THE BELL SYSTEM TECHNICAL JOURNAL
Vol. 58, No. 10, December 1979
Printed in U.S.A.

Coefficient Inaccuracy in Transversal Filtering

By A. GERSHO, B. GOPINATH, and A. M. ODLYZKO

(Manuscript received August 20, 1978)

Coefficient inaccuracy in transversal filters is known to degrade the frequency response, particularly in stopband regions. The magnitude of the stopband degradation increases with the number of stages n, the length of the impulse response. A widely used formula for the error in frequency response is proportional to $\sqrt{n}$. By extending recent results on random trigonometric polynomials, we show that for random additive coefficient errors with variance σ^2, the error $\Delta H(\omega)$ in frequency response for large n is such that

$$\max_{\omega} |\Delta H(\omega)| \simeq \sigma\sqrt{n \log n}$$

where log denotes the natural logarithm. This result leads to an absolute bound on attainable stopband rejection for any band-select transversal filter with given coefficient inaccuracy. In particular, the result places a definite limitation on the quality of band-select filtering that can be achieved using a CCD split-electrode filter. It also implies bounds for the peak sidelobes of random radar arrays.

I. INTRODUCTION

In recent years, the transversal filter has emerged as an essential signal-processing structure for a large variety of applications in communication systems. A few of these applications are matched filtering in radar or spread-spectrum systems, equalization in data receivers, echo cancellation for satellite communications, and band-select digital filters. The term "transversal filter" originally referred to the continuous-time tapped delay line structure where an output is formed from a weighted sum of the tap voltages. The same basic function has also been achieved using lumped networks to approximate the delay sections. More recently, transversal filters have been realized with digital circuitry using shift registers and digital multipliers, operating on a sampled and quantized input signal. The most recent development

is the emergence of two new technologies, charge-coupled devices (CCDs) and surface acoustic wave (SAW) devices which allow the realization of discrete-time transversal filters without the need for analog-to-digital conversion.

The new technological advances now offer the possibility of realizing transversal filters with hundreds and perhaps even thousands of tap-weight stages on a single integrated-circuit chip. These developments suggest that extremely sophisticated signal-processing functions can readily be obtained. Specifically, with a sufficient number of taps, a transversal filter can be designed to approximate virtually any specified frequency response as closely as desired. However, the inevitable inaccuracies in implementing the desired weighting coefficients result in a departure of the actual frequency response from the predesigned frequency response which increases with the number of tap-weight stages. In digital filtering, coefficient values can be made as accurate as needed, but at the price of increasing hardware costs. With the CCD or SAW technologies, there are fundamental limits on attainable accuracy. Also, in adaptive filtering, the weight-adjustment algorithm results in a steady-state coefficient inaccuracy. It is therefore necessary to have a quantitative knowledge of the degradation in performance of the transversal filter as a function of the coefficient inaccuracy and the number of stages.

For most applications, the appropriate performance measure for the realized transversal filter is the maximum deviation in frequency response magnitude from the desired values over the particular frequency band of interest. In this paper, we focus on this performance measure by examining the *error-frequency response* due to coefficient inaccuracy and show that under very general conditions the maximum magnitude is given asymptotically by $\sigma\sqrt{n \log n}$, where n is the number of stages, σ is the rms coefficient inaccuracy, and log denotes the natural logarithm. Several other closely related results and implications are also presented.

Since the attainable quality of a designed filter increases with n, the number of stages, and for a given coefficient inaccuracy the degradation increases with n, the question arises: Is it possible to realize a filter with arbitrarily high quality in spite of a given coefficient inaccuracy if n is made sufficiently large? We make this question more precise later and show that the answer is negative for low-pass filtering with a transversal filter structure when "quality" is measured by the amount of stopband rejection. In other words, a limit on filter accuracy implies a limit on attainable filter quality regardless of the number of stages used. The results of this paper provide a tool for determining the ultimate limitation on transversal filter performance associated with a particular technology or a particular adaptive algorithm for weight adjustment.

In CCD transversal filters, the split-electrode method requires that the tap weights be scaled so that the maximum magnitude of the coefficient values is unity. The pattern generator used in making the photomasks for CCD fabrication introduces a quantization error whose peak size is a fixed fraction of the maximum coefficient magnitude. Now, for most applications, increasing the number of stages to be realized corresponds to including additional coefficient values representing the tail of the desired impulse response. Consequently, increasing n does not alter the scaling of the coefficient values for CCD implementation. As a result, a coefficient error can indeed be modeled as an additive random variable whose variance does not depend on the desired coefficient value.

A problem that is very similar to that considered above occurs in the theory of random arrays.[1] These are arrays consisting of fewer elements than conventional phased arrays, with the locations of the elements in the array picked randomly. Such arrays are less costly than conventional phased arrays, but this advantage is gained at the cost of increasing the peak sidelobes. Our main result shows how big those sidelobes can be expected to become.

II. PROBLEM FORMULATION

Regardless of the particular application, the transversal filter may be described by its frequency response, $H(\omega)$, which has the general form

$$H(\omega) = \sum_{k=0}^{L-1} \alpha_k e^{jk\omega}, \tag{1}$$

where ω is the normalized frequency variable, L is the number of stages, $j = \sqrt{-1}$, and the coefficients α_k are real-valued numbers specified by the designer. Since $H(\omega)$ is periodic, only the frequency interval $0 \leq \omega \leq 2\pi$ need be considered.

We note, in passing, that (1) also describes the discrete Fourier transform, so that the results of this paper are also applicable to studying the effect of approximate representations of given data values on the Fourier transform of the data.

A special case of transversal filters, of particular interest in band-select filter design, arises when the coefficients are chosen to have the symmetry property:

$$\alpha_k = \alpha_{L-k-1} \qquad \text{for} \qquad 0 \leq k \leq (L-1). \tag{2}$$

When L is odd, this condition results in a linear phase transfer function having the form

$$H(\omega) = e^{+j\omega n} \sum_{k=0}^{n} b_k \cos k\omega, \tag{3}$$

with $n = (L-1)/2$, and

$$b_k = 2\,\alpha_{n-k} \qquad \text{for} \qquad k \neq 0, \quad b_0 = \alpha_n.$$

Implementation of the coefficients α_k for the general form (1) or b_k for the linear phase form (2) inevitably results in the introduction of errors or inaccuracies. We denote the actual (inaccurate) value realized as α'_k, or as b'_k for the linear phase case. Then the kth coefficient error is the difference $\epsilon_k = \alpha'_k - \alpha_k$, or $\epsilon_k = b'_k - b_k$ in the linear phase case. The realized transfer function then differs from the desired transfer function by the *error transfer function* defined as

$$f_L(\omega) = \sum_{k=0}^{L-1} \epsilon_k e^{jk\omega} \tag{4}$$

in the general case or, in the linear phase case:

$$g_n(\omega) = e^{+j\omega n} \sum_{k=0}^{n} \epsilon_k \cos k\omega. \tag{5}$$

Clearly, the error transfer function, if known, provides a full description of the degradation in performance of the realized filter from the desired performance in the absence of inaccuracies.

Since the errors, ϵ_k, are generally not known prior to fabrication of the filter, they are modeled most effectively as random variables whose distribution depends on the particular mechanism involved in fabricating the tap weights. In digital filtering, the errors are due to coefficient quantization and are usually modeled as uniformly distributed random variables. The error terms for different coefficients, being independently produced, can reasonably be assumed to be independent random variables.

Additive error components were used by Knowles and Olcayto[2] for modeling coefficient quantization in recursive filters. Chan and Rabiner[3] applied this approach for transversal filters and evaluated the rms values of $f_L(\omega)$ and $g_n(\omega)$ at a particular frequency. They assumed mutually independent and uniformly distributed errors ϵ_k resulting in rms values for the error transfer function proportional to $\sqrt{L}$, or $\sqrt{n}$ in the linear phase case. By taking the maximum over all frequencies of the rms deviation, a frequency-independent upper bound on the error transfer function is obtained which is valid at any particular frequency with high probability.

More recently, Heute[4,5] noted that the bounds of Chan and Rabiner underestimate the degradation due to the maximum of $|g_n(\omega)|$ over the frequency band. It is this latter measure of degradation that is meaningful in most applications. Chan and Rabiner's bound is not a high probability upper bound for the maximum ripple magnitude taken

on by $g_n(\omega)$. Heute proposed a heuristic upper bound for the maximum of $|g_n(\omega)|$ which has the form $Q[a + bn + (cn + d)^{1/2}]$, where a, b, c, and d are constants and Q is the peak amplitude of the uniformly distributed error terms ϵ_k. His bound gave an improved fit to simulated data for values of n up to 128. We shall see later that Heute's bound, which for large n grows linearly with number of stages n, grossly overestimates the degradation as n becomes much larger than 100.

Andrisano and Calandrino[6] assumed that the error transfer function is a Gaussian process and found an (implicit) bound on stopband rejection as the solution of a transcendental equation.

In this paper, we take as the measure of degradation due to coefficient inaccuracy,

$$D_L = \max_{\omega \in \Omega} |f_L(\omega)| \tag{6}$$

for the general transversal filter and

$$M_n = \max_{\omega \in \Omega} |g_n(\omega)| \tag{7}$$

for the linear phase transversal filter, where Ω is a particular frequency band of interest. We assume the errors ϵ_k are mutually independent random variables with a common distribution satisfying certain regularity conditions that include the uniform and normal distributions as special cases.

We establish here for the first time that the maximum frequency response errors D_n and M_n are asymptotically (for large n) given by $\sigma\sqrt{n \log n}$ where σ is the rms coefficient error. Although the result is asymptotic, Lawrence and Salazar[7] found that it was moderately accurate in one study of a low-pass filter with only 33 taps. Application of the result to low-pass filter performance is examined briefly in this paper and more extensively in Ref. 8 and 9. Until the report of our result,[8] the correct behavior of the error frequency response magnitude had apparently not been recognized in the digital filtering literature.

The existing mathematical results most closely related to our work are due to Halasz[10] who considered random trigonometric sums with coefficients that take on the values ± 1 with equal probability. While too restrictive to apply to transversal filters, his methodology was useful in deriving our more general upper bound on the maximum error frequency response.

Our main result is also applicable to the analysis of random arrays, and in particular to that of statistical arrays.[1] These are arrays consisting of k isotropic radiators placed among n positions ($n > k$) that are spaced $\lambda/2$ apart (λ = wavelength), with the k positions to be occupied by the k elements determined at random. The array factor of such an array is defined as

$$f(u) = \sum_{r=0}^{n-1} g_r \, e^{\pi jru}, \tag{8}$$

where $g_r = 1$ if the rth position is occupied by a radiator, and $g_r = 0$ otherwise. This can be rewritten as

$$f(u) = \frac{k}{n} \sum_{r=0}^{n-1} e^{\pi jru} + \sum_{r=0}^{n-1} \epsilon_r \, e^{\pi jru}, \tag{9}$$

where $\epsilon_r = 1 - k/n$ for the k values of r for which $g_r = 1$, and $\epsilon_r = -k/n$ otherwise. The first sum above represents (except for the k/n multiplier) the array factor of a conventional phased array. The random choice of the positions for the radiators corresponds to letting the ϵ_r be independent random variables, assuming the value $1 - k/n$ with probability k/n, and the value $-k/n$ with probability $1 - k/n$. If we assume that $k \sim \alpha n$ as $n \to \infty$, then our theorem shows that, with the probability approaching 1 as $n \to \infty$, the second sum in (9) will never be significantly larger than $\sqrt{1-\alpha}\,\sqrt{n \log n}$ and that, conversely, it will get that large on any subinterval. This result, which had been derived only heuristically before,[1] explains why random arrays are usually not very satisfactory.

III. STATEMENT OF MATHEMATICAL RESULTS

As we saw in Section II, the errors in the realized transfer functions are given by $\sum_{k=0}^{L-1} \epsilon_k \, e^{jk\omega}$, or $\sum_{k=0}^{n} \epsilon_k \cos k\omega$, or $\sum_{k=0}^{n} \epsilon_k \sin k\omega$. Hence the distribution of the random variables ϵ_k will depend on the model for the sources of inaccuracies. For digital implementation, the usual model assumes ϵ_k to be independently distributed uniformly between $-\Delta$, $+\Delta$, where Δ is the maximum error due to truncation of the coefficients of the filter. In other situations, a Gaussian distribution may be more appropriate. But, as we shall see, the asymptotic behavior of the maximum magnitude of the error is not dependent on the exact nature of the distribution. It depends only on a few functionals of the distribution.

The results presented here rely on an important assumption about the distribution of the ϵ_k. We assume throughout that the ϵ_k have mean zero and finite sixth moment, so that the characteristic function $E(e^{jx\epsilon_k})$ of ϵ_k is such that

$$E(e^{jx\epsilon_k}) = \exp\left[- \sum_{r=2}^{5} a_r x^r + O\,(x^6) \right] \tag{10}$$

for x in some nontrivial interval $[-d, d]$. (Note that $a_2 > 0$ if the ϵ_k are not identically zero.) Condition (10) is satisfied for most probability density functions of practical interest.

Now we are ready to state the main result:

Theorem: Let ϵ_k, $k = 1, 2, \cdots$ *be a sequence of independent identically distributed random variables satisfying* (10).

Then there exist constants C_1 *and* C_2, *not depending on* n, *such that*

$$\max_{0\le\theta\le 2\pi} \left| \sum_{k=1}^{n} \epsilon_k e^{kj\theta} \right| \le \sqrt{2a_2}\ \sqrt{n \log n} + C_1 \sqrt{\frac{n}{\log n}} \log \log n$$

holds with probability $\ge 1 - C_2 (\log n)^{-4}$. *Furthermore, if* Ω *is any subinterval of* $[0, 2\pi]$ *of length* $\ge (\log n)^{-1}$ *and* α *is any real number, then*

$$\max_{\theta\in\Omega} Re \left\{ e^{j\alpha} \sum_{k=1}^{n} \epsilon_k e^{kj\theta} \right\} \ge \sqrt{2a_2}\ \sqrt{n \log n} - C_1 \sqrt{\frac{n}{\log n}} \log \log n$$

holds with probability $\ge 1 - C_2 (\log n)^{-4}$.

Thus, with high probability, $\max |f(\theta)|$ is about $\sqrt{2a_2}\ \sqrt{n \log n}$. The proof is outlined in Section V.

Remark 1. By choosing α appropriately, we can conclude that each of $\sum \epsilon_k \cos(k\theta)$, $\sum \epsilon_k \sin(k\theta)$ becomes large on any long θ interval with high probability.

Remark 2: The estimates presented here are not the best possible ones. For example, the interval Ω in the lower bound proof can be of size $n (\log n)^{-\lambda}$ for any $\lambda > 0$.

IV. APPLICATION TO LOW-PASS FILTERS

The usual specifications for FIR low-pass filters are shown in Fig. 1.[11] A design problem is to find the smallest n such that

$$\left| \sum_{k=0}^{n-1} \alpha_k \cos k\theta \right|$$

lies between $1 - \delta_1$ and $1 + \delta_1$ in the passband, i.e., for $\theta \in [0, F_p]$ and between 0 and δ_2 in the stopband, i.e., for $\theta \in [F_s, \pi]$. Estimates for n given δ_1, δ_2, F_p, F_s are given in Ref. 12. However, the validity of the estimates in Ref. 12 for regions of practical interest is not proven. An empirical relationship is given in Ref. 11.

As n increases, smaller δ_2, δ_1 and $F_s - F_p$ are possible. Hence, a question that is usually raised is: Given that the α_k's cannot be realized exactly, what can be said about the minimum δ_2 possible if the distribution of ϵ_k, the error in α_k, is known. If α_k's could be realized exactly, arbitrarily small values of δ_2 can be obtained by making n large. However, ϵ_k's introduce errors that grow with n as seen from the theorem. So there is a trade-off between errors introduced by inaccur-

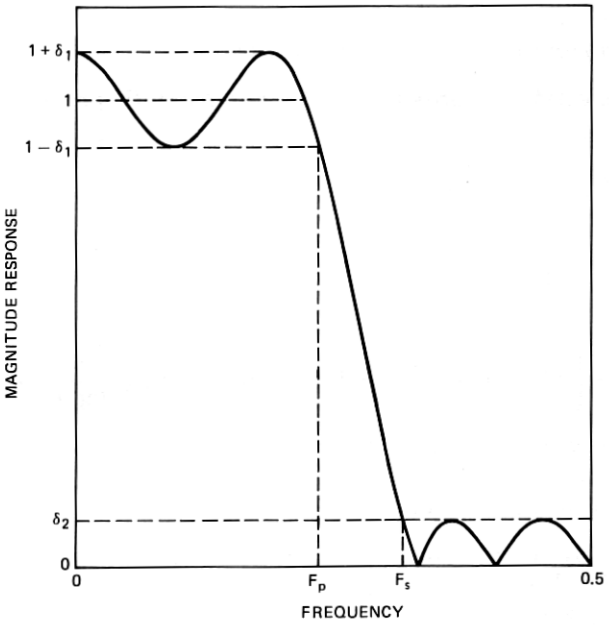

Fig. 1—Frequency as fraction of 2 π.

acies in α_k's and the improvement in performance with increasing n. For an example, we consider the stopband rejection, 20 $\log_{10} \delta_2$, as a figure of merit with given values of δ_1 and F_p, F_s. The empirical formula gives the following relation for $\hat{n}$, the minimum n required to achieve a stopband rejection of 20 $\log_{10} \delta_2$.

$$\hat{n} = c_1 \log \delta_2 + c_2, \tag{11}$$

where c_1 and c_2 are constants depending on δ_1 and $F_s - F_p$.[11] For fixed point digital implementations, if the coefficients of the filter, the α_k's, are truncated to d bits, then the "error" in α_k is generally modeled as a uniform random variable ϵ_k having values between -2^{-d} and $2^{-d} = \Delta$. For this model, a_2 of the theorem is $1/6\ \Delta^2$. Hence, the maximum error e_n due to these inaccuracies,

$$e_n = \max_{0 \le \theta \le 2\pi} \left| \sum_{k=0}^{n-1} \epsilon_k \cos k\theta \right|, \tag{12}$$

is such that

$$\frac{e_n}{\sqrt{n \log n}} \to \sqrt{2a_2} = \sqrt{\frac{\Delta^2}{3}} = \frac{\Delta}{\sqrt{3}} \quad \text{as} \quad n \to \infty \tag{13}$$

and

$$\left| e_n - \frac{\Delta}{\sqrt{3}} \sqrt{n \log n} \right| < c \sqrt{\frac{n}{\log n}} \log \log n \tag{14}$$

with probability $\geq 1 - O((\log n)^{-4})$.

Using the limit (13) to indicate expected deterioration in performance, we can arrive at a design rule. If coefficients are truncated to d bits, then the minimum achievable δ_2 before the random errors become comparable to δ_2 itself is given by:

$$\frac{2^{-d}}{\sqrt{3}} \sqrt{(c_1 \log \delta_2 + c_2) \log (c_1 \log \delta_2 + c_2)} = \delta_2. \tag{15}$$

Putting $\delta_2 = 2^{-s}$,

$$\frac{2^{-d}}{\sqrt{3}} = \frac{2^{-s}}{\sqrt{(-c_1's + c_2) \log (-c_1's + c_2)}}, \tag{16}$$

where $c_1' = c_1 \log 2$.

From the above formula, we can estimate the required precision for the coefficients for a given value of $\delta_2 = 2^{-s}$.

In design of CCD filters, a similar formula can be used. In situations where the tap-weight errors can be modeled by a Gaussian random variable with a standard deviation Δ, then a_2 for our theorem is $\Delta^2/2$. The minimum achievable δ_2 satisfies

$$\sqrt{(c_1 \log \delta_2 + c_2) \log (c_1 \log \delta_2 + c_2)} = \frac{\delta_2}{\Delta}. \tag{17}$$

Solving for δ_2, we can estimate the optimum value of n.

As an illustration of the effect of coefficient inaccuracy on limiting the stopband rejection of a low-pass filter, Fig. 2 shows how the best achievable rejection depends on the number of stages, n, for various values of the transition width $\Delta F = F_S - F_P$. These curves were calculated by solving the empirical formula of Ref. 11 for δ_2 and adding to it the maximum error $\sigma\sqrt{n} \log n$. This gives an expression for the best attainable stopband rejection in the presence of coefficient errors, as a function of n, δ_1, and ΔF. For additional curves obtained in this way, see Ref. 8. Computation also shows that varying the allowed passband ripple δ_1 has a negligible effect on the maximum attainable stopband rejection. It is evident that coefficient inaccuracy places an ultimate limitation on the attainable quality of a low-pass filter implemented with the transversal structure.

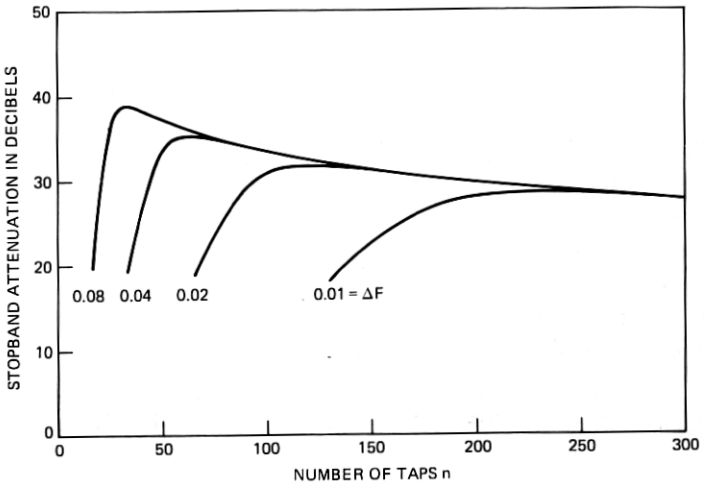

Fig. 2—Best obtainable stopband attenuation for a low-pass transversal filter in the presence of coefficient inaccuracies. Root-mean-square coefficient error = 0.001, passband ripple = 0.0122. Curves are shown for four values of the transition width. Note that, for each curve, an optimum number of taps exist. Reducing the passband ripple allowable has the effect of shifting these curves to the right while reducing the peak value of each curve.

V. ACKNOWLEDGMENTS

We wish to thank Prof. W. Schuessler who brought to our attention the work of Heute and the inadequacy of the $\sqrt{n}$ bound. Prof. P. Erdös kindly told us about the work of Halasz.[10] We also benefited from discussions with L. A. Shepp.

APPENDIX

Here we outline the main steps in the proof, which follows that of Halasz,[10] in which he assumed ϵ_k to be ± 1. (An earlier proof of a slightly weaker result had been outlined by Whittle.[13]) Results that are incidental to the main line of reasoning are collected at the end of this outline. Let

$$f(\alpha, \theta) = \sum_{k=1}^{n} \epsilon_k \cos(k\theta + \alpha) = \mathrm{Re}\left(e^{j\alpha} \sum_{k=1}^{n} \epsilon_k e^{jk\theta} \right).$$

(*i*) We construct a nonnegative function $u(x) \leq 1$ which can be used to indicate in an approximate sense the set of values of x that exceed given values. Let M_1, M_2, $D > 0$ be given numbers. Then $u(x)$ is zero for $-M_2 \leq x \leq M_1$, and $u(x) = 1$ for $x \geq M_1 + D$ or $x \leq -M_2 - D$. In the interval $[-M_2 - D, -M_2]$ and $[M_1, M_1 + D]$, $u(x)$ is 40 times differentiable and $u^{(r)}(x) = O(D^{-r})$ as $D \uparrow \infty$, for $0 \leq r \leq 40$.

For deriving the upper bound, we proceed as follows:

(*ii*) Put $M_1 = M_2 = M = \sqrt{2a_2}\sqrt{n \log n} + gD \log\log n$ where $D = \sqrt{n/\log n}$ and $g = 20\sqrt{a_2/2}$.

(*iii*) Let $v_1(t) = 1/2\pi \int_{-\infty}^{\infty} (1 - u(x))e^{-jtx}\,dx$. Then $|v_1(t)| = O(M)$ and $|t^r v_1(t)| = O(D^{-r+1})$, $1 \le r \le 40$.

(*iv*) Let $G = \int_0^{2\pi} d\alpha \int_0^{2\pi} d\theta\, u(f(\alpha, \theta))$ and $v(t) = \delta(t) - v_1(t)$, where $\delta(t)$ is the Dirac delta function. Then

$$G = \int_0^{2\pi} d\alpha \int_0^{2\pi} d\theta \int_{-\infty}^{\infty} \exp\left[\text{jt} \sum_1^n \epsilon_k \cos(k\theta + \alpha)\right] v(t)dt.$$

Using the properties of v_1, we can then show (for details, see the section at the end of this proof) that

$$E(G) = O(n^{-1}(\log n)^{-10}).$$

(*v*) Let $T = \max_{\alpha,\theta} |f(\alpha, \theta)|$. Then using the inequalities

$$\left|\frac{\partial}{\partial\theta} f(\alpha, \theta)\right| \le Tn, \qquad \left|\frac{\partial}{\partial\alpha} f(\alpha, \theta)\right| \le T$$

we can show that $G \le 1/(n \log^2 n) \Rightarrow T \le M + 2D$ for large enough n.

(*vi*) The result from step (*iv*) implies

$$Pr\left\{G \ge \frac{1}{n \log^2 n}\right\} = O\left(\frac{1}{(\log n)^8}\right).$$

Hence, using step (*v*)

$$\max_{\alpha,\theta} |f(\alpha, \theta)| \le \sqrt{2a_2}\sqrt{n \log n} + (g + 2)\sqrt{\frac{n}{\log n}} \log\log n$$

$$\text{with probability} \quad \ge 1 - O\left(\frac{1}{(\log n)^8}\right).$$

The derivation of the lower bound is more difficult, but similar. We will only outline the proof. We examine the values of $f(\alpha, \theta)$ at the points $\theta_m = 2\pi \cdot (2m - 1)/2n$, for $1 \le m \le n$.

(*vii*) Let $M_1 = M = \sqrt{2a_2}\sqrt{n \log n} - gD \log\log n$, and $M_2 = 2M$. Let S be a subset of the integers from 1 to n with cardinality greater than $n(\log n)^{-1}$ and put $F = \sum_{m \in S} u(f(\alpha, \theta_m))$.

As in the derivation of the upper bound, we can find the asymptotic behavior of the first two moments of F using the properties of u.

(*viii*) We can show $E(F) \ge c_3 |S| n^{-1}(\log n)^{21/2}$ for some constant $c_3 > 0$, and $E(F^2) - E^2(F) = O(|S| n^{-1}(\log n)^{21/2}) + O(|S|^2 n^{-2}(\log n)^7)$.

(*ix*) Now

$$\Pr\{\max_{m \in S} f(\alpha, \theta_m) \ge M \text{ or } \min_{m \in S} f(\alpha, \theta_m) \le -2M\} \ge 1 - \Pr\{F = 0\}$$

by the definition of n. But $\Pr\{\min f(\alpha, \theta_m) \le -2M\} = O((\log n)^{-8})$ from the upper bound, so that

$$\Pr\{\max_{m \epsilon S} f(\alpha, \theta_m) \ge M\} \ge 1 - \Pr(F = 0) - O((\log n)^{-8}).$$

Further, $\Pr\{F = 0\} \le \Pr\{(F - E(F))^2 \ge E^2(F)\}$, so by Chebyshev's inequality

$$\Pr\{F = 0\} \le \frac{E((F - E(F))^2)}{E^2(F)}.$$

(x) Using the bounds from step (*viii*), $\Pr\{F = 0\} = O((\log n)^{-9/2})$. Therefore, using the definition of M_1, M_2, and D, we have:

$$\begin{aligned}\Pr\left\{\max_{m \epsilon S} f(\alpha, \theta_m) \ge \sqrt{2a_2}\sqrt{n \log n} - g\sqrt{\frac{n}{\log n}} \log\log n\right\} \\ \ge 1 - \Pr\{F = 0\} - O((\log n)^{-8}) \\ \ge 1 - O((\log n)^{-9/2}).\end{aligned}$$

Details of Step (iv)

From the definition of $v(t)$ [see step (*iv*)], we can show that

$$\int_{-\infty}^{\infty} |t|^r |v(t)|\, dt = O(D^{-r}) \qquad 1 \le r \le 18 \tag{18}$$

$$\int_{|t|>d/2} |t|^r |v(t)|\, dt = O(D^{-19}). \tag{19}$$

Since the Fourier transform of $t^r v(t)$ is $j^{-r}u^{(r)}(x)$,

$$\begin{aligned}\left|\int_{-\infty}^{\infty} e^{-\beta t^2} t^r v(t)\, dt\right| &= \frac{1}{2\sqrt{\pi\beta}} \left|\int_{-\infty}^{\infty} u^{(r)}(x) e^{-x^2/4\beta} dx\right| \\ &= O\left(\beta^{-1/2} D^{-r} \int_{|x| \ge M} e^{-x^2/4\beta} dx\right) \\ &= O\left(\frac{\sqrt{\beta}}{MD^r} e^{-M^2/4\beta}\right) \end{aligned} \tag{20}$$

uniformly in $\beta > 0$, $0 \le r \le 18$.

Similarly, for $\beta > 0$

$$\int_{-\infty}^{\infty} \exp\left(-\beta t^2 \sum_{k=1}^{n} \cos^2(k\theta + \alpha)\right) t^r v(t)\, dt = O\left(\frac{\sqrt{\beta u}}{MD^2} \exp(-M^2/Q)\right), \tag{21}$$

where

$$Q = 4\beta \sum_{k=1}^{n} \cos^2 (k\theta + \alpha).$$

Further,

$$\sum_{k=1}^{n} \cos^2 (k\theta + \alpha) = \frac{n}{2} + \frac{1}{2} \sum_{k=1}^{n} \cos 2(k\theta + \alpha)$$

and

$$\sum_{k=1}^{n} \cos 2(k\theta + \alpha) \leq \frac{1}{|\sin \theta|}.$$

So

$$\sum_{k=1}^{n} \cos^2 (k\theta + \alpha) \leq \begin{cases} n & \forall \theta, \alpha \\ \dfrac{n}{2} + \dfrac{n}{2 \log n} & \text{for } \dfrac{\pi}{2} \dfrac{\log n}{n} \leq |\theta|, \end{cases}$$

since

$$|\sin \theta| \geq \frac{n}{\log n} \quad \text{for} \quad \frac{\pi}{2} \frac{\log n}{n} \leq |\theta| \leq \pi - \frac{\pi}{2} \frac{\log n}{n}.$$

Therefore

$$\int_0^{2\pi} \exp \frac{-M^2}{4\beta \sum_{k=1}^{n} \cos^2 (k\theta + \alpha)} \, d\theta = O(e^{-M^2/2\beta n}), \tag{22}$$

whence

$$\int_0^{2\pi} d\alpha \int_0^{2\pi} d\theta \int_{-\infty}^{\infty} \exp\left[-\beta t^2 \sum_{k=1}^{n} \cos^2 (k\theta + \alpha) \right] t^r v(t) dt$$

$$= O\left(\frac{\sqrt{\beta n}}{MD^r} e^{-M^2/2\beta n} \right). \tag{23}$$

Using (10) and (23) above, we can derive the result of step (*iv*) as follows:

Since ϵ_k are independent,

$$E\left(\exp\left[jt \sum_{k=1}^{n} \epsilon_k c_k \right] \right) = \prod_{k=1}^{n} E e^{jt\epsilon_k c_k}$$

$$= \exp\left[- \sum_{l=2}^{5} a_l t^l \sum_{k=1}^{n} c_k^l + O(nt^6) \right] \quad \text{for} \quad |t| \leq d$$

from assumption (10), where c_k denotes $\cos(k\theta + \alpha)$. Further,

$$E\left(\exp\left[jt\sum_{k=1}^{n}\epsilon_k c_k\right]\right) = \exp\left[-a_2t^2\sum_{k=1}^{n}c_k^2\right] + \sum_{j=3}^{5}a_jt^j\sum_{k=1}^{n}c_k^j\exp\left[-a_2t^2\sum_{k=1}^{n}c_k^2\right] + \frac{1}{2}\left\{\sum_{j=3}^{5}a_jt^j\sum_{1}^{n}e_k^j\right\}^2\exp\left[-a_2t^2\sum_{k=1}^{n}c_k^2\right] + O(nt^6) + O(n^3|t|^9) \quad \text{for} \quad |t| \le d, \tag{24}$$

since

$$e^{-a} = e^{-b} + (b-a)e^{-b} + \tfrac{1}{2}(b-a)^2e^{-b} + O(|b-a|^3)$$

uniformly for $b \in \mathscr{R}$, $b \ge 0$, $a \in \mathscr{C}$, $\mathrm{Re}(a) \ge 0$. Now we consider

$$E\left(\int_{-\infty}^{\infty}\exp\left[jt\sum_{1}^{n}\epsilon_k c_k\right]v(t)dt\right) = \int_{-\infty}^{\infty}E\left(\exp\left[jt\sum_{1}^{n}\epsilon_k c_k\right]\right)v(t)dt;$$

the expression on the right-hand side of (24) can be substituted for the integrand in the interval $|t| \le d$. Outside this interval, we can use the simple bound $|\exp[jt\sum_1^n \epsilon_k c_k]| \le 1$, and arrive at:

$$E\int_{-\infty}^{\infty}\exp\left[jt\sum_{1}^{n}\epsilon_k c_k\right]v(t)dt = \int_{-d}^{d}E\left(\exp\left[jt\sum_{1}^{n}\epsilon_k c_k\right]\right)v(t)dt + O\left(\int_{|t|\ge d}|v(t)|dt\right)$$

$$= \int_{-\infty}^{\infty}E\left(\exp\left[jt\sum_{1}^{n}\epsilon_k c_k\right]\right)v(t)dt + O\left(\int_{|t|\ge d}|v(t)|dt\right) + O\left(n\int_{|t|\ge d}|t|^5|v(t)|dt\right) + O\left(n^2\int_{|t|\ge d}t^{10}|v(t)|dt\right)$$

$$+ O\left(n \int_{-\infty}^{\infty} t^6 |v(t)|\,dt\right)$$

$$+ O\left(n^3 \int_{-\infty}^{\infty} |t|^9 |v(t)|\,dt\right). \tag{25}$$

From (18) and (19), we see that the right-hand side is

$$= \int_{-\infty}^{\infty} E\left(\exp\left[jt \sum_1^n \epsilon_k c_k\right]\right) v(t)dt + O(nD^{-6}) + O(n^2D^{-19}) + o(n^3D^{-9})$$

$$= \int_{-\infty}^{\infty} E\left(\exp\left[jt \sum_1^n \epsilon_k c_k\right]\right) v(t)dt + O\left(\frac{(\log n)^{9/2}}{n^{3/2}}\right). \tag{26}$$

To find the asymptotic behavior of $E(G)$, we use (21). After integrating with respect to α,θ, we have, for each of the terms in (24), with expressions in square brackets corresponding to those in (24),

$$\int_0^{2\pi} d\alpha \int_0^{2\pi} d\theta \int_{-\infty}^{\infty} \exp\left[-a_2 t^2 \sum_1^n c_k^2\right] v(t)dt = O\left(\frac{\sqrt{n}}{M} e^{-M^2/2a_2 n}\right)$$

$$\int_0^{2\pi} d\alpha \int_0^{2\pi} d\theta \sum_{j=3}^{5} a_j \sum_1^n c_k^j \int_{-\infty}^{\infty} t^j \exp[\ ]v(t)dt = O\left(n \frac{\sqrt{n}}{MD^3} e^{-M^2/2a_2 n}\right)$$

$$\int_0^{2\pi} d\alpha \int_0^{2\pi} d\theta \int_{-\infty}^{\infty} \left\{\sum_{j=3}^{5} a_j t^j \sum c_k^j\right\}^2 \exp[\ ]v(t)dt$$

$$= O\left(n^2 \frac{\sqrt{n}}{MD^6} e^{-M^2/2a_2 n}\right).$$

Therefore, collecting the previous results, we have

$$E(G) = \int_0^{2\pi} d\alpha \int_0^{2\pi} d\theta \int_{-\infty}^{\infty} E\left(\exp\left[jt \sum_1^n \epsilon_k c_k\right]\right) v(t)dt$$

$$= O\left(\frac{\sqrt{n}}{M} e^{-M^2/2a_2 n} + \frac{(\log n)^{9/2}}{n^{3/2}}\right) = O(n^{-1}(\log n)^{-10}).$$

REFERENCES

1. B. D. Steinberg, *Principles of Aperture and Array System Design*, New York: John Wiley & Sons, 1976.

2. J. B. Knowles and E. M. Olcayto, "Coefficient Accuracy and Digital Filter Response," IEEE Trans. on Circuit Theory, *CT-15* (March 1968), pp. 31–41.
3. D. S. K. Chan and L. R. Rabiner, "Analysis of Quantization Errors in the Direct Form for Finite Impulse Response Digital Filters," IEEE Trans. Audio and Electroacoustics, *AU-21* (August 1973), pp. 354–356.
4. U. Heute, "Koefficienten—Empfindlichkeit nicht—rekursiver Digitalfilter in direkter Struktur," Dissertation, Institut für Nachrichtentechnik Universität Erlangen, 1974.
5. U. Heute, "Necessary and Efficient Expenditure for Non-Recursive Digital Filters in Direct Structure," European Conf. on Circuit Theory and Design, IEEE Conf. Pub. No. 116 (July 1974), pp. 13–19.
6. O. Andrisano and L. Calandrino, "Tap Weight Tolerance Effects in CCD Transversal Filtering," Alta Frequenza, *45*, 1976, pp. 739–745.
7. V. B. Lawrence and A. C. Salazar, "Effects of Finite Coefficient Precision on FIR Filter Spectra," Proc. IEEE International Conf. Acoustics, Speech, Signal Processing, April 1979, pp. 378–379.
8. A. Gersho, "Charge Transfer Filtering," Proc. IEEE, *67*, No. 2 (February 1979), pp. 196–218.
9. A. Gersho, B. Gopinath, and A. Odlyzko, "Coefficient Inaccuracy in FIR Filters," Proc. Int'l. Symp. Acoustics Speech & Signal Processing, 1979, pp. 375–377.
10. G. Halasz, "On a Result of Salem and Zygmund Concerning Random Polynomials," Studia Scient. Math Hung., *8* (1973), pp. 369–377.
11. L. R. Rabiner and B. Gold, *Theory and Application of Digital Signal Processing*, Englewood Cliffs, N.J.: Prentice-Hall, 1975.
12. W. Fuchs, "A Problem on Approximation by Polynomials," unpublished work, 1975.
13. P. Whittle, "Sur la Distribution du Maximum d'un Polynome Trigonometrique à Coefficients Aléatories," Le Calcul des Probabilities et ses Applications, Centre National de la Recherche Scientifique, Paris, 1959 (Colloques Internationaux du C.N.R.S. 1958).