BOUNDEDNESS AND APERIODICITY OF COMMERCIAL SIGMA DELTA MODULATORS

Henri Huijberts[1], Alexey Pavlov[2], Josh Reiss[3]

[1]Department of Engineering, [3]Department of Electronic Engineering
Queen Mary, University of London
Mile End Road, London, E14NS U.K.
[1]H.J.C.Huijberts@qmul.ac.uk, [3]josh.reiss@elec.qmul.ac.uk

[2]Department of Engineering Cybernetics
Norwegian University of Science and Technology
Trondheim, NO-7491Norway
Alexey.Pavlov@itk.ntnu.no

Abstract: Sigma delta modulation is a popular form of A/D and D/A conversion. This nonlinear device exhibits a high degree of complex nonlinear behaviour, including chaotic dynamics. One of the main unsolved problems in the theory of sigma delta modulation concerns the ability to analytically derive conditions for the boundedness of solutions of a high order sigma delta modulator (SDM). In this work, we describe how a sigma delta modulator may be rephrased within the context of systems theory. We present several theoretical results concerning bounded solutions of general high order SDMs, including necessary and sufficient conditions for the lack of a finite escape time, necessary conditions for bounded solutions based on the nature of the output sequences, and topological properties of the solutions, which are a precursor to the study of chaotic solutions of SDMs. *Copyright © 2005 IFAC*

Keywords: Analog/Digital Converters, Modulators, Nonlinear Systems, Stability, Control Theory.

1. INTRODUCTION

1.1. Background

Sigma delta modulation is a popular form of A/D and D/A conversion. The technique has provided powerful means for converting analog to digital signals and vice versa with low circuit complexity and large robustness against circuit imperfections. As a result of this, 1-bit sigma–delta based analog-to-digital (A/D) and digital-to-analog (D/A) converters are widely used in audio applications, such as cellular phone technology and high-end stereo systems.

Sigma–delta modulation, originally conceived by De Jager (1952), is a well-established technique. However, theoretical understanding of the concept is very limited (Norsworthy, *et al.*, 1997). Important progress in the understanding of the dynamical systems properties of SDMs has lead to a description of their chaotic behaviours (Feely, 1997; Dunn and Sandler, 1996; Reiss and Sandler, 2001), a useful linearization technique (Ardalan and Paulos, 1987) and a framework for describing their periodic behaviour (Reefman, *et al.*, 2005). Yet in all these developments, there is no unified description of SDMs. Instead, several models are provided, each of which describes some aspects of an SDM to a certain accuracy.

In this work, we frame the dynamic behaviour of sigma delta modulators within the context of systems theory. The focus is on the characterization of the boundedness and aperiodicity of solutions. The analysis is intended to describe a large number of feedforward or interpolative SDM topologies (Norsworthy, *et al.*, 1997), as used in the design of commercial SDMs.

1.2. Motivation

Figure 1 depicts a 5th order, feedforward SDM. It may be implemented in either digital or analog circuitry. An input signal u is fed into the system, and passed through 5 discrete-time integrators. The output of each integrator is multipled by a coefficient c_i, the results are summed, and then quantised depending on the sign of the sum. This quantised output is fed back to the input. The coefficient vector $C=(c_1,\ldots,c_5)$ serves to shape the noise of quantisation away from the frequency band of interest. Thus the SDM acts as a filter, but with the quantiser introducing a high degree of nonlinearity.

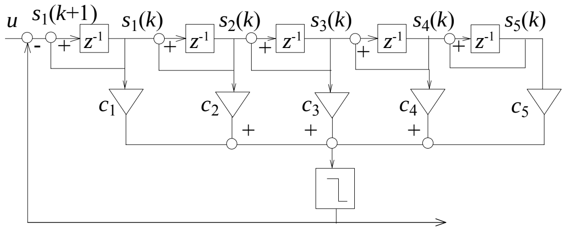

Figure 1. Implementation of a commercial fifth order SDM.

It is well-established that a first order SDM will produce bounded behaviour for input magnitudes less than 1, and similar results can be shown for some 2nd order SDM designs (Farrell and Feely, 1998). However, higher order SDMs may produce divergent behaviour, such that the magnitude of the input to the quantiser becomes exceedingly large and the output no longer tracks the input. To illustrate this, Figure 2 depicts the magnitude of the quantiser input for the implementation of a 5th order feedforward SDM (Reefman, *et al.*, 2005), intended to be used for analog to digital conversion in audio applications. The SDM is lowpass, and has a corner frequency of 80kHz, for a sample rate of 64x44.1 kHz. We have assumed constant input *u* with all initial conditions set to zero. It can easily be verified that the dynamics are bounded for an input of 0.7, unbounded for input 0.9, and unbounded with finite escape time (in a sense defined later in the paper) for an input of 1.1.

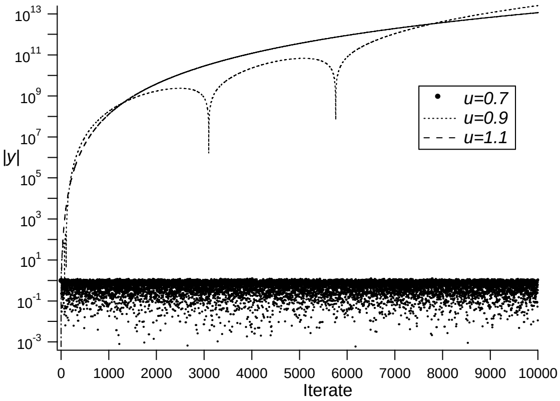

Figure 2. The magnitude of the quantiser input for different values of input signal *u*.

Besides boundedness of solutions, the avoidance of low-period solutions is also an important design consideration for commercial SDMs, primarily because they represent high frequency tones in the output bitstream that were not present in the input signal. It is therefore important to be able to identify initial conditions that will lead to aperiodic, preferably chaotic, behaviour. Our goal in this paper is to present a mathematical framework, based on systems theory, to describe the boundedness and aperiodicity of solutions of SDMs. We will focus on DC inputs, since this represents the most relevant practical situation.

The organisation of the rest of this paper is as follows. Section 2 introduces concepts from control theory, as applied to the general class of maps to which most SDMs belong. Such an introductory section is necessary in order to bridge the gap in terminology and understanding between the SDM designers and the systems theorists. It also introduces the concepts of finite escape time and well-posedness for SDMs and gives a canonical description of the type of SDMs that are studied in this paper. Section 3 provides several interesting results concerning the boundedness properties of solutions of arbitrary SDMs. Notably, we derive a proof that the dynamics of an SDM are bounded only if the output bitstream is also bounded. This allows the boundedness properties to be rephrased in terms of the constraints on the dynamics imposed by output sequences. Section 4 provides further results which lead to a topological understanding of the nature of bounded solutions and an indication of the existence of aperiodic and chaotic solutions. Finally, Section 5 provides conclusions and discussion of directions for future research.

2. SYSTEM THEORETIC PROPERTIES AND NORMAL FORM FOR SDMS

In this section we describe some system theoretic properties of SDMs which will be of importance when studying the boundedness of solutions of SDMs. For further system theoretic background we refer the reader to (Bernstein, 2005; Chen, 1998).

Consider an *n*-dimensional SDM of the form

$$\begin{cases} \sigma s = As + B(u - \mathrm{sgn}(y)) \\ \qquad y = Cs \end{cases} \tag{1}$$

where *u* denotes the input which is assumed to be constant with $|u| < 1$, *y* denotes the quantiser input, and σs denotes the forward shift of *s*, i.e., $\sigma s(k) = s(k+1)$ for all $k \in \mathbb{Z}$. We assume that *A* is a lower triangular matrix with diagonal elements equal to 1 and *B*,*C* are matrices of appropriate dimensions. Recall that the matrix pair (*A*,*B*) is called controllable if it satisfies one of the following equivalent conditions:

$$\mathrm{rank}(B \quad AB \quad \dots \quad \dots \quad A^{n-1}B) = n \tag{2}$$

or

$$\mathrm{rank}\left(\lambda I - A \quad B\right) = n \text{ for all eigenvalues } \lambda \text{ of } A \tag{3}$$

Further, recall that the matrix pair (*C*,*A*) is called observable if the matrix pair (A^T,C^T) is controllable. With a slight abuse of terminology, we will call the SDM (1) controllable if (*A*,*B*) is controllable while we will call it observable if (*C*,*A*) is observable. Write *B*= col(b_1,… , b_n), *C*= (c_1,… ,c_n), and let $a_{i,j}$ (*i*,*j*=1,…,*n*) denote the entries of *A*. By employing the condition (3) and using the properties of *A*, the following result on controllability and observability is almost immediate.

Proposition 2.1. *The SDM (1) is controllable if and only if* $b_1 \prod_{i=1}^{n-1} a_{i+1,i} \neq 0$, *and it is observable if and only if* $c_n \prod_{i=1}^{n-1} a_{i+1,i} \neq 0$.

In the rest of the paper we will assume that the SDM (1) is observable. This assumption can be made without loss of generality, because nonobservability would imply the presence of internal dynamics that do not contribute to the behavior of the quantiser input y.

Define the transfer function of the SDM (1) by $G(z)=C(zI-A)^{-1}B$ and recall that when (1) is observable we have that $G(z)=q(z)/p(z)$ where $p(z)=\det(zI-A)=(z-1)^n$ and $\deg(p)>\deg(q)$. If we then write $q(z)=\sum_{m=0}^{n-1}q_m z^m$, we have that the quantiser input y satisfies the following difference equation:

$$\sum_{m=0}^{n-1}(-1)^{n-m}\binom{n}{m}y(k+m)=\sum_{m=0}^{n-1}q_m(u-\mathrm{sgn}(y(k+m)))\quad(4)$$

We will say that a solution of (1) has *finite escape time* if there exists a $k^*\in\mathbb{N}$ such that $\mathrm{sgn}(y(k))=\mathrm{sgn}(y(k^*))$ for all $k\geq k^*$. Further, we will call the SDM (1) *well-posed* if it has no solutions with finite escape time. We then have the following result on well-posedness.

Proposition 2.2. *The SDM (1) is well-posed if and only if* $\beta:=\sum_{m=0}^{n-1}q_m>0$.

Proof. (Sketch) Assume that $\beta\leq 0$. Consider a solution of (4) with $y(k)>0$ for $k=0,...,N$, $N>n$. Define $y_k:=y(k)$ for $k=0,...,n-1$. Define $\varepsilon:=\beta(u-1)\geq 0$. It may then be shown that

$$y(k)=\sum_{m=0}^{n-1}P_m(k)y_m+Q(k)\varepsilon\quad(k=n,...,N)\qquad(5)$$

where

$$P_m(k)=\frac{(-1)^{n-m+1}}{m!(n-m-1)!}\prod_{\substack{r=1\\ r\neq n-m}}^{n}(k+r),$$
$$Q(k)=\frac{1}{n!}\prod_{r=0}^{n-1}(k+r)\qquad(6)$$

Thus we see that the solution is polynomial in k with coefficients of $k^0,\ldots,k^{n-1}$ depending linearly on $y_0,\ldots,y_{n-1}$ and coefficient of k^n equal to $\varepsilon\geq 0$. This means that we can choose $y_0,...,y_{n-1}$ in such a way that all coefficients of $y(k)$ are positive. This gives that $y(k)>0$ for all $k\geq 0$, and hence $y(k)$ has finite escape time. This establishes our claim. □

To check well-posedness, the following result can be used.

Lemma 2.3. *If (1) is observable, we have that* $\beta=b_1c_n\prod_{i=1}^{n-1}a_{i+1,i}$.

Proof. Recall that if (1) is observable, it follows from Cramer's Rule that $q(z)=C\,\mathrm{adj}(zI-A)B$ where adj(M) denotes the *adjoint* (see, e.g. (Bernstein, 2005)) of the square matrix M. This then gives that $\beta=q(1)=C\,\mathrm{adj}(I-A)B$. Using the properties of A, it is straightforwardly checked that $\mathrm{adj}(I-A)_{n1}=\prod_{i=1}^{n-1}a_{i+1,i}$, while all other entries of $\mathrm{adj}(I-A)$ are zero. This immediately establishes our claim. □

As a consequence of Proposition 2.1 and Lemma 2.3 we have that the SDM (1) is well-posed only if it is controllable.

In studying general properties of systems, it is often useful to consider canonical forms that are equivalent to a whole class of systems up to a coordinate transformation. In linear systems theory two of the most well-known canonical forms are the so-called controller canonical form and the observer canonical form. However, for SDMs these canonical forms are perhaps not the most insightful canonical forms in terms of the physical interpretation of the SDM. We therefore define a *controller canonical SDM form* as

$$\begin{cases}\sigma s_c=A_c s_c+B_c(u-\mathrm{sgn}(y_c))\\ \qquad y_c=C_c s_c\end{cases}\qquad(7)$$

where A_c is a matrix with diagonal elements equal to 1, $(A_c)_{i+1,i}=1$ ($i=1,\ldots,n-1$) and all other entries zero, $B_c=\mathrm{col}(1,0,\ldots,0)$, and C_c is arbitrary.

Proposition 2.4. *Every controllable SDM* (1) *admits a controller canonical SDM form in the sense that there exists an invertible matrix V such that* $s_c:=V^{-1}s$ *satisfies (7).*

Proof. Assume that (A,B) is controllable and define the matrix $V=(v_1,\ldots,v_n)$ where $v_i=(A-I)^{i-1}B(i=1,\ldots,n)$. By performing elementary column operations, it is straightforwardly shown that $\mathrm{rank}(V)=\mathrm{rank}(B\ \ldots A^{n-1}B)$, which implies that V is invertible. Note that $v_1=B$, which implies that $V^{-1}B=B_c$. It is further straightforwardly shown that $Av_i=v_{i+1}+v_i$ ($i=1,\ldots,n-1$). Also, it follows from the Cayley-Hamilton Theorem (see, e.g. (Bernstein, 2005)) that $(A-I)^{n-1}=0$, which allows one to show that $Av_n=v_n$. All identities obtained above then straightforwardly imply that $V^{-1}AV=A_c$ and hence that $s_c:=V^{-1}s$ satisfies (7).

3. RESULTS ON BOUNDEDNESS OF SOLUTIONS OF SDMS

Assuming $y(1)>0$, for a well-posed SDM it follows from Section 2 that there are infinite sequences N_m^+,N_m^- such that $\mathrm{sgn}(y(k))=1$ for $k\in\{P_r+1,...,P_r+N_{r+1}^+\}$ and $\mathrm{sgn}(y(k))=-1$ for $k\in\{P_r+N_{r+1}^++1,...,P_{r+1}\}$, where $r\in\mathbb{N}$ and $P_r:=\sum_{m=1}^{r}(N_m^++N_m^-)$. Similar sequences can be defined for $y(1)<0$. In this section we will give necessary conditions for boundedness of solutions in terms of these sequences.

Theorem 3.1. *Assume that the SDM (1) is well-posed. Then a solution of (1) is bounded only if the lengths of bit streams associated with the solution are bounded.*

Proof. Consider a solution of (4) with $y(k)>0$ for $k=0,...,N$, $N>n$. Define $y_k:=y(k)$ for $k=0,...,n-1$. Consider the expression for the solution given in (5) and (6), and define the polynomial $P(k)=\sum_{m=0}^{n-1}|P_m(k)|$. Note that for all $k\geq 0$ we have that $P(k) > |P_m(k)|$. Then the fact that $y(N)>0$ implies by (5) that

$$Q(N)\,|\varepsilon| < \left|\sum_{m=0}^{n-1} P_m(N) y_m\right| \leq \sum_{m=0}^{n-1} |P_m(N)|\,|y_m| < P(N)\sum_{m=0}^{n-1} |y_m| \tag{8}$$

which gives that

$$\sum_{m=0}^{n-1} |y_m| > \frac{Q(N)}{P(N)}|\varepsilon| \tag{9}$$

Since deg(Q)=N, deg(P)=n-1, this gives that $\sum_{m=0}^{n-1} |y_m| \to \infty$ as $N \to \infty$. In a similar way it may be shown that solutions become unbounded if the lengths of negative bit streams become unbounded. Thus our claim is established. □

The following result and its related corollaries provide more stringent necessary conditions for bounded solutions.

Theorem 3.2 *Consider a solution of a well-posed SDM with* $|u|<1$. *Define*

$$M_r := \sum_{m=1}^{r} (u-1)N_m^+ + (u+1)N_m^- \tag{10}$$

Then the solution is bounded only if *the sequence* M_k *(* $k \in \mathbb{N}$ *) is bounded. Moreover, for a one-dimensional SDM, this condition is also a sufficient condition for boundedness.*

Proof. Consider a bounded solution $s(k)$ of the SDM. Then obviously we have that $s_1(k)$ is bounded. It is straightforwardly checked that

$$s_1\left(\sum_{m=1}^{r}(N_m^+ + N_m^-)\right) = s_1(0) + \sum_{m=1}^{r}(u-1)N_m^+ + (u+1)N_m^- \tag{11}$$

which establishes our claim. Since for a one-dimensional SDM $s_1(k)$ is the only state space variable, sufficiency for one-dimensional SDMs is immediate. □

Corollary 3.2. *Consider a solution of a well-posed SDM with* $|u|<1$. *If the solution is bounded and* $u \in \mathbb{Q}$, *we have that*

$$(\forall r \in \mathbb{N})(\exists r_1, r_2 \in \mathbb{N} : r < r_1 < r_2)\left(\sum_{m=r_1}^{r_2}(u-1)N_m^+ + (u+1)N_m^- = 0\right) \tag{12}$$

Proof. From the Bolzano-Weierstrass Theorem it follows that the bounded sequence M_r has a convergent subsequence, i.e., there exists a sequence $(\rho_l)_{l\in\mathbb{N}}$ with $\rho_l \to \infty$ as $l \to \infty$ such that $\lim_{l\to\infty} M_{\rho_l}$ exists. As a consequence, we have that

$$(\forall \varepsilon > 0)(\exists l^* \in \mathbb{N})(\forall \sigma, \tau > l^*)(|M_{\rho_\sigma} - M_{\rho_\tau}| < \varepsilon) \tag{13}$$

Assume that $u \in \mathbb{Q}$, and write u = a/b , where $a \in \mathbb{Z}, b \in \mathbb{N}$ and $|a|$ < b. Choose ε< 1/b and let $r \in \mathbb{N}$ be given. Then according to (13) there exist $r<r_1:=\rho_\sigma<r_2:=\rho_\tau$ such that $|M_{r_1} - M_{r_2}| < \varepsilon$, which gives that $|bM_{r_1} - bM_{r_2}| < b\varepsilon < 1$. Noting that $bM_{r_1}, bM_{r_2} \in \mathbb{Z}$, this implies that in fact $|bM_{r_1} - bM_{r_2}| = 0$, which establishes (12). □

Corollary 3.3. *Consider a bounded solution of a well-posed SDM with* $|u|<1$ *and* $u \in \mathbb{Q}$. *Define* $\sigma_{0:}$= *0, and define* ρ_k, σ_k *recursively in the following way:*

$$\begin{aligned} \rho_{k+1} &:= \min\left\{r > \sigma_k \mid (\exists \bar{r} : r < \bar{r} < +\infty)\left(\sum_{m=r}^{\bar{r}}(u-1)N_m^+ + (u+1)N_m^- = 0\right)\right\} \\ \sigma_{k+1} &:= \min\left\{r > \rho_{k+1} \mid \sum_{m=\rho_{k+1}}^{r}(u-1)N_m^+ + (u+1)N_m^- = 0\right\} \end{aligned} \tag{14}$$

Note that from Corollary 4.2 the sequences ρ_k, σ_k *are well-defined and infinite. Then there exists a* $k^* \in \mathbb{N}$ *such that for all* $k \geq k^*$ *we have that* ρ_{k+1}=σ_k+1.

Proof. Define the index sets $\mathcal{I} := \bigcup_{k\in\mathbb{N}}\{\rho_k, ..., \sigma_k\}$, $\mathcal{J} := \mathbb{N} - \mathcal{I}$. Due to the boundedness of the sequence $(M_r)_{r\in\mathbb{N}}$, the sequence $(M_j)_{j\in\mathcal{J}}$ is also bounded. Assume that our claim does not hold. This implies that the index set $\mathcal{J}$is infinite. Using a similar argument as in the proof of Corollary 3.2, this implies that there exist j_1,j_2 with $j_1 < j_2 < +\infty$, such that $\sum_{m=j_1}^{j_2}(u-1)N_m^+ + (u+1)N_m^- = 0$. Since the sequences ρ_k, σ_k are infinite, we have that there exists a $k \in \mathbb{N}$ such that σ_k<j_1< ρ_{k+1}. However, this contradicts the definition of ρ_{k+1}. This establishes our result. □

4. TOPOLOGICAL RESULTS CONCERNING THE BOUNDEDNESS OF SDMS

Throughout this section, we consider an SDM in the controller canonical form (7). For brevity's sake, we will omit the c subscripts in the description of the sigma delta modulator. We denote the vector $\mathrm{col}(0,0,...,1)$ by e_n.

Theorem 4.1 *Let* $s(k), k \geq 0$ *be a bounded solution of the SDM and* $b(k) := \mathrm{sgn}(Cs(k)), k \geq 0$ *be the corresponding bit sequence. Then for any other bounded solution* $\tilde{s}(k)$ *with the same bit sequence, there exists* $\alpha \in \mathbb{R}$ *such that* $\tilde{s}(k) - s(k) = \alpha e_n$ *for all* $k \geq 0$.

Proof: By the conditions of the theorem, $s(k)$ and $\tilde{s}(k)$ satisfy

$$\begin{aligned} \sigma s &= As + B(u-b) \\ \sigma\tilde{s} &= A\tilde{s} + B(u-b) \end{aligned} \tag{15}$$

Therefore, the difference $\xi = s - \tilde{s}$ satisfies the difference equation $\sigma\xi = A\xi$, which gives that $\xi(k) = A^k\xi(0)$. It can be verified (Reefman, *et al.*, 2005) that for $k \geq n$, $A^k = I + \sum_{r=1}^{n-1}\beta_r(k)T_r$ where

$\beta_r(k)=\frac{k!}{r!(k-r)!}$ and the matrices T_r satisfy $(T_r)_{ij}=\delta_{i,j+r}$ (i,j = 1,…n) with δ_{ij} denoting the Kronecker delta. Thus the second component of the vector $\xi(k)$ is given by, $\xi_2(k)=\beta_1(k)\xi_1(0)+\xi_2(0)$. Notice that $\beta_i(k)\to\infty$ as $k\to\infty$, for i=1,…,n-1. Therefore, if $\xi_1(0)\neq 0$, then $\xi_2(k)$ is unbounded. At the same time, $\xi(k)$ is bounded because by the conditions of the theorem both $s(k)$ and $\tilde{s}(k)$ are bounded. Hence, we conclude that $\xi_1(0)=0$. Repeating this reasoning for the remaining components, we conclude that $\xi_i(0)=0$ for i=1,…,n-1, while the last component $\xi_n(0)$ can be arbitrary. Hence, $\xi(0)=\alpha e_n$ for some $\alpha\in\mathbb{R}$. Since $A^k e_n=e_n$, we obtain $\xi(k)=\alpha e_n$ for all $k\geq 0$. This completes the proof.

□

Theorem 4.2 *Consider an SDM in the controller canonical form. Let s*(k), $k\geq 0$ *be a bounded solution of the SDM and b*(k) := sgn(Cs(k)), *be the corresponding bit sequence. Suppose there is* δ> 0 *such that* |Cs(k)|>δ *for all* $k\geq 0$. *Then there exists* ε>0 *such that for all* |α|<ε *the sequence* $\tilde{s}(k):=s(k)+\alpha e_n$ *is a solution of the SDM and the corresponding bit sequence equals b*(k).

Proof: We will show that $\tilde{s}(k)$ satisfies (7). First note that due to the fact that Ae_n=e_n we have that

$$A(s(k)+\alpha e_n)=As(k)+\alpha e_n \qquad (16)$$

Next, assume that y(k) = Cs(k)>0. It then follows from the condition of the theorem that $y(k)>\delta$. Therefore, it follows from the choice of α and ε that $C(s(k)+\alpha e_n)>\delta-|\alpha||Ce_n|>\delta-\varepsilon|Ce_n|>0$. Thus,

$$\operatorname{sgn}(C(s(k)+\alpha e_n))=\operatorname{sgn}(y(k)) \qquad (17)$$

Combining (16) and (17) we then obtain

$$\begin{aligned}&A\tilde{s}(k)+B(u-\operatorname{sgn}(C\tilde{s}(k)))\\&=As(k)+\alpha e_n+B(u-\operatorname{sgn}(y(k))) \qquad (18)\\&=s(k+1)+\alpha e_n=\tilde{s}(k+1)\end{aligned}$$

for all $k\geq 0$. Hence, $\tilde{s}(k)$ is a solution of (7) and it has the bit sequence b(k). □

The conditions of Theorem 2 are satisfied, for example, for periodic solutions s that satisfy $Cs(k)\neq 0$, $k\geq 0$. Thus, we can formulate the following corollary.

Corollary 4.1 *Let s*(k) *be a periodic solution of an SDM such that* $Cs(k)\neq 0$, $k\geq 0$. *Then for all sufficiently small* α, *s*(k) + αe_n *is a periodic solution of the SDM with the same bit sequence.*

We note that a different proof of this corollary was provided in (Reefman, *et al.*, 2005), but that work *only* considered periodic solutions.

Next we present results on generic boundedness properties of solutions of SDMs.

Theorem 4.3 *Consider an SDM with* $n\geq 2$. *The set of all bounded solutions is isomorphic*[1] *to a subset of* $\mathbb{R}^2$.

Proof: We first introduce some notations. By bold font letters we will denote sequences. For example, a solution of the SDM s(k), k = 0,1,…, is denoted by **s**. The bit sequence b(k); k = 0,1,…, is denoted by **b**. Λ denotes the set of all bounded solutions of (7), and B_Λ denotes the set of all bit sequences **b** corresponding to the solutions $\mathbf{s}\in\Lambda$. By the construction of the set B_Λ, for any $\mathbf{b}\in B_\Lambda$ there is a solution $\mathbf{s}_\mathbf{b}^*\in\Lambda$ which has the bit sequence **b**. By Theorem 4.1, any other bounded solution of (7) $\mathbf{s}_\mathbf{b}$ with the same bit sequence **b** can be represented as $\mathbf{s}_\mathbf{b}$ = $\mathbf{s}_\mathbf{b}^*$+ αe_n for some $\alpha\in\mathbb{R}$. (Here, by adding the vector αe_n to the sequence $\mathbf{s}_\mathbf{b}^*$ we mean that this vector is added to every element of the sequence.) Denote $\mathcal{A}_\mathbf{b}$ to be the set of all $\alpha\in\mathbb{R}$ such that **s**= $\mathbf{s}_\mathbf{b}^*$ +αe_n is a bounded solution of (1).

Construct the set $M:=\{\mathbf{s}_\mathbf{b}^*+\alpha e_n,\alpha\in\mathcal{A}_\mathbf{b},\mathbf{b}\in B_\Lambda\}$. By the construction, M=Λ. Denote $\mathcal{P}:=\{(\alpha,\mathbf{b}):\alpha\in\mathcal{A}_\mathbf{b},\mathbf{b}\in B_\Lambda\}$. Thus $\Lambda\cong\mathcal{P}$.

For a bit sequence **b**, define the function r(**b**):=$\sum_{k=1}^{+\infty}2^{-k}\mathbf{b}(k)$. In other words, the bit sequence **b** is a binary representation of the decimal number d=r(**b**). By $D_\Lambda\subset\mathbb{R}$ we denote the set of all numbers d = r(**b**) such that $\mathbf{b}\in B_\Lambda$, i.e., $D_\Lambda=\{d\in\mathbb{R}:d=r(\mathbf{b}),\mathbf{b}\in B_\Lambda\}$. Notice that there is a one-to-one correspondence between a number and its binary representation. Therefore, $B_\Lambda\cong D_\Lambda$. Hence, $\Lambda\cong\mathcal{P}\cong\mathcal{P}^*:=\{(\alpha,d):\alpha\in\mathcal{A}_{r^{-1}(d)},d\in D_\Lambda\}$ where $r^{-1}(d)$ is the binary representation of d. Notice that $\mathcal{P}^*\subset\{(\alpha,d):\alpha\in\mathbb{R},d\in\mathbb{R}\}=\mathbb{R}^2$. This proves the claim of the theorem. □

Roughly speaking, Theorem 4.3 states that the set of all initial conditions corresponding to bounded solutions is not "thicker" than $\mathbb{R}^2$. This allows us to formulate the following corollary.

Corollary 4.2 *Consider an SDM with* $n\geq 3$. *Then for almost all initial conditions the corresponding solutions of SDM are unbounded.*

The result of Corollary 4.2 explains why for two-dimensional SDMs the bounded solutions are relatively abundant, while for higher-dimensional SDMs it is much more difficult to find bounded solutions. It also has the implication that for higher-order SDMs small perturbations may lead to an otherwise bounded solution becoming unbounded.

The next result concerns bounded solutions of SDMs with periodic or asymptotically periodic bit sequences. It is said that a bit sequence **b** is periodic if there exists T > 0 such that **b**(k) =**b**(k+T) for all

[1] We say two sets are isomorphic (denoted by $\cong$), when there is a one-to-one correspondence between the sets.

$k \geq 0$. A bit sequence **b** is called asymptotically periodic if it is periodic after some time instant $N > 0$, i.e., if $\mathbf{b}(k) = \mathbf{b}(k + T)$ for all $k \geq N$.

Theorem 4.4 *Consider an SDM, with* $n \geq 2$. *The set of all bounded solutions with bit sequences that are either periodic or asymptotically periodic is isomorphic to a subset of* $\mathbb{R} \times \mathbb{Q}$.

Proof: Similar to the proof of Theorem 4.3, we introduce the following notations. Let Ξ denote the set of all bounded solutions of (7) with bit sequences that are either periodic or asymptotically periodic. By B_Ξ denote the set of all bit sequences **b** corresponding to the solutions $\mathbf{s} \in \Xi$, and define $D_\Xi = \{d \in \mathbb{R} : d = r(\mathbf{b}), \mathbf{b} \in B_\Xi\}$. As in the proof of Theorem 4.3, it can be shown that $\Xi \cong \mathcal{P}_\Xi^* := \{(\alpha, d) : \alpha \in \mathcal{A}_{r^{-1}(d)}, d \in D_\Xi\}$. By the construction of the set D_Ξ, any number $d \in D_\Xi$ has a periodic (after some digit order number) binary representation. This can happen if and only if d is a rational number. Therefore, $D_\Xi \subset \mathbb{Q}$. Thus

$$\begin{aligned} \Xi \cong \mathcal{P}_\Xi^* &:= \{(\alpha, d) : \alpha \in \mathcal{A}_{r^{-1}(d)}, d \in D_\Xi\} \\ &\subset \{(\alpha, d) : \alpha \in \mathbb{R}, d \in \mathbb{Q}\} \end{aligned} \quad (19)$$

This completes the proof. □

Corollary 4.3 *Consider an SDM with* $n = 2$. *The set of solutions of the SDM starting in any open set of initial conditions contains either an unbounded solution or a bounded solution with a bit sequence that is neither periodic nor asymptotically periodic.*

Proof: Consider some open set of initial conditions $\Theta \subset \mathbb{R}^2$. Suppose that the set of all solutions of an SDM starting in Θ does not contain unbounded solutions (if it does, then the claim is proved). Since Θ is open, it contains an open subset Ω that is isomorphic to $\mathbb{R}^2$. Therefore, the set of solutions starting in Ω is isomorphic to $\mathbb{R}^2$. Assume that all solutions starting in Ω have bit sequences that are asymptotically periodic. By Theorem 4.4, we would then have that this set of solutions would be isomorphic to a subset of $\mathbb{R} \times \mathbb{Q}$ However, this would imply that Ω is also isomorphic to a subset of $\mathbb{R} \times \mathbb{Q}$, which gives a contradiction. □

It follows from Corollary 4.3 that if for a two-dimensional SDM one can identify an open set of initial conditions that lead to bounded solutions, there exist aperiodic solutions amongst these solutions. It is well-known (Farrell and Feely, 1998; Schreier, *et al.*, 1997) that indeed one can identify these sets of initial conditions. Thus, this indicates that there may be chaotic solutions for two-dimensional SDMs.

5. CONCLUSION

The work described herein is concerned with the boundedness of solutions of SDMs. Our approach has been to rephrase the sigma delta modulator as a controllable and observable discrete time system. We have shown that typical SDM designs fall into such a category. We further define an SDM as well-posed if it has no solutions which diverge to infinity in finite escape time. From an SDM design point of view, the input signal u is always confined to a magnitude less than 1, and the coefficient vector C is strictly positive for a lowpass SDM. Thus typical SDM designs are implicitly well-posed. This allows us to show that an SDM yields bounded behaviour only if the length of the associated output bit streams are also bounded. To the best of the authors' knowledge, this result has never before been derived. Its strength lies in that it allows the boundedness properties to be rephrased in terms of the constraints imposed by the output bits. Future research along this direction is concerned with proving the converse, and with identifying bounded solutions for commercial SDM designs.

ACKNOWLEDGMENTS

Research conducted by the second author was partly performed while visiting Queen Mary, University of London and supported by the EPSRC NETCA Network.

REFERENCES

Ardalan S.H. and J. J. Paulos (1987). An Analysis of Nonlinear Behavior in Delta-Sigma Modulators, *IEEE Trans. Circ. Syst. I,* **34**, pp. 593-603.

Bernstein, D.S. (2005). *Matrix Mathematics*. Princeton University Press, Princeton, NJ.

Chen C.-T. (1998). *Linear System Theory and Design*. Oxford University Press, New York.

De Jager, F. (1952). Delta modulation - a method of {PCM} transmission using the one unit code, *Philips Res. Rep.*, **7**, pp. 442-466.

Dunn, C. and M. Sandler (1996). A comparison of Dithered and Chaotic Sigma-Delta Modulators, *J. Audio Eng. Soc.*, **44**, pp. 227-244.

Farrell, R. and O. Feely (1998). Bounding the integrator outputs of second order sigma-delta modulators, *IEEE Trans. Circ. Syst. II,* **45**, pp. 691-702.

Feely O. (1997). A tutorial introduction to non-linear dynamics and chaos and their application to sigma-delta modulators, *Int. J. Circ. Theory Appl.*, **25**, pp. 347-367.

Norsworthy S., R. Schreier and G. Temes (1997). *Delta-Sigma Data Converters*, IEEE Press.

Reefman D., J. Reiss, E. Janssen and M. Sandler (2005). Description of limit cycles in sigma-delta modulators, *IEEE Trans. Circ. Syst. I*, **52**, pp. 1211-1223.

Reiss J. and M.B. Sandler (2001). The Benefits of Multibit Chaotic Sigma Delta Modulation, *CHAOS*, **11**, pp. 377-383.

Schreier R., M. Goodson and B. Zhang (1997). An algorithm for computing convex positively invariant sets for delta-sigma modulators, *IEEE Trans. Circ. Syst. I*, **44**, pp. 38-44.