N87-17952

Filter Distortion Effects on Telemetry Signal-to-Noise Ratio

R. Sadr and W. Hurd
Communications Systems Research Section

The effect of filtering on the Signal-to-Noise Ratio (SNR) of a coherently demodulated band-limited signal is determined in the presence of worst-case amplitude ripple. The problem is formulated mathematically as an optimization problem in the L_2-Hilbert space. The form of the worst-case amplitude ripple is specified, and the degradation in the SNR is derived in a closed form expression. It is shown that when the maximum passband amplitude ripple is 2Δ (peak-to-peak), the SNR is degraded by at most $(1 - \Delta^2)$, even when the ripple is unknown or uncompensated. For example, an SNR loss of less than 0.01 dB due to amplitude ripple can be assured by keeping the amplitude ripple to under 0.42 dB.

I. Introduction

Amplitude ripple is inherent in most physically realizable filters. We seek to determine the system performance degradation resulting from amplitude ripple. The Signal-To-Noise Ratio (SNR) is used as the performance criterion. Our results provide an easy method to determine the worst-case loss due to amplitude ripple. To derive the expression for the worst-case loss in the SNR, the class of worst-case amplitude ripple is explicitly found.

We consider filtering a signal $y(t)$ which is composed of a bandlimited signal $s(t)$, with bandwidth $(f_0 - W, f_0 + W)$ Hertz, added to a noise process $\{n(t)\}$, where f_0 denotes the center frequency and W is the half bandwidth of the signal. The passband of the filter covers the same band of frequency as the signal $s(t)$. This filter is shown in Fig. 1. Ideally, the transfer function for the filter is

$$H_0(j\omega) = e^{j\omega\tau}$$

for $|\omega| \epsilon (2\pi(f_0 - W), 2\pi(f_0 + W))$ where τ is a constant group delay. The shape of the filter response at frequencies outside of the signal band is shown to be immaterial.

The amplitude spectrum of the ideal filter, $H_0(j\omega)$, has constant gain and group delay in the passband frequency. This is never really true in practice where the gain of the filter exhibits a bounded ripple in the passband. The amplitude ripple is the deviation of $|H_0(j\omega)|$ from ideal and is denoted by $\Delta(j\omega)$. Characteristics of both an ideal and a nonideal bandpass filter are depicted in Fig. 1. The transfer function of the nonideal filter in the passband is represented as

$$H_0(j\omega) = (1 + \Delta(j\omega))\, e^{j(\phi(\omega)+\omega\tau)}, \qquad (1)$$

for $|\omega| \epsilon (2\pi(f_0 - W), 2\pi(f_0 + W))$ where $\phi(j\omega)$ represents the deviation in phase from constant group delay, τ.

A critical issue in the specification of the filter $h(t)$, for the design engineer, is to determine the impact of $\Delta(j\omega)$, and

$\phi(j\omega)$ on the SNR. The degradation of SNR for phase deviation $\phi(j\omega)$ has been studied previously. J. Jones (Ref. 1) in 1972 analyzed the filter distortion effects of the phase nonlinearity for BPSK and QPSK, and has shown that when phase deviation $\phi(j\omega)$ is bounded by ϕ_{max} in absolute value, the SNR is degraded by at most a factor of $\cos^2\phi_{max}$.

In this article, it is shown that if the amplitude ripple is bounded by Δ (i.e., $|\Delta(j\omega)| \leqslant \Delta$), the SNR in the presence of the worst case amplitude ripple waveform is degraded by at most $(1 - \Delta^2)$. This result holds even when the amplitude ripple is unknown, or known but not compensated.

To summarize the outline of the rest of this article, Section II describes the system under study and the underlying assumptions for which the SNR figure is analyzed. In Section III, the closed form SNR expression is derived for a coherently demodulated signal which is filtered by the nonideal transfer function characteristic defined in Eq. (1). In Section IV, the class and properties of the worst-case amplitude ripple are specified. Finally, in Section V, we make some concluding remarks based on our results.

II. Formulation

We consider a received waveform containing signal and noise, that is, $y(t) = s(t) + n(t)$. The signal amplitude spectrum $S(j\omega)$ is band-limited to $|f| \epsilon [f_0 + W, f_0 - W]$, and its waveform is completely known during each $t \epsilon [0,T]$. The noise process $\{n(t)\}$ in our analysis is assumed to be an Additive White Gaussian Noise (AWGN) process with single-sided spectral density N_0 W/Hz. The results generalize to the case where the noise is not white. The only restriction is that $\{n(t)\}$be a wide-sense stationary process. This implies that it has zero mean and autocorrelation function $R_n(\tau) = E[n(t)\, n(t + \tau)]$, where $E[\cdot]$ denotes the expectation operator.

The optimal receiver for the observed signal $y(t)$, which maximizes the SNR, is a matched filter (Ref. 2). This solution is expressed in the form of the Fredholm integral equation of the first kind. There are known methods to solve this integral equation explicitly to find the optimal matched filter solution $h_{MA}(t)$.

For an AWGN channel the matched filter solution is $h_{MA}(t) = s(T - t)$ or, equivalently, in the frequency domain it is $H_{MA}(j\omega) = S^*(j\omega)\, e^{-j\omega T}$. (Throughout the article, superscript * denotes complex conjugate while a midline $*$ denotes convolution).

For the case in which the noise is only wide sense stationary (not necessarily AWGN) with spectrum $S_n(j\omega)$, the matched filter solution may be expressed under certain assumptions (Ref. 2) as $H_{MA}(j\omega) = S^*(j\omega)\, e^{-j\omega T}/S_n(j\omega)$. This transfer function is recognized as the matched filter transfer function for the white noise case divided by the actual power density of the noise. Therefore, it is possible to generalize our result for wide sense stationary noise processes by simply using a matched filter which is matched to both the noise and the known signal. Thus, with no loss of generality, in our subsequent analysis we assume that the noise is white and Gaussian.

In digital communication systems, the signal $s(t)$ is modulated at the transmitter to a Radio Frequency (RF) by multiplying $s(t)$ by the carrier signal $\cos(\omega_0 t)$, where $2\pi\, \omega_0$ is the carrier frequency. At the receiver (Fig. 2), the observed signal is filtered and then demodulated by multiplying the observed signal by $2\cos(\omega_0 t)$. This signal is passed through a zonal low-pass filter to filter out the double frequency terms produced by the multiplication operation. The output of the low-pass filter is then fed to the matched filter.

The sampled output of the matched filter each T s is denoted by M_i.

The SNR is defined as the ratio of the square of the expected value to the variance of the random variable M_i. In the following section, a closed form expression for the SNR is derived. This expression is formulated in the form of a functional. We minimize this functional over the ensemble of all possible amplitude waveform ripples in the passband of the filter $h_{BF}(t)$, as shown in Fig. 1.

III. Signal-to-Noise Ratio Expression

Since the filtering processes are linear, we can consider the system response to signal and noise separately. We need to determine the mean value of the signal and the variance of the noise, both at the matched filter output. We denote the response of each stage of the system to the signal by $e_i(t)$, as shown in Fig. 2. The signal is represented by amplitude spectrum throughout the following analysis. Throughout this article, the square of a complex function is meant to be the magnitude square of that function.

Neglecting the noise response, the amplitude spectrum of $y(t)$, $Y(j\omega)$, is expressed as

$$Y(j\omega) = S(j(\omega - \omega_0)) + S(j(\omega + \omega_0)) \qquad (2)$$

We denote the bandpass filter $H'(j\omega)$ as

$$H'(j\omega) = H_{BF}(j(\omega - \omega_0)) + H^*_{BF}(j(-\omega + \omega_0))$$

where $H_{BF}(\cdot)$ is the transfer function of a complex low-pass filter. The spectrum of $e_1(t)$, the output of the bandpass filter under study (Fig. 2), is

$$E_1(j\omega) = H'(j\omega)\, Y(j\omega) \tag{3}$$

The demodulated waveform $e_2(t)$ has the spectrum containing the sum and difference frequencies,

$$\begin{aligned} E_2(j\omega) &= E_1(j(\omega - \omega_0)) + E_1(j(\omega + \omega_0)) \\ &= H'(j(\omega - \omega_0))\, Y(j(\omega - \omega_0)) \\ &\quad + H'(j(\omega + \omega_0))\, Y(j(\omega + \omega_0)) \end{aligned} \tag{4}$$

The demodulator is followed by an ideal low-pass filter, which filters out the double frequency terms, and the resulting output spectrum $E_3(j\omega)$ is $E_3(j\omega) = E_2(j\omega)\, H_{LP}(j\omega)$, where

$$H_{LP}(j\omega) = \begin{cases} 1, & |\omega| \leqslant 2\pi(f_0 - W) \\ 0, & \text{otherwise} \end{cases}$$

Thus, the output of the low-pass filter from Eq. (4) is

$$E_3(j\omega) = (H_{BF}^{*}(-j\omega) + H_{BF}(j\omega))\, S(j\omega) \tag{5}$$

Note that if the low-pass filter is not ideal, its deviation from ideal should be included in the filter under study.

Neglecting deviation in $H_{MA}(j\omega)$ from ideal, the matched filter that maximizes the SNR has the transfer function (Ref. 2) $h_{MA}(t) = s(T - t)$, or equivalently

$$H_{MA}(j\omega) = S^{*}(j\omega)\, e^{-j\omega T} \tag{6}$$

Let for simplicity $H(j\omega)$ denote $H_{BF}^{*}(-j\omega) + H_{BF}(j\omega)$. Then the matched filter output $e_4(t)$ can be represented as

$$E_4(j\omega) = E_3(j\omega)\, S^{*}(j\omega)\, e^{-j\omega T}$$

and substituting $E_3(j\omega)$ using Eq. (5), $e_4(t)$ may be expressed as

$$e_4(t) = \frac{1}{2\pi} \int_I H(j\omega)\, |S(j\omega)|^2\, e^{-j\omega(T-t)}\, d\omega \tag{7}$$

where $I = [-W, W]$. The output of the matched filter is sampled at the end of every time interval T. Thus at $t = T$ we have

$$e_4(T) = \frac{1}{2\pi} \int_I H(j\omega)\, |S(j\omega)|^2\, d\omega$$

The system noise response is denoted by $z(t)$. The random process $\{M_i\}$ (taking values in $\mathbb{R}^1$) is the sum of the filtered signal plus the filtered noise component. Hence, we can write

$$M_i = e_4(T) + z(T) \tag{8}$$

Taking expectation of Eq. (8), and noting that the noise is assumed to be zero mean, we get

$$E[M_i] = \frac{1}{2\pi} \int_I H(j\omega)\, |S(j\omega)|^2\, d\omega \tag{9}$$

To compute $\mathrm{Var}[M_i]$, let $z(t) = n(t) * x(t)$. Note that $z(t)$ is a filtered white noise process, which is filtered by the filter under study and the matched filter. The cascaded filter is denoted by $x(t)$.

Let $X(j\omega) = S(j\omega)\, H(j\omega)\, e^{-j\omega T}$, thus we have

$$\mathrm{Var}[M_i] = \mathrm{Var}[z(t)]\, \Big|_{t=T} \tag{10}$$

and

$$\begin{aligned} z(t) &= n(t) * x(t) \\ E[n(t)\, n(t+\tau)] &= N_2\, \delta(\tau) \end{aligned} \tag{11}$$

From Eq. (11), the variance of $z(t)$ can be expressed as

$$\mathrm{Var}[z(t)] = \frac{N_0}{4\pi} \int_I |X(j\omega)|^2\, d\omega \tag{12}$$

Combining Eqs. (12) and (10) and evaluating these expressions at $t = T$, results in

$$\mathrm{Var}[M_i] = \frac{N_0}{4\pi} \int_I |H(j\omega)\, S(j\omega)|^2\, d\omega \tag{13}$$

Therefore since $\mathrm{SNR} = (E[M_i])^2 / \mathrm{Var}[M_i]$ we have

$$\mathrm{SNR} = \frac{2}{N_0} \frac{\left(\int_I H(j\omega)\, S^2(j\omega)\, d\omega \right)^2}{\int_I |H(j\omega)\, S(j\omega)|^2\, d\omega} \tag{14}$$

The filter transfer function is $H(j\omega) = H_{BF}(j\omega) + H^{*}_{BF}(-j\omega)$ and furthermore $H(j\omega)$ is

$$H(j\omega) = (1 + \Delta(j\omega))\, e^{j\omega\tau} \;\forall\, |\omega| \leqslant 2\pi W$$

We assume that the group delay is negligible (i.e., $e^{j\omega\tau} \approx 1$), or it is compensated in the matched filter, hence, by substituting $H(j\omega)$ in the SNR expression (14), the SNR is

$$\text{SNR} = \frac{2}{N_0} \frac{\left| \int_I (1 + \Delta(j\omega))\, S^2(j\omega)\, d\omega \right|^2}{\int_I (1 + \Delta(j\omega))^2\, S^2(j\omega)\, d\omega} \tag{15}$$

IV. Worst Case Amplitude Ripple

With no loss of generality assume $I = [0,1]$ and let

$$\xi = \int_0^{\xi} S^2(j\omega)\, d\omega$$

Using Eq. (15) we can formulate the minimization problem

$$F(g) = \inf_{g(\xi) \in L_{2[0,1]}} \frac{\left(\int_0^1 (1 + g(\xi))\, d\xi \right)^2}{\int_0^1 (1 + g(\xi))^2\, d\xi} \tag{16}$$

subject to the constraint

$$|g(\xi)| < \Delta < 1 \tag{16a}$$

The integrals are understood in the Lebesgue sense.

In the appendix we prove the following theorem. In the proof, the class of the functions for which the minimum occurs is explicitly exhibited.

Theorem 1. There is a continuum of measurable step functions which minimize Eq. (16), and at the minimum $F_{MIN}(\tilde{g}) = 1 - \Delta^2$, and $\tilde{g}(\xi)$ is

$$\tilde{g}(\xi) = \begin{cases} \Delta, & \text{for } 0 < \xi \leqslant \frac{1-\Delta}{2} \\ -\Delta, & \text{for } \frac{1-\Delta}{2} < \xi < 1 \end{cases}$$

Using the result of Theorem 1, we can state the following corollary.

Corollary 1. Define interval I_1, as any subset of the interval I such that

$$\int_{I_1} S^2(j\omega)\, d\omega = \frac{1-\Delta}{2}$$

Then Eq. (15) is minimized by

$$\Delta(j\omega) = \begin{cases} \Delta, & \text{for } \omega \in I_1 \\ -\Delta, & \text{for } \omega \notin I_1 \end{cases}$$

Corollary 2. The minimum SNR for Eq. (15) is achieved by $\Delta(j\omega)$ of the form specified by corollary (1), and furthermore, the total SNR for the worst case ripple distortion is $(1 - \Delta^2)\, \text{SNR}_{\text{ideal}}$.

Corollaries (1) and (2) are direct consequences of Theorem 1. The point a, in Fig. 3, indicates the point at which the integrated power of the signal $s(t)$ is $(1 - \Delta)/2$.

In general, the SNR is minimized when the ripple is $+\Delta$ for frequencies containing $(1 - \Delta)/2$ of the energy of $s(t)$, and $-\Delta$ for frequencies containing $(1 + \Delta)/2$ of the energy of $s(t)$.

The shape of the amplitude ripple is not unique, and it is the whole continuum of step functions which satisfies the condition stated in corollary (1). To construct another amplitude ripple waveform which satisfies the conditions of corollary (1), one can take the waveform of Fig. 3 and move a segment from $[0,a]$ to $(a,1]$, and move an equal energy segment from $(a,1]$ to $[0,a]$. Conceptually, this method may be thought of as juggling equal energy line segments from each interval. This method results in obtaining a new step function from the basic function of Fig. 3 which satisfies the statement of corollary (1).

V. Discussion and Conclusion

We have shown that the worst-case loss in SNR due to amplitude ripple is $(1 - \Delta^2)$. This result can be used to specify filters for communication receivers, such as the Advanced Receiver for the NASA's Deep Space Network.

To express the maximum allowable ripple, for a given loss (L_{dB}) in dB, we expand the expression for dB of ripple and keep the first terms. Thus for small ripple and loss, the ripple loss in dB may be expressed as

$$\text{Ripple}_{\text{dB}}\,(L_{\text{dB}}) = 4.168\sqrt{L_{\text{dB}}}$$

Consider a simple example: If the system can tolerate a loss of 0.01 dB due to amplitude ripple, the maximum allowable ripple in dB would be 0.4168 dB. The loss is much less than the ripple and decreases as the square of the ripple.

From a mathematical standpoint, we solved the minimization problem stated in Eq. (16). And we showed that the optimal solution lies on the boundary, and it is a continuum of measurable step functions in the interval [0,1]. This is explicitly exhibited in the appendix.

Acknowledgment

Special thanks are due to Dr. Mehrdad Shahshahani for his help in proving theorem 1 and Dr. Eugene R. Rodemich for initially pointing out the solution.

References

1. Jones, J. J., "Filtered Distortion and Intersymbol Interference Effects on PSK Signals," *IEEE Trans. Comm. Tech., COM-19*, No. 2, April 1971: 121–132.
2. Van Trees, H. L., *Detection, Estimation, and Modulation Theory*. New York: John Wiley & Sons, Part I, 1968; Part II, 1971.

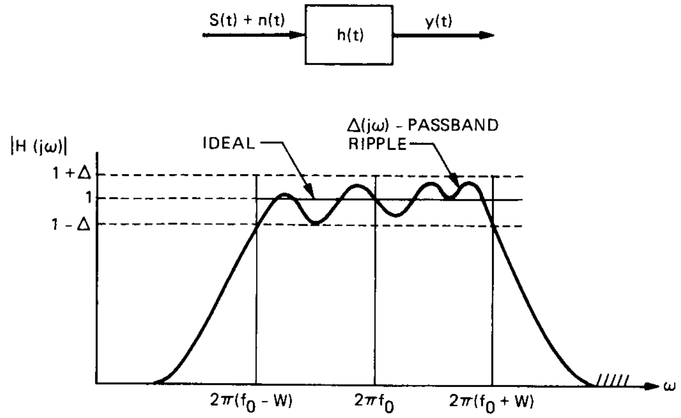

Fig. 1. Filtering

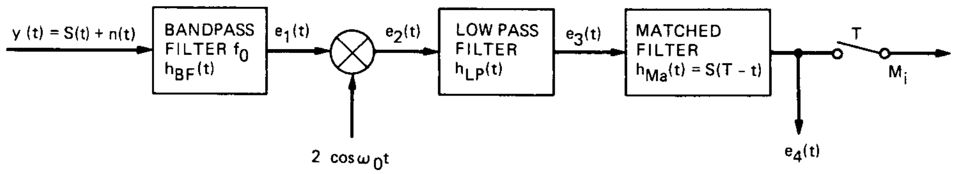

Fig. 2. Communication model

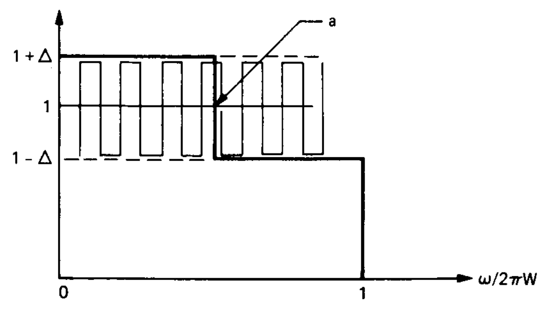

Fig. 3. Ripple waveform

Appendix

Let $0 < \Delta < 1$ and

$$C_\Delta = \{ f\colon [0,1] \to \mathbb{R} \mid f \text{ measurable}$$
$$\text{and } |f(z) - 1| < \Delta \text{ for almost all } u \}$$

We equip C_Δ with the induced metric from $L^1(I)$ $(I = [0,1])$ so that C_Δ becomes a complete metric space (we identify functions equal almost everywhere). Consider the continuous functional F on C_Δ defined by

$$F(f) = \frac{\left(\int_0^1 f(x)\,dx\right)^2}{\int_0^1 f^2(x)\,dx}$$

In the appendix we investigate the existence and nature of minima of F. Let

$$C_n = \{(a_0, \ldots, a_n; c_1, \ldots, c_n) \mid$$
$$0 = a_0 \leqslant a_1 \leqslant \ldots \leqslant a_n = 1, c_j \in [1 - \Delta, 1 + \Delta]\}$$

Then C_n is naturally identified with a compact subset of $[0,1]^{n-1} \times [1 - \Delta, 1 + \Delta]^n$. To every $\gamma \in C_n$ we assign the function $f_\gamma \in C_\gamma$ which takes value c_i on the open interval (a_{i-1}, a_i) provided it is nonempty. (Note that f_γ is undefined on a finite set, and can be assigned arbitrary values there.) The mapping

$$\gamma \longrightarrow F(f_\gamma)$$

is a continuous function on C_n and hence achieves a minimum on C_n.

Lemma. Let $\gamma = (a_0, \ldots, a_n; c_1, \ldots, c_n)$ and assume $F(f_\gamma)$ is a minimum on C_n. Then $c_i = 1 \pm \Delta$ whenever $a_{i-1} < a_i$.

Proof. Let C'_{n+2} be the compact subset of C_{n+2} consisting of $(a_0, \ldots, a_{n+2}; c_1, \ldots, c_{n+2})$ such that

$$c_1 = 1 + \Delta \quad \text{and} \quad c_{n+2} = 1 - \Delta$$

We embed C_n in C'_{n+2} by

$$(a_0, \ldots, a_n; c_1, \ldots, c_n)$$
$$\longrightarrow (a_0, a_0, \ldots, a_{n-1}, a_n, a_n; 1 + \Delta, c_1, \ldots, c_n, 1 - \Delta)$$

and prove the assertion for C'_{n+2}. Let

$$M_\gamma = \frac{\int_0^1 f_\gamma^2(x)\,dx}{\int_0^1 f_\gamma(x)\,dx}$$

Assume f_γ assumes value $c \neq 1 \pm \Delta$ on a nonempty open interval (a_{i-1}, a_i). Either $c \geqslant M_\gamma$ or $c < M_\gamma$ and assume for definiteness that the former possibility occurs. We may assume $c_2 = c \geqslant M_\gamma$. In fact it is easy to see that there is $\gamma' \in C_{n+2}$ such that $M_{\gamma'} = M_\gamma$, $F(f_\gamma) = F(f_{\gamma'})$ and $c'_2 = c$ where $\gamma' = (a'_0, a'_1, \ldots, a'_{n+2}; c'_1, \ldots, c'_{n+2})$. Let $0 \leqslant \delta \leqslant |a_2 - a_1|$ and define $\Theta \in C'_{n+2}$ by

$$\Theta = (a_0, a_1 + \delta, a_2, \ldots, a_{n+2}; 1 + \Delta, c_2, \ldots, c_{n+1}, 1 - \Delta)$$

Then by a simple computation

$$F(f_\Theta) = \frac{A + \alpha}{B + \beta}$$

where

$$A = \left(\int_0^1 f_\gamma(x)\,dx\right)^2$$

$$B = \int_0^1 f_\gamma^2(x)\,dx$$

and

$$\frac{\alpha}{\beta} = \frac{\int_0^1 f_\gamma(x)\,dx + \frac{1}{2}(1 + \Delta - c)\,\delta}{c + \frac{1}{2}(1 + \Delta - c)}$$

Since $c \geqslant M_\gamma$, by taking $\delta > 0$ sufficiently small we can ensure $\alpha/\beta < A/B$ and consequently $F(f_\Theta) < F(f_\gamma)$. This proves the lemma.

Let $\{f_n\}$ be a sequence in C_Δ such that

$$\lim_{n\uparrow\infty} F(f_n) = \inf_{f\in C_\Delta} F(f)$$

Since

$$\bigcup_n \{f_\gamma \mid \gamma \in C_n\}$$

is dense in C_Δ there is sequence $\{\gamma_k\}$with $\gamma_k \epsilon C_{n_k}$ such that

$$F(f_{\gamma_k}) < F(f_k) + \frac{1}{k}$$

In view of the lemma we may assume f_{γ_k} takes only values $1 \pm \Delta$ and hence there is $\Theta_k \in C_2$ such that $F(f_{\gamma_k}) = F(f_{\Theta_k})$.

Therefore inf $F(f)$ is achieved for some $F(f_\Theta)$ with $\Theta = (0, a, 1; 1 + \Delta, 1 - \Delta)$. It is a simple exercise to show that $a = (1 - \Delta)/2$ and $F(f_\Theta) = 1 - \Delta^2$.

Finally we note that for any partition $I = E_1 \cup E_2$ with $\text{meas}(E_1) = (1 - \Delta)/2$, $\text{meas}(E_2) = (1 + \Delta)/2$ the function

$$f(x) = \begin{cases} 1 + \Delta, & \text{if } x \in E_1 \\ 1 - \Delta, & \text{if } x \in E_2 \end{cases}$$

is also a minimum for F. The argument above also shows that all minima of F are of this form.