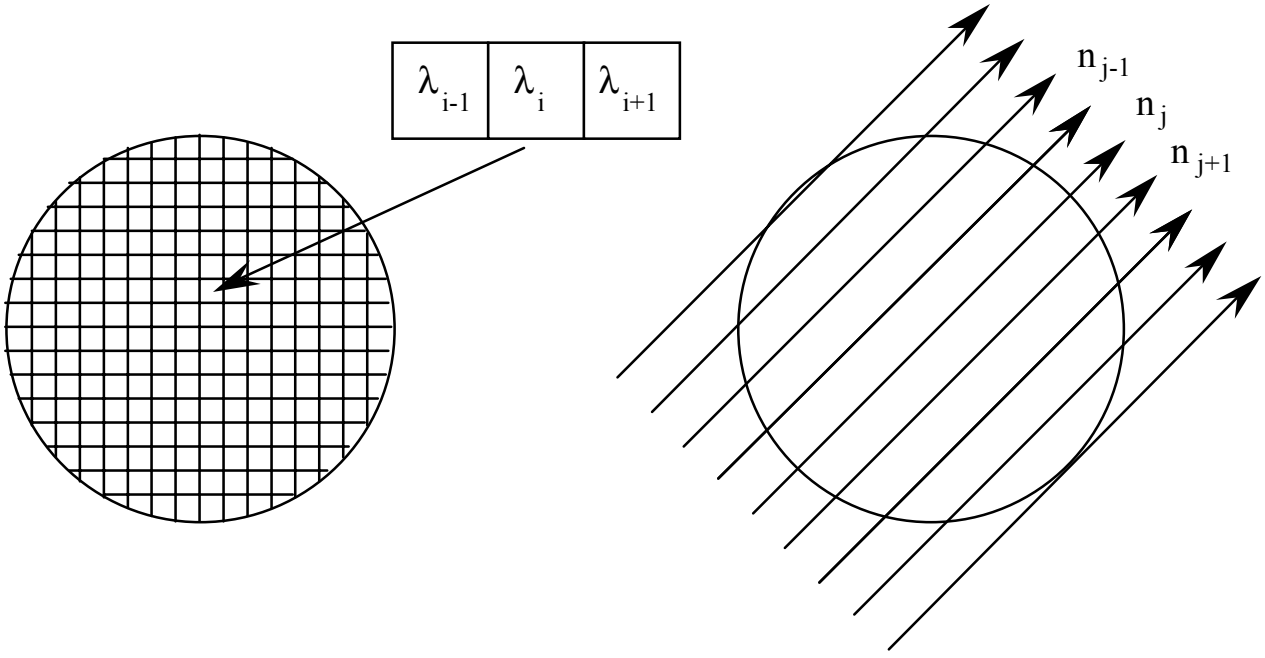


Image represents spatial distribution of radionuclide concentration in the object

Projection data represent the data recorded by each detector element

method of doing this calculation is by means of an iterative reconstruction algorithm. As before, we begin by assuming that we are imaging an object having a radionuclide distribution represented by the matrix $\{\lambda_b\}$, where the index b denotes the location of an individual volume element ("voxel") in the object. The purpose of the tomographic reconstruction algorithm is to calculate the radionuclide image $\{\lambda_b\}$ from the set of counts $\{n_d\}$ recorded by the external detector array surrounding the patient, where the index d identifies each detector element used in making our measurement from the patient. In this mathematical derivation, we represent the radionuclide image $\{\lambda_b\}$ as a "row" vector, but in fact it generally is represented as a two-dimensional (or even a three-dimensional) matrix in which each matrix element λ_b represents the radionuclide concentration at a point identified by the index b. Similarly, we also represent the recorded radionuclide projection data $\{n_d\}$ as a "row" vector, but in fact it generally is represented as a two-dimensional (or even a three-dimensional) matrix in which each matrix element n_d identifies both the angle and the location on the detector, where the data are recorded.

For the sake of argument, let's assume that we completely understand the physics of the medical imaging system. In particular, we assume that we know the detection probability P_{db} that a photon emitted in voxel b is detected in detector d. The value of the detection probability P_{db} is dictated by physical phenomena such as the geometric efficiency of the collimator, the attenuation provided by the material surrounding the radioactive voxel b, and the detection efficiency of the imaging system. We then can express the behavior of the imaging system using a matrix equation coupling the radionuclide concentration $\{\lambda_b\}$ in the object (i.e., in the set of voxels comprising the object) to the complete set of detector measurements $\{n_d\}$, using the known value of the probability elements P_{db} :

$$\begin{pmatrix} n_1 \\ n_2 \\ \cdots \\ n_d \\ \cdots \\ n_D \end{pmatrix} = \begin{pmatrix} P_{11} & P_{12} & \cdots & P_{1b} & \cdots & P_{1B} \\ P_{21} & P_{22} & \cdots & P_{2b} & \cdots & P_{2B} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ P_{d1} & P_{d2} & \cdots & P_{db} & \cdots & P_{dB} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ P_{D1} & P_{D2} & \cdots & P_{Db} & \cdots & P_{DB} \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \cdots \\ \lambda_b \\ \cdots \\ \lambda_B \end{pmatrix} \quad (1)$$

In an ideal world, since we measure the values of $\{n_d\}$ with our detector array, and since we presumably know the value of the probability elements P_{db} from the physics of the imaging system, we could invert the matrix equation (Eq 1) mathematically to solve for the radionuclide distribution (i.e., generate the cross-sectional image) λ_b . There are, however, several important factors that keeps us from taking this direct approach.

(1) The size of the probability matrix (P_{db} } is immense! A typical nuclear medicine study might acquire a set of 64 projection images spread over 180°, where each projection image is recorded in a 64 x 64 matrix format. Thus, the number of measured detector values is $(64)^3 = 262{,}144$. Thus, the matrix $\{n_d\}$ has 262,144 elements, with each element representing the number of counts recorded by a point on the detector at one specific angular projection. Similarly, the radionuclide distribution is reconstructed, for example, as a set of 32 slices in which each slice is represented as a 64 x 64 matrix. Thus, the radionuclide concentration $\{\lambda_b\}$ must be determined at N separate points, where $N = 32(64)^2 = 131{,}072$. Since each detector bin measurement is connected to a

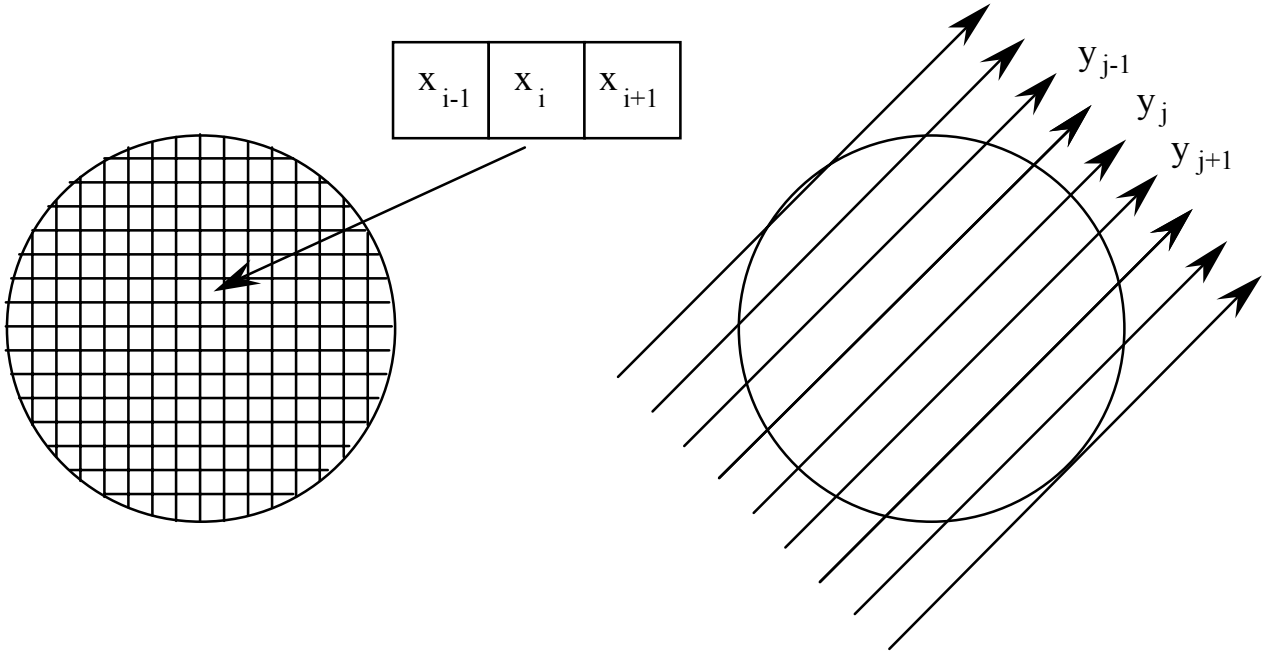
radionuclide concentration element through the probability matrix P, the probability matrix P has 131,072 columns and 262,144 rows, or over 34 billion matrix elements.

(2) The probability matrix $\{P_{db}\}$ is very sparse, i.e., most of the matrix elements are zero. This is especially true in nuclear medicine where data are acquired with a highly collimated detector such that only a few detector elements record photons emitted by a given radioactive point in the object.

(3) The detector values $\{n_d\}$ are noisy (in a statistical sense), and are not known exactly. In fact, in a typical nuclear medicine image, each detector element right record a few hundred counts (at most), so that the percentage statistical uncertainty in the value of each detector value n_d will be in the range of 5% to 10%. In mathematical terminology, we say that the probability matrix $\{P_{db}\}$ is nearly singular and the matrix equation (Eq. 1) cannot be solved uniquely for the radionuclide distribution $\{\lambda_b\}$.

These factors make direct calculation of the probability matrix P_{db} difficult, and make it impossible to invert the matrix equation (Eq. 1) to solve for the spatial distribution of radionuclide concentration $\{\lambda_b\}$, even if we have carefully measured the number of counts recorded by each detector element $\{n_d\}$.

Fortunately, there are several methods to solve for the radionuclide concentration $\{\lambda_b\}$ for a given set of detector measurements $\{n_d\}$. One of the most satifying approaches (we will elaborate this point later) is the "maximum-likelihood expectation-maximization" iterative reconstruction algorithm, generally abbreviated "ML-EM". Referring back to matrix Eq. 1, since we do not know the radionuclide concentration $\{\lambda_b\}$, we can assume a solution which we will denote by the matrix $\{x_b\}$. That is, $\{x_b\}$ is our best estimate of the true radionuclide distribution $\{\lambda_d\}$; we will use successive estimation of the radionuclide distribution $\{x_b\}$ to converge toward the actual radionuclide distribution $\{\lambda_d\}$.



Estimated radionuclide distribution

Estimated projection distribution calculated from estimated radionuclide distribution

We being the algorithm by assuming that $\{x_b\}$ is the radionuclide distribution in the object – that is, that x_b represents the radionuclide concentration at voxel b. Since we know that the matrix $\{P_{bd}\}$ represents the probability that photons emitted from voxel (bin) b is detected in detector d, we can estimate the detector counts $\{y_d\}$ that would be obtained for the estimated radionuclide distribution $\{x_d\}$ using a matrix equation similar to that given in Eq. 1:

$$\begin{pmatrix} y_1 \\ y_2 \\ \cdots \\ y_d \\ \cdots \\ y_D \end{pmatrix} = \begin{pmatrix} P_{11} & P_{12} & \cdots & P_{1b} & \cdots & P_{1B} \\ P_{21} & P_{22} & \cdots & P_{2b} & \cdots & P_{2B} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ P_{d1} & P_{d2} & \cdots & P_{db} & \cdots & P_{dB} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ P_{D1} & P_{D2} & \cdots & P_{Db} & \cdots & P_{DB} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \cdots \\ x_b \\ \cdots \\ x_B \end{pmatrix} \tag{2}$$

which is equivalent to the summation equation

$$y_d = \sum_{b=1}^{B} P_{db}\, x_b \tag{3}$$

At this point, we have assumed that the object can be represented by the radionuclide distribution $\{x_b\}$, for which we would obtain detector counts $\{y_d\}$. However, we also have measured the "true" detector counts $\{n_d\}$. Since the measured detector counts $\{n_d\}$ are Poisson-distributed, we know that the probability of observing a value n_d, assuming an underlying Poisson distribution and given an estimated mean value y_d for any one detector bin is

$$p(n_d \,|y_d) = \frac{e^{-y_d}\, y_d{}^{n_d}}{n_d!} \tag{4}$$

Furthermore, the probability of observing the measured projection data $\{n_d\}$ for all detector bins, given the assumed radionuclide distribution $\{x_d\}$ is

$$L(n|x) = \prod_{d=1}^{D} P(y_d \,|n_d) = \prod_{d=1}^{D} \frac{e^{-y_d} y_d{}^{n_d}}{n_d!} \tag{5}$$

where L is called the "likelihood function". In other words, L(n|x) represents the probabilistic likelihood that the measured data $\{n_d\}$ is statistically equivalent to the estimated detector data $\{y_d\}$. In the ML-EM algorithm we wish to maximize the likelihood function L, to obtain the estimated detector data $\{y_d\}$ that best fits the measured detector data $\{n_d\}$. Moreover, we do this in a way that provides us with the "maximum-likelihood" estimate $\{x_b\}$ of the true radionuclide distribution $\{\lambda_b\}$. Since the maximization of the likelihood function L is tedious mathematically, we calculate the log-likelihood function $\Lambda = \ln(L)$ such that from Eq. 3 and 5,

$$\Lambda = \sum_{d=1}^{D} [\, -y_d + n_d \ln y_d - \ln(n_d!)\,] = \sum_{d=1}^{D} \left\{ -\sum_{b=1}^{B} P_{db} x_b + n_d \ln\left[\sum_{b=1}^{B} P_{db} x_b \right] - \ln[n_d!] \right\} \tag{6}$$

The natural logarithm is monotonically increasing. Therefore, we can maximize the likelihood function L by maximizing the log-likelihood function Λ

$$\frac{\partial \Lambda}{\partial x_k} = \sum_{d=1}^{D} \left[-\frac{\partial y_d}{\partial x_k} + \frac{n_d}{y_d} \frac{\partial y_d}{\partial x_k} \right] = 0 \tag{7}$$

From Eq. 3, we know that

$$\frac{\partial y_d}{\partial x_k} = \frac{\partial}{\partial \partial x_k} \left[\sum_{b=1}^{B} P_{db} x_b \right] = \mathrm{P_{dk}} \tag{8}$$

Substituting Eq. 8 into Eq. 7 gives us

$$\sum_{d=1}^{D} \left[-\mathrm{P_{dk}} + \frac{\mathrm{n(d)}}{\mathrm{y_d}} \mathrm{P_{dk}} \right] = 0 \tag{9}$$

into which we can substitute Eq. 3 to obtain

$$\sum_{d=1}^{D} P_{dk} = \sum_{d=1}^{D} \left[\frac{n_d}{\sum_{b=1}^{B} P_{db} x_b} P_{dk} \right] \tag{10}$$

which provides

$$\frac{1}{\sum_{d=1}^{D} P_{dk}} \sum_{d=1}^{D} \left[\frac{n_d}{\sum_{b=1}^{B} P_{db} x_b} P_{dk} \right] = 1 \tag{11}$$

thus satisfying the condition for maximizing the log-likelihood function (and thereby maximizes the likelihood function Eq. 5).

Our goal is to find the best estimate of the radionuclide distribution $\{x_b\}$ which satisfies the condition specified by Eq. 11. A common way of doing this is using an iterative technique known as the maximum-likelihood expectation-maximization algorithm in which we assume a radionuclide distribution $\{x_b^{(n)}\}$ that is our best estimate of the true radionuclide distribution $\{\lambda_b\}$, where the superscript "(n)" denotes the n^{th} iterative estimate of the radionuclide distribution

$$x_k^{(n+1)} = x_k^{(n)} \left\{ \frac{1}{\sum_{d=1}^{D} P_{dk}} \sum_{d=1}^{D} \left[\frac{n_d}{\sum_{b=1}^{B} P_{db} x_b^{(n)}} P_{dk} \right] \right\} \tag{12}$$

We start with the estimate $\{x_k^{(0)}\}$ (i.e., iteration n = 0) and use Eq. 12 to obtain an improved estimate $\{x_k^{(1)}\}$. We then use $\{x_k^{(1)}\}$ to obtain $\{x_k^{(2)}\}$ which represents an improved estimate of the radionuclide distribution $\{\lambda_k\}$, and so on. We continue this iterative process until our estimate $\{x_k^{(n+1)}\}$ satisfies some convergence criterion, at which point we stop the calculation and use $\{x_k^{(n+1)}\}$ to

represent the true (but unknown) radionuclide distribution $\{\lambda_k\}$. We can, for example, decide to stop the iterative calculation when the calculated value of the likelihood-function (Eq. 5) reaches some predetermined value. Another possible stopping condition would be imposed, for example, if the mean-squared difference between the pixel values from one iteration were sufficiently close to those obtained from the next iterative estimate, i.e., when

$$S = \sum_{b=1}^{B} (x_{b+1}^{(n)} - x_b^{(n)})^2 \tag{13}$$

is smaller than some predetermined level. The most common stopping condition invokes the practical criterion that the calculation (Eq. 12) is performed for a preset number (e.g., 20 to 30) of iterations, at which point the estimated radionuclide distribution $\{x_b^{(n)}\}$ represents our true radionuclide distribution $\{\lambda_b\}$. This latter stopping condition, and our choice of the "optimal" number of iterations, can be based on how many iterations gave reasonable estimates of the radionuclide distribution in previous reconstructions of the same or similar objects.

The maximum-likelihood expectation-maximization (ML-EM) algorithm (Eq. 12) has several important properties that we will state without proof. First, the ML-EM algorithm converges to the maximum-likelihood solution, and it converges monotonically. Second, the ML-EM algorithm automatically imposes a nonnegativity constraint; the estimated values $x_b^{(n)}$ never assume negative values, which corresponds to our observation that the radionuclide concentration λ_b cannot be negative. Third, the estimated detector measurement $\{y_d^{(n)}\}$ give a maximum-likelihood solution that approximate the true detector measurement $\{n_d\}$ in a Poisson-statistical sense. Finally, the ML-EM algorithm preserves counts; the total number of counts in the estimated detector measurement equals the total number of counts in the actual detector measurement

$$\sum_{d=1}^{D} x_d^{(n)} = \sum_{d-1}^{D} n_d \tag{14}$$

All of these properties are ones that agree with physically observed properties of realistic radionuclide distributions and detector measurements, and thereby make the ML-EM solution one that is satisfying conceptually.

Note that the ML-EM algorithm (Eq. 12) contains the elements of our generic iterative algorithm including a projector, a comparator, and a corrector. That is, the sum

$$y_d^{(0)} = \sum_{b=1}^{B} P_{db}\, x_b^{(0)} \tag{15}$$

in the denominator of Eq. 12 also represents the backprojection of the estimated radionuclide distribution $\{x_b^{(0)}\}$, while the expression

$$\frac{n_d}{\sum_{b=1}^{B} P_{db} x_b^{(0)}} \tag{16}$$

compares the estimated projection (Eq. 15) with the measured projection value n_d. Next, the term

$$\sum_{d=1}^{D}\left[\frac{n_d}{\sum_{b=1}^{B} P_{db}x_b^{(n)}}P_{dk}\right] \tag{17}$$

backprojects the correction factors obtained in Eq. 16, which then are used to update the previous iteration $\{x_k^{(n)}\}$ to obtain the next estimate $\{x_k^{(n+1)}\}$ of the radionuclide distribution

$$x_k^{(n+1)} = x_k^{(n)}\left\{\frac{1}{\sum_{d=1}^{D} P_{dk}}\sum_{d=1}^{D}\left[\frac{n_d}{\sum_{b=1}^{B} P_{db}x_b^{(n)}}P_{dk}\right]\right\} \tag{18}$$

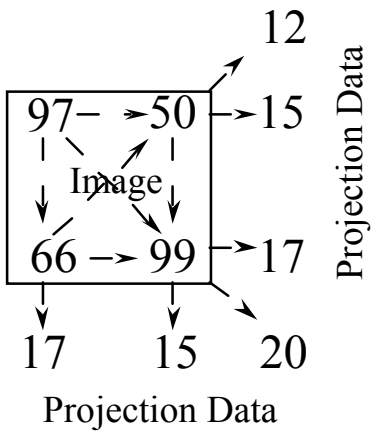
Finally, the demoninator in Eq. 12

$$\sum_{d=1}^{D} P_{dk} \tag{19}$$

is the "sensitivity term" and represents the probability that a photon emitted from radioactive voxel k, will be detected in one of the detector elements (d=1 to D).

Example: Let's assume that a simple object can be represented by a 2x2 matrix $\{\lambda_j\}$ from which we obtain six projection measurements $\{n_i\}$ (figure at right). Furthermore, we know that each detector measurement is obtained as the sum of counts along the ray intersecting two pixels, where the probability of detection from each voxel lying along the ray is

$$P_{ij} = \begin{cases} 0.1 & \text{if voxel j lines along the ray intersecting detector i} \\ 0 & \text{if voxel j does not lie along the ray intersecting detector i} \end{cases} \tag{20}$$



If we apply this model to the following radionuclide distribution that is represented spatially as

$$\{\lambda_i\} = \begin{pmatrix} \lambda_1 & \lambda_2 \\ \lambda_3 & \lambda_4 \end{pmatrix} = \begin{pmatrix} 97 & 50 \\ 66 & 99 \end{pmatrix} \tag{21}$$

the imaging process can be represented in matrix form as

$$\begin{pmatrix} n_1 \\ n_2 \\ n_3 \\ n_4 \\ n_5 \\ n_6 \end{pmatrix} = \begin{pmatrix} P_{11} & P_{12} & P_{13} & P_{14} \\ P_{21} & P_{22} & P_{23} & P_{24} \\ P_{31} & P_{32} & P_{33} & P_{34} \\ P_{41} & P_{42} & P_{43} & P_{44} \\ P_{51} & P_{52} & P_{53} & P_{54} \\ P_{61} & P_{62} & P_{63} & P_{64} \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \lambda_4 \end{pmatrix} \tag{22}$$

For our specific example, we obtain the projection data

$$\begin{pmatrix} 12 \\ 15 \\ 17 \\ 20 \\ 15 \\ 17 \end{pmatrix} = \begin{pmatrix} 0 & 0.1 & 0.1 & 0 \\ 0.1 & 0.1 & 0 & 0 \\ 0 & 0 & 0.1 & 0.1 \\ 0.1 & 0 & 0 & 0.1 \\ 0 & 0.1 & 0 & 0.1 \\ 0.1 & 0 & 0.1 & 0 \end{pmatrix} \begin{pmatrix} 97 \\ 50 \\ 66 \\ 99 \end{pmatrix} \tag{23}$$

Now, let's take the case where we only know the projection data from Eq. 23. We will use these data to perform the first iteration of the ML-EM algorithm. We start with the zeroth iteration by estimating the radionuclide distribution in the object as being uniform and equal to one

$$\{x_i^{(0)}\} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \tag{24}$$

We calculate the first iteration using Eq. 12 for the estimated radionuclide distribution k=1:

$$x_1^{(1)} = x_1^{(0)} \left\{ \frac{1}{\sum_{d=1}^{D} P_{d1}} \sum_{d=1}^{D} \left[\frac{n_d}{\sum_{b=1}^{B} P_{db} x_b^{(0)}} P_{d1} \right] \right\} \tag{25}$$

Note that

$$\left\{ y_d^{(0)} \right\} = \left\{ \sum_{b=1}^{B} P_{db} x_b^{(0)} \right\} = \begin{pmatrix} 0 & 0.1 & 0.1 & 0 \\ 0.1 & 0.1 & 0 & 0 \\ 0 & 0 & 0.1 & 0.1 \\ 0.1 & 0 & 0 & 0.1 \\ 0 & 0.1 & 0 & 0.1 \\ 0.1 & 0 & 0.1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0.2 \\ 0.2 \\ 0.2 \\ 0.2 \\ 0.2 \\ 0.2 \end{pmatrix} \tag{26}$$

$$\left\{ \frac{n_d}{\sum_{b=1}^{B} P_{db} x_b^{(0)}} \right\} = \begin{pmatrix} \frac{12}{0.2} \\ \frac{15}{0.2} \\ \frac{17}{0.2} \\ \frac{20}{0.2} \\ \frac{15}{0.2} \\ \frac{17}{0.2} \end{pmatrix} \tag{16}$$

so that from Eq. 25,

$$\sum_{d=1}^{D}\left[\frac{n_d}{\sum_{b=1}^{B} P_{db}x_b^{(n)}}P_{dk}\right]=\begin{pmatrix}0 & 0.1 & 0.1 & 0\\ 0.1 & 0.1 & 0 & 0\\ 0 & 0 & 0.1 & 0.1\\ 0.1 & 0 & 0 & 0.1\\ 0 & 0.1 & 0 & 0.1\\ 0.1 & 0 & 0.1 & 0\end{pmatrix} \tag{27}$$

Similarly, we can calculate the other elements of the estimated radionuclide distribution $\{x_i^{(1)}\}$, which we can show is equal to

$$\begin{pmatrix}x_1^{(1)} & x_2^{(1)}\\ x_3^{(1)} & x_4^{(1)}\end{pmatrix}=\begin{pmatrix}86.67 & 70\\ 76.67 & 86.67\end{pmatrix} \tag{28}$$

which is our first-estimate of the radionuclide distribution given exactly in Eq. 21.