ON THE DETERMINATION OF THE TOTAL FLUX OF A RADIO SOURCE

by

Ralph J. Gagnon

NRAO Internal Report

September 1965

ABSTRACT

Several methods for determining the flux density of a radio-source from measurements made with a high-gain radio telescope are described. The relationships between the total flux and the telescope record are compared for source distributions in both Gaussian and disc forms. These are presented in a manner to display the wavelength dependence. Comparisons are also made between methods of flux estimation based on the peak received power and the integral of the received power.

The statistical problem of estimating the flux in the presence of random-noise fluctuations is considered. The first case considered is when the source position is known but the amplitude is unknown. It is shown that a good way to estimate the received flux is to multiply the received power record by the antenna pattern and then integrate. The resulting quantity is proportional to the desired flux estimate. The variance of this estimate is also computed.

The second case considered is when both the amplitude and position are unknown. It is shown that a good method for estimating in this case is to use the same method as used when the source position is known, and to compute this estimate for all possible positions. The desired estimate is then the maximum value resulting from these computations.

INTRODUCTION

One of the problems in radio astronomy is the determination of the total flux density of a radio source from measurements made with a high-gain antenna. The measurement made with the receiver system typically consists of a curve which is formed as the antenna beam moves through the source region. For point sources, this curve is the antenna pattern itself, and the flux can be obtained quite simply. However, quite often this smooth antenna-pattern curve is obscured by random noise fluctuations, and it becomes a statistical problem to estimate the total flux. Also, when the source size becomes commensurate with that of the antenna pattern, the relationship between the total flux and the response curve depends on the shape of the source distribution as well as the amplitude. It is the purpose of this paper to discuss some of these statistical and shape-dependent considerations which arise in the determination of radio-noise flux.

THE GAUSSIAN SOURCE

Let us consider a Gaussian source. The brightness, or flux density available per unit solid angle, is given by

$$S(x, y) = \frac{S_o}{a^2 \pi} e^{-\frac{x^2 + y^2}{a^2}} , \qquad (1)$$

where a can be thought of as the source radius, x and y are linearized angular coordinates centered on the source which measure solid angle in steradians, and S_o is the total flux density in watts/m^2/cps, i.e.,

$$S_o = \int S(x, y)\, dxdy. \qquad (2)$$

Let the normalized main-beam power response be approximated by

$$F(x, y) = e^{-\frac{x^2 + y^2}{b^2}} , \tag{3}$$

where b can be considered the beam radius. The main-beam solid angle Ω_M is defined to be

$$\Omega_M = \int_{\substack{\text{MAIN}\\ \text{BEAM}}} F(x, y)\, dxdy = \pi b^2 , \tag{4}$$

a result which follows if we allow the integration limits to become infinite. The sidelobe or stray-region solid angle Ω_S is defined to be

$$\Omega_S = \int_{\substack{\text{STRAY}\\ \text{REGION}}} F\, d\Omega, \tag{5}$$

where we have written the solid angle as $d\Omega$ because the linear range (x, y) is exceeded. The total antenna solid angle Ω_T is the sum of these:

$$\Omega_T = \Omega_S + \Omega_M. \tag{6}$$

We define the beam efficiency η_B to be the ratio of the main-beam area to the total beam area, i.e.,

$$\eta_B = \frac{\Omega_M}{\Omega_T}. \tag{7}$$

Let A_e be the effective aperture of the antenna and A the physical aperture. We define the aperture efficiency η_A to be the ratio of these:

$$\eta_A = \frac{A_e}{A} . \tag{8}$$

According to Ko (1964), for lossless antennas these quantities are related by

$$\frac{\eta_A}{\eta_B} \Omega_M A = \lambda^2 , \tag{9}$$

where λ is the wavelength. By substituting (4) we get

$$b^2 = \frac{\eta_B}{\pi A \eta_A} \lambda^2 . \tag{10}$$

The response of the antenna to the source per unit bandwidth, when it is oriented towards ($x = 0$, $x = x_o$), is given by

$$P(x_o) = \frac{\eta_A A}{2} \int F(x - x_o, y) \quad S(x, y) \, dxdy, \tag{11}$$

which in our case is equal to

$$P(x_o) = \frac{\eta_A A}{2} \frac{S_o}{1 + a^2/b^2} e^{-\frac{x_o^2}{a^2+b^2}} .$$

The factor $\frac{1}{2}$ arises because the feed is polarized and accepts only $\frac{1}{2}$ the incident power. By use of (10) this is

$$P(x_o) = \frac{\eta_A A S_o}{2\left(1 + \frac{\pi A \eta_A a^2}{\eta_B \lambda^2}\right)} e^{-\frac{x_o^2}{a^2 + \frac{\eta_B \lambda^2}{\pi \eta_A A}}} . \tag{12}$$

To find S_o, it may be appropriate to integrate this function and then solve for S_o. We get

$$S_o = \frac{2}{\sqrt{A \eta_A \eta_B}} \frac{\sqrt{\lambda^2 + \frac{\pi A \eta_A a^2}{\eta_B}}}{\lambda^2} \int P(x_o) \, dx_o . \tag{13}$$

In this form (13) shows the wavelength dependence of the relationship between total flux and integrated power response. In terms of the beam radius b which might be empirically determined as a function of wavelength, this can also be written

$$S_o = \frac{2\sqrt{a^2 + b^2}}{\sqrt{\pi}\, A\ \eta_A b^2} \int P(x_o)\, dx_o. \tag{14}$$

For very short wavelengths (narrow beamwidth), (13) becomes approximately

$$S_o \cong \frac{2\sqrt{\pi}\, a}{\eta_B\, \lambda^2} \int P(x_o)\, d\, x_o, \tag{15}$$

while for long wavelengths (wide beamwidth) it becomes

$$S_o \cong \frac{2}{\sqrt{A\ \eta_A \eta_B}\,\lambda} \int P(x_o)\, d\, x_o. \tag{16}$$

The long-wavelength limit corresponds to a common case when the source may be considered a point source. We see that when this is true the dimensions of the source are not required in order to determine the total flux density.

We have not said anything about the wavelength dependence of η_A and η_B. If there is such a dependence over the wavelength range of interest, then of course this must be taken into account.

An alternative method for estimating the total flux is in terms of the peak value. From (12) we get

$$S_o = \frac{2\left(1 + \frac{\pi A\ \eta_A a^2}{\eta_B\, \lambda^2}\right)}{\eta_A A}\ P(o). \tag{17}$$

The short wavelength limit is

$$S_o \sim \frac{2\pi a^2}{\eta_B\, \lambda^2} P(o), \tag{18}$$

while the long wavelength limit is

$$S_o \sim \frac{2}{\eta_A A} P(o). \tag{19}$$

Both (15) and (18) require an estimate of the source size a in order to determine the total flux density. However, for small errors, the fractional error in S is equal to the fractional error in a in (15), while for (18) it is twice as large. For this situation the integral method may be preferable to minimize the error due to unknown source size. However, suppose it is brightness temperature rather than flux density which is desired. We write

$$T(x, y) = T(o, o)\, e^{-\frac{x^2+y^2}{a^2}} \tag{20}$$

and, from the Rayleigh-Jeans law,

$$S(x, y) = \frac{2k}{\lambda^2} T(x, y), \tag{21}$$

where k is Boltzmann's constant. We get, in place of (18) for short wavelengths,

$$T(o, o) \sim \frac{P(o)}{\eta_B kB} . \tag{22}$$

If we write $P(x_o) = kT_A(x_o)$ where $T_A(x_o)$ is the antenna temperature, this becomes

$$T(o, o) \sim \frac{T_A(o)}{\eta_B} .$$

Thus for short wavelengths we can obtain temperature and for long wavelengths we can obtain total flux density without knowing the source size.

THE UNIFORM DISC

Let us consider the case where

$$S(x, y) = \frac{S_o}{\pi a^2}, \quad x^2 + y^2 \leqslant a^2$$

$$= o, \text{ otherwise.} \tag{23}$$

As in the previous case, the integral of the observed response is given by

$$\int P(x_o)\, d\, x_o = \frac{\eta_A A}{2} \int d\, x_o \iint F(x-x_o, y)\, S(x, y)\, dxdy.$$

Therefore,

$$S_o = \frac{2\pi a^2 \int P(x_o)\, d\, x_o}{\eta_A\, A\, K}, \tag{24}$$

where

$$K = \int_{-\infty}^{\infty} d\, x_o \iint_{\substack{\text{source}\\ \text{region}}} F(x-x_o, y)\, dxdy.$$

We substitute (3) for F(x, y), convert to polar coordinates $x = r \cos \varphi$ and $y = r \sin \varphi$, and integrate with respect to φ to get

$$K = 2\pi \int_{-\infty}^{\infty} e^{-\frac{x_o^2}{b^2}}\, dx_o \int_0^a r\, e^{-\frac{r^2}{b^2}}\, J_o\left(-\frac{2ri}{b^2}\, x_o\right) dr,$$

where i is the imaginary unit $\sqrt{-1}$.

We first evaluate the integral with respect to x_o by the relationship (Erdelyi et al., 1953)

$$\int_0^{\infty} J_\mu\, (at)\, e^{-\gamma^2 t^2}\, dt = \frac{1}{2}\, \pi^{1/2}\, \gamma^{-1}\, \exp(-2^{-3}\, a^2\, \gamma^{-2})\, I_{\frac{1}{2}\mu}\, (2^{-3}\, a^2\, \gamma^{-2}) \quad ,$$

from which we find that

$$K = 2\pi^{3/2}\, b \int_{0}^{a} e^{-r^2/2b^2}\, I_o\left(\frac{-r^2}{2b^2}\right) r\, dr.$$

With the change of variable $t = \dfrac{-r^2}{2b^2}$ this becomes

$$K = -2\pi^{3/2}\, b^3 \int_{0}^{-\frac{a^2}{2b^2}} e^t\, I_o(t)\, dt.$$

This definite integral is evaluated by Luke (1962):

$$\int_{0}^{Z} e^t\, I_o(t)\, dt = Z\, e^Z\, (I_o(Z) - I_{-1}(Z)) \quad ,$$

from which we conclude

$$K = \pi^{3/2}\, a^2\, b\, e^{-\frac{a^2}{2b^2}} \left[I_o\left(\frac{a^2}{2b^2}\right) + I_1\left(\frac{a^2}{2b^2}\right) \right] .$$

We substitute this result back into (24) to get

$$S_o = \frac{2e^{\frac{a^2}{2b^2}} \int P(x_o)\, d\, x_o}{\eta_A\, A\, \sqrt{\pi}\, b \left[I_o\left(\frac{a^2}{2b^2}\right) + I_1\left(\frac{a^2}{2b^2}\right) \right]} \quad . \tag{25}$$

To find the short wavelength (narrow beamwidth) limit, we use the asymptotic form

$$I_o(Z) + I_1(Z) \rightarrow \frac{2e^Z}{\sqrt{2\pi Z}} \quad .$$

When we substitute this into (25) we get for the short-wavelength limit

$$S_o \sim \frac{a \int P(x_o)\, d\, x_o}{\eta_A\, A\, b^2}$$

or

$$S_o \sim \frac{\pi a \int P(x_o)\, dx_o}{\eta_B \lambda^2} , \qquad (26)$$

which is approximately, but not identically, equal to the Gaussian result (15).

For the long wavelength (wide beamwidth) limit we use

$$I_o(o) + I_1(o) = 1$$

to get

$$S_o \sim \frac{2 \int P(x_o)\, dx_o}{\eta_A A \sqrt{\pi}\, b} , \qquad (27)$$

which is, of course, identical with (16).

To find the flux density in terms of the peak values we evaluate

$$P(o) = \frac{\eta_A A}{2} \int F(x, y)\, S(x, y)\, dxdy$$

$$= \frac{\eta_A A S_o b^2}{2a^2} \left[1 - e^{-\frac{a^2}{b^2}} \right] .$$

Therefore

$$S_o = \frac{2a^2}{\eta_A A b^2 \left(1 - e^{-a^2/b^2} \right)} P(o). \qquad (28)$$

The short wavelength limit is

$$S_o \sim \frac{2a^2}{\eta_A A b^2} P(o)$$

or

$$S_o \sim \frac{2\pi a^2}{\eta_B \lambda^2} P(o), \qquad (29)$$

while the long wavelength limit is

$$S_o \sim \frac{2}{\eta_A A} P(o). \tag{30}$$

These are the same results as (18) and (19) for the Gaussian case, although the equations do not coincide for intermediate wavelengths.

One important point to note is that we have been implicitly assuming that very little of the source contribution to the received power is by way of the sidelobes. If this condition is violated, then there will be some error in the results we have derived.

ESTIMATING THE TOTAL FLUX IN THE PRESENCE OF NOISE

We shall confine ourselves to a discussion of point sources, so that the general form of the received distribution is known, i.e., it is the antenna pattern $e^{-x_o^2/b^2}$ multiplied by an amplitude factor which we shall call c, i.e., $P(x_o) = c\, e^{-x_o^2/b^2}$. We shall assume the location of the source is known, but not the amplitude. We can formulate the problem considering the position as an unknown also, but the equations then are more difficult to work with. We shall assume that the amplitude of the source is a Gaussian random variable with zero mean, a fiction designed for mathematical convenience.

Since the observed power response is contaminated with noise, we may write

$$Q(x_o) = P(x_o) + N(x_o),$$

where

$Q(x_o)$ = observed response

$P(x_o)$ = noise-free response

$N(x_o)$ = noise component.

We assume the noise is a Gaussian variable with a correlation function $R(x_o, y_o)$ defined by

$$R(x_o, y_o) = E\{N(x_o)\, N(y_o)\}, \qquad (31)$$

where E stands for expected value. For example, white noise passed through an RC filter with time constant t has the correlation function

$$\sigma_n^2 \, e^{-\frac{|x_o - y_o|}{t}}$$

where σ_n is the rms value of the output noise. If this noise is transferred to a written record, then t should be replaced by ℓ, where ℓ is the length of record transversed in time t.

If we use sample values of the output function, then it is appropriate to express the set of sample values in matrix form. Let $(x_1, x_2, \ldots x_n)$ be the n values of x_o for which samples are taken. Then we define a 1 by n observation matrix $\underline{Q}$,

$$\underline{Q} = [Q(x_1)\; Q(x_2)\; \ldots\; Q(x_n)], \qquad (32)$$

a noise-free matrix $\underline{P}$,

$$\underline{P} = [P(x_1)\; P(x_2)\; \ldots\; P(x_n)], \qquad (33)$$

and a noise matrix $\underline{N}$,

$$\underline{N} = [N(x_1)\; N(x_2)\; \ldots\; N(x_n)]. \qquad (34)$$

We also define an n by n correlation matrix $\underline{R}$ whose ij^{th} element is given by

$$\underline{R}_{ij} = R(x_i, x_j). \qquad (35)$$

The joint probability density function for the noise is given by

$$p(\underline{N}) = \frac{1}{\sqrt{(2\pi)^n \; |R|}} \; e^{-\frac{1}{2} \underline{N} \; \underline{R}^{-1} \; \underline{N}^t} \quad , \tag{36}$$

where $\underline{N}^t$ is the transpose of $\underline{N}$.

The true distribution is given by

$$P(x_o) = c \; e^{-\frac{x_o^2}{b^2}} \quad , \tag{37}$$

where c is unknown. We define the matrix $\underline{F}$ such that

$$\underline{F} = \left[e^{-\frac{x_1^2}{b^2}} \; e^{-\frac{x_2^2}{b^2}} \; \ldots \; e^{-\frac{x_n^2}{b^2}} \right] . \tag{38}$$

Then $\underline{F}$ and $\underline{P}$ are related by

$$\underline{P} = c \; \underline{F} \; . \tag{39}$$

We assume c to be Gaussian distributed with standard deviation σ_c, i.e., with density function p(c) given by

$$p(c) = \frac{1}{\sqrt{2\pi} \; \sigma_c} \; e^{-\frac{1}{2} \; \frac{c^2}{\sigma_c^2}} \; . \tag{40}$$

The deviation σ_c must be chosen at least as big as the largest anticipated value of c. Physically, large values of σ_c imply very little a priori knowledge of the magnitude of c.

The joint probability density function for noise-free response and noise then is

$$p(c, \underline{N}) = \frac{1}{\sqrt{(2\pi)^{n+1} \; \sigma_c^2 \; |\underline{R}|}} \; e^{-\frac{c^2}{2\sigma_c^2} \; - \frac{1}{2} \underline{N} \; \underline{R}^{-1} \; \underline{N}^t} \quad . \tag{41}$$

This is the function which we can use to determine the best mean-square estimate to our integral and the error which results from our choice.

Let $\hat{q}$ be an estimate of the integral $\int P(x_o)\, dx_o$. We wish to choose $\hat{q}$ such that

$$\text{MSE} = E[\hat{q} - \int P(x_o)\, dx_o]^2$$

is minimum, where MSE stands for mean square error. This can be expanded to

$$\text{MSE} = E\{\hat{q}^2 - 2\hat{q}^2 \int P(x_o)\, dx_o + [\int P(x_o)\, dx_o]^2\} .$$

The middle term contains

$$\int P(x_o)\, dx_o = \sqrt{\pi}\, b\, c,$$

while the third term is

$$E[\int P(x_o)\, dx_o]^2 = \pi\, b^2\, \sigma_c^2 .$$

Thus

$$\text{MSE} = \pi\, b^2\, \sigma_c^2 + E\{\hat{q}^2 - 2\sqrt{\pi}\, b\, c\, \hat{q}\}. \tag{42}$$

To minimize MSE we must choose $\hat{q}$ such that

$$E\{\hat{q}^2 - 2\sqrt{\pi}\, b\, c\, \hat{q}\}$$

is minimum. The expected value is given by

$$E\{\hat{q}^2 - 2\sqrt{\pi}\, b\, c\, \hat{q}\} = \iint [\hat{q}^2 - 2\sqrt{\pi}\, b\, c\, \hat{q}]\; p(c, \underline{N})\, dc\, d\underline{N} .$$

We make the change of variables

$$\underline{N} = \underline{Q} - c\,\underline{F}$$

$$d\underline{N} = d\,\underline{Q}$$

to get

$$E\{\hat{q}^2 - 2\sqrt{\pi}\, b\, c\, \hat{q}\} = \iint [\hat{q}^2 - 2\sqrt{\pi}\, b\, c\, \hat{q}]\; p(c, \underline{Q} - c\underline{F})\, dc\, d\underline{Q}.$$

Now we integrate with respect to c to get (note that $\hat{q}$ is a function of the observed response $\underline{Q}$)

$$E\{\hat{q} - 2\sqrt{\pi}\, b\, c\, \hat{q}\} = \frac{1}{\sqrt{(2\pi)^n\, |\underline{R}|\, (1 + \sigma_c^2\, \underline{F}\, \underline{R}^{-1}\, \underline{F}^t)}}$$

$$\int \left[\hat{q}^2 - 2\sqrt{\pi}\, b \frac{\underline{F}\, \underline{R}^{-1}\, \underline{Q}^t}{\frac{1}{\sigma_c^2} + \underline{F}\, \underline{R}^{-1}\, \underline{F}^t} \hat{q} \right] e^{-\frac{1}{2} \underline{Q}\, \underline{R}^{-1}\, \underline{Q}^t + \frac{1}{2} \frac{(\underline{F}\, \underline{R}^{-1}\, \underline{Q}^t)^2}{\frac{1}{\sigma_c^2} + \underline{F}\, \underline{R}^{-1} \underline{F}^t}} d\underline{Q}.$$

Now the integrand is a function of $\underline{Q}$ only, so that we may differentiate with respect to $\hat{q}$ to minimize the expression. We get

$$\boxed{\hat{q} = \frac{\sqrt{\pi}\, b\, \underline{F}\, \underline{R}^{-1}\, \underline{Q}^t}{\frac{1}{\sigma_c^2} + \underline{F}\, \underline{R}^{-1}\, \underline{F}^t}\,.} \tag{43}$$

SPECIAL CASE

Suppose that the noise samples are independent and identically distributed, so that $\underline{R}^{-1} = \frac{1}{\sigma_n^2} \underline{I}$, where $\underline{I}$ is the identity matrix. Our estimate becomes

$$\hat{q} = \frac{\sqrt{\pi}\, b\, \underline{F}\, \underline{Q}^t}{\frac{\sigma_n^2}{\sigma_c^2} + \underline{F}\, \underline{F}^t}\,.$$

If, in addition, σ_c is large, we have

$$\hat{q} = \sqrt{\pi}\, b \frac{\underline{F}\, \underline{Q}^t}{\underline{F}\, \underline{F}^t}\,,$$

which can be written alternatively as

$$\hat{q} = \sqrt{\pi}\, b \frac{\sum e^{-\frac{x_i^2}{b^2}} Q(x_i)}{\sum e^{-\frac{2x_i^2}{b^2}}} . \tag{44}$$

The limiting form for this as an infinite number of samples are taken is

$$\hat{q} = \sqrt{2} \int e^{-\frac{x_o^2}{b^2}} Q(x_o)\, dx_o.$$

This is an integral form which is somewhat to be expected since we originally set out to estimate an integral. However, we originally set out to estimate an integral. However, we would have obtained this form even if we attempted to estimate the peak value. Thus we have another reason for preferring an estimate based on the integral rather than the peak value. Of course, the estimate based on this integral is only approximately best in the least-square sense, because the assumption that each noise sample is uncorrelated with every other noise sample is not true for samples more closely spaced than the time constant of the receiver.

Note that (43) is a biased estimate, while (44) is unbiased. It is up to the observer to decide whether to remove the bias from (43) by eliminating the term $\frac{1}{\sigma_c^2}$, which increases the r.m.s. error somewhat. Since σ_c is a quantity which can only be guessed at anyhow, it is probably well to remove it. Also note that the assumption σ_c is large means that

$$\sigma_c^2 \gg \frac{\sigma_n^2}{\sum e^{-\frac{2x_i^2}{b^2}}}$$

in the case that $\underline{R} = \sigma_n^2\, \underline{I}$.

THE EXPECTED MEAN-SQUARE ERROR

Once we have the estimate $\hat{q}$ given by (43) we can continue to integrate and evaluate (42) to find the expected mean-square error. We get

$$\mathrm{MSE} = \pi b^2 \sigma_c^2 - \frac{\pi b^2 \, \underline{F} \left[\underline{R} - \dfrac{\underline{F}^t \, \underline{F}}{\dfrac{1}{\sigma_c^2} + \underline{F} \, \underline{R}^{-1} \, \underline{F}^t} \right]^{-1} \underline{F}^t}{\sigma_c \sqrt{\left| \underline{I} - \dfrac{\underline{R}^{-1} \, \underline{F}^t \, \underline{F}}{\dfrac{1}{\sigma_c^2} + \underline{F} \, \underline{R}^{-1} \, \underline{F}^t} \right| \left(\dfrac{1}{\sigma_c^2} + \underline{F} \, \underline{R}^{-1} \, \underline{F}^t \right)^{5/2}}} . \qquad (45)$$

Rather than go through the tedious steps of deriving this expression, we shall show that the results are reasonable in the special case when $\underline{R} = \sigma_n^2 \, \underline{I}$. Then the expression becomes

$$\mathrm{MSE} = \pi b^2 \sigma_c^2 - \frac{\dfrac{\pi b^2}{\sigma_n^2} \, \underline{F} \left(\underline{I} - \dfrac{\underline{F}^t \, \underline{F}}{\dfrac{\sigma_n^2}{\sigma_c^2} + \underline{F} \, \underline{F}^t} \right)^{-1} \underline{F}^t}{\sigma_c \sqrt{\left| \underline{I} - \dfrac{\underline{F}^t \, \underline{F}}{\dfrac{\sigma_n^2}{\sigma_c^2} + \underline{F} \, \underline{F}^t} \right| \left(\dfrac{1}{\sigma_c^2} + \dfrac{1}{\sigma_n^2} \underline{F} \, \underline{F}^t \right)^{5/2}}} . \qquad (46)$$

By direct matrix multiplication we can show that

$$\left(\underline{I} - \frac{\underline{F}^t \, \underline{F}}{\dfrac{\sigma_n^2}{\sigma_c^2} + \underline{F} \, \underline{F}^t} \right)^{-1} = \underline{I} + \frac{\sigma_c^2}{\sigma_n^2} \underline{F}^t \, \underline{F} . \qquad (47)$$

By expanding the determinant according to its definition we can show that

$$\left| \underline{I} - \frac{\underline{F}^t \, \underline{F}}{\dfrac{\sigma_n^2}{\sigma_c^2} + \underline{F} \, \underline{F}^t} \right| = \frac{\dfrac{\sigma_n^2}{\sigma_c^2}}{\dfrac{\sigma_n^2}{\sigma_c^2} + \underline{F} \, \underline{F}^t} . \qquad (48)$$

We substitute (47) and (48) into (46) to get

$$\text{MSE} = \pi\, b^2\, \sigma_c^2 \left(1 - \frac{\underline{F}\ \underline{F}^t}{\underline{F}\ \underline{F}^t + \frac{\sigma_n^2}{\sigma_c^2}} \right) . \tag{49}$$

For very large values of σ_c we get

$$\underset{\sigma_c \to \infty}{\text{MSE}} = \frac{\pi\, b^2\, \sigma_n^2}{F\ F^t} . \tag{50}$$

This is exactly the result which one would obtain using simplified arguments from the start.

For small expected signal-to-noise ratios we get

$$\underset{\sigma_n^2 \to \infty}{\text{MSE}} = \pi\, b^2\, \sigma_c^2 , \tag{51}$$

which is the variance due to the flux density itself. This is also to be expected. As the noise component gets hopelessly large, the best estimate for the flux density is its mean value. The expected r.m.s. error then is simply the variance.

As a computational procedure, we might recommend sampling the received curve at intervals spaced as least as close as the time constant of the receiver. Then compute the estimate

$$\hat{q} = \text{estimate of} \int P(x_o)\, d\, x_o$$

$$= \sqrt{\pi}\, b \quad \frac{\sum e^{-\frac{x_i^2}{b^2}}\, Q(x_i)}{\sum e^{-\frac{2x_i^2}{b^2}}} . \tag{52}$$

Next compute an estimate of the mean-square error by

$$\text{MSE} \sim \frac{\pi\, b^2\, \sigma_n^2}{e^{-\frac{2x_i^2}{b^2}}}, \qquad (53)$$

where in this equation the sample values are spaced by an interval equal to the time constant of the system.

Usually, an estimate is given as the estimate plus or minus so many standard deviations. Equation (52) gives the estimate, while the square root of (53) gives the standard deviation. If the standard deviation is too large, then either the results may be discarded or more samples may be taken and the more exact equations (43) and (45) used to get the maximum significance from the data.

ESTIMATING THE FLUX WHEN THE SOURCE POSITION IS UNKNOWN

It is rather difficult to get an exact mean-square estimate for the integral $\int P(x_o)\, d\, x_o$ of the type (43) when the source position is also unknown. Nevertheless, some useful approximate expressions can be derived.

First, let us point out that (43) is a stationary expression for $\int P(x_o)\, d\, x_o$ with respect to position, provided sample points are selected equally about the origin $x_o = 0$. For example, suppose that $P(x_o) = c\, e^{-\frac{(x_o-\delta)^2}{b^2}}$, where δ is the true position of the center of the source. Then the estimate (43), in which it is assumed that the source is located at the origin, is given by

$$\hat{q} = \frac{\sqrt{\pi}\, b\, c \sum_{ij} e^{-\frac{x_i^2}{b^2}} [\underline{R}^{-1}]_{ij}\, e^{-\frac{(x_j-\delta)^2}{b^2}}}{\frac{1}{\sigma_c^2} + \underline{F}\, \underline{R}^{-1}\, \underline{F}^t} + \text{noise contribution.}$$

The derivative with respect to δ, evaluated at $\delta = 0$, is given by

$$\left.\frac{\partial \hat{q}}{\partial \delta}\right|_{\delta=0} = \frac{\sqrt{\pi}\, b\, c \sum_{ij} e^{-\frac{x_i^2}{b^2}} [\underline{R}^{-1}]_{ij} \frac{2x_j}{b^2} e^{-\frac{x_j^2}{b^2}}}{\frac{1}{\sigma_c^2} + \underline{F}\, \underline{R}^{-1}\, \underline{F}^t} .$$

If the sample values are distributed equally on both sides of the origin, then

$$\left.\frac{\partial \hat{q}}{\partial \delta}\right|_{\delta=0} = 0, \qquad (54)$$

which is the result we are looking for. Thus one need know the source position only moderately well in order to get a good estimate of the flux density.

To get an idea of how good our knowledge of the source position must be, let us use the approximate integral form

$$\hat{q} = \sqrt{2} \int e^{-\frac{x_o^2}{b^2}} Q(x_o)\, d\, x_o$$

and substitute $P(x_o) = c\, e^{-\frac{(x_o-\delta)^2}{b^2}}$ to get

$$\hat{q} = \sqrt{\pi}\, b\, c\, e^{-\frac{\delta^2}{2b^2}} + \text{noise}.$$

This expression can be used to determine the allowed limits on the ratio $(\frac{\delta}{b})$. Suppose, for example, that the flux density must be known to an accuracy of 1/2%. Then δ need only be known to lie within the limits of $\pm$ 10% of b.

Let us proceed to a more exact estimation of the flux density when the assumption that the source is at the origin is no longer satisfactory. We define the matrix $\underline{F}$ (δ) now as

$$\underline{F}(\delta) = \left[e^{\frac{(x_1-\delta)^2}{b^2}} \quad e^{\frac{-(x_2-\delta)^2}{b^2}} \quad \ldots \quad e^{\frac{-(x_n-\delta)^2}{b^2}} \right] . \tag{55}$$

We assume that δ is distributed according to a distribution function $p(\delta)$ such that

$$\int p(\delta)\, d\delta = 1.$$

We define the following estimate $\hat{q}(\delta)$:

$$\hat{q}(\delta) = \frac{\sqrt{\pi}\, b\, \underline{F}(\delta)\, \underline{R}^{-1}\, \underline{Q}^t}{\frac{1}{\sigma_c^2} + F(\delta)\, \underline{R}^{-1}\, \underline{F}^t(\delta)} . \tag{56}$$

This is the same as (43), except that $\underline{F}$ is replaced by $\underline{F}(\delta)$. It is thus the estimate which would be used if it were known that the source was centered at $x_o = \delta$ instead of at the origin.

To find the best mean-square estimate we now have to minimize the expression

$$E\{\hat{q}^2 - 2\sqrt{\pi}\, b\, c\, \hat{q}\} = \int (\hat{q}^2 - 2\sqrt{\pi}\, b\, c\, \hat{q})\, p(\underline{N}, c, \delta)\, d\underline{N}\, dc d\delta .$$

We again make the variable change $\underline{N} = \underline{Q} - c\, \underline{F}(\delta)$, differentiate with respect to $\hat{q}$, and set equal to zero to get

$$\hat{q} = \frac{\sqrt{\pi}\, b \int c\, p(\underline{Q} - c\, \underline{F}(\delta), c, \delta)\, dc\, d\delta}{\int p(\underline{Q} - c\, \underline{F}(\delta), c, \delta)\, dc\, d\delta} .$$

We note that

$$\frac{\sqrt{\pi}\, b \int c\, p(\underline{Q} - c\,\underline{F}(\delta),\, c)\, dc}{\int p(\underline{Q} - c\,\underline{F}(\delta))\, dc} = \hat{q}(\delta)$$

so that

$$\hat{q} = \frac{\int \hat{q}(\delta)\, p(\delta)\, d\delta \int p(\underline{Q} - c\,\underline{F}(\delta),\, c)\, dc}{\int p(\delta)\, d\delta \int p(\underline{Q} - c\,\underline{F}(\delta),\, c)\, dc} .$$

But $\int p(\underline{Q} - c\,\underline{F}(\delta),\, c)\, dc$ is proportional to

$$e^{\frac{\hat{q}^2(\delta)\left(\frac{1}{\sigma_c^2} + \underline{F}(\delta)\,\underline{R}^{-1}\,\underline{F}^t(\delta)\right)}{2\pi b^2}} \left(\frac{1}{\sigma_c^2} + \underline{F}(\delta)\,\underline{R}_{11}^{-1}\,\underline{F}^t\right)^{-1} .$$

If we assume $\underline{F}(\delta)\, R_{11}^{-1}\, \underline{F}^t(\delta)$ is more or less constant, which we must justify, our estimate becomes

$$\hat{q} = \frac{\int \hat{q}(\delta)\, e^{\frac{\hat{q}^2(\delta)\left[\frac{1}{\sigma_c^2} + \underline{F}\,\underline{R}^{-1}\,\underline{F}^t\right]}{2\pi b^2}}\, p(\delta)\, d\delta}{\int e^{\frac{\hat{q}^2(\delta)\left[\frac{1}{\sigma_c^2} + \underline{F}\,\underline{R}^{-1}\,\underline{F}^t\right]}{2\pi b^2}}\, p(\delta)\, d\delta} . \tag{57}$$

Let us suppose that δ is uniformly distributed over a finite interval. Let us approximate the integrals by sums, so that

$$\hat{q} = \frac{\sum_{i=1}^{m} \hat{q}(\delta_i)\, e^{\frac{\hat{q}^2(\delta_i)\left[\frac{1}{\sigma_c^2} + \underline{F}\,\underline{R}^{-1}\,\underline{F}^t\right]}{2\pi b^2}}}{\sum_{i=1}^{m} e^{\frac{\hat{q}^2(\delta_i)\left[\frac{1}{\sigma_c^2} + \underline{F}\,\underline{R}^{-1}\,\underline{F}^t\right]}{2\pi b^2}}} , \tag{58}$$

where we take the δ_i's at intervals equal to the sampling interval of the x_i's. In this case we see that $\underline{F}(\delta)\ \underline{R}^{-1}\ \underline{F}^t(\delta)$ represents a statistical operation between the noise and a displaced Gaussian curve. If the matrix dimensions are infinite, we might just as well consider the Gaussian curve to be at the origin and the noise displaced by δ. If the noise statistics do not depend on position, which we shall assume, then $\underline{F}(\delta_i)\ \underline{R}^{-1}\ \underline{F}^t(\delta_i) = \underline{F}(o)\ \underline{R}^{-1}\ \underline{F}^t(o)$.

SPECIAL CASE

We again assume that $\sigma_c^2 \to \infty$ and $\underline{R}^{-1} = \frac{1}{\sigma_n^2}\ \underline{I}$. Then we get, for (58)

$$\hat{q} = \frac{\sum_{i=1}^{m} q(\delta_i)\ e^{\frac{q^2(\delta_i)}{2\pi b^2\ \sigma_n^2} \sum_{j=1}^{n} e^{-\frac{2x_j^2}{b^2}}}}{\sum_{i=1}^{m} e^{\frac{q^2(\delta_i)}{2\pi b^2\ \sigma_n^2} \sum_{j=1}^{n} e^{-\frac{2x_j^2}{b^2}}}}\ , \tag{59}$$

where n is proportional to the range for which data samples are taken, and m is proportional to the range for which a source position is suspected. We see that (59) is a weighted average over all computed estimates $\hat{q}(\delta_i)$, in which large estimates are emphasized and small estimates de-emphasized. In its crudest form, (59) suggests computing $\hat{q}(\delta)$ for all possible δ, and then selecting the maximum value for the estimate of $\int P(x_o)\ d\ x_o$.

REFERENCES

Erdelyi, A., Magnus, W. Oberhettinger, F. and Tricomi, F. G. 1953. Higher Transcendental Functions, Vol. II (McGraw Hill Book Co., New York), p. 50.

Ko, H. C. 1964. IEEE Trans. Antennas and Propagation AP-12, 891-898 (December, 1964).

Luke, Y. L. 1962. Integrals of Bessel Functions. (McGraw-Hill Book Co., New York), p. 122.