Estimation of differential Doppler shifts

Peter M. Schultheiss

Yale University, New Haven, Connecticut 06520

Ehud Weinstein

Woods Hole Oceanographic Institute, Woods Hole, Massachusetts 02556
(Received 18 May 1979; accepted for publication 2 August 1979)

If the acoustic signal radiated by a moving source is observed at two or more locations, the received signals exhibit differential Doppler shifts which provide information about source motion. This paper calculates minimum mean-square errors for the estimates of differential Doppler shifts, which can be obtained in a given observation interval. Both Gaussian and sinusoidal signals are considered. The noise is assumed to be Gaussian and independent from sensor to sensor. Dependence of the estimation errors on observation time, signal-to-noise ratio, and size of the receiving array are studied. The estimation of differential Dopplers is found to be uncoupled from the estimation of differential delays and from the estimation of signal parameters, such as center frequency and bandwidth. A comparison is made between two possible procedures of differential Doppler estimation: coherent processing of the signals received at two sensors and subtraction of separate frequency estimates obtained at each sensor. The two are equivalent for sinusoidal signals, but for large *TW* Gaussian signals the coherent procedure yields substantially smaller errors.

PACS numbers: 43.60.Gk, 43.30.Cq, 43.28.Bj, 43.60.Cg

INTRODUCTION

The location of a radiating source can be determined by observing its signal at an array of sensors. In the absence of detailed knowledge concerning the signal waveshape, all information about source location is contained in the differential delays between the signal components received by various sensor pairs. Differential delay estimation has therefore attracted a great deal of interest in recent literature.[1-3]

When the signal source is moving relative to the array, the observed differential delays are functions of time. Suppose the source is moving at constant velocity and is sufficiently remote from the receiving array so that differential delays vary essentially linearly over the observation interval. In that case, the differential delay $\Delta\tau_i(t)$ for the ith sensor pair is given by the relation

$$\Delta\tau_i(t)=\Delta\tau_i(0)-(\Delta v_i/c)t\,. \tag{1}$$

c is the velocity of signal propagation and Δv_i is the difference between the radial components of source velocity towards the two members of the pair. In this situation the sensor outputs contain information not only about the source location at $t=0$ [specified by the $\Delta\tau_i(0)$], but also about source track. The measurement of the Δv_i in addition to the $\Delta\tau_i(0)$, is therefore of considerable practical interest. If an adequate number of sensors (at least five) is available, the Δv_i alone are sufficient to determine source location in the plane and can therefore be used for localization when the differential delay measurements are unreliable (perhaps because the narrow bandwidth of the signal makes the differential delay observations highly ambiguous).

This paper is concerned with the estimation of the Δv_i. When the radiated signal is a sinusoidal, the linearly time varying delay of Eq. (1) leads to a differential Doppler shift between the signals received by two sensors. Estimation of Δv_i is equivalent to the estimation of this differential Doppler shift. The problem under consideration is therefore often described loosely as "differential Doppler estimation" and we shall adhere to this convention even though our results are not confined to sinusoidal, or even very narrowband, signals. Our objective will be to evaluate the performance of the best possible estimator for the Δv_i.

If one defines the "best possible" estimator as one which achieves zero bias and minimizes mean-square error, a framework for the necessary computations is provided immediately by the following well known results of statistical theory:

(1) Let $\mathbf{r}$ be the available data vector and θ the vector of unknown parameters. Define J as the matrix (Fisher information) of elements J_{ij} given by

$$J_{ij}=-E\left\{\frac{\partial^2 \log p(\mathbf{r}/\theta)}{\partial\theta_i\partial\theta_j}\right\}. \tag{2}$$

$p(\mathbf{r}/\theta)$ is the conditional probability of $\mathbf{r}$ given θ and $E\{\ \}$ stands for the expectation of the bracketed quantity. If $\hat{\theta}_i$ is any unbiased estimate of θ_i, then its variance $D^2(\hat{\theta}_i)$ satisfies the Cramer–Rao inequality

$$D^2(\hat{\theta}_i)\geqslant(J^{-1})_{ii}\,. \tag{3}$$

(2) When the observation time T increases without bound, there exists a physically realizable estimator (the maximum likelihood estimator) whose performance approaches the lower bound specified by Eq. (3). The practical meaning of the "long observation time" condition is that T must be large compared with the correlation time of the noise (and with that of the signal, if the latter is also a random process). These conditions will be satisfied in the problems of interest here. One can therefore use the right side of Eq. (3) to specify the mean-square error of the best estimator.

I. BASIC THEORY

The relevant geometry is shown in Fig. 1. For the greater part of this paper we shall assume that the source radiates a stationary Gaussian random process $s(t)$ with zero mean, autocorrelation function $R_s(\tau)$, and

 0001-4966/79/111412-08$00.80 © 1979 Acoustical Society of America

power spectrum $S(\omega)$, related to $R_s(\tau)$ through the equation

$$S(\omega)=\int_{-\infty}^{\infty} R_s(\tau)e^{-j\omega\tau}d\tau . \tag{4}$$

We shall further assume that the observation time T is large compared with the correlation time (or the inverse bandwidth) of the signal. In passive sonar applications these assumptions are quite likely to be satisfied with two exceptions: The signal may be a pure sinusoid, or it may be a random process of a bandwidth so narrow that its correlation time exceeds T. In the latter case the typical sample function looks like a sinusoid of random amplitude and phase. By treating the problem of sinusoidal signals in a separate section at the end we are, therefore, accomodating most signals likely to be encountered in practice.

If the source signal is Gaussian and transmission through the medium is a linear process, the signal components received at the M sensors are jointly Gaussian random processes with zero means. We shall assume that the noise $n_i(t)$ received at the ith sensor is also Gaussian with zero mean and that the noise components received by different sensors are statistically independent. (We are interested in differential Doppler shifts. If these are to be significant, the sensors cannot be spaced very closely. The assumptions of noise incoherent from sensor to sensor is therefore not very restrictive.) We shall generate a data vector $\mathbf{r}$ by Fourier analyzing each sensor output $r_i(t)$ and concatenating the M sets of Fourier coefficients. Since the computation of Fourier coefficients is a linear operation, $\mathbf{r}$ is a (complex) Gaussian random variable and the conditional probability density $p(\mathbf{r}/\theta)$ required in Eq. (2) assumes the form

$$p(\mathbf{r}/\theta)=[\mathrm{Det}(\pi K)]^{-1/2}\exp(-\mathbf{r}^*K^{-1}\mathbf{r}) . \tag{5}$$

$\mathbf{r}^*$ denotes the conjugate-transpose of $\mathbf{r}$ and K is the covariance matrix

$$K=E\{\mathbf{r}\mathbf{r}^*/\theta\} . \tag{6}$$

Substituting Eq. (5) into Eq. (2) one can obtain the elements of the Fisher information matrix by direct computation. The result is available in the literature[4]:

$$J_{ij}=\mathrm{Tr}\left(K^{-1}\frac{\partial K}{\partial\theta_i}K^{-1}\frac{\partial K}{\partial\theta_j}\right). \tag{7}$$

Tr( ) stands for the trace of the bracketed matrix.

The components of our data vector $\mathbf{r}$ are the Fourier coefficients

$$R_i(\omega_n)=\frac{1}{\sqrt{T}}\int_{-T/2}^{T/2} r(t)\exp(-j\omega_n t)\,dt , \tag{8}$$

where $\omega_n=2\pi n/T$.

The elements of K are therefore given by

$$E\{R_i(\omega_n)R_j^*(\omega_k)\}=\frac{1}{T}\int_{-T/2}^{T/2}dt\int_{-T/2}^{T/2}d\sigma\, E\{r_i(t)r_j(\sigma)\} \times\exp[-j(\omega_n t-\omega_k\sigma)] . \tag{9}$$

If the travel time of the signal from the source to the ith sensor is

$$\tau_i(t)=\tau_i(0)-(v_i/c)t , \tag{10}$$

the received waveshape at the ith sensor assumes the form

$$r_i(t)=s[t-\tau_i(0)+(v_i/c)t]+n_i(t) . \tag{11}$$

The correlation functions required in Eq. (9) can now be written down by inspection

$$E\{r_i(t)r_j(\sigma)\}=R_s\left\{\left(1+\frac{v_i}{c}\right)t-\left(1+\frac{v_j}{c}\right)\sigma-\tau_i(0)+\tau_j(0)\right\}+\frac{1}{\beta_i}R_n(t-\sigma)\delta_{ij} . \tag{12}$$

$(1/\beta_i)R_n(\tau)$ is the autocorrelation of the noise received at the ith sensor. The β_i therefore measures the relative noise level at various sensors, with large β_i indicating relatively noisefree sensors. For convenience (and without loss of generality) the β_i will be normalized so that

$$\sum_{i=1}^{M}\frac{1}{\beta_i}=M . \tag{13}$$

In that case $R_n(0)$ is the average noise power for the array. The symbol δ_{ij} denotes the Kroneker delta function

$$\delta_{ij}=\begin{cases}1 & i=j\\ 0 & i\neq j\end{cases} . \tag{14}$$

The most striking feature of Eq. (12) is that it depends explicitly on both t and σ, not only on their difference. The existence of differential Doppler shifts causes the data covariance matrix to become nonstationary, even though the radiated signal, and each sensor output, is stationary.[5] The nonstationarity generates formidable

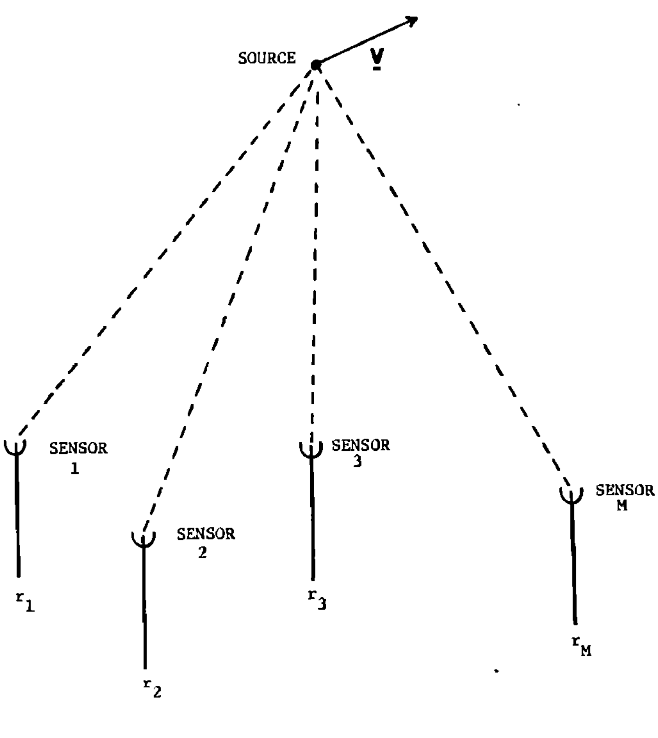

FIG. 1. Array-source geometry.

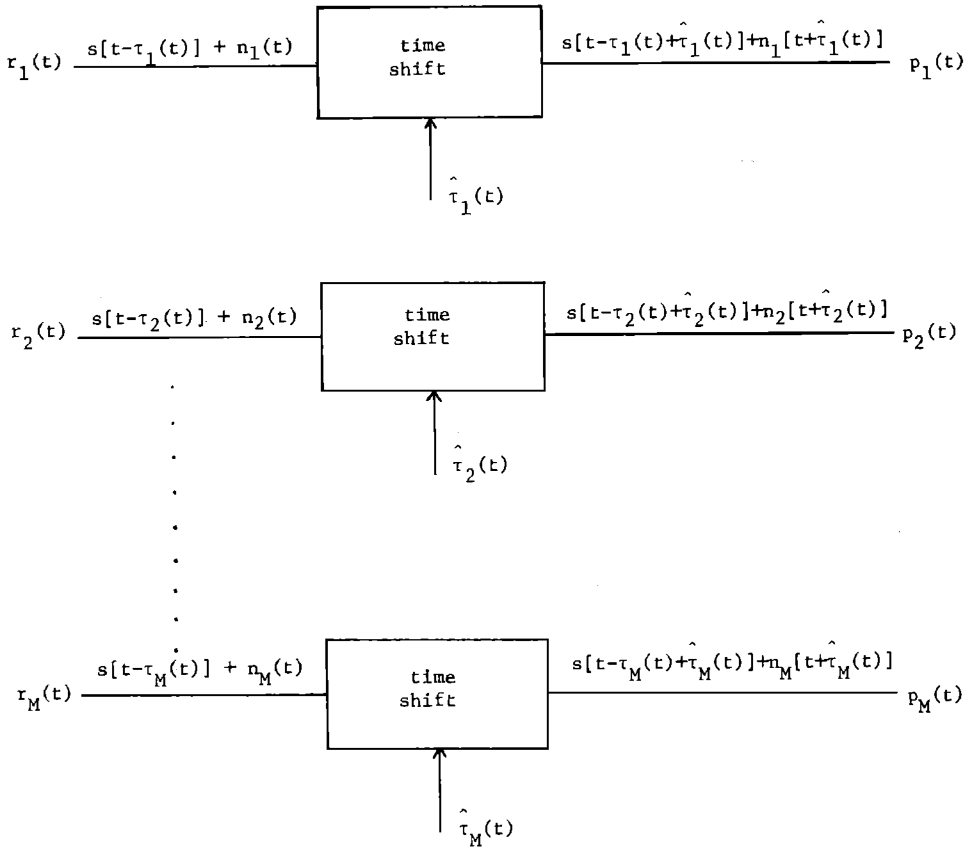

FIG. 2. Effect of an arbitrary set of known delays.

computational problems and makes the brute-force approach of substituting Eq. (9) into Eq. (7) quite intractable. We therefore resort to an alternative procedure.

Consider the arrangement shown in Fig. 2. At the left are the actual sensor outputs $r_i(t)$. The output of the ith sensor is shifted in time by an arbitrary, known amount $\hat{\tau}_i(t)$, generating the time function $p_i(t)$.

We assert that the best estimator of signal parameters (θ) using the $p_i(t)$ as data, performs exactly as well as the best estimator using the $r_i(t)$ as data. This appears intuitively obvious because the $\hat{\tau}_i(t)$ are known, so that the $p_i(t)$ can be constructed from the $r_i(t)$ and vice versa. If one set were preferable to the other the best estimator would simple carry out the necessary conversion. It is then equally obvious that identical Fisher information matrices for θ are generated by the data sets $p_i(t)$ and $r_i(t)$. The Appendix contains a formal proof of that assertion. The line of reasoning can then proceed as follows: Since the $r_i(t)$ are totally independent of the $\hat{\tau}_i(t)$, the Fisher information matrix for θ is independent of the $\hat{\tau}_i(t)$, even if $p_i(t)$ is used as the basic data set. In calculating the Cramer–Rao bound one can therefore use *any* value of $\hat{\tau}_i(t)$ that is computationally convenient without affecting the result. The obvious choice is $\hat{\tau}_i(t)=\tau_i(t)$. The fact that the time delays $\tau_i(t)$ are not known *a priori* is irrelevant: we are simply asserting that all choices of $\hat{\tau}_i(t)$ yield the same θ accuracy. Among these choices is $\hat{\tau}_i(t)=\tau_i(t)$, for which the computation happens to be relatively simple. If our only aim is to evaluate performance, we can therefore use the $p_i(t)$ as data and evaluate all components of Eq. (7) at $\hat{\tau}_i(t)=\tau_i(t)$. (Use of the reversibility property is clearly not confined to the analysis of linearly time varying delays. It remains equally applicable when the source is capable of more general maneuvers.)

The procedure of Fig. 2 creates one difficulty: while making the signal components stationary, it makes the noise components nonstationary. This turns out to be a minor problem under the assumption that the noise is uncorrelated from sensor to sensor. In that case, noise statistics are completely characterized by the autocorrelation and the effect of the $\tau_i(t)$ on that statistic is small as long as the noise bandwidth is broad compared to the maximum Doppler shift.[6] We assume that this condition is satisfied and proceed with the solution of the stationary problem.

The components of our data vector are now

$$p_i(\omega_n)=\frac{1}{\sqrt{T}}\int_{-T/2}^{T/2} p(t)\exp(-j\omega_n t)\,dt\,, \tag{15}$$

and the equivalent of Eq. (12) becomes

$$E\{p_i(t)p_j(\sigma)\}\Big|_{\hat{\tau}_i(t)=\tau_i(t)} =R_s(t-\tau)+(\beta_i)^{-1}R_n(t-\sigma)\delta_{ij}\,. \tag{16}$$

The computation equivalent to Eq. (9) is now straightforward. For T large compared with the correlation time of signal and noise it leads to the well-known result

$$E\{P_i(\omega_n)P_j{}^*(\omega_k)\}\big|_{\hat{\tau}_i(t)=\tau_i(t)} =[S(\omega_n)+(\beta_i)^{-1}N(\omega_n)\delta_{ij}]\delta_{nk}\,. \tag{17}$$

$S(\omega)$ is the signal spectrum defined in accordance with Eq. (4) and $N(\omega)$ is the average noise spectrum for the array, the Fourier transform of $R_n(\tau)$. If one arranges the data vector in the order

$$\mathbf{P}=\{P_1(\omega_1)\cdots P_M(\omega_1), P_1(\omega_2)\cdots P_M(\omega_2), \cdots P_1(\omega_N)\cdots P_M(\omega_N)\}^T , \tag{18}$$

the covariance matrix K is now block diagonal and its analytical inversion becomes feasible.

II. ESTIMATION OF DIFFERENTIAL DOPPLER SHIFTS

In the present section we shall assume that only the Δv_i are unknown. We use sensor M as a reference, so that $\Delta v_i = v_i - v_M$. We must then calculate the various derivatives $\partial K/\partial \Delta v_i$ and evaluate them at $\hat{\tau}_i(t)=\tau_i(t)$. The computation is tedious but straightforward (see Ref. 6 for details). The result is

for $i \neq j$

$$\left(\frac{\partial}{\partial \Delta v_i} E\{P_i(\omega_n)P_j{}^*(\omega_n)\}\right)_{\hat{\tau}_i=\tau_i}$$
$$=\frac{(-1)^{n-k}}{4\pi c(n-k)}[\omega_n S(\omega_n)+\omega_k S(\omega_k)], \quad n\neq k ,$$
$$=-\frac{1}{2cT}[S(\omega_n)+\omega_n S'(\omega_n)], \quad n=k , \tag{19}$$

and for $i=j$

$$\left(\frac{\partial}{\partial \Delta v_i} E\{P_i(\omega_n)P_i{}^*(\omega_k)\}\right)_{\hat{\tau}_i=\tau_i}$$
$$=-\frac{1}{cT}[S(\omega_n)+\omega_n S'(\omega_n)]\delta_{nk} . \tag{20}$$

The various components of Eq. (20) have very different orders of magnitude. In particular

$$\frac{1}{T}[S(\omega_n)+\omega_n S'(\omega_n)]=\left\{\frac{1}{T}\frac{d}{d\omega}[\omega S(\omega)]\right\}_{\omega=\omega_n} . \tag{21}$$

This is the change in $\omega S(\omega)$ over a frequency interval $1/T$ near ω_n. If the signal spectrum varies smoothly and its bandwidth W is large compared with $1/T$ (large TW product) Eq. (20) and the second line of Eq. (19) are therefore small compared with the first line of Eq. (19) and one can write to an excellent approximation

$$\left(\frac{\partial}{\partial \Delta v_i} E\{P_i(\omega_n)P_j{}^*(\omega_k)\}\right)_{\hat{\tau}_i=\tau_i}$$
$$=0 , \quad n=k$$
$$=\frac{(-1)^{n-k}}{4\pi c(n-k)}[\omega_n S(\omega_n)+\omega_k S(\omega_k)](1-\delta_{ij}), \quad n\neq k . \tag{22}$$

The evaluation of Eq. (7) and the inversion of the Fisher information matrix J are now straightforward. Again we refer to Ref. 6 for computational details. The resulting Cramer–Rao bound on the variance of $\widehat{\Delta v}_i$ is

$$D^2(\widehat{\Delta v}_i)=(J^{-1})_{ii}$$
$$=\frac{12\pi c^2(1/\beta_i+1/\beta_M)}{T^3\sum_{i=1}^{M}\beta_i\int_0^\infty \frac{\omega^2 S^2(\omega)/N^2(\omega)}{1+\sum_{i=1}^{M}\beta_i S(\omega)/N(\omega)}\,d\omega} . \tag{23}$$

Equation (23) has several features of practical interest.

(a) The mean-square error in differential velocity varies as T^{-3}. Comparable results for differential delay estimation show a variation with T^{-1}.[7] Thus the accuracy of the differential velocity estimate improves much more rapidly than that of the differential delay estimate. One must be careful not to read into this observation more than it actually asserts: If the source location is known, a pair of differential Doppler measurements provides an estimate of source velocity whose mean-square error indeed varies as T^{-3}. If the source location is unknown, two differential Doppler measurements do not determine the source velocity at all. One is now forced to estimate location and velocity jointly and in the translation from velocity to position, one multiplies by time, thus increasing the mean-square errors by the factor T^2. The mean-square error in location therefore varies as T^{-1}, even if one uses differential Dopplers as the basic data.

(b) When β_i for all i (same noise power at each sensor), Eq. (23) assumes the simpler form

$$D^2(\widehat{\Delta v}_i)=24\pi c^2\left(T^3 M\int_0^\infty \frac{\omega^2 S^2(\omega)/N^2(\omega)}{1+MS(\omega)/N(\omega)}\,d\omega\right)^{-1} . \tag{24}$$

Equations (23) and (24) hold for signals and noises of arbitrary bandwidth, subject only to the large TW assumption. $MS(\omega)/N(\omega)$ is simply the output signal-to-noise ratio of a conventional beamformer. If it is much larger than unity throughout the signal band, Eq. (24) becomes independent of M so that the use of additional sensors does not improve the estimate of Δv_i. Under these conditions the mean-square error depends on the inverse first power of the input signal-to-noise ratio. If the post-beamforming signal-to-noise ratio falls below unity, the mean-square error varies as the inverse square of the input signal-to-noise ratio.

(c) If the noise levels at different sensors are unequal, the situation is more complicated. The physical meaning of the term $\beta_i S(\omega)/N(\omega)$ in Eq. (23) is the signal-to-noise ratio at the ith sensor. The sum $\sum_{i=1}^{M}\beta_i S(\omega)/N(\omega)$ is, therefore, the sum of the individual signal-to-noise ratios, not the signal-to-noise ratio at the output of the conventional beamformer. Useful insight can be gained, however, by considering the following special case:

$$\frac{S(\omega)}{N(\omega)}=\begin{cases} S/N \text{ (a constant)}, & \omega_1\leq\omega\leq\omega_2 ,\\ 0, & \text{elsewhere} .\end{cases}$$

The integral in Eq. (23) is now easily evaluated. We focus on the improvement in the Δv_i estimate obtainable by using information from sensors other than i and M.

A straightforward computation yields

$$\frac{D^2(\hat{\Delta}v_i)\ \text{using all } M \text{ sensors}}{D^2(\hat{\Delta}v_i)\ \text{using only sensors } i \text{ and } M}$$

$$=\frac{\beta_i+\beta_M}{\sum_{i=1}^{M}\beta_i}\ \frac{1+\frac{S}{N}\sum_{i=1}^{M}\beta_i}{1+\frac{S}{N}(\beta_i+\beta_M)}\ . \tag{25}$$

Equation (25) assumes a value of approximately unity if $(\beta_i+\beta_M)S/N \gg 1$. If the sum of the signal-to-noise ratios at the ith and Mth sensors exceeds unity, information provided by other sensors is therefore essentially useless in the estimation of Δv_i. If such a statement holds for all $(M-1)$ linearly independent sensor pairs, the Δv_i estimates can be carried out pairwise without loss of performance. This condition can always be met if there is at least one sensor at which the signal-to-noise ratio exceeds unity. This sensor can then be chosen as a reference and the $(M-1)$ pairwise measurements result in essentially optimal estimates of all Δv_i. Conversely, if no sensor has a signal-to-noise ratio in excess of unity, Δv_i estimates obtained from the ith sensor pair only are distinctly suboptimal.

Differential velocity estimation has one other important property: It is uncoupled from estimation of spectral parameters such as signal center frequency and bandwidth. This is almost immediately evident from Eq. (7). In accordance with Eq. (17), K is block diagonal, so that K^{-1} is block diagonal. For any spectral parameter α, $\partial K/\partial\alpha$ has the same block diagonal property. It follows that $K^{-1}(\partial K/\partial\alpha)K^{-1}$ is block diagonal. However, according to Eq. (22), $\partial K/\Delta v_i$ has only zero entries in the $M\times M$ blocks along the principal diagonal. Hence Eq. (7) yields $J_{ij}=0$ for $\theta_i=\alpha$, $\theta_j=\Delta v_i$. Thus lack of prior knowledge concerning spectral parameters of the signal does not degrade the differential velocity estimate. A very similar argument can be used to establish lack of coupling between differential velocity and differential delay (position) estimates.

III. NARROWBAND SIGNALS

In many practical applications the signal has a very small bandwidth compared with the center frequency ω_0 of the radiated signal spectrum. If the bandwidth is so small that the signal correlation time exceeds the observation time, the signal is essentially a sinusoid of random amplitude and phase and the discussion in Sec. IV becomes applicable. Here we assume that the observation time is still large compared with the signal correlation time. To minimize algebraic complexity we deal only with a two-sensor receiving array.

The narrowband assumption implies that each frequency component of the signal is Doppler shifted by approximately the same amount so that there is a well-defined differential Doppler shift given by

$$\Delta\omega=(\Delta v/c)\omega_0\,. \tag{26}$$

It follows immediately from Eqs. (26) and (23) that

$$D^2(\hat{\Delta}\omega)=\frac{\omega_0^2}{c^2}D^2(\hat{\Delta}v)$$

$$\approx\frac{12\pi}{T^3\beta_1\beta_2\int_0^{\infty}\frac{S^2(\omega)/N^2(\omega)}{1+(\beta_1+\beta_2)S(\omega)/N(\omega)}d\omega}\,. \tag{27}$$

The narrowband assumption is invoked to justify replacement of ω^2 by ω_0^2 in the integral of Eq. (23).

In the narrowband case there is an instrumentationally attractive alternative for the estimation of $\Delta\omega$: One can make separate estimates of center frequency at each sensor and subtract. Consider the frequency estimation problem for the ith sensor, where the unknown center frequency is ω_i. According to Eq. (7), the best estimate of ω_i has the variance

$$D^2(\hat{\omega}_i)=\left[\mathrm{Tr}\left(K^{-1}\frac{\partial K}{\partial\omega_i}\right)^2\right]^{-1}. \tag{28}$$

According to Eq. (17), K is a diagonal matrix for the single sensor case $(i=j)$ and the diagonal entries are $S(\omega_n)+\beta_i^{-1}N(\omega_n)$. Hence $\partial K/\partial\omega_i$ is a diagonal matrix with the diagonal entries $\partial S(\omega_n)/\partial\omega_i$. It follows that $(K^{-1}\partial K/\partial\omega_i)^2$ is a diagonal matrix with the diagonal entries

$$\left(\frac{\partial S(\omega_n)/\partial\omega_i}{S(\omega_n)+\beta_i^{-1}N(\omega_n)}\right)^2=\left(\frac{\beta_i S(\omega_n)/N(\omega_n)}{1+\beta_i S(\omega_n)/N(\omega_n)}\ \frac{\partial S(\omega_n)/\partial\omega_i}{S(\omega_n)}\right)^2. \tag{29}$$

Summing over n one obtains the required trace. Under the large TW assumption the n sum can then be converted into an integral, so that one obtains finally

$$D^2(\hat{\omega}_i)=\frac{2\pi}{T}\left\{\int_0^{\infty}\left(\frac{\beta_i S(\omega)/N(\omega)}{1+\beta_i S(\omega)/N(\omega)}\ \frac{\partial}{\partial\omega_i}\ \log S(\omega)\right)^2 d\omega\right\}^{-1}_{i=1,2}. \tag{30}$$

Comparison of Eq. (27) with Eqs. (30) yields several interesting insights:

(a) Eq. (27) varies as T^{-3} whereas Eq. (30) varies as T^{-1}. Even though the signal waveshapes are unknown, the direct differential Doppler measurement exploits the fact that the waveshapes at the two sensors differ only by a fixed delay and frequency shift. It is basically a coherent procedure. Center frequency measurement is an incoherent procedure. It is basically a power measurement and obeys the well known T^{-1} dependence of such measurements.

(b) For high signal-to-noise ratios at each sensor, Eq. (30) assumes a limiting form independent of the noise spectrum. Eq. (27) decreases with the first power of the signal-to-noise ratio even for very high signal-to-noise ratios. The fact that the center frequency estimates has an absolute lower bound greater than zero is another property common to power measurements: even if there is no noise at all the randomness of the signal prevents error-free determination of its spectral distribution in a finite time T. A coherent procedure, on the other hand, can establish $\Delta\omega$ without error in the absence of noise by matching any finite segment of the Doppler shifted output from one sensor against the output of the other sensor.

(c) The quality of the center frequency estimate, and hence that of the differential Doppler estimate computed from it, depends critically on the slope of the signal spectrum. This is reasonable, for the best estimator of center frequency compares the average power in filters above and below the center frequency. If there is a frequency range where this power changes rapidly, a more sensitive measurement can be performed.

(d) There is an apparent anomaly in the comments made under a): the observation time $T=1$ appears to have special significance, making the differential Doppler estimate more accurate for large T and the center frequency estimate more accurate for small T. If this were true, one could improve differential Doppler estimates for small T by subtracting center frequency estimates. This is clearly nonsense because we have established the absolute optimality of the direct differential Doppler measurement (at least for large TW products). It is not difficult to find an explanation for the apparent paradox. Suppose the signal spectrum has the form

$$S(\omega)=S_1[(\omega-\omega_0)/\Omega], \tag{31}$$

where $S_1(x)$ does not depend on ω_0 and Ω. In other words $S_1()$ specifies the basic spectral form while the parameters ω_0 and Ω ($=2\pi W$ rad/s) adjust the center frequency and bandwidth, respectively. Suppose further that the noise spectrum is essentially flat with spectral level N_1 over the signal band. Then Eqs. (27) and (30) can be written in the form

$$D^2(\hat{\Delta}\omega)=\frac{12\pi}{\beta_1\beta_2T^2(\Omega T)}\left\{\int_{-\infty}^{\infty}\frac{S_1^2(x)/N_1^2}{1+(\beta_1+\beta_2)S_1(x)/N_1}dx\right\}^{-1}, \tag{32}$$

$$D^2(\hat{\omega}_i)=\frac{2\pi(\Omega T)}{T^2}\left\{\int_{-\infty}^{\infty}\left(\frac{\beta_iS_1(x)/N_1}{1+\beta_iS_1(x)/N_1}\times\frac{d}{dx}[\log S_1(x)]\right)^2dx\right\}^{-1}. \tag{33}$$

The two integrals are independent of ω_0 and Ω. One therefore obtains immediately

$$D^2(\hat{\Delta}\omega)/D^2(\hat{\omega}_i)\propto 1/(T\Omega)^2. \tag{34}$$

Thus the significant parameter is not the observation time T but the time-bandwidth product $T\Omega$. We have assumed that $T\Omega\gg1$ so that the advantage rests very clearly with the differential Doppler measurement.

The only remaining question is whether the difference of the center frequency estimates, $\hat{\omega}_1-\hat{\omega}_2$, has a mean-square error of a form similar to Eq. (33) so that Eq. (34) represents a reasonable comparison. The answer is very obvious when the signal-to-noise ratio at each sensor is low. Because of the postulated lack of noise coherence from sensor to sensor, the errors in the two center frequency measurements are then essentially uncorrelated so that their mean square values add. The relative efficiency of direct differential Doppler measurements and subtraction of center frequencies is, therefore, well characterized by Eq. (34). When the signal-to-noise ratio is high, the errors in the two center frequency estimates are correlated because of the common signal component. Calculation of the correlation coefficient is quite tedious and the resulting expression is algebraically cumbersome. For our purposes the only significant feature of the result is that the T and Ω dependence of $D^2(\hat{\omega}_1-\hat{\omega}_2)$ is precisely the same as that of Eq. (33) so that Eq. (34) remains valid.

IV. SINUSOIDAL SIGNALS

If the source radiates a pure sinusoid the signal components received at various sensors are sinusoids differing in frequency (because of Doppler shifts) phase (because of differential signal travel time) and possibly amplitude. Using the signal component at sensor M as a reference we can write the waveshape received at sensor i as follows:

$$r_i(t)=A_i\sin[(\omega_M+\Delta\omega_i)t-\phi_m-\Delta\phi_i]+n_i(t). \tag{35}$$

$\Delta\omega_i$ is the differential Doppler shift, $\Delta\phi_i$ the differential phase delay, and A_i the signal amplitude at sensor i. We shall assume that the noise is Gaussian, spatially incoherent, and spectrally white with bandwidth W. Nyquist rate samples of the received waveshapes are then statistically independent Gaussian random variables with mean values provided by the signal samples. The conditional probability density $p(\mathbf{r}/\boldsymbol{\theta})$ required by Eq. (2) therefore assumes the form

$$p\left(\frac{\mathbf{r}}{\boldsymbol{\theta}}\right)=C\exp\left\{-\sum_{p=1}^{M}\sum_{q=1}^{n}\frac{1}{2N_pW}[r_p(t_q)-s_p(t_q)]^2\right\}. \tag{36}$$

$s_p(t)$ is the signal component of the waveshape received by the pth sensor, N_p is the spectral level of the noise at the pth sensor (so that N_pW is the average noise power at that sensor), and C is a normalizing constant. The unknown parameters $\boldsymbol{\theta}$ are contained only in the $s_p(t_q)$, so that the elements of the Fisher information matrix are readily computed from Eq. (2).

$$J_{ij}=\sum_{p=1}^{M}\sum_{q=1}^{n}\frac{1}{N_pW}\frac{\partial s_p(t_q)}{\partial\theta_i}\frac{\partial s_p(t_q)}{\partial\theta_j}. \tag{37}$$

If the signal waveshape does not change significantly over an interval of $1/2W$ s, we can convert the q sum into an integral and obtain results in very compact form. Furthermore, we may make this assumption without loss of generality, for we can always obtain a sufficiently large W by a prewhitening operation which has negligible effect on the signal as long as the noise TW product is large. The remainder of the calculation is straightforward and we only state the results.

We arrange the vector of unknown parameters in the following order

$$\boldsymbol{\theta}^T=\{\Delta\omega_1\cdots\Delta\omega_{M-1},\omega_M,\phi_M,\Delta\phi_1\cdots\Delta\phi_{M-1},A_1\cdots A_M\}. \tag{38}$$

With this ordering, the Fisher information matrix turns out to be block diagonal, one block consisting of the frequency and differential frequency parameters, the second of the phase and differential phase parameters and the third of the amplitude parameters. The three groups are therefore uncoupled and, since phases and amplitudes are only nuisance parameters in our problem, we can confine our attention to the Fisher information matrix for the subset $\{\Delta\omega_1, \Delta\omega_2, \cdots \Delta\omega_{M-1}, \omega_M\}$.

$$J(\Delta\hat{\omega}_1, \cdots \Delta\hat{\omega}_{M-1}, \hat{\omega}_M) = \frac{T^3}{12}\begin{bmatrix} A_1^2/N_1 & 0 & \cdots & 0 & A_1^2/N_1 \\ 0 & A_2^2/N_2 & & & \\ \cdot & \cdot & & & \\ \cdot & & \cdot & 0 & \\ \cdot & & & \cdot & \\ 0 & 0 & & A_{M-1}^2/N_{M-1} & A_{M-1}^2/N_{M-1} \\ A_1^2/N_1 & & & A_{M-1}^2/N_{M-1} & \sum_{i=1}^{M} A_i^2/N_i \end{bmatrix}. \tag{39}$$

If the frequency ω_M at the reference sensor is known, the Fisher matrix for $\{\Delta\omega_1 \cdots \Delta\omega_{M-1}\}$ consists only of the first $(M-1)$ rows and columns of Eq. (39). It is therefore diagonal, the various differential Doppler estimates are uncoupled, and their variance is given by

$$D^2(\Delta\hat{\omega}_i) = (12/T^3)(N_i/A_i^2). \tag{40}$$

The A_i are the signal amplitudes defined in connection with Eq. (35).

In practive ω_M is almost certainly unknown. This lack of knowledge introduces coupling between the differential Doppler estimates and reduces their accuracy. To study this effect one must formally invert Eq. (39), an operation which is not difficult to perform because of the special form of the equation. The diagonal elements of J^{-1} are given by

$$D^2(\hat{\omega}_i) = (J^{-1})_{ii} = \frac{12}{T^3}\left(\frac{N_i^2}{A_i^2} + \frac{N_M^2}{A_M^2}\right), \quad i = 1, 2, \cdots (M-1), \tag{41}$$

$$D^2(\hat{\omega}_M) = (J^{-1})_{MM} = (12/T^3)(N_M^2/A_M). \tag{42}$$

The second term in Eq. (41) represents the increment in the mean square error due to lack of *a priori* knowledge of ω_M. Note that the mean-square error is independent of the array size M. Information from other sensors is not useful in improving the differential Doppler estimate for a given pair. This is consistent with the result obtained for Gaussian signals. There we obtained optimal performance with a single sensor pair when the signal-to-noise ratio in the signal band was high for each sensor. In the sinusoidal case the signal bandwidth is nominally zero so that the high signal-to-noise ratio condition is always satisfied.

Equation (41) also confirms the obvious desirability of using, for the reference, the sensor with the highest available signal-to-noise ratio. The second term in the equation is simply the inverse of the signal-to-noise ratio at the reference sensor. With the optimal choice of reference this term is never larger than the first.

V. COMPARISON OF SINUSOIDAL AND NARROWBAND GAUSSIAN CASE

What, if any, advantage is to be gained from the knowledge that the signal is a true sinusoid rather than a sample function of a narrowband Gaussian process? For an answer to this question we examine our results for differential Doppler estimation using two sensors. The result for narrowband Gaussian signal is given by Eq. (27), that for sinusoidal signal by Eq. (41) (with $i=1$, $M=2$). Because signal-to-noise ratio in the signal band is automatically high for the sinusoidal case, we make a similar assumption for the Gaussian signal. Equation (27) now becomes

$$D^2(\Delta\hat{\omega}) = \frac{12\pi}{T^3}\left\{\left(\int_0^\infty \frac{\beta_1 S(\omega)}{N(\omega)}\,d\omega\right)^{-1} + \left(\int_0^\infty \frac{\beta_2 S(\omega)}{N(\omega)}\,d\omega\right)^{-1}\right\}. \tag{43}$$

In the sinusoidal case we defined N_i as the white noise power level at the ith sensor measured as power/Hertz; Eq. (43) implies spectral levels measured as power/(rad/s). The appropriate equivalent is therefore

$$N(\omega)/\beta_i = N_i/2\pi. \tag{44}$$

Then Eq. (43) assumes the form

$$D^2(\Delta\hat{\omega}) = 6(N_1 + N_2)/T^3 P_s, \tag{45}$$

where $P_s = \int_0^\infty S(\omega)\,d\omega$, the total signal power at each sensor. In that terminology the sinusoidal signal power at each sensor is

$$A_1^2/2 = A_2^2/2 = P_s. \tag{46}$$

Expressed in terms of P_s, Eq. (41) becomes

$$D^2(\Delta\hat{\omega}) = 6(N_1 + N_2)/T^3 P_s. \tag{47}$$

This is identical with Eq. (45). The best attainable accuracy of the differential Doppler estimate is therefore the same, no matter whether the signal is a sinusoid or a narrowband Gaussian process.

There is, however, one practically important difference between the two situations. In the Gaussian case, the estimate of center frequency was of much poorer quality than the differential Doppler estimate. This is not the case when the signal is sinusoidal. For a single sensor the mean-square estimation error of center frequency at one sensor output is given by Eq. (42). Since the frequency estimates at different sensors are statistically independent

$$D^2(\hat{\omega}_2 - \hat{\omega}_1) = \frac{12}{T^3}\left(\frac{N_1}{A_1^2} + \frac{N_2}{A_2^2}\right). \tag{48}$$

Equation (48) is identical with Eq. (41) (with $i=1, M=2$).

When the signals are sinusoidal, one therefore does not suffer any performance degradation from the instrumentationally attractive procedure of estimating frequencies at separate sensors, and then subtracting to generate the differential Doppler estimate.

ACKNOWLEDGMENTS

The work reported in this paper was supported by the Naval Ocean Systems Center, San Diego under contracts N66001-76-C-1082 and N66001-77-C-0396 and by NORDA under contract N00014-77-C-0196.

APPENDIX

Theorem: Let $\mathbf{r}$ be the data vector and let

$$\mathbf{q} = \mathbf{F}(\mathbf{r}), \tag{A1}$$

be a continuous one-to-one transformation with continuous partial derivatives $\partial q_i/\partial r_j$. Then

$$E\left\{\frac{\partial^2}{\partial\theta_k\partial\theta_l}\log P_Q\left(\frac{\mathbf{q}}{\theta}\right)\right\} = E\left\{\frac{\partial^2}{\partial\theta_k\partial\theta_l}\log P_R\left(\frac{\mathbf{r}}{\theta}\right)\right\}. \tag{A2}$$

Proof: From Eq. (A1) it follows[8] that $\mathbf{q}$ has a probability density simply related to that of $\mathbf{r}$

$$P_Q(\mathbf{q}/\theta) = (1/d)P_R(\mathbf{r}/\theta)\big|_{\mathbf{r}=\mathbf{F}^{-1}(\mathbf{q})}. \tag{A3}$$

d, the Jacobian of the transformation in Eq. (A1), is independent of θ. Taking the natural logarithm of both sides of Eq. (A3), and differentiating with respect to θ_k and θ_l, one obtains

$$\frac{\partial^2}{\partial\theta_k\partial\theta_l}\log P_Q\left(\frac{\mathbf{q}}{\theta}\right) = \frac{\partial^2}{\partial\theta_k\partial\theta_l}\log P_R\left(\frac{\mathbf{r}}{\theta}\right)\Bigg|_{\mathbf{r}=\mathbf{F}^{-1}(q)}. \tag{A4}$$

The left hand side of Eq. (A2) reads

$$E\left\{\frac{\partial^2}{\partial\theta_k\partial\theta_l}\log P_Q\left(\frac{\mathbf{q}}{\theta}\right)\right\} = \iint_{\Omega_Q}\left[\frac{\partial^2}{\partial\theta_k\partial\theta_l}\log P_Q\left(\frac{\mathbf{q}}{\theta}\right)\right]P_Q\left(\frac{\mathbf{q}}{\theta}\right)d\mathbf{q}. \tag{A5}$$

Substituting Eq. (A4) in Eq. (A5) and making the change of variables

$$\mathbf{r} = \mathbf{F}^{-1}(\mathbf{q}), \tag{A6}$$

Eq. (A5) becomes

$$E\left\{\frac{\partial^2}{\partial\theta_k\partial\theta_l}\log P_Q\left(\frac{\mathbf{q}}{\theta}\right)\right\} = \int\int_{\Omega_R}\left[\frac{\partial^2}{\partial\theta_k\partial\theta_l}\log P_R\left(\frac{\mathbf{r}}{\theta}\right)\right]P_R\left(\frac{\mathbf{r}}{\theta}\right)d\mathbf{r}. \tag{A7}$$

The right side of Eq. (A7) is equivalent to the right side of Eq. (A2), thus proving the theorem.

[1]G. C. Carter and C. H. Knapp, "Time Delay Estimation," Proc. IEEE 1976 Int. Conf. Acoust. Speech and Signal Proc. (ICASSP), p. 357.

[2]W. R. Hahn, "Optimum Signal Processing for Passive Sonar Range and Bearing Estimation," J. Acoust. Soc. Am. 58, 201 (1975).

[3]V. H. MacDonald and P. M. Schultheiss, "Optimum Passive Bearing Estimation in a Spatially Incoherent Noise Environment," J. Acoust. Soc. Am. 46, 37 (1969).

[4]W. J. Bangs, "Array Processing with Generalized Beamformers" Ph.D. dissertation, Yale University 1971 (also NUSC Tech. Rep.).

[5]C. H. Knapp and G. C. Carter, "Estimation of Time Delay in the Presence of Source or Receiver Motion," J. Acoust. Soc. Am. 61, 1545 (1977).

[6]E. Weinstein and P. M. Schultheiss, "Localization of a Moving Source Using Passive Array Data" Nav. Ocean Syst. Center Tech. Rep. December 1978.

[7]Reference 3 Eq. (17).

[8]H. Cramer, *Mathematical Methods of Statistics* (Princeton U. P., Princeton, NJ, 1951), Chap. 22.