IMPROVED ERROR ESTIMATES FOR FIRST ORDER SIGMA-DELTA SYSTEMS

C. Sinan Güntürk

Princeton University
Program in Applied and Computational Mathematics
Fine Hall, Princeton NJ, 08544
cgunturk@math.princeton.edu

ABSTRACT

A sigma-delta modulator converts a sequence of real numbers bounded by 1 into a digital sequence with entries $+1$ and -1 only. When the input sequence consists of samples of a bandlimited function, oversampled by a factor λ with respect to its Nyquist frequency, we prove that filtering the digital output with an appropriate kernel reproduces the original bandlimited function up to an error of at most $O(\lambda^{-4/3+\epsilon})$.

1. INTRODUCTION

We consider the class $\mathcal{B}_\Omega$ of real-valued functions x defined on the real line with Fourier transforms $\hat{x}$ that are finite Borel measures supported on a finite interval $[-\Omega, \Omega]$. We call such a function *bandlimited*; it is the restriction to $\mathbf{R}$ of an entire function of exponential type:

$$x(z) = \frac{1}{2\pi} \int e^{i\xi z} d\hat{x}(\xi). \tag{1}$$

We fix Ω, the *Nyquist frequency*; by rescaling we can assume Ω to be π. The *sampling theorem* states that if the *filter* g is such that $\hat{g}$ is a continuous cut-off function satisfying

$$\hat{g}(\xi) = \begin{cases} 1 & \text{if } |\xi| \leq \pi \text{ and} \\ 0 & \text{if } |\xi| > \lambda\pi, \end{cases} \tag{2}$$

then

$$x(t) = \frac{1}{\lambda} \sum_{n \in \mathbf{Z}} x(\tfrac{n}{\lambda}) g(t - \tfrac{n}{\lambda}), \tag{3}$$

where $\lambda > 1$. The equality holds pointwise, at least for all values of t for which the right hand side converges, and in general for all t if the summation method is the Cesàro means.

Clearly, the bandlimited functions we have defined are in L^∞. We further restrict to those with L^∞ norms strictly less than a given number which, for normalization purposes, we choose to be 1. Given a function x in this class, we study a class of approximations of the form

$$\tilde{x}_\lambda(t) = \frac{1}{\lambda} \sum_{n \in \mathbf{Z}} \mathbf{q}_n h(t - \tfrac{n}{\lambda}) \tag{4}$$

where the $\mathbf{q}_n$ are drawn from the discrete set $\{+1, -1\}$. Since $\tilde{x}_\lambda$ is only an approximation, we shall not always require $\hat{h}$ to be a cut-off function of the kind previously described; by allowing more general $\hat{h}$ than in (2), we can, e.g., also consider the case where h is supported on a finite interval. We describe and bound the error introduced in this case in the next section. Note that h may also depend implicitly on λ; we will write h_λ whenever we wish to emphasize this dependence.

Of course, the heart of the approximation is the construction of $\{\mathbf{q}_n\}$. The simplest construction would be to quantize each $x(\frac{n}{\lambda})$ by keeping only its sign. However, this fails due to the fact that there are many bandlimited functions without sign change. The so called $\Sigma - \Delta$ *quantization* (or $\Sigma - \Delta$ *modulation*) is a special way of computing $\mathbf{q}_n$. The references [1] and [2] cover the history as well as recent advances in the theory and applications of $\Sigma - \Delta$ quantization in electrical engineering.

A first order $\Sigma - \Delta$ quantizer runs the following recursion:

$$\begin{aligned} \mathbf{q}_n &= Q(x(\tfrac{n}{\lambda}), u_{n-1}) & (5) \\ u_n &= u_{n-1} + x(\tfrac{n}{\lambda}) - \mathbf{q}_n, & (6) \end{aligned}$$

where (u_n) is an auxiliary sequence that is recursively computed as given by (6), and Q is a ± 1-valued function that needs to be specified. We will take $Q(x_1, x_2) = \text{sign}(x_1 + x_2)$ with the convention that $\text{sign}(0) = 1$. The more familiar version in the electrical engineering literature is that $\mathbf{q}_n = \text{sign}(u_{n-1})$, but this system is equivalent to the one above up to an index shift in $\mathbf{q}_n$. The reason for the name "first order" is that (6) is a difference equation of order 1 in u_n. An N-th order scheme would involve N-th order differences.

The author's research was supported by a Francis Robbins Upton honorific graduate fellowship from Princeton University; his participation in this conference was supported by ONR grant N00014-96-1-0367.

It is not immediate that such an approximation may become arbitrarily close to x as λ increases. Moreover, the sense of the convergence, as measured by several possible functional norms, and the rate of convergence are also questions of study. There are few mathematical results in the literature; Gray in [3] derives results for the mean square error when the input is a random constant, with its value uniformly distributed in $[-1, 1]$; the mean is taken both in time and over the random constant input. Gray *et al.* have a similar analysis for sinusoidal functions in [4]. Recently, Daubechies and DeVore in [5] have built a family of arbitrary order sigma-delta modulators and have obtained results for the same class of bandlimited functions we consider. They show that $\tilde{x}_\lambda$ converges uniformly to x as λ increases and that the error $\|x - \tilde{x}_\lambda\|_\infty$ is $O(\lambda^{-N})$ for the N-th order sigma-delta modulator in their family.

In this work, we will restrict ourselves to the first-order scheme; we will derive pointwise estimates for the approximation error that decay faster than the $O(\lambda^{-1})$ L^∞-norm estimate of [5]. The result we will prove may be formulated as the following:

Theorem 1 *For all $\eta > 0$, there exists a family $\{h_\lambda\}_{\lambda \geq 1}$ of filters such that, for all $x \in \mathcal{B}_\Omega$, and all t for which $x'(t)$ does not vanish, we have*

$$|x(t) - \tilde{x}_\lambda(t)| \leq C(\eta, x'(t))\lambda^{-4/3+\eta} \tag{7}$$

for some constant $C(\eta, x'(t))$.

We note that the condition of the derivative to be nonvanishing at a point is not very restrictive. Since bandlimited functions are analytic, the derivative x' of a non constant band-limited function has at most countably many zeros, with no accumulation point. We describe in [6] how to proceed at zeros of x'. Moreover, if the function is a constant, we will separately show the following stronger result:

Theorem 2 *Consider the constant functions with values in $[-1, 1]$. Then, for almost all constants, using λ consecutive $\mathbf{q}_n$ and the same method of reconstruction, it is possible to approximate this constant up to an error less than $O(\lambda^{-2+\eta})$ for any $\eta > 0$.*

The proofs of these two theorems are essentially based on the quantification of the observation that under generic conditions and in a finite neighborhood of every point, the corresponding part of the sequence (u_n) will be "asymptotically uniformly distributed" in $[-1, 1]$. To show this, we first need to switch to an equivalent form of the modulator recursion in terms of rotations of the unit circle and then analyze the error expression more carefully. Next section deals with this:

2. TRANSFORMATION TO THE ROTATION MAP AND THE ERROR TERM

An equivalent form of the modulator recursion, as also derived in [3], is the following: When $\|x\|_\infty < 1$, the u_n stay bounded in $[-1, 1)$ and upon defining $w_n = (u_n + 1)/2$ which scales the range of the sequence to $[0, 1)$, we get

$$\begin{aligned} w_n &= \ < w_{n-1} + \tfrac{1}{2}(x(\tfrac{n}{\lambda}) + 1) > & (8)\\ \mathbf{q}_n &= x(\tfrac{n}{\lambda}) - 2(w_n - w_{n-1}), & (9) \end{aligned}$$

where $<\alpha> = \alpha - \lfloor \alpha \rfloor$ denotes the fractional part of α. Note that (8) defines a time-varying rotation map on the unit circle $\mathbf{T} = \mathbf{R}/\mathbf{Z}$ identified with the unit interval $[0, 1)$.

Clearly, (8) and (9) allow for defining w_n and $\mathbf{q}_n$ for all $n \in \mathbf{Z}$ (including negative n) although this was not explicit in (5) and (6). We also see that it is possible to find the "state" u_{n-1} of the modulator just by knowing u_n and $x(\frac{n}{\lambda})$.

Let us define the following error terms:

$$\begin{aligned} \epsilon_\lambda(t) &= x(t) - \tilde{x}_\lambda(t), & (10)\\ \epsilon_\lambda^{(1)}(t) &= x(t) - \frac{1}{\lambda}\sum_n x(\tfrac{n}{\lambda})h(t - \tfrac{n}{\lambda}), & (11)\\ \epsilon_\lambda^{(2)}(t) &= \frac{1}{\lambda}\sum_n (x(\tfrac{n}{\lambda}) - \mathbf{q}_n)h(t - \tfrac{n}{\lambda}). & (12) \end{aligned}$$

Then, clearly $\epsilon_\lambda = \epsilon_\lambda^{(1)} + \epsilon_\lambda^{(2)}$.

Note that $\epsilon_\lambda^{(1)}$ is identically zero if $\hat{h}$ is in the class of cut-off functions described by (2). For more general h, the Fourier pair

$$\frac{1}{\lambda}\sum_n x(\tfrac{n}{\lambda})h(\cdot - \tfrac{n}{\lambda}) \leftrightarrow \hat{h}(\cdot)\sum_n \hat{x}(\cdot + 2\pi\lambda n) \tag{13}$$

provides

$$\|\epsilon_\lambda^{(1)}\|_\infty \leq \frac{1}{2\pi}\|\hat{x}(\cdot) - \hat{h}(\cdot)\sum_n \hat{x}(\cdot + 2\pi\lambda n)\|_{\mathbf{M}(\hat{\mathbf{R}})}, \tag{14}$$

where we use $\|\mu\|_{\mathbf{M}(\hat{\mathbf{R}})} := \int |\mathrm{d}\mu|$ to denote the total mass of a measure μ on $\hat{\mathbf{R}}$. Noting that supp $\hat{x} \subset [-\pi, \pi]$, the right hand side of (14) is bounded by

$$\frac{1}{2\pi}\|\hat{x}\|_{\mathbf{M}(\hat{\mathbf{R}})}\Big(\|(1-\hat{h})\chi_{[-\pi,\pi]}\|_\infty + \sum_{n\neq 0}\|\hat{h}\chi_{[-\pi,\pi]-2\pi\lambda n}\|_\infty\Big). \tag{15}$$

This bound is sharp in the sense that it is equal to zero for the cut-off filters defined by (2). For other h, (15) gives an estimate of $\epsilon_\lambda^{(1)}$. Note that from a "filter design" point of view, minimizing (15), together with other constraints, can be considered to be a design criterion. We will see in the following how the second error term $\epsilon_\lambda^{(2)}$ is related to h.

Let the forward difference operator Δ_δ be defined by $\Delta_\delta f(u) := f(u+\delta) - f(u)$. Then, using (9) and summing by parts,

$$\begin{aligned} \epsilon_\lambda^{(2)}(t) &= \frac{2}{\lambda}\sum_n (w_n - w_{n-1}) h(t - \tfrac{n}{\lambda}) && (16)\\ &= \frac{2}{\lambda}\sum_n w_n \Delta_{1/\lambda} h(t - \tfrac{n+1}{\lambda}) && (17)\\ &= \frac{2}{\lambda}\sum_n (w_n - \tfrac{1}{2}) \Delta_{1/\lambda} h(t - \tfrac{n+1}{\lambda}). && (18)\end{aligned}$$

We note that, in the above calculation, there is no contribution from the boundary terms of the partial summation because w_n is bounded and the filters h we consider are zero at infinity with sufficient decay. Now and for the rest of the paper we fix t, the point at which we evaluate the error. Define $N_0 = \lfloor \lambda t \rfloor$ and an associated sequence $(W_n)_{n\in\mathbf{Z}}$ by

$$W_n - W_{n-1} = w_n - \tfrac{1}{2}, \qquad W_{N_0} = 0. \qquad (19)$$

Substituting this into (18) and after another summation by parts, we get

$$\begin{aligned} \epsilon_\lambda^{(2)}(t) &= \frac{2}{\lambda}\sum_n W_n \, \Delta_{1/\lambda}^2 h(t - \tfrac{n+2}{\lambda}) && (20)\\ &= \frac{2}{\lambda}\sum_n W_{N_0+n} \Delta_{1/\lambda}^2 h(t - t_0 - \tfrac{n+2}{\lambda}), \end{aligned}$$

where we have defined $t_0 = N_0/\lambda$. Using the bound

$$|\Delta_\delta^2 h(u)| \le 2\delta \int_u^{u+2\delta} |h''(s)| \, ds, \qquad (21)$$

we get

$$|\epsilon_\lambda^{(2)}(t)| \le \frac{4}{\lambda^2}\sum_n |W_{N_0+n}| \int_{t-t_0-\frac{n+2}{\lambda}}^{t-t_0-\frac{n}{\lambda}} |h''(s)| \, ds. \qquad (22)$$

This expresses how the choice of h influences the second error term. Note that, by the definition of $t_0 = \lfloor \lambda t \rfloor/\lambda$, we have $|t - t_0| \le 1/\lambda$. Thus, each of the integrals in (22) is taken over an $O(1/\lambda)$ neighborhood of $-n/\lambda$ and will be of order $\frac{1}{\lambda}|h''(-n/\lambda)|$. To get a good estimate of $|\epsilon_\lambda^{(2)}(t)|$, we need to know the order of growth of $|W_{N_0+n}|$ in n. This is studied in the next section:

3. BOUNDING W_{N_0+n}: THE UNIFORM DISTRIBUTION OF (w_m)

In this section, we will prove the following lemma:

Lemma 1 *Assume $x'(t_0) \neq 0$. Then*

$$|W_{N_0+n}| \le \min\left\{ \frac{|n|}{2}, C_1\left(\lambda^{2/3} + \frac{\lambda^{1/2}}{|x'(t_0)|^{1/2}}\right)\right\} \qquad (23)$$

for all $|n| \le C_2|x'(t_0)|\lambda$ and $\lambda \ge C_3|x'(t_0)|^{-1}$, where C_1, C_2 and C_3 are absolute positive constants.

Note: As is customary, $C, C', C_1, C_2, \ldots$ are notations for generic constants that may change value from one proof to another; constants of different value occuring in the same argument will be distinguished by different indices.

Proof. We will prove only the case $n > 0$. The other case is very similar. From the definition of W_n in (19) one has the following expression for W_{N_0+n}:

$$W_{N_0+n} = \sum_{m=N_0+1}^{N_0+n} (w_m - \frac{1}{2}). \qquad (24)$$

That $|W_{N_0+n}| \le |n|/2$ is trivial. The other estimate has to do with how well the sequence (w_m) for the given range of indices is distributed in $[0,1]$. Note that if x were a constant function with an irrational value, then the recursion (9) would result in a uniformly distributed sequence. We expect to have a similar result when $x(\frac{n}{\lambda})$, hence the amount of rotation at each time step, is changing slowly. The machinery to quantify this idea comes from the *theory of uniform distribution*.

The first inequality we use is the following combination of *Koksma's inequality* and *Erdős-Turán inequality* [7]: For any finite sequence $(u_n)_{n\in I}$ in $[0,1]$,

$$\Big|\sum_{n\in I} u_n - \frac{1}{2}|I|\Big| \le \frac{1}{2}\min_{K>0}\Big(\frac{|I|}{K+1} + 3\sum_{k=1}^{K}\frac{1}{k}\Big|\sum_{n\in I} e^{2\pi i k u_n}\Big|\Big). \qquad (25)$$

We apply this inequality to the finite sequence $w_{N_0+1}, \ldots, w_{N_0+n}$. The next step is to bound the exponential sums that appear in this expression. Define

$$S_k(n) = \sum_{m=N_0+1}^{N_0+n} e^{2\pi i k w_m}. \qquad (26)$$

A trivial bound on such an exponential sum is given simply by the number of terms; when the exponents of the terms in an exponential sum are (the fractional parts of) the samples of a smooth function it is usually possible to find better bounds, by using the stationary phase methods of van der Corput. To use this approach, we first need to write the w_n as the fractional parts of the samples of some smooth function. That is,

$$w_m = \,<\varphi(m)>, \qquad m = N_0, N_0+1, \ldots, N_0+n. \qquad (27)$$

We construct such a φ in [6]; see also Proposition 1 below. Once w_m is rewritten in this form, we replace the exponential sum by something that is easier to estimate. We use the following truncated Poisson summation formula from [7], which is originally due to van der Corput:

Theorem 3 *(Truncated Poisson) Let f be a real-valued function and suppose that f' is continuous and increasing on*

$[a,b]$. *Put* $\alpha = f'(a)$, $\beta = f'(b)$. *Then*

$$\sum_{a\le m\le b} e^{2\pi i f(m)} = \sum_{\alpha-1\le\nu\le\beta+1} \int_a^b e^{2\pi i (f(\tau)-\nu\tau)} d\tau + O(\log(2+\beta-\alpha)). \tag{28}$$

Remark 1: If f' is decreasing, then by taking the complex conjugate of the above expression applied to $-f$, one finds the same expression with α and β switched.

Remark 2: We shall apply Theorem 3 to $f = k\varphi$, where k will range from 1 to K, and $a = N_0 + 1$, $b = N_0 + n$. Note that the requirement of the monotonicity of φ' will restrict the size of n.

Let's look at the defining relation for φ. The recursion (8) suggests

$$\varphi(m) = \varphi(m-1) + y(\tfrac{m}{\lambda}), \tag{29}$$

with the initial condition $\varphi(N_0) = w_{N_0}$ and where $y := (x+1)/2$. This defines φ at integers, and is consistent with (27). We prove the following two propositions in [6]:

Proposition 1 *There exists an analytic function* φ *defined on the whole real line which satisfies* $<\varphi(m)> = w_m$ *for all* $m \ge N_0$. *Moreover,*

$$|\varphi^{(s+1)}(N_0+\tau) - \tfrac{1}{\lambda^s} y^{(s)}(t_0+\tfrac{\tau}{\lambda})| \le 2\big(\tfrac{\pi}{\lambda}\big)^{s+1} e^{\tau\pi/\lambda} \tag{30}$$

for all $s \ge 0$ *and* $\tau \ge 0$.

Proposition 2 *Let* $x'(t_0) \ne 0$. *If* $\lambda \ge 4\pi^2 e^{\pi/2} |x'(t_0)|^{-1}$, *then*

$$\frac{|x'(t_0)|}{8\lambda} \le |\varphi''(N_0+\tau)| \le \frac{|x'(t_0)|}{\lambda}, \tag{31}$$

for $0 \le \tau \le \lambda |x'(t_0)|/4\pi$.

Finally, we borrow the following well-known theorem from [8], which also goes back to van der Corput:

Theorem 4 *Suppose* ϕ *is real-valued and smooth in* (a,b), *and that* $|\phi^{(r)}(t)| \ge \mu$ *for all* $t \in (a,b)$ *and for a positive integer* r. *If* $r=1$, *suppose additionally that* ϕ' *is monotonic. Then there exists an absolute constant* c_r *such that*

$$\Big| \int_a^b e^{i\phi(t)}\, dt \Big| \le c_r \mu^{-1/r}. \tag{32}$$

We now have the tools to prove Lemma 1. First, we bound $S_k(n)$ using Theorem 3: For each k, we set $f = k\varphi$, $a = N_0 + 1$, and $b = N_0 + n$. Then $\alpha = k\varphi'(N_0+1)$ and $\beta = k\varphi'(N_0+n)$. Suppose $n \le \lambda |x'(t_0)|/4\pi$ as in the hypothesis of the Lemma. It follows from Proposition 2 that the number of terms in the right hand side of (28) is bounded by

$$\begin{aligned} |\beta - \alpha + 3| &= |k\varphi'(N_0+n) - k\varphi'(N_0+1) + 3| \\ &\le 3 + k(n-1) \sup_{1\le\tau\le n} |\varphi''(\tau)| \\ &\le 3 + k(n-1)\frac{|x'(t_0)|}{\lambda}. \end{aligned} \tag{33}$$

Now, we estimate the size of each exponential integral in (28). The second derivative of the function in the exponent is $2\pi k\varphi''$, which by Proposition 2 is bounded away from zero. Thus, we set $r = 2$ and $\mu = 2\pi k|x'(t_0)|/(8\lambda)$ in Theorem 4. This bounds each of the integrals by

$$C(k|x'(t_0)|/\lambda)^{-1/2}$$

for some absolute constant C. Combining this with the bound on the number of terms that we have just found in (33), we get

$$\begin{aligned} |S_k(n)| &\le C(k|x'(t_0)|/\lambda)^{-1/2} |\beta-\alpha+3| \\ &\quad + O(\log(2+\beta-\alpha)) \\ &\le C_1 n \Big(\frac{k}{\lambda}\Big)^{1/2} + C_2 \Big(\frac{k}{\lambda}\Big)^{-1/2} |x'(t_0)|^{-1/2} \\ &\quad + O(\log(2+k)), \end{aligned} \tag{34}$$

where in the first term we have made use of the fact that $|x'| \le \pi$ and in the logarithmic term that $|\beta - \alpha| \le k$ for the given range of n.

We proceed by carrying out the optimization of the right hand side of (25). The expression to be minimized over integer K is

$$\frac{n}{K+1} + 3\sum_{k=1}^{K} \frac{1}{k} |S_k(n)|, \tag{35}$$

which by (34) is dominated by

$$\frac{n}{K} + C_1 n \lambda^{-1/2} K^{1/2} + C_2 \lambda^{1/2} |x'(t_0)|^{-1/2} + C_3 (\log K)^2. \tag{36}$$

One can find a sub-optimal minimizer value of K for this expression by carrying out the minimization ignoring the logarithmic term. This results in $K \asymp \lambda^{1/3}$. To simplify the expressions, we further bound n by λ, which will not affect our result, and also absorb the logarithmic term in the rest. We also prefer to use a single constant (say the maximum of them) for convenience. The final result is

$$|W_{N_0+n}| \le C\Big(\lambda^{2/3} + \frac{\lambda^{1/2}}{|x'(t_0)|^{1/2}}\Big). \tag{37}$$

This finishes the proof of the lemma of this section.

4. THE PROOF OF THEOREM 1 FOR TWO DIFFERENT TYPES OF FILTERS

4.1. Finite support in frequency

Fix a filter $\hat{h}_0$ that is C^∞, supported on $[-c_0\pi, c_0\pi]$ for some fixed $c_0 > 1$ and which equals 1 on $[-\pi, \pi]$. This makes h_0 a Schwartz function, i.e., h_0 is also C^∞ and its derivatives have rapid decay: There are constants $C_N^{(l)}$ for all $N \geq 0$ and $l \geq 0$ such that

$$|h_0^{(l)}(t)| \leq \frac{C_N^{(l)}}{(1+|t|)^N}. \quad (38)$$

We define h_λ by

$$h_\lambda(t) = \Omega_\lambda h_0(\Omega_\lambda t) \quad (39)$$

for $\lambda \geq c_0$ and some as yet unspecified $\Omega_\lambda \in [1, \lambda/c_0]$. Then $\hat{h}_\lambda(\xi) = \hat{h}_0(\xi/\Omega_\lambda)$ and hence supp $\hat{h}_\lambda \subset [-\lambda\pi, \lambda\pi]$, and $\hat{h}_\lambda(\xi) = 1$ for $\xi \in [-\Omega_\lambda\pi, \Omega_\lambda\pi]$. Certainly, this makes $\epsilon_\lambda^{(1)}(t) = 0$ for all t and we are left with $\epsilon_\lambda^{(2)}(t)$.

We start with the bound we found in (22). For $|W_{N_0+n}|$, we have two estimates: One is $|n|/2$, which is always valid, and the other is the one given by Lemma 1, which is valid for $|n| \leq C_2|x'(t_0)|\lambda$, where C_2 is as given in the lemma. Set $c = C_2|x'(t_0)|$. Thus,

$$\begin{aligned}|\epsilon_\lambda^{(2)}(t)| \quad &\leq \quad \frac{4}{\lambda^2}\Big(\sum_{|n|\leq c\lambda} + \sum_{|n|>c\lambda}\Big) \\ &\qquad \Big(|W_{N_0+n}| \int_{t-t_0-\frac{n+2}{\lambda}}^{t-t_0-\frac{n}{\lambda}} |h''_\lambda(s)|\,ds\Big). \quad (40)\end{aligned}$$

We bound $|W_{N_0+n}|$ by $O(c^{-1/2}\lambda^{2/3})$ in the first sum and by $|n|/2$ in the second sum in (40). h''_λ relates to h''_0 by scaling: $h''_\lambda(t) = \Omega_\lambda^3 h''_0(\Omega_\lambda t)$. This says that $\|h''_\lambda\|_1 = \Omega_\lambda^2\|h''_0\|_1$ and also lets us use (38) to estimate the decay of h''_λ. Thus, for any $N \geq 3$, it follows that

$$\begin{aligned}|\epsilon_\lambda^{(2)}(t)| \quad &\leq \quad \frac{4}{\lambda^2}\Big(O(c^{-1/2}\lambda^{2/3})\sum_n \int_{t-t_0-\frac{n+2}{\lambda}}^{t-t_0-\frac{n}{\lambda}} |h''_\lambda(s)|\,ds \\ &\qquad + \sum_{|n|>c\lambda} \frac{|n|}{2}\cdot\frac{2}{\lambda}\cdot\frac{\Omega_\lambda^3 C_N^{(2)}}{(1+\Omega_\lambda|n|\lambda^{-1})^N}\Big) \\ &\leq \quad \frac{4}{\lambda^2}\Big(O(c^{-1/2}\lambda^{2/3})\Omega_\lambda^2\|h''_0\|_1 \\ &\qquad + C_N^{(2)}\Omega_\lambda^{-N+3}\lambda^{N-1}O((c\lambda)^{-N+2})\Big) \\ &\leq \quad C(N, |x'(t_0)|)\Big(\Omega_\lambda^2\lambda^{-4/3} + \Omega_\lambda^{-N+3}\lambda^{-1}\Big).\end{aligned}$$

For any arbitrarily small fixed $\eta > 0$, we set $\Omega_\lambda = \lambda^{\eta/2}$ and choose $N = N(\eta)$ such that $(N-3)\eta/2 > 1/3$. This results in the estimate

$$\epsilon_\lambda(t) \leq C(\eta, |x'(t_0)|)\lambda^{-4/3+\eta}. \quad (41)$$

Remark: Since $|t - t_0| \leq 1/\lambda$, and $\lambda > C|x'(t_0)|^{-1}$, it follows that $x'(t) \asymp x'(t_0)$ if the constant C is chosen to be sufficiently large. Hence one can replace t_0 by t above and also in the non-vanishing derivative condition.

This proves the main theorem of this paper.

4.2. Finite support in time

In this subsection, we construct filters that have finite supports in time, i.e., that are FIR, for which we will prove the same result. Similar to the previous subsection, we set h_λ as in (39) but now define h_0 to be

$$\hat{h}_0(\xi) = 1 - \big[1 - \big(\frac{\sin\xi/2}{\xi/2}\big)^{K_0}\big]^{K_1}, \quad (42)$$

where K_0 and K_1 are integers to be determined. The filter h_λ has a finite support in time; in fact, h_λ has the following closed form in terms of B-splines:

$$h_\lambda(t) = \sum_{k=1}^{K_1}(-1)^{k-1}\binom{K_1}{k}\Omega_\lambda B_{kK_0}(\Omega_\lambda t), \quad (43)$$

where

$$B_1 = \chi_{[-\frac{1}{2},\frac{1}{2}]} \quad (44)$$

and

$$B_r = B_1 * ... * B_1, \quad (45)$$

the r-fold convolution of B_1 with itself. Then

$$\text{supp } h_\lambda = \Big[-\frac{K_0K_1}{2\Omega_\lambda}, \frac{K_0K_1}{2\Omega_\lambda}\Big].$$

Choose K_0 to be an even integer greater than or equal to 2 and fix it. Using the elementary inequality

$$\Big(\frac{\sin\xi/2\Omega_\lambda}{\xi/2\Omega_\lambda}\Big)^{K_0} \quad \geq \quad 1 - \frac{K_0}{6}\Big(\frac{\xi}{2\Omega_\lambda}\Big)^2, \quad (46)$$

we get

$$\big\|(1-\hat{h}_\lambda)\chi_{[-\pi,\pi]}\big\|_\infty \leq \Big(\frac{\pi^2K_0}{24\Omega_\lambda^2}\Big)^{K_1}. \quad (47)$$

It also easily follows that

$$\hat{h}_\lambda(\xi) \quad \leq \quad K_1\Big(\frac{2\Omega_\lambda}{\xi}\Big)^{K_0}$$

so that

$$\sum_{n\neq 0}\big\|\hat{h}_\lambda\chi_{[-\pi,\pi]+2\pi\lambda n}\big\|_\infty \quad \leq \quad 2K_1\zeta(K_0)\Big(\frac{2\Omega_\lambda}{\pi\lambda}\Big)^{K_0}, \quad (48)$$

where $\zeta(s) = \sum_n n^{-s}$ is the Riemann zeta function. So, setting $\Omega_\lambda = \lambda^{\eta/2}$ for $0 < \eta < 1$ and using (15), we get

$$\begin{aligned}\big\|\epsilon_\lambda^{(1)}\big\|_\infty \quad &\leq \quad \frac{1}{2\pi}\|\hat{x}\|_{\mathbf{M}(\hat{\mathbf{R}})}\Big(C_1(K_0,K_1)\lambda^{-\eta K_1} \\ &\qquad + C_2(K_0,K_1)\lambda^{-(1-\eta/2)K_0}\Big). \quad (49)\end{aligned}$$

Given K_0 and K_1, the support of h_λ is decreasing to zero. Hence, for all sufficiently large λ, we can use the nontrivial part of the estimate of Lemma 1 that is valid for $|n| \le C|x'(t_0)|\lambda$. Using it in (22), we get

$$|\epsilon_\lambda^{(2)}(t)| \le C(x'(t_0))\left\|h''_\lambda\right\|_1 \lambda^{-4/3}. \qquad (50)$$

Let $C_3(K_0, K_1)$ denote the L^1-norm of h''_0. Then, clearly $\left\|h''_\lambda\right\|_1 = C_3(K_0, K_1)\lambda^\eta$. Hence

$$|\epsilon_\lambda^{(2)}(t)| \le C_3(K_0, K_1) C(x'(t_0))\lambda^{-4/3+\eta}. \qquad (51)$$

Fix a small $\eta > 0$, and choose $K_1 = K_1(\eta) \ge 4/3\eta$. Due to (49), this bounds $\left\|\epsilon_\lambda^{(1)}\right\|_\infty$ by $C(K_0, K_1)\lambda^{-4/3}$. Thus, combining the two bounds and following the remark in §4.1, it follows that

$$\epsilon_\lambda(t) \le C(\eta, |x'(t)|)\lambda^{-4/3+\eta}. \qquad (52)$$

5. THE PROOF OF THEOREM 2

The proof will be an easy corollary of the analysis in Section 2. Let x be the constant function with value α in $[-1, 1]$. Then, the sequence $(\mathbf{q}_n)$ becomes a special "binary expansion" of α. Using the notation of (45), we set $h = B_2$, that is,

$$h(t) = \begin{cases} 1 - |t| & \text{if } |t| \le 1 \\ 0 & \text{otherwise.} \end{cases} \qquad (53)$$

Since h satisfies

$$\sum_n h(t - \tfrac{n}{\lambda}) = \lambda \qquad (54)$$

for all t and integer λ, we see that $\epsilon_\lambda^{(1)}$ is identically zero. We look at (22). It follows from (8) that (w_n) is the sequence of fractional parts of $w_0+\beta n$, where $\beta := (\alpha+1)/2$. The exact order of the growth of the quantity W_{N_0+n} usually depends on the Diophantine properties of β, and has been studied in great detail in the literature (see, e.g. [9]). However, for the sake of simplicity, we prefer to use the following bound that is valid for almost all numbers with respect to the Lebesgue measure [9, 10]:

$$\Big|\sum_{m=1}^{n} \big(<a + \beta m> - \frac{1}{2}\big)\Big| \le C_\beta (\log n)^{2+\gamma} \qquad (55)$$

for any $\gamma > 0$ and a, where the constant depends on β.

Now, the integral expression on the right hand side of (22) is zero for $|n| > \lambda + 3$ and we bound W_{N_0+n} by $C_\beta(\log\lambda)^{2+\gamma}$ for all $|n| \le \lambda + 3$. Note that (21) continues to hold once we regard h'' as a measure, which, in our case is a linear combination of three Dirac masses. Summing over n and replacing $\left\|h''\right\|_{\mathbf{M}(\mathbf{R})}$ with an absolute constant, we finally get

$$\left\|\epsilon_\lambda^{(2)}\right\|_\infty \le C_\beta(\log\lambda)^{2+\gamma}\lambda^{-2}, \qquad (56)$$

which proves Theorem 2.

6. ACKNOWLEDGMENTS

The author would like to thank I. Daubechies for her support and helpful advice during this work, and P. Sarnak for suggesting [7] for number-theoretic results of possible relevance to this problem. The author would also like to thank W. Schlag for pointing out the steps (18) through (20), which simplify an earlier version of this paper and which led to a generalization of the class of filters that can be used.

7. REFERENCES

[1] S.R. Norsworthy, R. Schreier, G.C. Temes, eds., *Delta-Sigma Data Converters*, IEEE Press, 1997.

[2] J.C. Candy, G.C. Temes, eds., *Oversampling Delta-Sigma Data Converters*, IEEE Press, 1992.

[3] R.M. Gray, "Spectral Analysis of Quantization Noise in a Single-Loop Sigma-Delta Modulator with dc Input", *IEEE Transactions on Communications, Vol. 37, June 1989*.

[4] R.M. Gray, W. Chou, P.W. Wong, "Quantization Noise in Single-Loop Sigma-Delta Modulation with Sinusoidal Inputs", *IEEE Transactions on Communications, Vol. 37, Sept. 1989*.

[5] I. Daubechies, R. DeVore, "Reconstructing a Bandlimited Function From Very Coarsely Quantized Data: I. A Family of Stable Sigma-Delta Modulators of Arbitrary Order", preprint.

[6] C.S. Güntürk, "Reconstructing a Bandlimited Function From Very Coarsely Quantized Data: II. Improving the Error Estimate for First Order Sigma-Delta Modulators", preprint.

[7] H.L. Montgomery, *Ten Lectures on the Interface Between Analytic Number Theory and Harmonic Analysis*, AMS, 1994.

[8] E.M. Stein, *Harmonic Analysis: Real-Variable Methods, Orthogonality, and Oscillatory Integrals*, Princeton University Press, 1993.

[9] M. Drmota, R.F. Tichy, *Sequences, Discrepancies and Applications*, Springer, 1997.

[10] S. Lang, *Introduction to Diophantine Approximation*, Springer-Verlag, 1995.