

Analysis of the Automatic Frequency Control in Heterodyne Optical Receivers

Alberto Bononi, Pierpaolo Ghiggino, and Giorgio Picchi, *Member, IEEE*

Abstract—A theoretical analysis of the automatic frequency control (AFC) in binary DPSK and FSK heterodyne optical receivers is presented and its impact on the system performance is evaluated. The effects of the shot-noise AM/FM conversion in the frequency discriminator are analyzed. It is shown that with commercial DFB lasers a common analysis of the frequency control loop can be done for the three modulations considered. Design guidelines are given that account for the presence of the AFC loop by properly using some performance results derived in the assumption of perfect frequency tracking.

I. INTRODUCTION

SEMICONDUCTOR lasers used in coherent optical communications suffer from a frequency noise [1]–[3] whose spectral density has two main components: 1) a high-frequency flat or white component, which has the effect of broadening the spectral linewidth of the emitted field, and 2) a low-frequency flicker component which causes a slow drift of the spectrum itself.

Optical heterodyne systems with noncoherent demodulation are robust against the degradation due to the white component [4], [5], but need an automatic frequency control (AFC) loop to keep the heterodyned spectrum within the passband of the intermediate frequency (IF) stages. Without any frequency control, the flicker component causes the IF spectrum to drift away from the center of the IF passband, thus causing distortions, fading, and possibly the complete loss of the signal.

This paper analyzes on a common basis the AFC design and its effects on the receiver performance for three binary modulation formats with noncoherent demodulation, i.e., DPSK, narrow-deviation FSK and wide-deviation FSK. The form and impact of the shot noise on the AFC tracking ability is also investigated. In Section II a model of the receiver is given. Section III contains the AFC loop analysis, with focus on the main design parameters. In Sections IV and V the theoretical system bit-error-rate (BER) degradation as a function of the AFC parameters is derived for DPSK and FSK receivers, respectively. Conclusions are drawn in Section VI.

Manuscript received July 30, 1991. This work was supported by The National Research Council (CNR) and Marconi S.p.A. in the frame of the Telecommunication Project.

A. Bononi and G. Picchi are with the Dipartimento di Ingegneria dell'Informazione, Università di Parma, Viale delle Scienze, I-43100 Parma, Italy.

P. Ghiggino was with Marconi S.p.A., Genova Cornigliano, Italy. He is now with GEC Hirst Research Centre, East Lane, Middx. HA 97PP, United Kingdom.

IEEE Log Number 9107428.

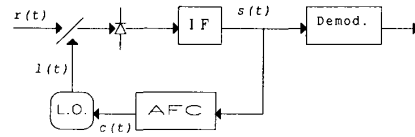


Fig. 1. Heterodyne receiver with general AFC scheme.

II. RECEIVER MODEL

A general scheme of a heterodyne receiver including the AFC loop is drawn in Fig. 1.

Since angle modulation only is considered in this paper, the optical received signal is expressed by

$$r(t) = \sqrt{2P_S} \cos[2\pi f_T t + \theta_T(t) + \phi_m(t)] \quad (1)$$

where $\theta_T(t)$ is the phase noise of the transmitting laser, P_S is the received optical power, and $\phi_m(t)$ is the digital angle modulation.

The local laser oscillator generates the optical signal

$$l(t) = \sqrt{2P_L} \cos[2\pi f_L t + \theta_L(t) + \phi_c(t)] \quad (2)$$

where P_L is the optical power, $\theta_L(t)$, is the phase noise, and $\phi_c(t)$ is a control phase driven by the AFC loop through an electrical signal $c(t)$. If the control characteristic of the laser is ideal, as assumed here, the control phase is

$$\phi_c(t) = 2\pi K_c \int_0^t c(u) du \quad (3)$$

where K_c is a constant depending on the laser drive. The received and the locally generated optical beams are mixed in a coupler and the resulting field is converted into a current by a photodiode. The IF is an ideal, bandpass filter centered at $f_{IF} \triangleq f_T - f_L$, with bandwidth B_{IF} , that does not affect the signal. The voltage at the output of the IF filter is then

$$s(t) = V \cos[2\pi f_{IF} t + \theta_{IF}(t) - \phi_c(t) + \phi_m(t)] + w(t) \quad (4)$$

where V is the signal amplitude

$$V = 2Z_{IF}\mathcal{R}\sqrt{P_L P_S} \quad (\text{volt}) \quad (5)$$

Z_{IF} is the IF stage transimpedance, \mathcal{R} the photodiode responsivity, $\theta_{IF}(t) \triangleq \theta_T(t) - \theta_L(t)$ is the IF phase noise and $w(t)$ is the shot noise generated in the photodetection process.

The time derivative, divided by 2π , of the phase noise process $\theta_{IF}(t)$ is the corresponding frequency noise process which will be denoted by $n_{IF}(t)$ and assumed as a zero-mean Gaussian process [3].

In shot noise limited condition, the shot noise $w(t)$ in (4) is a zero-mean additive Gaussian process, with spectral density $S_w(f)$ flat on the IF bandwidth

$$S_w(f) = (q\mathcal{R}P_L)Z_{IF}^2[\mathcal{G}_{B_{IF}}(f - f_{IF}) + \mathcal{G}_{B_{IF}}(f + f_{IF})] \quad (6)$$

where q is the electron charge and $\mathcal{G}_{B_{IF}}(f)$ is a normalized gate centered at zero frequency, extending from $-B_{IF}/2$ up to $B_{IF}/2$. From (5) and (6) we obtain the IF signal-to-noise ratio

$$\text{SNR} \triangleq \frac{V^2/2}{2S_w(f_{IF})B_{IF}} = \frac{\mathcal{R}P_S}{qB_{IF}}$$

The spectral density $S_{n_{IF}}(f)$ of the frequency noise at the IF filter output can be found from the spectral density of the same noise at the IF filter input, by using some results from [6], [7] on the effect of filtering on the frequency noise. The following one-sided model will be used for $S_{n_{IF}}(f)$ [8]:

$$S_{n_{IF}}(f) = \begin{cases} \Delta\nu/\pi + K/f & \text{for } f < B_{IF}/2 \\ 0 & \text{elsewhere} \end{cases} \quad (7)$$

where $\Delta\nu$ is the 3-dB linewidth of the IF unmodulated spectrum and K is the flicker noise coefficient, ranging from 10^{11} to 10^{13} Hz² for semiconductor lasers [8]. The model is acceptable for $B_{IF} \gg \Delta\nu$ [6].

III. AFC LOOP ANALYSIS

All the AFC schemes considered in the paper require the modulation to be removed from the IF signal. Since the way of removing the modulation depends on the modulation format itself, the general scheme of Fig. 1 specializes into one of the forms shown in Fig. 2 accordingly. The instantaneous frequency of the cosine term in (4), with modulation removed, is

$$f_{IF} + n_{IF}(t) - K_c c(t) \quad (8)$$

where $c(t)$ is the control signal generated in all the schemes by a hard-limiter/frequency-discriminator followed by a low-pass loop filter $F(f)$. It is evident from (8) that a closed-loop frequency error in tracking f_{IF} can be defined as

$$e(t) \triangleq n_{IF}(t) - K_c c(t). \quad (9)$$

Note that $c(t)$ has low-frequency components only, while $n_{IF}(t)$ has both low- and high-frequency components, i.e., flicker and white respectively. Goal of the AFC is to track the low frequency component of $n_{IF}(t)$ only. The untracked component has the effect of broadening the IF spectrum.

The variance σ_e^2 of the tracking error $e(t)$ is a key parameter to determine the BER in FSK receivers and to judge the overall goodness of the AFC loop. This variance will be evaluated by integrating its spectral density $S_e(f)$ obtained by solving the AFC loop equation.

With modulation removed, the passband IF signal (4) can be expressed in its Rice expansion with respect to ω_{IF} , using (8) and (9), as

$$s(t) \triangleq \text{Re}[\tilde{s}(t)e^{j\omega_{IF}t}] = \text{Re}\left[\left(V e^{j\alpha(t)} + \tilde{w}(t)\right)e^{j\omega_{IF}t}\right] \quad (10)$$

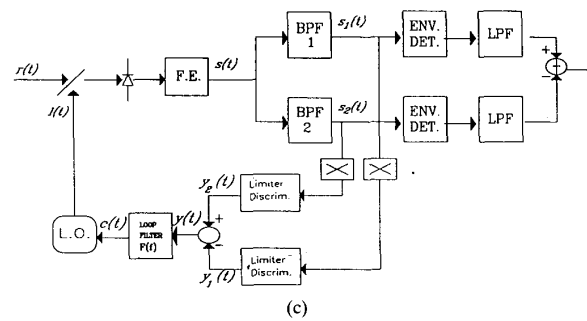
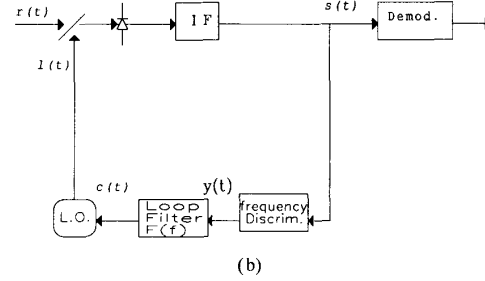
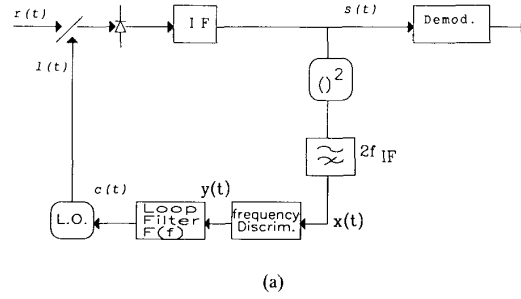


Fig. 2. (a) DPSK receiver with AFC scheme. (b) ND-FSK receiver with AFC scheme. (c) WD-FSK dual branch heterodyne receiver with envelope detection.

where $\alpha(t) \triangleq 2\pi \int_0^t e(u) du$, $\tilde{w}(t)$ is the complex envelope of the shot noise process $w(t)$, and $\tilde{s}(t)$ is the complex envelope of $s(t)$, that can be also written as

$$\tilde{s}(t) = [V + \tilde{w}'(t)]e^{j\alpha(t)} \quad (11)$$

where

$$\tilde{w}'(t) \triangleq \tilde{w}(t)e^{-j\alpha(t)}. \quad (12)$$

Writing the term in square brackets in (11) in its magnitude and phase gives

$$V + \tilde{w}'(t) = E(t)e^{j\psi(t)} \quad (13)$$

where $\psi(t)$ and $E(t)$ are real. The IF complex envelope then becomes

$$\tilde{s}(t) = E(t)e^{j[\alpha(t) + \psi(t)]}. \quad (14)$$

An ideal limiter-discriminator centered at ω_{IF} outputs a signal proportional to the derivative of the argument $[\alpha(t) + \psi(t)]$ of $\tilde{s}(t)$, divided by 2π , i.e., to the instantaneous frequency offset

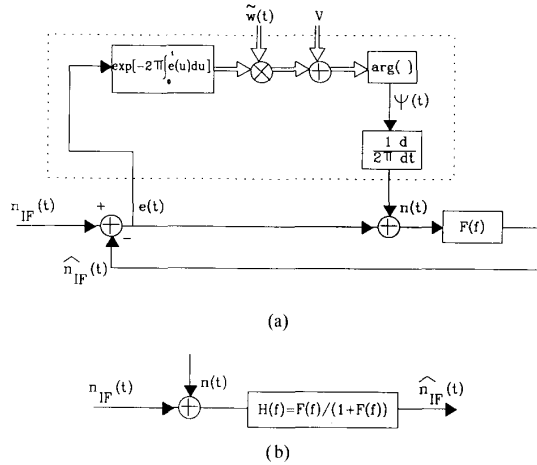


Fig. 3. (a) AFC loop block diagram, including the generation of the shot noise AM to FM conversion $n(t)$. (b) AFC loop as an estimator of $n_{IF}(t)$ affected by $n(t)$.

from the discriminator center frequency. This signal can thus be written as

$$y(t) = K_d[e(t) + n(t)] \quad (15)$$

where

$$n(t) \triangleq \frac{\dot{\psi}(t)}{2\pi} \quad (16)$$

is the shot noise AM-to-FM conversion in the discriminator. The loop equation is then easily written from (9), by observing that $c(t)$ is a filtered version of $y(t)$. Therefore

$$e(t) = n_{IF}(t) - [e(t) + n(t)] \otimes f(t) \quad (17)$$

where \otimes indicates convolution, $f(t)$ is the impulse response of the loop filter $F(f)$, and K_c , K_d have been adsorbed into $f(t)$. Assuming that at steady state all the processes in the loop are wide-sense stationary, the spectral density of both sides of (17) is obtained

$$S_e(f) = C_1(f)S_{n_{IF}}(f) + C_2(f)S_n(f) \quad (18)$$

where $S_n(f)$ is the spectral density of $n(t)$ and

$$C_1(f) \triangleq \left| \frac{1}{1+F(f)} \right|^2, \quad C_2(f) \triangleq \left| \frac{F(f)}{1+F(f)} \right|^2. \quad (19)$$

Since $F(f)$ is usually low-pass, $C_1(f)$ is high-pass and $C_2(f)$ is low-pass. The frequency error is then composed of the high-frequency part of $n_{IF}(t)$ and of the low-frequency part of $n(t)$.

Two block diagrams representing (17) are given in Fig. 3. Fig. 3(a) shows the AFC loop and the blocks that generate $n(t)$. Fig. 3(b) shows the AFC loop as an estimator of the input $n_{IF}(t)$ disturbed by $n(t)$. The function $H(f) \triangleq F(f)/(1+F(f))$ is the closed-loop transfer function of the AFC and its 3-dB bandwidth B_{cl} is the closed-loop bandwidth. Note that, since $|H(f)|^2 = C_2(f)$, the shot noise component in (18) is filtered directly by $H(f)$.

To find $S_e(f)$ from (18) a loop filter $F(f)$ must be specified and an expression of the spectral density $S_n(f)$ must be found.

The first component of $S_e(f)$, i.e., $S_{n_{IF}}(f)$, has the typical $1/f$ low-frequency behavior. Therefore a finite value of σ_e^2 can be obtained only if the product $C_1(f)S_{n_{IF}}(f)$ remains integrable as f approaches zero. The simplest filter satisfying this condition is the ideal integrator, i.e.,

$$F(f) = \frac{B_{cl}}{jf} \quad (20)$$

and consequently

$$C_1(f) = \frac{f^2}{B_{cl}^2 + f^2}, \quad C_2(f) = \frac{B_{cl}^2}{B_{cl}^2 + f^2}. \quad (21)$$

The form of the second component of $S_e(f)$, i.e., $S_n(f)$, depends on the modulation format. In Appendices A and B, $S_n(f)$ is evaluated for the various schemes under consideration and it is shown that, in the assumption

$$\frac{\Delta\nu}{\pi} \gg \frac{B_{cl}^2}{B_{IF}(\text{SNR})}$$

the contribution of $S_n(f)$ can be neglected. This allows a simplified common analysis of the AFC loop based on the assumption

$$S_e(f) \cong C_1(f)S_{n_{IF}}(f) \quad (22)$$

that reduces $e(t)$ to a simple Gaussian process.

Based on (22), the one-sided expression of $S_e(f)$ for the ideal integrator loop filter is obtained from (7) and (21)

$$S_e(f) \cong \begin{cases} \frac{f^2}{B_{cl}^2 + f^2} \left[\frac{\Delta\nu}{\pi} + \frac{K}{f} \right] & f \leq B_{IF}/2 \\ 0 & f > B_{IF}/2 \end{cases} \quad (23)$$

This is plotted in Fig. 4 for various values of the closed loop bandwidth B_{cl} . Choosing for the loop filter a more realistic first-order low-pass filter

$$F(f) = \frac{G_{ol}}{1 + j(f/B_{ol})} \quad (24)$$

of dc gain G_{ol} and bandwidth B_{ol} , the coefficients C_1 and C_2 become

$$C_1(f) = \frac{B_{ol}^2 + f^2}{[(G_{ol} + 1)B_{ol}]^2 + f^2}$$

$$C_2(f) = \frac{(G_{ol}B_{ol})^2}{[(G_{ol} + 1)B_{ol}]^2 + f^2}. \quad (25)$$

For high gain $G_{ol} \gg 1$ the closed loop bandwidth is then

$$B_{cl} = G_{ol}B_{ol}.$$

A plot of $S_e(f)$ in the approximation (22), using the coefficients (25), is given in Fig. 5 for a fixed value of B_{cl} .

As G_{ol} gets higher and B_{ol} gets smaller, the curves tend to match the ideal integrator curve better and better, as expected. Therefore the ideal integrator will be the only filter considered in the analysis.

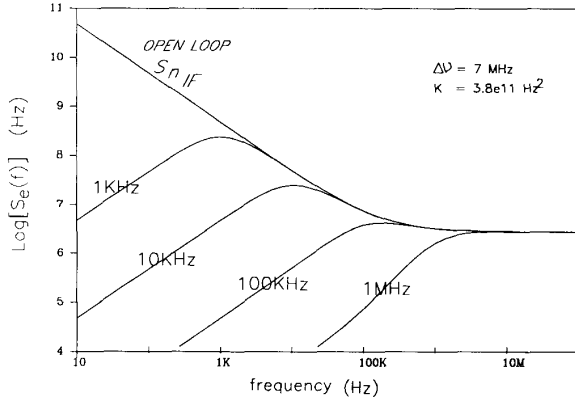


Fig. 4. Closed loop frequency error spectral density for several values of closed-loop bandwidth B_{cl} and ideal integrator loop filter $F(f)$. Drop at $B_{IF}/2$ not shown ($B_{IF} > 200$ MHz).

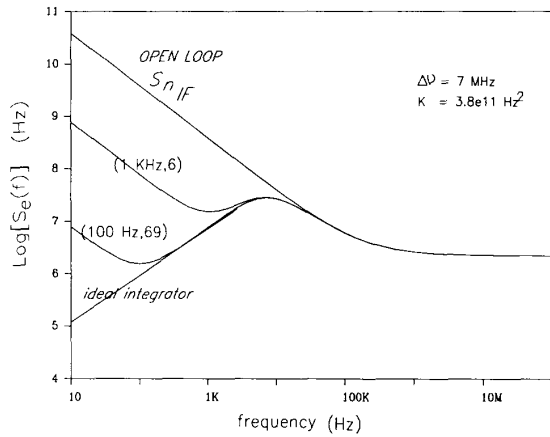


Fig. 5. Closed loop frequency error spectral density for fixed closed loop bandwidth $B_{cl} = 7$ MHz and first order low-pass loop filter $F(f)$ for two pairs of open loop bandwidth-gain (B_{ol}, G_{ol}).

Integration of $S_e(f)$ in (23) yields, for $B_{IF} \gg B_{cl}$

$$\sigma_e^2 = \left[\frac{\Delta\nu}{2\pi} (B_{IF} - \pi B_{cl}) \right] + K \ln \left(\frac{B_{IF}}{2B_{cl}} \right). \quad (26)$$

In Fig. 6 a plot of the closed-loop error frequency rms—square root of (26)—versus the closed loop bandwidth is given for two pairs ($K, \Delta\nu$) of the input frequency noise parameters.

Although the system bit error rate depends on σ_e^2 as a whole, each term of (26) will now be analyzed for a better insight.

Without AFC, the variance due to the white component would be [4]

$$\sigma_{white,o}^2 \triangleq \frac{\Delta\nu B_{IF}}{2\pi}. \quad (27)$$

The AFC has usually a narrow bandwidth B_{cl} and does not track much of the white component. Thus, in the presence of the AFC, the variance due to the white component is

$$\sigma_{white}^2 \triangleq \frac{\Delta\nu}{2\pi} (B_{IF} - \pi B_{cl}) \cong \frac{\Delta\nu B_{IF}}{2\pi} \quad \text{for } B_{IF} \gg B_{cl}.$$

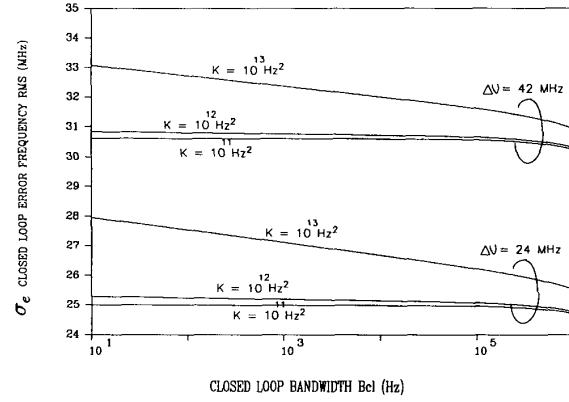


Fig. 6. AFC: closed loop frequency error r.m.s. for an ideal integrator loop filter. IF bandwidth 140 MHz.

The “fast” white component originates the lorentzian-like lineshape of the unmodulated IF spectrum. The effect of the “slow” flicker component is to make the lineshape slowly drift in the optical frequency domain. The effect can be directly observed by monitoring the signal at the output of the frequency discriminator. Due to the finite bandwidth of real discriminators, the rms of this signal is mainly due to the square root of the second term of (26), i.e.,

$$\sigma_{slow} \triangleq \sqrt{K \ln \left(\frac{B_{IF}}{2B_{cl}} \right)}. \quad (28)$$

For a standard value of $K = 10^{12} \text{ Hz}^2$, σ_{slow} is of the order of a few megahertz even with values of B_{cl} as low as 10 Hz [13].

This AFC loop analysis has been carried out assuming that the modulation has been perfectly removed from the IF signal. In Appendix B it is shown how this removal can be practically implemented for the various modulations considered and what the actual spectra are. It is found, however, that the results are, in all cases, quite close to the simplified analysis given in this section.

IV. DPSK PERFORMANCE EVALUATION

As shown in Fig. 2(a) a delay demodulator is used to demodulate the IF DPSK signal $s(t)$ given in (B.1).

It is shown in [9] that the system BER can be expressed as a function of both the IF SNR and the variance $\sigma_{\Delta\theta_e}^2$ of the Gaussian phase noise wander over the symbol interval T

$$\Delta\theta_e \triangleq 2\pi \int_{t-T}^t e(u) du \quad (29)$$

where $e(t)$ is the closed-loop frequency error process in (9).

Fixing $\text{BER} = 10^{-9}$, the corresponding values of the IF SNR versus $\Delta\theta_e$ give the sensitivity curve shown in Fig. 7.

To evaluate $\sigma_{\Delta\theta_e}^2$ from (29), observe that $\Delta\theta_e$ can be seen as a convolution between a gate T seconds long and $e(t)$. Therefore

$$\sigma_{\Delta\theta_e}^2 = 4 \int_0^\infty S_e(f) \frac{\sin^2(\pi f T)}{f^2} df. \quad (30)$$

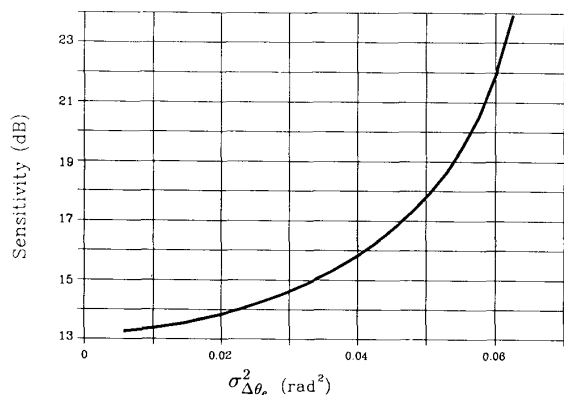


Fig. 7. DPSK Sensitivity for BER = 10^{-9} versus phase error variance.

In DPSK the IF bandwidth equals the bit rate R . In the assumption $B_{IF} = R \gg B_{cl}$ and using (23) for $S_c(f)$ in (30), we get the approximation

$$\sigma_{\Delta\theta_e}^2 \cong \sigma_{\Delta\theta_{white}}^2 + \left(\frac{2\pi}{R}\right)^2 \left[K \ln\left(\frac{R/2}{B_{cl}}\right) - \frac{\Delta\nu B_{cl}}{2} \right] \quad (31)$$

where

$$\sigma_{\Delta\theta_{white}}^2 \triangleq 4 \int_0^{B_{IF}/2} \frac{\Delta\nu}{\pi} \frac{\sin^2(\pi fT)}{f^2} df \cong 1.54 \frac{\Delta\nu}{R} \quad (32)$$

is the open-loop variance due to the white component only. Computation of (30) with typical parameters shows that the main portion of $\sigma_{\Delta\theta_e}^2$ is $\sigma_{\Delta\theta_{white}}^2$, and that the difference ($\sigma_{\Delta\theta_e}^2 - \sigma_{\Delta\theta_{white}}^2$) is in the range 10^{-5} to 10^{-3} rad². Comparing these values with the typical values of $\sigma_{\Delta\theta_e}^2$ in Fig. 7 shows that an increase of the variance due to the nonwhite component results in little degradation of the sensitivity.

In this evaluation the IF signal distortion due to the residual "vibration" of the signal spectrum within the IF bandwidth has been neglected. This is because typical values for the rms of the residual low-frequency drift are in the range of a few megahertz, which is much less than the IF bandwidth. In fact from (31) and Fig. 7 it is clear that DPSK coherent systems require linewidths $\Delta\nu/R < 0.01$ to have a sensitivity degradation of less than 1 dB. Since commercial DFB lasers have linewidths $\Delta\nu$ in the range 10 to 50 MHz, this imposes bit rates $R > 1$ GHz and, as a consequence, IF bandwidths of the same order.

V. FSK PERFORMANCE EVALUATION

Garrett and Jacobsen use a model in [4] and [11] to evaluate the performance of both wide- and narrow-deviation FSK (WD-FSK and ND-FSK, respectively). As a result, curves of the BER versus the variance of the frequency noise at the IF filter output (at the detection filter output in the case of WD-FSK) are given. Considering the white component only of the

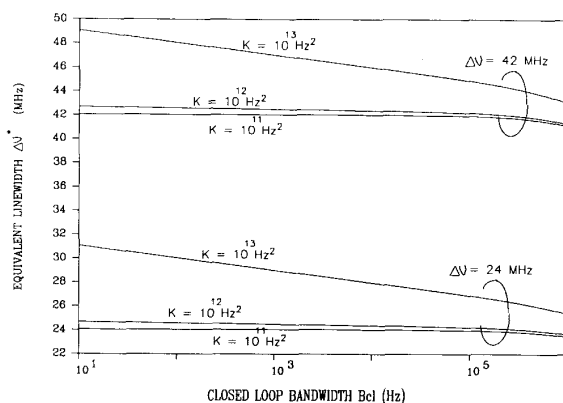


Fig. 8. Equivalent linewidth for Garrett and Jacobsen data. Ideal integrator loop filter, IF bandwidth 140 MHz.

laser frequency noise, this variance is shown to be

$$\sigma_{GJ}^2 = \frac{\Delta\nu^* B_{IF}}{2\pi} \quad (33)$$

where $\Delta\nu^*$ is the IF linewidth. In the WD-FSK case B_{IF} must be interpreted as the bandwidth B_D of the bandpass filters (BPF) in Fig. 2(c).

Equation (26) gives the same variance in the presence of an AFC. By equating (26) and (33) and solving for $\Delta\nu^*$, an equivalent linewidth is found which can be used, instead of the true linewidth $\Delta\nu$, to find the performance in the presence of an AFC by using the curves available in [4], [11].

The equivalent linewidth is

$$\Delta\nu^* = \Delta\nu + \frac{2\pi}{B_{IF}} \left[K \ln\left(\frac{B_{IF}/2}{B_{cl}}\right) - \frac{\Delta\nu B_{cl}}{2} \right] \quad (34)$$

A plot of $\Delta\nu^*$ versus the closed loop bandwidth is given in Fig. 8.

Using $\Delta\nu$ instead of $\Delta\nu^*$ when designing an FSK receiver may result in a penalty due the presence of the AFC. For example, in WD-FSK systems the main parameter to be optimized to minimize the effects of the nonzero linewidth is the IF bandwidth expansion. In the final design, an optimum expansion factor must be chosen for an "effective" linewidth $\Delta\nu^*$ instead of the smaller $\Delta\nu$. The penalty that results if the bandwidth expansion is not chosen in this way, can be seen from Fig. 9. Here the same curves as in Fig. 10 of [4] are given for a WD-FSK receiver.

In the case of an IF linewidth $\Delta\nu = 24$ MHz, ($\Delta\nu/R = 0.17$ at $R = 140$ Mb/s), the design for the minimum penalty, neglecting the flicker noise, would suggest an IF bandwidth of 237 MHz. If the actual flicker coefficient is $K = 4 \cdot 10^{12}$ Hz², and the receiver has an AFC of bandwidth $B_{cl} = 50$ Hz, the chosen bandwidth would give a penalty of 2.5 dB with respect to the true minimum. This can be reached choosing a 267 MHz bandwidth, which is the optimum value found using the effective linewidth $\Delta\nu^*/R = 0.19$ evaluated from (34).

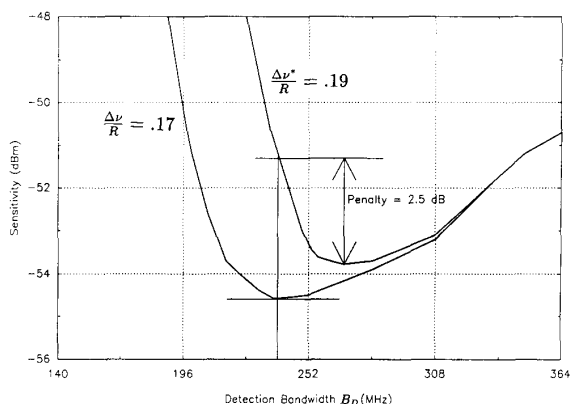


Fig. 9. WD-FSK receiver sensitivity (for BER = 10^{-9}) as a function of detection bandwidth for $R = 140$ Mb/s [after Fig. 10 in [4]]. The penalty incurred when the AFC presence is neglected in the design is pointed out.

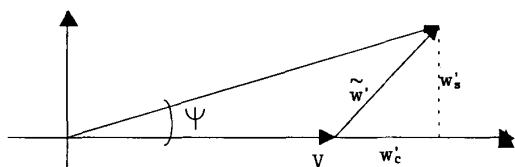


Fig. 10. Complex plane diagram of (13).

VI. CONCLUSIONS

An analysis of the automatic frequency control in coherent heterodyne optical systems has been given. In these systems, the tracking of the laser frequency noise is disturbed by shot noise AM-to-FM conversion in the frequency discriminator. Neglecting this disturb, a common analysis of the loop can be done. The approximation is justified by the use of commercial semiconductor DFB lasers, which feature linewidths between 1 and 50 MHz, and flicker factor K in the range 10^{11} to 10^{13} Hz². If the linewidth $\Delta\nu$ is decreased in the KHz range, then the frequency noise contribution due to the shot noise—and to the electronic noise—in the AFC loop becomes more important, especially at low signal to noise ratios. In this case a simulation of the loop can be done to check the various spectra in the loop.

In heterodyne schemes, the intermediate frequency is often chosen higher than 1 GHz. Thus classical frequency discriminators, such as the Foster-Seeley, cannot be used any more. For our purposes, a simple non-linearity giving FM-to-AM conversion can do the job; to this end, a delay line discriminator is usually employed.

After discrimination and modulation removal, the processing in the AFC loop is done at very low frequency and the design is thus a simple matter. However, care must be taken to allow sufficient IF bandwidth expansion to accommodate the residual "vibration" of the IF spectrum. This fact can be of some importance for IF bandwidths narrower than say 100 MHz, where the residual rms of some MHz is not negligible. With presently available semiconductor lasers, the

DPSK modulation must be operated at bit rates exceeding 1 GHz, with IF bandwidth of the same order, and the AFC design is not a problem. With WD-FSK modulation, where lower rates and thus lower IF bandwidths are possible, we have shown the penalty that can be incurred if a nonoptimal bandwidth expansion is chosen. Similar results can be obtained for the ND-FSK case, again taking advantage of the method of Garrett and Jacobsen, using the results given in [11] and our formula (34).

APPENDIX A

In this appendix the term $S_n(f)$ and its effects on $S_e(f)$ will be considered. In Fig. 10 a diagram of (13) on the complex plane is given. When the IF SNR is high, $V \gg |\dot{w}'(t)|$ and thus (16) can be approximated as

$$n(t) \cong \frac{1}{2\pi V} \dot{w}'_s(t). \quad (\text{A.1})$$

where w'_s is the quadrature component of $\dot{w}(t)$. This means $\psi(t)$ is almost everywhere small ($\ll 1$ rad) and the FM click rate is negligible. Taking the power spectrum of (A.1) yields

$$S_n(f) \cong \frac{f^2}{V^2} S_{w'_s}(f) = \frac{f^2}{(\text{SNR})B_{\text{IF}}} \frac{S_{w'_s}(f)}{S_{\dot{w}}(0)}. \quad (\text{A.2})$$

The second equality follows by making use of (5) and (6). The spectrum of $\dot{w}'(t)$ in (12) will now be considered. Two simplifying assumptions will be made and then relaxed. Assuming first that $\dot{w}(t)$ and $\alpha(t)$ are independent processes, the spectrum of $\dot{w}'(t)$ is the convolution of two terms

$$S_{\dot{w}'}(f) = S_{\dot{w}}(f) \otimes \Lambda(f) \quad (\text{A.3})$$

where $S_{\dot{w}}(f)$ is the spectrum of $\dot{w}(t)$, in this case a real gate of height $4qRPLZ_{\text{IF}}^2$, and $\Lambda(f)$ is the spectrum of the process

$$\lambda(t) \triangleq e^{-j\alpha(t)} = e^{-j2\pi \int_0^t e(u) du}. \quad (\text{A.4})$$

This is recognized as a frequency modulation. If, as a second assumption, $e(t)$ is zero-mean normal, the above spectrum is [10]

$$\Lambda(f) = \mathcal{F}[E[\lambda(\tau)]] = \mathcal{F}\left[e^{-\frac{\text{var}[\alpha(\tau)]}{2}}\right] \quad (\text{A.5})$$

where \mathcal{F} indicates the Fourier transform. Since $\alpha(t)$ is the convolution of $e(t)$ with a gate τ seconds long, its variance is given in (30), with T replaced by τ .

When the term $C_1(f)S_{n_{\text{IF}}}(f)$ dominates in $S_e(f)$, the two above assumptions hold, and the shape of $\Lambda(f)$ can be obtained analytically. Fig. 11(a) shows $\Lambda(f)$ using an ideal integrator as the loop filter (23), for various values of the closed-loop bandwidth B_{cl} and a low value of the linewidth $\Delta\nu = 7$ MHz. Fig. 11(b) is for a large $\Delta\nu$ value.

A Lorentzian of bandwidth $\Delta\nu/2$, corresponding to the case $S_e(f) = \frac{\Delta\nu}{\pi}$, is also shown in Fig. 11 as a reference. Since $\Lambda(f - f_{\text{IF}})$ is the spectrum of the IF signal when no modulation is applied, this means that, with the AFC locked, on a spectrum analyzer the IF spectrum will appear completely fixed when $\Delta\nu$ is as large as 100 MHz, while it will move from scan to scan when $\Delta\nu$ is less than 10 MHz. The actual drift

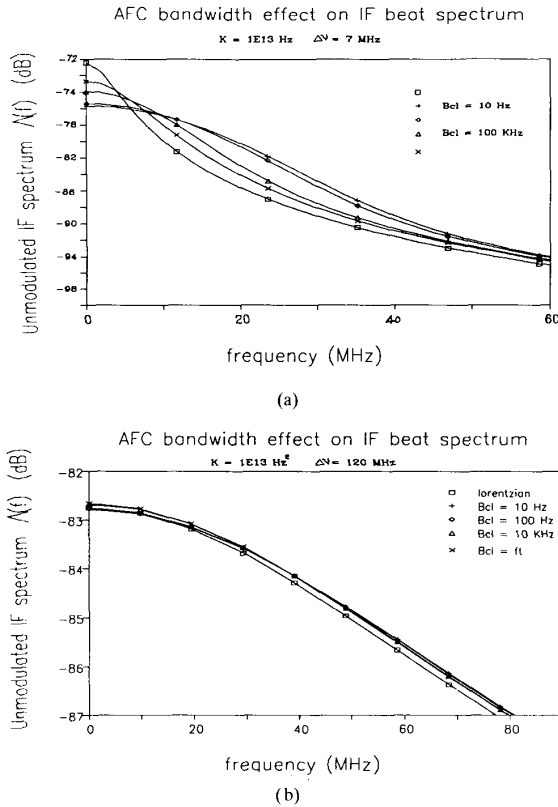


Fig. 11. Unmodulated IF beat spectrum with AFC locked for various values of the closed loop bandwidth. For large linewidths the spread is minimum.

will be in both cases of \pm some megahertz. An explanation of this fact is given in (28).

When the term $C_2(f)S_n(f)$ is not negligible in $S_e(f)$, the spectrum $S_n(f)$ depends on itself through (A.2), (A.3), and (A.4). Although an analytical evaluation is not possible, still some considerations can be made. Assuming for simplicity that $w'_c(t)$ and $w'_s(t)$ are uncorrelated, then

$$S_{w'_s}(f) = \frac{1}{2} S_{\tilde{w}'(f)} \quad (\text{A.6})$$

and therefore

$$S_n(f) = \frac{f^2}{2(\text{SNR})B_{\text{IF}}} \left(\left[\frac{S_{\tilde{w}'(f)}}{S_{\tilde{w}'(0)}} \right] \otimes \Lambda(f) \right). \quad (\text{A.7})$$

If an ideal integrator is used as the loop filter, then the loop equation (18) with the coefficients (21) can be rewritten by factoring the f^2 term in $S_n(f)$

$$S_e(f) = \frac{f^2}{B_{\text{cl}}^2 + f^2} \left[S_{n_{\text{IF}}}(f) + \frac{B_{\text{cl}}^2}{2(\text{SNR})B_{\text{IF}}} \left(\left[\frac{S_{\tilde{w}'(f)}}{S_{\tilde{w}'(0)}} \right] \otimes \Lambda(f) \right) \right]. \quad (\text{A.8})$$

$\Lambda(f)$ integrates to one, as $\int_{-\infty}^{\infty} \Lambda(f) df = E[\lambda(0)] = 1$. If $\Lambda(f)$ is far narrower than the gate $S_{\tilde{w}'(f)}$, then most of its energy is inside the band of the gate and its convolution

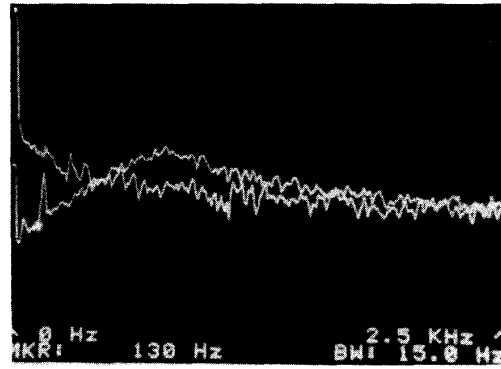


Fig. 12. Frequency noise with AFC locked and unlocked. Second-order loop filter.

with the gate will have a flat part of height $S_{\tilde{w}'(0)}$ at low frequencies, and will then fall down at higher frequencies. Otherwise, even the top level of the convolution will be less than $S_{\tilde{w}'(0)}$. It is then enough to check the two-sided white level of $S_{n_{\text{IF}}}$, $\Delta\nu/2\pi$, against the top level of the shot component, i.e., $B_{\text{cl}}^2/(2\text{SNR}B_{\text{IF}})$. When

$$\frac{\Delta\nu}{\pi} \gg \frac{B_{\text{cl}}^2}{(\text{SNR})B_{\text{IF}}} \quad (\text{A.9})$$

the contribution of the shot noise can be neglected. For small values of the closed loop bandwidth and typical values of the IF linewidth for commercial semiconductor lasers, condition (A.9) is always met. If $w'_c(t)$ and $w'_s(t)$ are correlated, a heavier notation must be used, but the quantitative results are still valid.

If the high SNR approximation does not hold, a white term arises in $S_n(f)$, due to the FM clicks, that tends to dominate.

It is worth mentioning that the use of higher order filters in the loop can lead to a fairly different behavior. For instance, the use of a second-order loop filter can create a resonance in the coefficients $C_1(f)$ and $C_2(f)$. This causes the closed-loop frequency error $S_e(f) \simeq C_1(f)S_{n_{\text{IF}}}(f)$ to dominate over the open loop $S_{n_{\text{IF}}}(f)$ in the frequency range corresponding to the resonance. This accounts for the experimentally observed hump in $S_e(f)$ in the AFC loop of the FSK system reported in [14] and shown in Fig. 12. The same hump can be observed in [12].

APPENDIX B

In this appendix it is shown how the modulation is removed in the various schemes under analysis. The similarities with the unmodulated case treated in Appendix A will be stressed in each case.

A. DPSK

Refer to Fig. 2(a). In the case of binary DPSK (4) can be written as

$$s(t) = Vm(t) \cos[2\pi f_{\text{IF}}t + \alpha(t)] + w(t) \quad (\text{B.1})$$

where $m(t) = \sum_i a_i p(t - iT)$, $a_i \in \{1, -1\}$, $p(t)$ is a NRZ impulse T seconds long and

$$\alpha(t) = 2\pi \int_0^t e(u) du. \quad (\text{B.2})$$

Expanding $w(t)$ in its Rice components yields

$$s(t) = [m(t)V + w'_c(t)] \cos[2\pi f_{\text{IF}}t + \alpha(t)] - w'_s(t) \sin[2\pi f_{\text{IF}}t + \alpha(t)]. \quad (\text{B.3})$$

In the DPSK receiver, the modulation is removed by cascading a square-law device and a bandpass filter centered at $2f_{\text{IF}}$.

The signal $x(t)$ at the filter output (see Fig. 2(a)) is

$$x(t) = E(t) \cos[2(2\pi f_{\text{IF}}t + \alpha(t)) - \psi(t)] \quad (\text{B.4})$$

where

$$\psi(t) = \arctan \left[\frac{2(m(t)V + w'_c(t))w'_s(t)}{(m(t)V + w'_c(t))^2 - w'_s(t)^2} \right]. \quad (\text{B.5})$$

The limiter-discriminator yields a signal $y(t)$ proportional to the offset of the instantaneous frequency of (B.4) from $2f_{\text{IF}}$

$$y(t) = 2K_d[e(t) - n(t)] \quad (\text{B.6})$$

where the shot noise contribution is defined here as

$$n(t) \triangleq \frac{\dot{\psi}(t)}{4\pi}. \quad (\text{B.7})$$

In high IF SNR conditions, it is $V \gg w'_c(t)$, $w'_s(t)$, $\dot{w}'_c(t)$, $\dot{w}'_s(t)$ for most of the time; thus from (B.5)

$$n(t) \cong \frac{1}{2\pi V} (m(t)\dot{w}'_s(t) + w'_s\dot{m}(t)). \quad (\text{B.8})$$

The power spectrum of (B.8) is found to be

$$S_n(f) = \frac{f^2}{V^2} (S_{w'_s}(f) \otimes S_m(f)) \quad (\text{B.9})$$

where $S_m(f)$ is the spectrum of the modulating PAM signal $m(t)$. Using (5) and (6) as before, and (A.3) and (A.6), gives

$$S_n(f) = \frac{f^2}{2(\text{SNR})B_{\text{IF}}} \left\{ \left[\frac{S_{\tilde{w}}(f)}{S_{\tilde{w}}(0)} \right] \otimes \Lambda(f) \otimes S_m(f) \right\}. \quad (\text{B.10})$$

This is the same as in (A.7) except for the convolution with $S_m(f)$, which has little effect on the low-frequency part of $S_n(f)$. For DPSK we use $\Delta\nu/R \ll 1$ and an IF bandwidth equal to the bit rate R . Thus $\Lambda(f)$ is far narrower than the gate $S_{\tilde{w}}(f)$ and their convolution is very close to $S_{\tilde{w}}(f)$ itself. A plot of (B.10) is given in Fig. 13(a) and (b) in linear and log-log scale.

We note that the low-frequency part of the curve can be approximated by a parabola of equation

$$S_n(f) \cong 0.77 \frac{f^2}{2(\text{SNR})B_{\text{IF}}} \quad |f| < B_{\text{IF}}/2. \quad (\text{B.11})$$

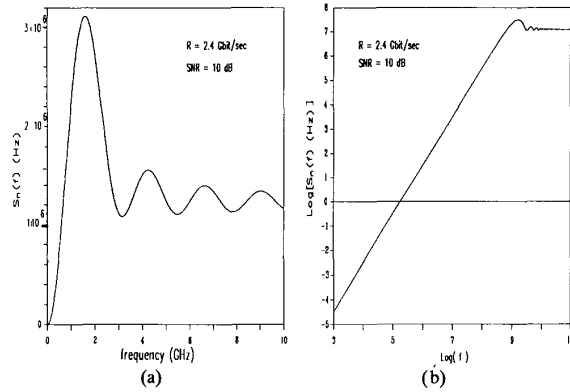


Fig. 13. DPSK: Spectral density of the shot noise contribution to the frequency noise $n(t)$ in both linear (a) and log-log scale (b).

The floor for $f \rightarrow \infty$ is due to the NRZ assumption. Actually, the $m(t)$ NRZ signal is IF filtered. It is easy to verify that, considering this filtering, the floor disappears.

B. Narrow-Deviation FSK

The scheme is shown in Fig. 2(b). Here the modulation is swept off by the loop filter itself.

In this case (4) gives for the IF signal

$$s(t) = V \cos\{2\pi f_{\text{IF}}t + \alpha(t)\} + w(t) \quad (\text{B.12})$$

where

$$\begin{cases} a(t) = \alpha_e(t) + \alpha_m(t) \\ \alpha_e(t) \triangleq 2\pi \int_0^t e(u) du \\ \alpha_m(t) \triangleq 2\pi \int_0^t m(u) du \end{cases}$$

and the PAM modulation is

$$m(t) = (m_0 R/2) \sum_i a_i p(t - iT), \quad a_i \in \{1, -1\}.$$

$p(t)$ is a NRZ impulse T seconds long; m_0 is the modulation index, $R = 1/T$ is the bit rate. Equation (B.12) can be recast as in (11), using (13). The output of the limiter-discriminator is then

$$y(t) = K_d[e(t) + m(t) + n(t)] \quad (\text{B.13})$$

where $n(t)$ has the same form as in (16). In high SNR conditions the same formula (A.2) is obtained. The only difference is in the form of $S_{\tilde{w}}(f)$. Equation (12) can be written here as

$$\tilde{w}'(t) = \tilde{w}(t)e^{-j\alpha_e(t)}e^{-j\alpha_m(t)}. \quad (\text{B.14})$$

If the spectrum of $m(t)$ is outside the AFC bandwidth, $e(t)$ is little affected by $m(t)$ —which means no pattern dependency on $e(t)$ is present—and $\alpha_e(t)$ can be assumed independent of $\alpha_m(t)$. The spectrum of $\tilde{w}'(t)$ then becomes

$$S_{\tilde{w}'}(f) = S_{\tilde{w}}(f) \otimes \Lambda_e(f) \otimes \Lambda_m(f) \quad (\text{B.15})$$

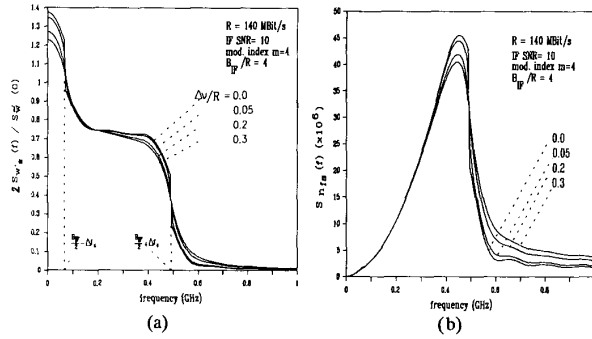


Fig. 14. ND-FSK modulation (a) Normalized w'_s spectrum for various values of $\Delta\nu/R$. (b) Spectrum of $n(t)$ for various values of $\Delta\nu/R$.

where $\Lambda_e(f)$ is the spectrum of $\lambda_e(t) = \exp(-j\alpha_e(t))$, the same as $\Lambda(f)$ in Appendix A, and $\Lambda_m(f)$ is the spectrum of $\lambda_m(t) = \exp(-j\alpha_m(t))$. Using an NRZ PAM signal, with integer and even modulation index m_0 , it can be shown that [10]

$$E[\lambda_m(\tau)] = \begin{cases} \left(1 - \frac{|\tau|}{2T}\right) \cos(\pi m_0 R \tau) + \frac{\sin(\pi m_0 R |\tau|)}{2\pi m_0} & |\tau| < T \\ \frac{1}{2} \cos(\pi m_0 R \tau) & |\tau| > T \end{cases} \quad (\text{B.16})$$

For $m_0 \geq 4$ the $\sin(\cdot)$ term in (B.16) can be neglected and the Fourier transform of (B.16) simplifies to

$$\Lambda_m(f) = \frac{T}{2} \left\{ \text{Sa}^2 \left[\pi \left(f + \frac{m_0 R}{2} \right) T \right] + \text{Sa}^2 \left[\pi \left(f - \frac{m_0 R}{2} \right) T \right] \right\} + \frac{1}{4} \left[\delta \left(f + \frac{m_0 R}{2} \right) + \delta \left(f - \frac{m_0 R}{2} \right) \right] \quad (\text{B.17})$$

where $\delta(\cdot)$ is the Delta function and $\text{Sa}(x) \triangleq \sin(x)/x$.

Again assuming $\lambda_e(\tau)$ is real, then $\Lambda_e(f)$ is even and $S_{w'_s}(f) = \frac{1}{2} S_{\bar{w}'_s}(f)$. A plot of $S_{w'_s}(f)/S_{\bar{w}'_s}(0)$ is given in Fig. 14(a), where $e(t)$ is evaluated neglecting $n(t)$. A plot of $S_n(f)$ as in (A.2) is given in Fig. 14(b). For low frequencies, this curve is approximated by a parabola

$$S_n(f) \cong \frac{f^2}{(\text{SNR})B_{\text{IF}}} \frac{S_{w'_s}(0)}{S_{\bar{w}'_s}(0)} \quad |f| < B_{\text{IF}}/2 \quad (\text{B.18})$$

where from Fig. 14(a) it is seen that the ratio $[S_{w'_s}(0)/S_{\bar{w}'_s}(0)]$ is of the order of 1 for all values of $\Delta\nu/R$. Therefore for low frequencies $S_n(f)$ in (B.18) is essentially the same as in (A.2). Again the floor is due to the NRZ assumption in $m(t)$.

C. Wide-Deviation FSK

The double-branch envelope detected wide-deviation FSK scheme is shown in Fig. 2(c). To drive the AFC loop, the IF signals s_1 and s_2 out of the two detection bandpass filters (BPF) of bandwidth B_D are sent to two limiter-discriminators centered around the two nominal heterodyned frequencies f_1 and f_2 , and the two output signals y_1 and y_2 are then summed. The crossed boxes in the picture indicate squelch

circuits. Thus, when frequency f_1 is sent, an error signal from discriminator 1 and—ideally—zero signal from the other channel are summed. The reverse is true when f_2 is sent.

The signal $s(t)$ out of the front-end (F.E.) is the same as in (B.12). Neglecting the crosstalk between the two branches, the two BPF filtered signals can be expressed as

$$\begin{cases} s_1(t) = \text{ook}(t) V_1 \cos\{2\pi f_1 t + \alpha(t)\} + w_1(t) \\ s_2(t) = \overline{\text{ook}(t)} V_2 \cos\{2\pi f_2 t + \alpha(t)\} + w_2(t) \end{cases} \quad (\text{B.19})$$

where

$$\alpha(t) = 2\pi \int_0^t e(u) du$$

w_1 and w_2 are independent Gaussian processes, white on the BPF bandwidth B_D ; $\text{ook}(t) = \sum_i a_i p(t - iT)$; $a_i \in \{1, 0\}$; $p(t)$ is a NRZ pulse T -seconds long filtered by the lowpass equivalent of the BPF; $\overline{\text{ook}(t)}$ is the filtered version of the boolean complement of the NRZ stream in the other channel.

$s_1(t)$ and $s_2(t)$ are ON-OFF signals plus noise. Splitting $w_1(t)$ and $w_2(t)$ in their Rice components $w_{c_1}, w_{s_1}, w_{c_2}, w_{s_2}$ with respect to the whole argument of the cosine in (B.19), the expression for $s_1(t)$ and $s_2(t)$ during their ON bit can be compacted as

$$\begin{cases} s_i(t) = E_i(t) \cos[2\pi f_i t + \alpha(t) - \psi_i(t)] & i = 1, 2 \\ \psi_i(t) = \arctan \left[\frac{w'_{s_i}(t)}{(V_i + w'_{c_i}(t))} \right] \end{cases} \quad (\text{B.20})$$

The signal $y(t)$ out of the discriminator can be written

$$y(t) = K_d[e(t) - n(t)] \quad (\text{B.21})$$

where

$$\begin{cases} n(t) \triangleq \frac{\dot{\psi}(t)}{2\pi} \\ \dot{\psi}(t) \triangleq \begin{cases} \dot{\psi}_1(t) & \text{when } a_i = 1 \\ \dot{\psi}_2(t) & \text{when } a_i = 0. \end{cases} \end{cases} \quad (\text{B.22})$$

In high SNR conditions, when signal is present in the two IF bandpass filters, the AM-to-FM conversion of the shot noise can be approximated as

$$n(t) \cong \begin{cases} \frac{1}{2\pi V_1} \dot{w}_{s_1}(t) & \text{when } a_i = 1 \\ \frac{1}{2\pi V_2} \dot{w}_{s_2}(t) & \text{when } a_i = 0. \end{cases}$$

If the two demodulation branches are identical, then $V_1 = V_2 = V$; also w'_{s_1}, w'_{s_2} and w'_{c_1}, w'_{c_2} have the same statistics. This further approximation can be made in this case

$$n(t) \triangleq \frac{1}{2\pi V} \dot{w}'_s(t) \quad (\text{B.23})$$

where $w'_s(t)$ is an equivalent process, with same statistics as $w'_{s_1}(t)$ and $w'_{s_2}(t)$.

Expression (B.23) is the same as (A.1). Therefore (A.2) applies here too. B_{IF} is now the demodulation bandwidth B_D .

ACKNOWLEDGMENT

Special thanks to P. Ciammaichella and F. Forghieri for their early work on the AFC for the DPSK modulation.

REFERENCES

- [1] I. Garrett and G. Jacobsen, "Statistics of laser frequency fluctuations in coherent optical receivers," *Electron. Lett.*, p. 168, Jan. 1986.
- [2] Y. Yamamoto, "AM and FM quantum noise in semiconductor lasers, Parts I and II," *IEEE J. Quantum Electron.*, vol. QE-19, no. 1, pp. 34-57, Jan. 1983.
- [3] B. Daino *et al.*, "Phase noise and spectral line shape in semiconductor lasers," *IEEE J. Quantum Electron.*, vol. QE-19, no. 3, pp. 266-270, Mar. 1983.
- [4] I. Garrett and G. Jacobsen, "Theoretical analysis of heterodyne optical receivers for transmission systems using (semiconductor) lasers with nonnegligible linewidth," *J. Lightwave Technol.*, vol. LT-4, no. 3, pp. 323-334, Mar. 1986.
- [5] I. Garrett and G. Jacobsen, "The effect of laser linewidth on coherent optical receivers with nonsynchronous demodulation," *J. Lightwave Technol.*, vol. LT-5, no. 4, pp. 551-560, Apr. 1987.
- [6] I. Garrett, G. Jacobsen, E. Bødtker, and R. J. S. Pedersen, "Filtered laser beat-frequency fluctuations," *Proc. IEEE*, vol. 135, pp. 408-412, Dec. 1988.
- [7] E. Bedrosian and S. O. Rice, "Distortion and crosstalk of linearly filtered, angle-modulated signals," *Proc. IEEE*, vol. 56, pp. 2-13, Jan. 1968.
- [8] K. Kikuchi, "Impact of $1/f$ -type FM noise on coherent optical communications," *Electron. Lett.*, vol. 23, no. 17 Aug. 1987.
- [9] K. Kikuchi *et al.*, "Degradation of bit-error rate in coherent optical communications due to spectral spread of the transmitter and the local oscillator," *J. Lightwave Technol.*, vol. LT-2, no. 6, p. 1027, Dec. 1984.
- [10] A. Papoulis, *Probability, Random Variables and Stochastic Processes*. New York: McGraw-Hill, 1984, second ed., pp. 321-329.
- [11] I. Garrett and G. Jacobsen, "Theory for optical heterodyne narrow-deviation FSK receivers with delay demodulation," *J. Lightwave Technol.*, vol. 6, no. 9, pp. 1415-1423, Sept. 1988.
- [12] F. G. Walther, S. D. Lowney, and J. E. Kaufmann, "FC3 frequency tracking for heterodyne optical communications using semiconductor lasers," in *Proc. CLEO '83*.
- [13] S. D. Lowney and D. V. L. Marquis, "Frequency acquisition and tracking for optical heterodyne communication systems," *J. Lightwave Technol.*, vol. LT-5, no. 4, pp. 538-550, Apr. 1987.
- [14] A. Bononi, P. Ghiggino, I. Blanchflower, and M. Scoones, "A 565 Mbit/sec FSK heterodyne system using commercial DFB lasers," in *Proc. IEE Electronics Division Colloquium on Coherent Opt. Commun.*, (London), June 1990.



Alberto Bononi was born in La Spezia, Italy on December 7, 1963. He received the "Laurea in Ingegneria Elettronica," cum laude, from the University of Pisa, Pisa, Italy, in 1988. Since September 1990 he has been pursuing the Ph.D. degree (course) at Princeton University, Princeton, NJ.

In 1989 he was with the University of Parma doing research on optical coherent communications. In 1990 he worked at GEC-Hirst Research Laboratory in Wembley, UK, on a Marconi S.p.A. Project on coherent FSK systems.



Pierpaolo Ghiggino (M'88) was born in Alessandria, Italy on December 31, 1959. He received the "Laurea in Ingegneria Elettronica," cum laude, from the University of Pavia, Pavia, Italy, in 1984.

In 1985 he joined the Research and Development Laboratories of Marconi S.p.A. in Genoa, Italy, where he worked on fiber-optic digital transmission systems. After a short period, he was seconded to GEC-Marconi Hirst Research Centre in Wembley, England, where he worked on advanced optical sources and systems and especially on coherent optical systems. Presently he is still based at Hirst Research Centre, coordinating development and studies for Marconi S.p.A. on coherent optical systems, optical amplification, fiber-optic nonlinear propagation as well as on LAN and DQDB-based MAN implementation and inter-networking. Dr. Ghiggino holds one patent and is author or co-author of several technical papers.



Giorgio Picchi received the "Laurea in Ingegneria Elettronica" from the University of Pisa, Pisa, Italy, in 1974.

From 1975 to 1990 he was Research Associate and then Associate Professor of Electrical Engineering at the University of Pisa. In 1988 he was Visiting Fellow in the Department of Electrical Engineering, Princeton University, Princeton, NJ. Since November 1990 he has been Professor of Telecommunications Engineering at the University of Parma, Parma, Italy. His research interests are in digital signal processing for communications, and optical communication systems. He is corecipient of the 1987 IEEE "W. Bennett" Prize Paper Award.