MULTILEVEL AND NON-IDEAL QUANTIZATION IN Σ—Δ MODULATION

by

Orla Feely and Leon O. Chua

Memorandum No. UCB/ERL M91/54

23 June 1991

MULTILEVEL AND NON-IDEAL QUANTIZATION IN Σ—Δ MODULATION

by

Orla Feely and Leon O. Chua

Memorandum No. UCB/ERL M91/54

23 June 1991

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

Multilevel and non-ideal quantization in $\Sigma-\Delta$ modulation

Orla Feely and Leon O Chua
Department of Electrical Engineering and Computer Sciences, University of California, Berkeley CA 94720, USA

June 23, 1991

Abstract

Oversampled sigma-delta analog-to-digital conversion has been the focus of much attention in recent years. Converters based on this principle have been found to be particularly suited to integrated circuit implementation, due to their simple structure and robustness to circuit imperfection. The analysis of the sigma-delta system is complicated by the presence of a discontinuous nonlinear element, with the result that few rigorous studies of such converters have appeared in the literature.

In this paper we study the operation of the single-loop sigma-delta modulator, consisting of a quantizer and a discrete-time integrator in a feedback loop, with constant input. Our analysis uses the tools of nonlinear dynamics to derive rigorous descriptions of system behavior. In particular we focus on the effects of the discontinuous nonlinear element—the quantizer. In the most basic case the quantizer has two levels. We show that system behavior in this case is unaffected by quantizer offsets only if the integrator is ideal. We then study the effect on system operation of increasing the number of quantization levels, and explain in quantitative terms the structure of the resultant input versus average output characteristic.

1 Introduction

Oversampled Sigma-Delta ($\Sigma - \Delta$) modulation[1] as a method of analog-to-digital conversion has been the focus of much attention in recent years, finding application in digital signal processing systems, voiceband telecommunication systems and commercial compact disc players[2]. These modulators convert an analog input into a low-resolution digital representation at a very high sampling rate. A decoder then decimates and increases the resolution of the output bit stream—in other words, the decoder uses several low-precision digital signal representations to generate each high-precision representation. It has been found that analog-to-digital converters based on this principle are particularly suited to integrated circuit implementation— the analog circuitry required is simple, performance is insensitive to circuit imperfections, and the high speeds required by the oversampling process with audio and other low frequency signals are easily attainable by VLSI technology.

Despite the simple structure of the $\Sigma - \Delta$ modulator, the presence of a discontinuous nonlinear element—a quantizer—in the system means that the analysis is highly non-trivial. As a result, few rigorous analyses of the system have been attempted. Most researchers who study the problem begin by linearizing the nonlinearity and then apply standard linear theory, while others avoid the analysis altogether and concentrate on simulations.

In the simplest oversampled $\Sigma - \Delta$ modulator—the single-loop $\Sigma - \Delta$ system—a one-bit quantizer and a discrete-time integrator are connected in a feedback loop. This basic structure can be modified by adding more feedback loops or increasing the number of quantization levels. In this paper we concentrate only on the single feedback loop structure.

In a recent paper[3] we apply the theory of nonlinear dynamics—in particular that of symbolic dynamics[4]—to the single loop $\Sigma - \Delta$ system with integrator imperfections. The results describe in an exact and rigorous manner the effect of these common circuit nonidealities on the behavior of the modulator, and allow designers to translate resolution specifications into circuit constraints. In this paper we add to that investigation by considering the effect of nonideal and multilevel quantization on the operation of the single-loop system with constant input.

With ideal integration and two quantization levels, quantizer offsets have no significant effect on the operation of the single-loop system. If the integrator is nonideal, however, quantizer offsets translate to an effective offset of the system input. Similarly, increasing the number of quantization levels in the ideal system has no significant effect. When combined with imperfect integration, however, this increase changes the form of the input versus average output characteristic.

After a brief description of the ideal single-loop $\Sigma-\Delta$ system in Section 2, in Section 3 we summarize the results of Reference 3 concerning the effect of integrator imperfection on this system. In Section 4 we show how comparator offsets in the presence of imperfect integration translate to an effective input offset. Section 5 introduces multilevel quantization and explains its influence on the ideal system. In Sections 6, 7 and 8 we consider the effect of multilevel quantization combined with imperfect integration, first for the case of four quantization levels and later for the general case. Lengthy proofs are confined to the Appendices.

2 Ideal single-loop $\Sigma-\Delta$ modulator

The single-loop $\Sigma-\Delta$ modulator consists of a discrete-time integrator and a quantizer inside a feedback loop as shown in Figure 1.

Figure 1

In the most basic implementation the quantizer has just two levels, with

$$\begin{aligned} Q(u_n) = sgn(u_n) &= \quad 1 \quad \text{for } u_n \geq 0 \\ &= -1 \quad \text{for } u_n < 0 \end{aligned}$$

The system is then described by the first order difference equation

$$u_{n+1} = u_n + x - sgn(u_n) \tag{1}$$

The integrator state u and the system input x are discrete-time signals. In this paper we assume the input to be a constant, so we drop the subscript on x. This $\Sigma-\Delta$ model has been studied by Friedman[5] and Gray[6,7]. They show that for rational x the output bit stream is periodic with average x, whereas for irrational x the output is quasiperiodic. From the perspective of dynamical systems theory, this follows immediately from the fact that the dynamics of (1) in the region of interest are just those of the well-known rotation of the circle[8]. This system is not structurally stable, however, and any infinitesimal perturbation can produce a qualitative change in the dynamics. In terms of the original $\Sigma-\Delta$ system, this means that if any approximations were made in defining the model (1) then we cannot claim that (1) captures the dynamical behavior of the real system. A better model is required—one which takes into account the effect of circuit nonidealities.

3 Single-loop $\Sigma-\Delta$ modulator with leaky integration

In Reference 3 we studied the effect of integrator imperfection on the single-loop $\Sigma-\Delta$ system with two quantization levels. The system is now modeled by the map[9,10]

$$u_{n+1} = pu_n + g.(x - sgn(u_n)) \tag{2}$$

In the ideal case $p = g = 1$. g represents the effect of component mismatch in the integrator implementation— it can be greater than or less than one. p represents integrator leak, typically due to finite op-amp gain, and is always less than one. (The dynamics of (2) with $p > 1$ are of interest to students of nonlinear dynamics, with chaotic motion and an interesting self-similarity in the bifurcation diagram[11].)

A rescaling of the state variable u in (2) eliminates the mismatch factor g, so without loss of generality g can be set to 1. The results of the study of the dynamics of (2) for $p < 1$ can be summarized as follows:

- For a given p, almost all (in the sense of probability theory) inputs $x \in (-1, 1)$ give rise to a periodic sequence (or limit cycle) at the quantizer output. If $x \geq 1$ (resp. ≤ -1) the quantizer output is fixed at 1 (resp. -1).

- Each limit cycle is globally asymptotically stable—i.e. for a given x and p all initial integrator states lead to trajectories which converge to the same limit cycle.

- For a given p each limit cycle persists over an interval of x values. The width of this interval decreases with increasing period of the limit cycle.

- The exact bounds of these intervals can be determined using a simple procedure, as described in Reference 3.

- Since these intervals have non-zero width, any admissible limit cycle corresponds to a range of input values. The consequent loss of resolution is a highly nonlinear function of the input.

Figure 2 plots average output over a limit cycle versus input for $p = 0.8$.

Figure 2

The graph contains a "step" corresponding to each rational output. Despite the complex self-similarity of the graph—given the name "devil's staircase" in dynamical systems theory—it is completely described by a simple algorithmic procedure. Figure 3 plots the region of admissibility in the $p-x$ plane of the 27 shortest limit cycles—i.e. those with periods ≤ 9.

Figure 3

Remember that the modulator output consists of a bit stream which will subsequently be averaged (in some sense) and decimated to retrieve a high resolution digital representation of x. This decoder output will not in general equal the average output over a limit cycle, but as the decimation factor increases (with a simple averaging decoder) these quantities will approach each other. The loss of resolution due to non-zero step width is clearly decoder-independent, since all inputs within a step give rise to the same limit cycle at the output. Given a resolution requirement, a circuit designer can use the results of Reference 3 to determine the minimum value of p, and hence the minimum op-amp gain, consistent with this requirement.

4 Comparator offsets

In this section we extend the results of Section 3 to consider the effects of offset in the comparator. We will find that comparator offsets translate to an effective offset in the input x, and that the component mismatch g can no longer be neglected. The comparator $Q(.)$ is now given by

$$\begin{aligned} Q(u_n) &= a \quad \text{for } u_n \geq \Delta \\ &= b \quad \text{for } u_n < \Delta \end{aligned}$$

with $a > b$. In practice a and b represent error in the digital-to-analog converter which lies in the feedback path, but it is convenient for the analysis to consider this as part of the comparator. Since

$$Q(u_n) = \tfrac{1}{2}\Big((a+b) + (a-b)sgn(u_n - \Delta)\Big)$$

(2) is now replaced by

$$u_{n+1} = p\,u_n + g\Big(x - \tfrac{1}{2}((a+b) + (a-b)sgn(u_n - \Delta))\Big) \tag{3}$$

$$u_{n+1} - \Delta = p(u_n - \Delta) + (p-1)\Delta + \frac{g(a-b)}{2}\left(\frac{x - \frac{a+b}{2}}{\frac{a-b}{2}} - sgn(u_n - \Delta)\right)$$

$$\tilde{u}_{n+1} = p\,\tilde{u}_n + \frac{g(a-b)}{2}\left(\frac{x - \frac{a+b}{2} + \frac{(p-1)\Delta}{g}}{\frac{a-b}{2}} - sgn(\tilde{u}_n)\right)$$

where the new state variable $\tilde{u}_n = u_n - \Delta$. Thus we find that the dynamics of (3) are identical to those of (2) with x replaced by

$$\frac{x - \frac{a+b}{2} + \frac{(p-1)\Delta}{g}}{\frac{a-b}{2}}$$

Note that the mismatch factor g, which had no effect in (2) of Section 3, now affects the effective input offset.

5 Ideal single-loop Σ–Δ modulator with multilevel quantization

One commonly proposed extension to the basic modulator topology of Sections 2–4 involves increasing the number of reference levels in the quantizer. In this section we examine the effect of this procedure on the ideal system. We consider first the case of 4 quantization levels and then generalize.

With 4 equally spaced quantization levels, the system is modeled by the mapping

$$u_{n+1} = u_n + x - Q(u_n) \tag{4}$$

where

$$\begin{aligned} Q(u_n) &= -1 \quad \text{for } u_n < -\tfrac{2}{3} \\ &= -\tfrac{1}{3} \quad \text{for } -\tfrac{2}{3} \leq u_n < 0 \\ &= \tfrac{1}{3} \quad \text{for } 0 \leq u_n < \tfrac{2}{3} \\ &= 1 \quad \text{for } \tfrac{2}{3} \leq u_n \end{aligned}$$

This map is plotted in Figure 4.

Figure 4

The quantization studied here is of rounding type rather than truncation—it can easily be verified that the results obtained are independent of this choice.

Assume $x \in (-1, 1)$—this is the only region of interest since all other inputs lead to output streams which are fixed at 1 (for $x \geq 1$) or -1 (for $x \leq -1$). It is clear from Figure 4 that all trajectories originating outside the interval $[x - \frac{1}{3}, x + \frac{1}{3})$ will eventually enter this interval and that no trajectories can leave the interval. We can confine our interest, therefore, to the range $[x - \frac{1}{3}, x + \frac{1}{3})$ of u_n.

There are 5 possibilities for x:

(i) $x = \frac{1}{3}$

(ii) $-\frac{1}{3} < x < \frac{1}{3}$

(iii) $\frac{1}{3} < x < 1$

(iv) $x = -\frac{1}{3}$

(v) $-1 < x < -\frac{1}{3}$

The one-dimensional maps in the region $[x - \frac{1}{3}, x + \frac{1}{3})$ for cases (i), (ii), and (iii) are plotted in Figure 5.

Figure 5

Case (i): All states $\in [0, \frac{2}{3})$ are fixed points of the map, giving rise to an output stream fixed at $Q(u) = \frac{1}{3}$.

Case (ii): The dynamics in this case are identical to those of the ideal system with binary quantization (or the rotation of the circle) with input $\frac{1}{2}((3x + 1) + (3x - 1)) = 3x$, the only difference being that the output takes on values $\pm\frac{1}{3}$ instead of ± 1. It follows that for rational input x the output is periodic with average $\frac{1}{3}(3x) = x$; for irrational input the output is quasiperiodic.

Case (iii): Shifting the state variable by $\frac{2}{3}$ yields that the dynamics of this system are identical to those of the ideal system of Section 2 with input $(3x - 2)$, the only difference being that the output takes on values 1 and $\frac{1}{3}$

instead of 1 and -1 respectively. It follows that for rational input x the output is periodic with average x; for irrational input the output is quasiperiodic.

Since cases (iv) and (v) follow by symmetry from (i) and (iii), the dynamics of (4) have been completely explained by comparison with the binary quantization system (1).

It is clear that this method of analysis generalizes instantly to systems where the number of quantization levels is other than 4. The one-dimensional map for the general case (m levels) is sketched in Figure 6.

Figure 6

The only point which is not immediately obvious is the generalization of the analysis of case (iii). Considering the one-dimensional map in the vicinity of any general breakpoint q, such as that circled in Figure 6, and repeating the analysis of case (iii) yields that the output is periodic with average value

$$\frac{1}{m-1}((m-1)x-2q)+\frac{2q}{m-1}=x$$

As in the 2 or 4 level case, it follows that for rational $x \in (-1,1)$ the output is periodic with average equal to x; for irrational $x \in (-1,1)$ the output is quasiperiodic. Thus the effect of multilevel quantization on the ideal $\Sigma-\Delta$ system (1) has been explained. For a given rational input x, increasing the number of quantization levels changes the form but not the average value of the limit cycles at the output. For irrational x the output is quasiperidic with average tending to x.

6 Four level quantization with leaky integration

Now that the dynamics of the $\Sigma\!-\!\Delta$ modulator with multilevel quantization are completely understood in the ideal case, we study the effect on this system of leaky integration. We begin by examining the case where the quantizer has four levels and the system is described by the equation

$$u_{n+1} = pu_n + x - Q(u_n) \tag{5}$$

where

$$\begin{aligned} Q(u_n) &= -1 \quad \text{for } u_n < -\tfrac{2}{3} \\ &= -\tfrac{1}{3} \quad \text{for } -\tfrac{2}{3} \leq u_n < 0 \\ &= \tfrac{1}{3} \quad \text{for } 0 \leq u_n < \tfrac{2}{3} \\ &= 1 \quad \text{for } \tfrac{2}{3} \leq u_n \end{aligned}$$

with $p < 1$. Note that we have not yet included the mismatch factor g—this will be added in Section 8. The one-dimensional map for (5) is plotted in Figure 7.

Figure 7

The dynamics of the map (5) are studied in Appendix A, where it is shown that over certain ranges of x the average output is constant, while over the remaining ranges the dynamics are just those of the binary quantization system studied in detail in Reference 3. Using the results of Appendix A, we can make the following predictions regarding the form of the average output as a function of the input:

- for $x \in [-\frac{1}{3}, \frac{1}{3})$ the graph takes on the form of the devil's staircase of Reference 3 for the appropriate value of p, scaled by a factor of $\frac{1}{3}$ in both directions;

- for $x \in [\frac{1}{3}, 1-\frac{2}{3}p)$ (resp. $[-1+\frac{2}{3}p, -\frac{1}{3})$) the average output is a constant $\frac{1}{3}$ (resp. $-\frac{1}{3}$);

- for $x \in [1-\frac{2}{3}p, \frac{5}{3}-\frac{2}{3}p)$ (resp. $[-\frac{5}{3}+\frac{2}{3}p, -1+\frac{2}{3}p)$ the graph once again takes the form of the devil's staircase of Reference 3 scaled by a factor of $\frac{1}{3}$;

- for $x \geq \frac{5}{3}-\frac{2}{3}p$ (resp. $< -\frac{5}{3}+\frac{2}{3}p$) the output is a constant 1 (resp. -1).

Figure 8 plots the dependence of the average output over a limit cycle on the input with $p = 0.8$ and 4 levels of quantization.

Figure 8

This graph was obtained by iterating the map (5) with 20000 input values equally spaced in the interval $[-1.5, 1.5]$. Note that, as predicted by the analysis, the graph consists of three scaled versions of Figure 2 (the corresponding plot from the binary quantization system) joined by steps at output values $\pm\frac{1}{3}$. The regions of admissibility in the $p - x$ plane of all limit cycles with period ≤ 5 are plotted in Figure 9.

Figure 9

Between the regions of admissibility of the fixed points we see skewed versions of Figure 3 (the corresponding plot from the binary quantization system). At $p = 1$ the widths of all steps shrink to zero, and the analysis of Section 5 is applicable. The difference between input and average output is seen to be due to two features—non-zero step width and divergence of the step centers from their ideal values. Figure 10 shows the resultant input-error plot for $p = 0.99$.

Figure 10

The Taylor series method of Reference 3 can be used to explain the underlying piecewise-linear nature of the graph. Note that one effect of increasing the number of quantization levels has been to decrease the width of the widest step, and therefore improve the resolution, by a factor of $\frac{3}{1+p}$. Also, the greatest loss of resolution is no longer in the neighborhood of the origin, but rather in the neighborhood of $\pm\frac{1}{3}$.

The importance of this analysis is that it provides an exact relationship between integrator leak and loss of resolution for a single-loop system with four quantization levels. A system designer can, given a certain resolution requirement, use these results to find the minimum allowable p, and hence the minimum op-amp gain, if four quantization levels are to be used. The next section generalizes these results to the case where m quantization levels are used.

7 Multilevel quantization with leaky integration

In this section we generalize the analysis of the last section to cover the case where the number of quantization levels is greater than 4. Since the concepts involved in the proof are essentially those of Section 6, we state the results here and give a summary of the proof in Appendix B. The quantizer $Q(.)$ now has m levels, so the system is described by

$$u_{n+1} = pu_n + x - Q(u_n) \tag{6}$$

where

$$
\begin{aligned}
Q(u_n) &= \tfrac{1}{m-1} && \text{for } 0 \le u_n < \tfrac{2}{m-1} \\
&= \tfrac{3}{m-1} && \text{for } \tfrac{2}{m-1} \le u_n < \tfrac{4}{m-1} \\
&= \tfrac{5}{m-1} && \text{for } \tfrac{4}{m-1} \le u_n < \tfrac{6}{m-1} \\
&\vdots \\
&= 1 && \text{for } \tfrac{m-2}{m-1} \le u_n \\
&= -\tfrac{1}{m-1} && \text{for } -\tfrac{2}{m-1} \le u_n < 0 \\
&\vdots \\
&= -1 && \text{for } u_n < -\tfrac{m-2}{m-1}
\end{aligned}
$$

with $p < 1$. The one-dimensional map for (6) is plotted in Figure 11.

Figure 11

Using the results of Appendix B, we can make the following predictions regarding the form of the average output as a function of the input:

- The output is fixed at

$$
\begin{aligned}
1 &\quad \text{for} \quad 1 + (1-p)\tfrac{m-2}{m-1} \le x \\
-1 &\quad \text{for} \quad x < -1 - (1-p)\tfrac{m-2}{m-1}
\end{aligned}
$$

- The widest steps occur at average values of

$$
\begin{aligned}
\tfrac{1}{m-1} &\quad \text{for} \quad \tfrac{1}{m-1} \le x < \tfrac{3-2p}{m-1} \\
\tfrac{3}{m-1} &\quad \text{for} \quad \tfrac{5-2p}{m-1} \le x < \tfrac{7-4p}{m-1} \\
\tfrac{5}{m-1} &\quad \text{for} \quad \tfrac{9-4p}{m-1} \le x < \tfrac{11-6p}{m-1}
\end{aligned}
$$

$$\vdots$$

$$-\tfrac{1}{m-1} \quad \text{for} \quad -\tfrac{3-2p}{m-1} \le x < -\tfrac{1}{m-1}$$

$$\vdots$$

- Between these steps the graph takes on the form of the devil's staircase of Reference 3 for the appropriate value of p, scaled by a factor of $\frac{1}{m-1}$ in both directions.

Figure 12 plots the dependence of the average output over a limit cycle on the input with $p = 0.8$ and 8 levels of quantization.

Figure 12

Once again, this graph should be compared with Figure 2. Note that with m quantization levels instead of 2 the width of the widest step has been decreased by a factor of $\frac{m-1}{1+p}$. The greatest loss of resolution now occurs at output values $\pm\frac{2k-1}{m-1}$, $k = 1, 2, \ldots, \frac{m}{2} - 1$.

8 Effect of component mismatch factor

We conclude with a brief discussion of the effect of the component mismatch factor g. With the inclusion of this factor the single-loop $\Sigma - \Delta$ system is described by the difference equation

$$u_{n+1} = pu_n + g.(x - Q(u_n)) \tag{7}$$

With just two levels of quantization the state variable u_n can be scaled to eliminate g from the equation, so the dynamics are independent of g. This is no longer the case when the number of quantization levels exceeds two.

We begin by studying the dynamics of (7) with $g < 1$ and four levels of quantization. The generalization to m levels is straightforward, as in Section 7 where the $g = 1$ case was studied, and for this reason will not be carried out here. Once again we state the results in this section and give a summary of the proof in Appendix C.

With $g < 1$ and four quantization levels we find once again that over certain ranges of x the average output is constant, while over the remaining ranges the dynamics are just those of the binary quantization system. The graph of average output over a limit cycle versus input takes the following form:

- for $x \in [-\frac{1}{3}, \frac{1}{3})$ the form of the graph is that of the devil's staircase of Reference 3 for the appropriate value of p, scaled by a factor of $\frac{1}{3}$ in both directions;
- for $x \in [\frac{1}{3}, \frac{2-2p+g}{3g})$ (resp. $[\frac{-2+2p-g}{3g}, -\frac{1}{3})$) the average output is a constant $\frac{1}{3}$ (resp. $-\frac{1}{3}$);
- for $x \in [\frac{2-2p+g}{3g}, \frac{2-2p+3g}{3g})$ (resp. $[-\frac{2-2p+3g}{3g}, -\frac{2-2p+g}{3g})$ the graph once again takes the form of the devil's staircase of Reference 3 scaled by a factor of $\frac{1}{3}$ and shifted;
- for $x \geq \frac{2-2p+3g}{3g}$ (resp. $< -\frac{2-2p+3g}{3g}$) the output is a constant 1 (resp. -1).

With $g > 1$ and four levels of quantization the dynamics are identical to those of the $g < 1$ system for all inputs outside the range

$$\left[-\frac{1}{3}, -\frac{2-g}{3g}\right) \cup \left[\frac{2-g}{3g}, \frac{1}{3}\right) \cup \left[-\frac{2-2p+g}{3g}, -\frac{-2p+3g}{3g}\right) \cup \left[\frac{-2p+3g}{3g}, \frac{2-2p+g}{3g}\right) \qquad (8)$$

For inputs within this range, the analysis is complicated by the appearance of two qualitatively new features. Recall that for $g \leq 1$ the output for any given input x (after a possible transient) takes on at most two values. Recall also that for $g \leq 1$ more than one limit cycle can be admissible for certain values of x, but that all must have the same average value. With $g > 1$ and x in the range (8) neither of these facts is true.

$\boxed{\textbf{Figure 13}}$

Figure 13 plots average output over a limit cycle versus input for $p = 0.8$ and $g = .7$, 1 and 1.3. (These values are not, of course, typical of those encountered in any real $\Sigma-\Delta$ system.) In Figure 14 regions of Figure 13 are magnified in order to display the form of the $g = 1.3$ characteristic.

$\boxed{\textbf{Figure 14}}$

This graph cannot be described by a simple self-similar algorithmic structure as in the $g \leq 1$ case. We can, however, derive bounds on the characteristic — this is again explained in Appendix C.

Finally, we consider the effect of using truncation-type quantization rather than rounding. In Section 5 it was pointed out that the behavior of the ideal single-loop $\Sigma-\Delta$ system with multilevel quantization is independent of the nature of the quantization—i.e. if truncation is used instead of rounding the dynamics are unaffected. With leaky integration this is no longer the case.

$\boxed{\textbf{Figure 15}}$

As can be seen from Figure 15, if $Q_t(.)$ represents m-level truncation and $Q(.)$ represents m-level rounding, then

$$Q_t(u_n) = Q(u_n - \tfrac{1}{m-1})$$

Using truncation instead of rounding, therefore, the system is represented by the difference equation

$$u_{n+1} = p u_n + g.(x - Q_t(u_n))$$

$$u_{n+1} = p u_n + g.\left(x - Q(u_n - \tfrac{1}{m-1})\right)$$

$$(u_{n+1} - \tfrac{1}{m-1}) = p(u_n - \tfrac{1}{m-1}) + \tfrac{p-1}{m-1} + g.\left(x - Q(u_n - \tfrac{1}{m-1})\right)$$

$$\tilde{u}_{n+1} = p\tilde{u}_n + g.\left(\left(x + \tfrac{p-1}{g(m-1)}\right) - Q(\tilde{u}_n)\right)$$

The effect of using truncation instead of rounding has been to introduce an effective input offset of $\frac{p-1}{g(m-1)}$.

9 Concluding remarks

Using techniques from the field of nonlinear dynamics, we have described the effect of nonideal and multilevel quantization on the single-loop $\Sigma-\Delta$ modulator with constant input. Offset in a two-level quantizer has been shown to affect the operation of the modulator only in conjunction with leaky integration. The net result of such offset is to introduce an effective offset of the system input. With perfect integration, increasing the number of quantization levels beyond two changes the form but not the average value of the limit cycles observed at the output. Once again, the behavior changes when integrator imperfections are taken into account. When the component mismatch factor g is less than or equal to one almost all (in the sense of probability theory) inputs give rise to limit cycles at the quantizer output. These limit cycles take on at most two values. In certain cases it is possible for two or more stable limit cycles to be admissible for a given input, but in each of these cases the average value of every such admissible limit cycle is the same. The graph plotting average output over a limit cycle versus input consists of a number of scaled versions of the corresponding curve from the binary quantization system. It can easily be derived by combining the analysis of Section 7 with the algorithmic procedure of Reference 3. With $g > 1$ the above analysis applies over all but four intervals of the input x. Over these intervals it is not possible to describe the input versus average output characteristic by a simple self-similar structure. We can, however, bound the characteristic as explained in Section 8.

10 Acknowledgements

This work is supported in part by the Intel Foundation Graduate Fellowship program and by the Office of Naval Research under Grant N00014-89-J-1402.

11 Appendix A

In this Appendix we study the dynamics of the map (5), plotted in Figure 7, i.e.

$$u_{n+1} = pu_n + x - Q(u_n)$$

where

$$\begin{aligned} Q(u_n) &= -1 \quad \text{for } u_n < -\tfrac{2}{3} \\ &= -\tfrac{1}{3} \quad \text{for } -\tfrac{2}{3} \le u_n < 0 \\ &= \tfrac{1}{3} \quad \text{for } 0 \le u_n < \tfrac{2}{3} \\ &= 1 \quad \text{for } \tfrac{2}{3} \le u_n \end{aligned}$$

Once again, it is natural to consider several different ranges of the input x:

Case (i): $x \ge \frac{5}{3} - \frac{2}{3}p$

All initial conditions give rise to a fixed word stream of 1 at the quantizer output.

Case (ii): $-\frac{1}{3} \leq x < \frac{1}{3}$

All trajectories eventually enter the interval $[x - 1 + \frac{2}{3}p, x + 1 - \frac{2}{3}p)$ and, having entered, can never leave. We therefore confine our interest to this interval, without loss of generality.

Given an initial condition in the range $[-\frac{2}{3}, \frac{2}{3})$, the dynamics are just those of the binary quantization system with the same value of p and input $3x$. The analysis of Reference 3 is, therefore, immediately applicable. The output takes on values $\pm\frac{1}{3}$ instead of ± 1, so for these initial conditions the average output is one-third of the average output of the binary quantization system with input $3x$.

If $p \geq 0.5$ then all trajectories enter the region $[-\frac{2}{3}, \frac{2}{3})$, and the above analysis is completely general. With $p < 0.5$, however, the dynamics are more complicated. If

$$p(x + 1 - \tfrac{2}{3}p) + x - 1 < -\tfrac{2}{3} \quad \text{and} \quad p(x - 1 + \tfrac{2}{3}p) + x + 1 \geq \tfrac{2}{3}$$

$$\text{i.e.} \quad \frac{-\frac{1}{3} + p - \frac{2}{3}p^2}{1 + p} \leq x < \frac{\frac{1}{3} - p + \frac{2}{3}p^2}{1 + p}$$

then it is easily verified that all points in the region $[x - 1 + \frac{2}{3}p, -\frac{2}{3}) \cup [\frac{2}{3}, x + 1 - \frac{2}{3}p)$ tend to a fixed point of period 2 giving the output sequence $1, -1, 1, -1, \ldots$. The second iterate of the one-dimensional map in this case is sketched in Figure A1.

Figure A1

Note that since

$$\frac{\frac{1}{3} - p + \frac{2}{3}p^2}{1+p} < \frac{\frac{1}{3}(1-p)}{1+p}$$

all points in the interval $[-\frac{2}{3}, \frac{2}{3})$ eventually yield the period 2 output sequence $\frac{1}{3}, -\frac{1}{3}, \ldots$ for these values of x. This is the first case where two asymptotically stable limit cycles, with different basins of attraction, occur in our system. The average output is the same—zero—over both limit cycles.

Finally, consider the case where $p < 0.5$ and $x \in [-\frac{1}{3}, \frac{1}{3}) \setminus [\frac{-\frac{1}{3}+p-\frac{2}{3}p^2}{1+p}, \frac{\frac{1}{3}-p+\frac{2}{3}p^2}{1+p})$. Suppose $x < \frac{-\frac{1}{3}+p-\frac{2}{3}p^2}{1+p}$—the other case can be handled by the same argument. Since all points in $[x - 1 + \frac{2}{3}p, -\frac{2}{3})$ map into the interval $[-\frac{2}{3}, x + 1 - \frac{2}{3}p)$, and all trajectories originating in the interval $[\frac{2}{3}, x + 1 - \frac{2}{3}p)$ eventually enter the interval $[-\frac{2}{3}, \frac{2}{3})$, the relevant dynamics in this case are just those of the region $[-\frac{2}{3}, \frac{2}{3})$, as studied above.

Case (iii): $\frac{1}{3} \leq x < 1 - \frac{2}{3}p$

We can without loss of generality consider the dynamics only in the interval $[x - 1 + \frac{2}{3}p, x + \frac{1}{3})$. An analysis similar to that of Case (ii) yields that there is a stable fixed point in the range $[0, \frac{2}{3})$ which attracts all trajectories if $x \notin [\frac{\frac{1}{3}+p-\frac{2}{3}p^2}{1+p}, \frac{1-\frac{1}{3}p}{1+p})$. For all other values of x trajectories originating in the range $[0, \frac{2}{3})$ tend to the fixed point, while those originating in $[x - 1 + \frac{2}{3}p, 0) \cup [\frac{2}{3}, x + \frac{1}{3})$ tend to a fixed point of period 2 giving output $1, -\frac{1}{3}, \ldots$. The average output is $\frac{1}{3}$ in either case.

Case (iv) $1 - \frac{2}{3}p \leq x < \frac{5}{3} - \frac{2}{3}p$

The one dimensional map in the region of interest is plotted in Figure A2

Figure A2

By a translation and scaling, we find that the dynamics are equivalent to those of the binary quantization system with input $3(x-\frac{4}{3}+\frac{2}{3}p)$, with the output taking on values 1 and $\frac{1}{3}$ instead of 1 and -1 respectively. A little algebraic manipulation yields that the average output of this system is $\frac{1}{3}(2+$ average output of the binary quantization system with input $(3x+2p-4))$.

All other values of x can be handled by symmetry, so we have a complete understanding of the dynamics of (5).

12 Appendix B

In this Appendix we derive the dependence of average output on input for the general multilevel quantization system of Section 7. We will present a sketch of the proof, as it is similar to that of the four level system of Appendix A. The difference equation to be studied is equation (6), plotted in Figure 11, i.e.

$$u_{n+1} = pu_n + x - Q(u_n)$$

where

$$\begin{aligned}
Q(u_n) &= \tfrac{1}{m-1} && \text{for } 0 \leq u_n < \tfrac{2}{m-1} \\
&= \tfrac{3}{m-1} && \text{for } \tfrac{2}{m-1} \leq u_n < \tfrac{4}{m-1} \\
&= \tfrac{5}{m-1} && \text{for } \tfrac{4}{m-1} \leq u_n < \tfrac{6}{m-1} \\
&\vdots && \\
&= 1 && \text{for } \tfrac{m-2}{m-1} \leq u_n \\
&= -\tfrac{1}{m-1} && \text{for } -\tfrac{2}{m-1} \leq u_n < 0
\end{aligned}$$

$$\vdots$$

$$= \quad -1 \qquad \text{for } u_n < -\tfrac{m-2}{m-1}$$

with $p < 1$.

It is clear that for $x \geq 1 + \frac{(1-p)(m-2)}{m-1}$ (resp. $x < -1 - \frac{(1-p)(m-2)}{m-1}$) the output word stream is fixed at 1 (resp. -1).

Consider now the case $-\frac{1}{m-1} \leq x < \frac{1}{m-1}$. For initial conditions $\in [-\frac{2}{m-1}, \frac{2}{m-1})$ the dynamics are just those of the system with two quantization levels and input $(m-1)x$. For all remaining initial conditions, the only other possibilities are the limit cycles

$$\begin{array}{lll}
\frac{3}{m-1} \quad -\frac{3}{m-1} & \text{for} & -\frac{1-2p}{m-1}\frac{1-p}{1+p} \leq x < \frac{1-2p}{m-1}\frac{1-p}{1+p} \\
\frac{5}{m-1} \quad -\frac{5}{m-1} & \text{for} & -\frac{1-4p}{m-1}\frac{1-p}{1+p} \leq x < \frac{1-4p}{m-1}\frac{1-p}{1+p} \\
 & \vdots & \\
1 \quad -1 & \text{for} & -\frac{1-(m-2)p}{m-1}\frac{1-p}{1+p} \leq x < \frac{1-(m-2)p}{m-1}\frac{1-p}{1+p}
\end{array}$$

The average output is the same in any of these cases, so the average output as a function of x over this interval takes on the by now familiar devil's staircase form, scaled by $\frac{1}{m-1}$.

By shifting the state variable, this analysis can be applied to the case $\frac{4k-1-2kp}{m-1} \leq x < \frac{4k+1-2kp}{m-1}$. Simply transform the state variable u to $u - \frac{2k}{m-1}$. The above analysis can then be applied, with x shifted by $\frac{4k-2kp}{m-1}$. For initial conditions $\in [\frac{2k-2}{m-1}, \frac{2k+2}{m-1})$ the dynamics are those of the binary quantization system, with output values $\frac{2k+1}{m-1}$ and $\frac{2k-1}{m-1}$. For all remaining initial conditions, the only other possibilities are the limit cycles

$$\frac{2k+3}{m-1} \quad \frac{2k-3}{m-1} \quad \text{for} \quad -\frac{1-2p}{m-1}\frac{1-p}{1+p} + \frac{4k-2kp}{m-1} \leq x < \frac{1-2p}{m-1}\frac{1-p}{1+p} + \frac{4k-2kp}{m-1}$$

$$\vdots$$

$$
\begin{array}{llll}
1 & \frac{4k-m+1}{m-1} & \text{for} & -\frac{1-(m-2-2k)p}{m-1}\frac{1-p}{1+p}+\frac{4k-2kp}{m-1} \le x < \frac{1-(m-2-2k)p}{m-1}\frac{1-p}{1+p}+\frac{4k-2kp}{m-1} \text{ (if } k>0) \\
\frac{4k+m-1}{m-1} & -1 & \text{for} & -\frac{1-(m-2+2k)p}{m-1}\frac{1-p}{1+p}+\frac{4k-2kp}{m-1} \le x < \frac{1-(m-2+2k)p}{m-1}\frac{1-p}{1+p}+\frac{4k-2kp}{m-1} \text{ (if } k<0)
\end{array}
$$

Once again the average output over this interval takes on the devil's staircase form, with x shifted by $\frac{4k-2kp}{m-1}$ and the output shifted by $\frac{2k}{m-1}$.

Now consider the case $\frac{1}{m-1} \le x < \frac{3-2p}{m-1}$. There is then a fixed point $\in [0, \frac{2}{m-1})$ which attracts all trajectories in this interval. For all remaining initial conditions, the only other possibilities are the limit cycles

$$
\begin{array}{llll}
\frac{3}{m-1} & -\frac{1}{m-1} & \text{for} & \frac{1+3p-2p^2}{(m-1)(1+p)} \le x < \frac{3-p}{(m-1)(1+p)} \\
\frac{5}{m-1} & -\frac{3}{m-1} & \text{for} & \frac{1+5p-4p^2}{(m-1)(1+p)} \le x < \frac{3-3p+2p^2}{(m-1)(1+p)} \\
 & & \vdots & \\
1 & \frac{3-m}{m-1} & \text{for} & \frac{1+(m-1)p-(m-2)p^2}{(m-1)(1+p)} \le x < \frac{3-(m-3)p+(m-4)p^2}{(m-1)(1+p)}
\end{array}
$$

The average output is $\frac{1}{m-1}$ for any of these possibilities.

To complete the analysis, consider the case $\frac{4k+1-2kp}{m-1} \le x < \frac{4k+3-(2k+2)p}{m-1}$, $k = \pm 1, \pm 2, \ldots \pm \frac{m-4}{2}$.

Once again, we apply a shift of $\frac{2k}{m-1}$ to find that the above analysis is applicable, giving an output of period 1 or 2 with average $\frac{2k+1}{m-1}$.

13 Appendix C

In this Appendix we study the dynamics of the map (7), i.e.

$$u_{n+1} = pu_n + g.(x - Q(u_n))$$

where

$$\begin{aligned} Q(u_n) &= -1 \quad \text{for } u_n < -\tfrac{2}{3} \\ &= -\tfrac{1}{3} \quad \text{for } -\tfrac{2}{3} \le u_n < 0 \\ &= \tfrac{1}{3} \quad \text{for } 0 \le u_n < \tfrac{2}{3} \\ &= 1 \quad \text{for } \tfrac{2}{3} \le u_n \end{aligned}$$

We begin with the case $g < 1$:

Case (i): $x \geq \frac{2-2p+3g}{3g}$

All initial conditions give rise to a fixed word stream of 1 at the quantizer output.

Case (ii): $-\frac{1}{3} \leq x < \frac{1}{3}$

All trajectories eventually enter the interval $[g(x-1)+\frac{2}{3}p, g(x+1)-\frac{2}{3}p)$ and, having entered, can never leave. We therefore confine our interest to this interval, without loss of generality.

Given an initial condition in the range $[-\frac{2}{3}, \frac{2}{3})$, the dynamics are just those of the binary quantization system with the same value of p and input $3x$. The analysis of Reference 3 is, therefore, immediately applicable. The output takes on values $\pm\frac{1}{3}$ instead of ± 1, so for these initial conditions the average output is one-third of the average output of the binary quantization system with input $3x$.

If $p \geq \frac{3g}{2} - 1$ then all trajectories enter the region $[-\frac{2}{3}, \frac{2}{3})$, and the above analysis is completely general. With $p < \frac{3g}{2} - 1$, however, the dynamics are more complicated. If

$$-\frac{-2+3g-3gp+2p^2}{3g(1+p)} \leq x < \frac{-2+3g-3gp+2p^2}{3g(1+p)}$$

then all initial conditions in the region $[g(x-1)+\frac{2}{3}p, -\frac{2}{3}) \cup [\frac{2}{3}, g(x+1)-\frac{2}{3}p)$ tend to a fixed point of period 2 giving the output sequence $1, -1, 1, -1, \ldots$.

Note that since

$$\frac{-2+3g-3gp+2p^2}{3g(1+p)} < \frac{\frac{1}{3}(1-p)}{1+p}$$

all points in the interval $[-\frac{2}{3}, \frac{2}{3})$ eventually yield the period 2 output sequence $\frac{1}{3}, -\frac{1}{3}, \ldots$ for these values of x. Once again, although there are two admissible limit cycles for these values of x, the average output is the same—zero—over both limit cycles.

Finally, consider the case where $p < \frac{3g}{2} - 1$ and $x \in [-\frac{1}{3}, \frac{1}{3}) \setminus [-\frac{-2+3g-3gp+2p^2}{3g(1+p)}, \frac{-2+3g-3gp+2p^2}{3g(1+p)})$ Suppose $x < -\frac{-2+3g-3gp+2p^2}{3g(1+p)}$ —the other case can be handled by the same argument. Since all points in $[g(x-1)+\frac{2}{3}p, -\frac{2}{3})$ map into the interval $[-\frac{2}{3}, g(x+1)-\frac{2}{3}p)$, and all trajectories originating in the interval $[\frac{2}{3}, g(x+1)-\frac{2}{3}p)$ eventually enter the interval $[-\frac{2}{3}, \frac{2}{3})$, the relevant dynamics in this case are just those of the region $[-\frac{2}{3}, \frac{2}{3})$, as studied above.

Case (iii): $\frac{1}{3} \leq x < \frac{2-2p+g}{3g}$

We can without loss of generality consider the dynamics only in the interval $[g(x-1)+\frac{2}{3}p, g(x+\frac{1}{3}))$. An analysis similar to that of Case (ii) yields that there is a stable fixed point in the range $[0, \frac{2}{3})$ which attracts all trajectories if $x \notin [\frac{2-g+3gp-2p^2}{3g(1+p)}, \frac{1-\frac{1}{3}p}{1+p})$. For all other values of x trajectories originating in the range $[0, \frac{2}{3})$ tend to the fixed point, while those originating in $[g(x-1)+\frac{2}{3}p, 0) \cup [\frac{2}{3}, g(x+\frac{1}{3}))$ tend to a fixed point of period 2 giving output $1, -\frac{1}{3}, \ldots$. The average output is $\frac{1}{3}$ in either case.

Case (iv) $\frac{2-2p+g}{3g} \leq x < \frac{2-2p+3g}{3g}$

By a translation and scaling, we find that the dynamics are equivalent to those of the binary quantization system with input $\frac{3}{g}(\frac{2}{3}p + gx - \frac{2}{3}g - \frac{2}{3})$ with the output taking on values 1 and $\frac{1}{3}$ instead of 1 and -1 respectively. A little algebraic manipulation yields that the average output of this system is $\frac{1}{3}(2+$ average output of the binary quantization system with input $\frac{3}{g}(\frac{2}{3}p + gx - \frac{2}{3}g - \frac{2}{3})$.

All other values of x can be handled by symmetry, so we have a complete understanding of the dynamics of (7) for $g < 1$.

With $g > 1$ and four levels of quantization the dynamics are identical to those of the $g \leq 1$ system for all inputs outside the range

$$\left[-\frac{1}{3}, -\frac{2-g}{3g}\right) \cup \left[\frac{2-g}{3g}, \frac{1}{3}\right) \cup \left[-\frac{-2p+3g}{3g}, -\frac{2-2p+g}{3g}\right) \cup \left[\frac{2-2p+g}{3g}, \frac{-2p+3g}{3g}\right)$$

We will examine the dynamics for $x \in [\frac{2-g}{3g}, \frac{1}{3})$ — the dynamics in the other three intervals follow from this by translation and symmetry. The one-dimensional map in the region of interest is shown in Figure C1.

$\boxed{\textbf{Figure C1}}$

The difference between this and previous examples is that it is now possible for $g(x+\frac{1}{3})$ to exceed $\frac{2}{3}$, giving 3 possible output values in the region of interest. If $g > 1 + \frac{p}{2}$ it is possible that $g(x+\frac{1}{3}) > \frac{2}{3}$ and $\frac{2p}{3} + g(x-1) < -\frac{2}{3}$, giving 4 possible output values in the region of interest. Since such parameter values would be unreasonable in the context of our single-loop $\Sigma-\Delta$ system, we restrict our attention to the case $g \leq 1 + \frac{p}{2}$.

We can explain the structure of the orbits in the following manner. With two quantization levels the orbits at the output are of the form

$$\text{-}\tfrac{1}{3}\,(\tfrac{1}{3})^{\alpha_1}\,\text{-}\tfrac{1}{3}\,(\tfrac{1}{3})^{\alpha_2}\,\text{-}\tfrac{1}{3}\,(\tfrac{1}{3})^{\alpha_3}\ldots \tag{9}$$

with any two of the α_i differing by no more than 1. The details of this structure are contained in Reference 3.

- For $\frac{2-g}{3g} \le x < \frac{2-g+gp}{3g(1+p)}$ some of the $\frac{1}{3}$-$\frac{1}{3}\frac{1}{3}$ blocks of these orbits are replaced by $\frac{1}{3}$-$\frac{1}{3}1$-$\frac{1}{3}$ blocks, as sketched in Figure C2.

Figure C2

- For $\frac{2-g+gp}{3g(1+p)} \le x < \frac{2-g+3gp-gp^2}{3g(1+p+p^2)}$ the $\frac{1}{3}$-$\frac{1}{3}\frac{1}{3}$ sequence is no longer admissible, and all orbits are of the form (9) with the $\frac{1}{3}$-$\frac{1}{3}\frac{1}{3}$ blocks replaced by $\frac{1}{3}$-$\frac{1}{3}1$-$\frac{1}{3}$ blocks.

- For $\frac{2-g+3gp-gp^2}{3g(1+p+p^2)} \le x < \frac{2-g+3gp-gp^2+gp^3}{3g(1+p+p^2+p^3)}$ the 1-$\frac{1}{3}1$ sequence has become admissible, and some of the $\frac{1}{3}$-$\frac{1}{3}1$-$\frac{1}{3}\frac{1}{3}$ blocks from the previous case are replaced by $\frac{1}{3}$-$\frac{1}{3}1$-$\frac{1}{3}1$-$\frac{1}{3}$ blocks. Note that this case occurs only when $\frac{2-g+3gp-gp^2}{3g(1+p+p^2)} < \frac{1}{3}$, i.e. when $g > \frac{1}{p^2-p+1}$.

- For $\frac{2-g+3gp-gp^2+gp^3}{3g(1+p+p^2+p^3)} \le x < \frac{2-g+3gp-gp^2+3gp^3-gp^4}{3g(1+p+p^2+p^3+p^4)}$ the $\frac{1}{3}$-$\frac{1}{3}1$-$\frac{1}{3}\frac{1}{3}$ sequence is no longer admissible, and all orbits are of the form (9) with the $\frac{1}{3}$-$\frac{1}{3}\frac{1}{3}\frac{1}{3}$ blocks replaced by $\frac{1}{3}$-$\frac{1}{3}1$-$\frac{1}{3}1$-$\frac{1}{3}$ blocks.

- For $\frac{2-g+3gp-gp^2+3gp^3-gp^4}{3g(1+p+p^2+p^3+p^4)} \le x < \frac{2-g+3gp-gp^2+3gp^3-gp^4+gp^5}{3g(1+p+p^2+p^3+p^4+p^5)}$ the 1-$\frac{1}{3}1$-$\frac{1}{3}1$ sequence has become admissible, and some of the $\frac{1}{3}$-$\frac{1}{3}1$-$\frac{1}{3}1$-$\frac{1}{3}\frac{1}{3}$ blocks from the previous case are replaced by $\frac{1}{3}$-$\frac{1}{3}1$-$\frac{1}{3}1$-$\frac{1}{3}1$-$\frac{1}{3}$ blocks. Note that this case occurs only when $\frac{2-g+3gp-gp^2+3gp^3-gp^4}{3g(1+p+p^2+p^3+p^4)} < \frac{1}{3}$, i.e. when $g > \frac{1}{p^4-p^3+p^2-p+1}$.

The pattern is clear, and allows us to place bounds on the input versus average output characteristic. The inequality analysis of Reference 3 can be applied to the orbits of form (9) with the $\frac{1}{3}$-$\frac{1}{3}\frac{1}{3}$ blocks replaced by $\frac{1}{3}$-$\frac{1}{3}1$-$\frac{1}{3}$

(or higher order substitutions, as above). In this way the bounds of admissibility of each of these orbits can be obtained. Corresponding to each such substitution the input versus average output characteristic can be plotted in the appropriate input interval, as derived above. The resultant plot provides bounds on the average value of a limit cycle observed at the output.

14 References

1. J C Candy, "A use of limit cycle oscillations to obtain robust analog-to-digital conversion". *IEEE Trans. Comm.*, vol. COM-22, pp. 298–305, Mar 1974.

2. D. Goedhart, R. van de Plassche and E. Stikvoort, "Digital-to-analog conversion in playing a compact disc". *Philips Tech. Rev.*, vol. 40, pp. 174-179, 1982.

3. Orla Feely and Leon O Chua, "The effect of integrator leak in $\Sigma - \Delta$ modulation". University of California, Berkeley, Electronics Research Laboratory Memorandum No. M90/116, Dec 1990. To appear in *IEEE Trans. Circuits and Systems.*

4. Hao Bai-Lin, *Elementary Symbolic Dynamics.* Singapore: World Scientific, 1989.

5. Vladimir Friedman, "The structure of the limit cycles in sigma-delta modulation". *IEEE Trans Comm.*, vol. COM-36, pp. 972-979, August 1988.

6. R. M. Gray, "Oversampled sigma-delta modulation". *IEEE Trans. Comm.*, vol. COM-35, pp. 481–489, May 1987.

7. R. M. Gray, "Spectral analysis of quantization noise in a Single Loop Sigma Delta Modulator with dc Input". *IEEE Trans. Comm*, vol. COM-37, pp. 588–599, June 1989.

8. Robert Devaney, *An Introduction to Chaotic Dynamical Systems.* Redwood City, California: Addison-Wesley, 1989.

9. Bernhard Boser, "Design and implementation of oversampled analog-to-digital converters". PhD thesis, Stanford University, October 1988.

10. N. N. Leonov, "Map of the line onto itself". *Radiofisica*, vol. 2, pp. 942-956, 1959.

11. Orla Feely and Leon O Chua, "Nonlinear dynamics of a class of analog-to-digital converters". University of California, Berkeley, Electronics Research Laboratory Memorandum No. M91/30, Apr 1991. Also submitted to *Int. Journal of Bifurcation and Chaos.*

List of Figures

Figure 12: Average output over a limit cycle plotted versus dc input x. This graph corresponds to the single-loop system with $p = 0.8$, $g = 1$ and eight quantization levels. Compare to Figure 2.

Figure 13: Average output over a limit cycle plotted versus dc input $x \in [-1.6, 1.6]$. This graph corresponds to the single-loop system with four quantization levels, $p = 0.8$ and $g = 1$, 0.7 and 1.3.

Figure 14: Average output over a limit cycle plotted versus dc input $x \in [0.15, 0.4]$. This graph corresponds to the single-loop system with four quantization levels, $p = 0.8$ and $g = 1$ and 1.3.

Figure 15: m-level (a) rounding and (b) truncation quantization functions.

Figure A1: Second iterate of the one-dimensional map (5) for $\frac{-\frac{1}{3}+p-\frac{2}{3}p^2}{1+p} \leq x < \frac{\frac{1}{3}-p+\frac{2}{3}p^2}{1+p}$.

Figure A2: One-dimensional map of (5) for $1 - \frac{2}{3}p \leq x < \frac{5}{3} - \frac{2}{3}p$.

Figure C1: One-dimensional map of (7) for $g > 1$ and $\frac{2-g}{3g} \leq x < \frac{1}{3}$.

Figure C2: One-dimensional map of (7) for $g > 1$ showing the appearance of the $\frac{1}{3}$-$\frac{1}{3}$1-$\frac{1}{3}$ blocks.

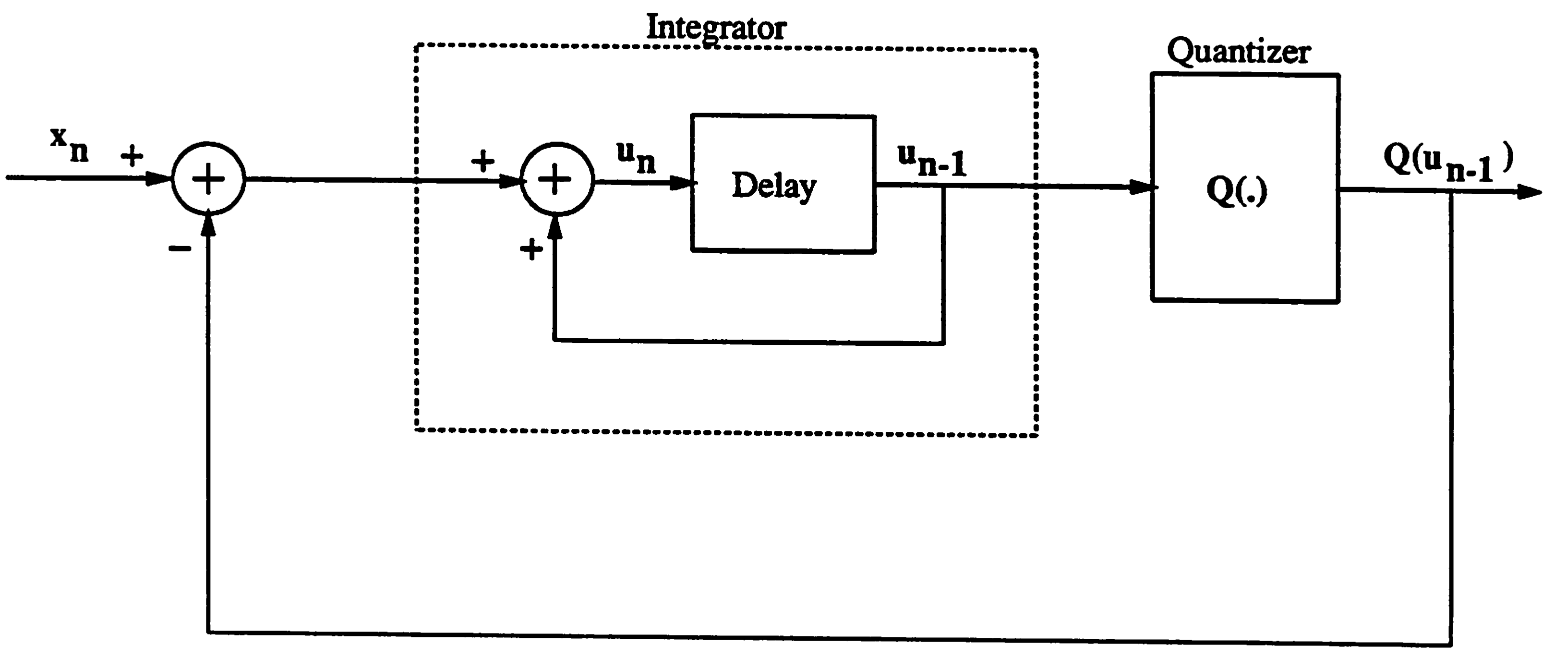
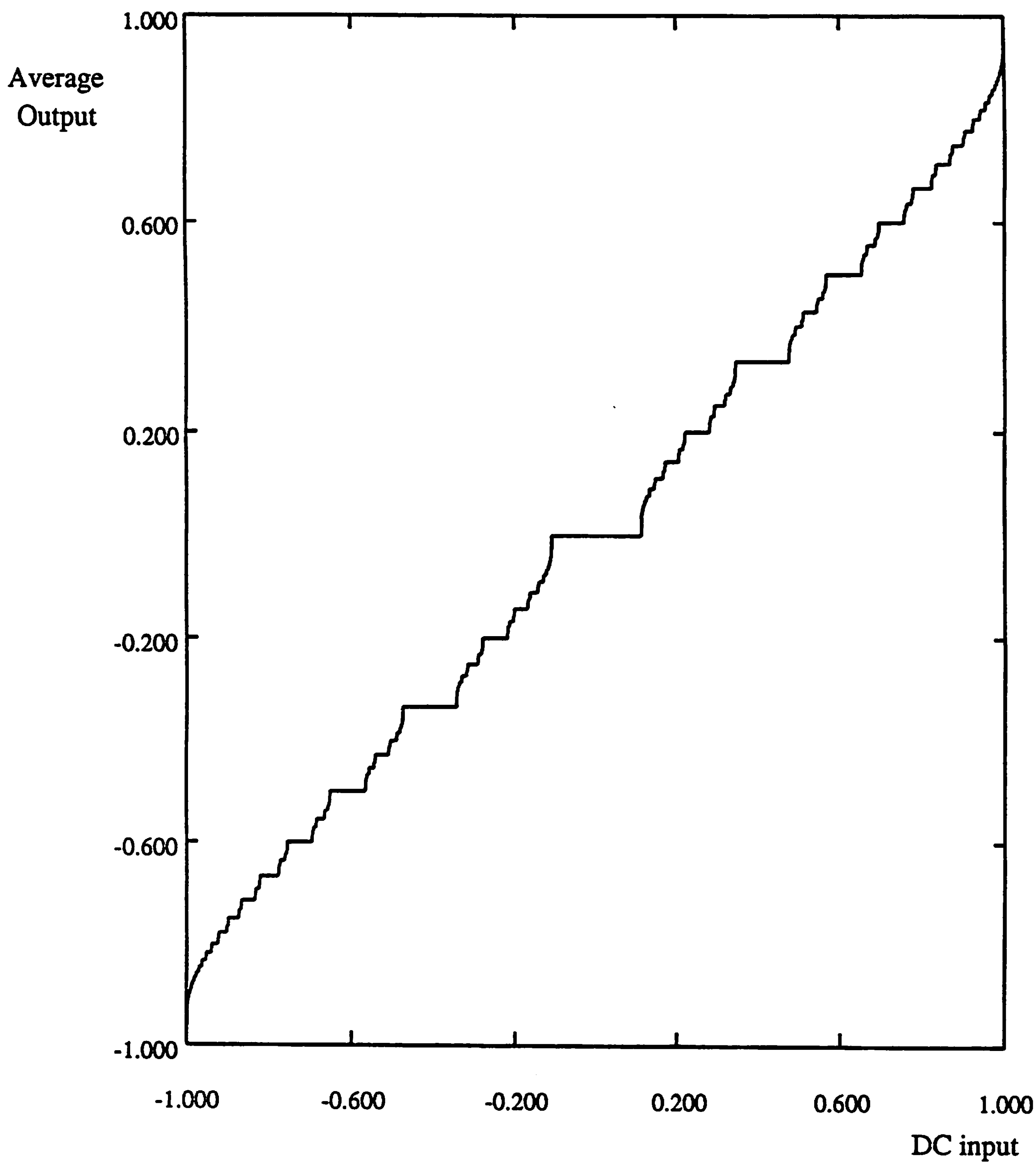


Figure 1



Average
Output
1.000
0.600
0.200
-0.200
-0.600
-1.000
-1.000
-0.600
-0.200
0.200
0.600
1.000
DC input

Figure 2

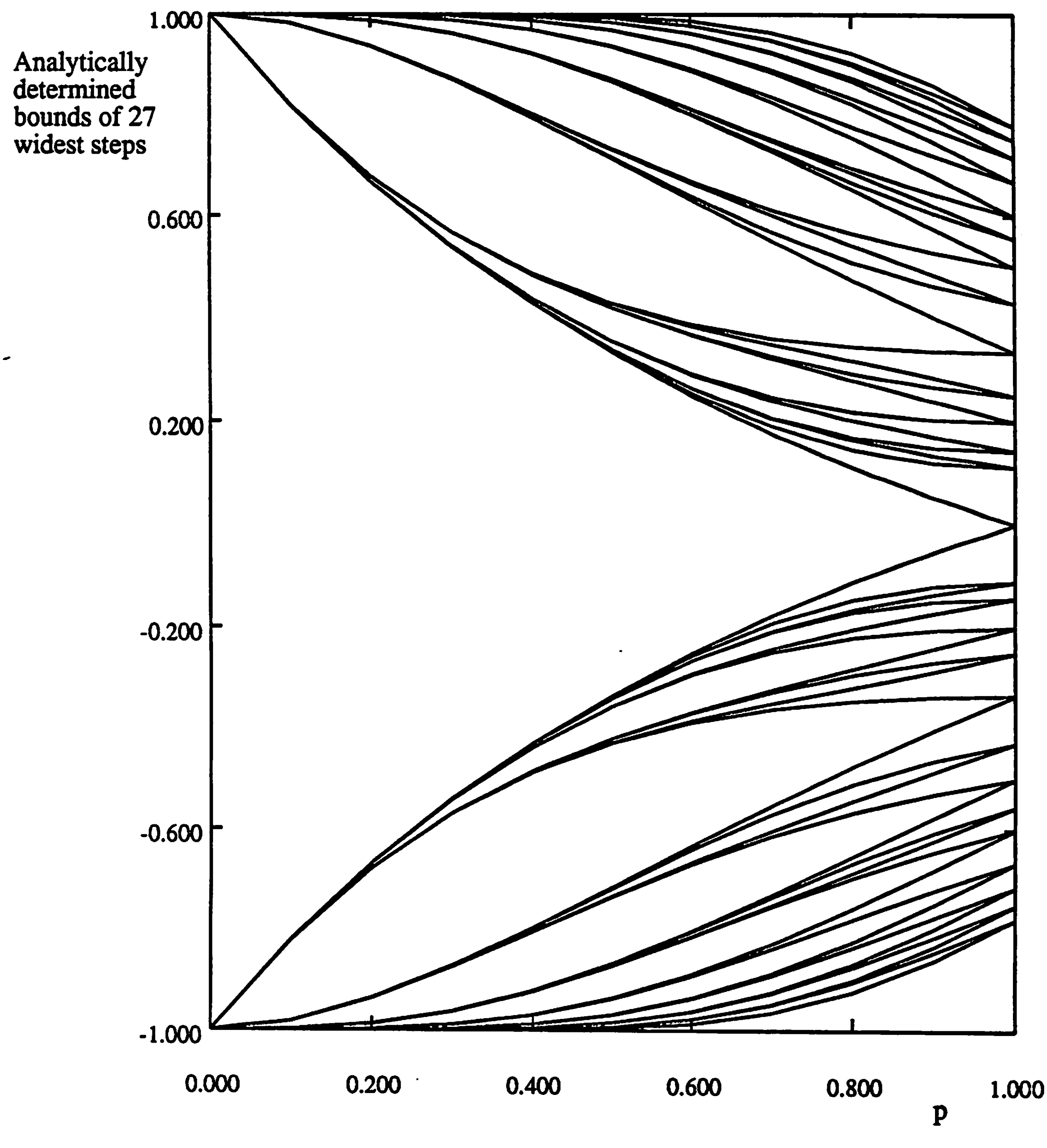

1.000
Analytically
determined
bounds of 27
widest steps
0.600
0.200
-0.200
-0.600
-1.000
0.000
0.200
0.400
0.600
0.800
1.000
p

Figure 3

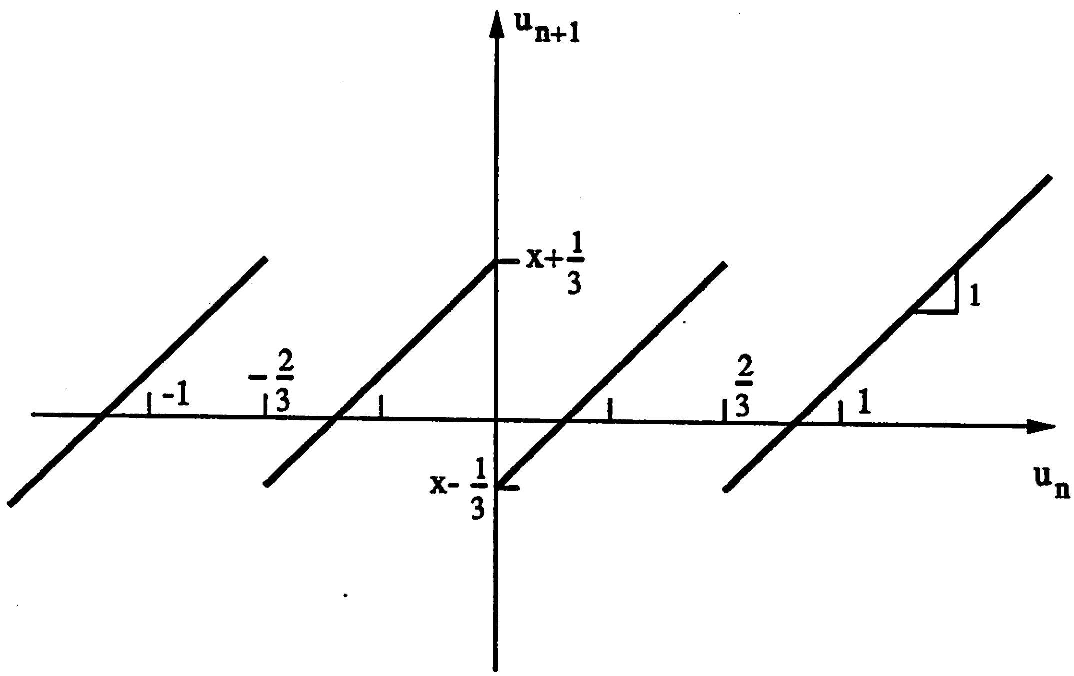


Figure 4

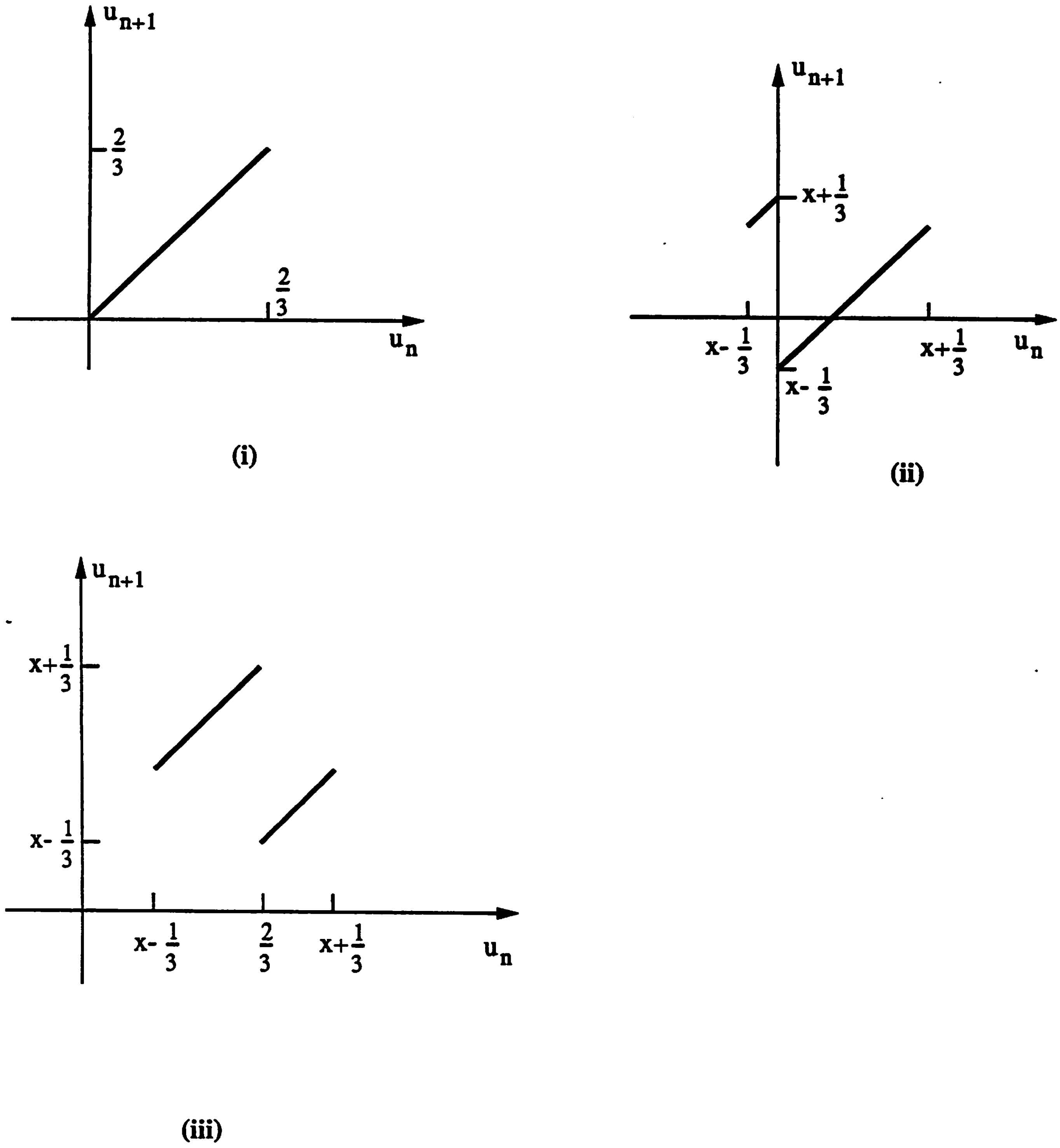
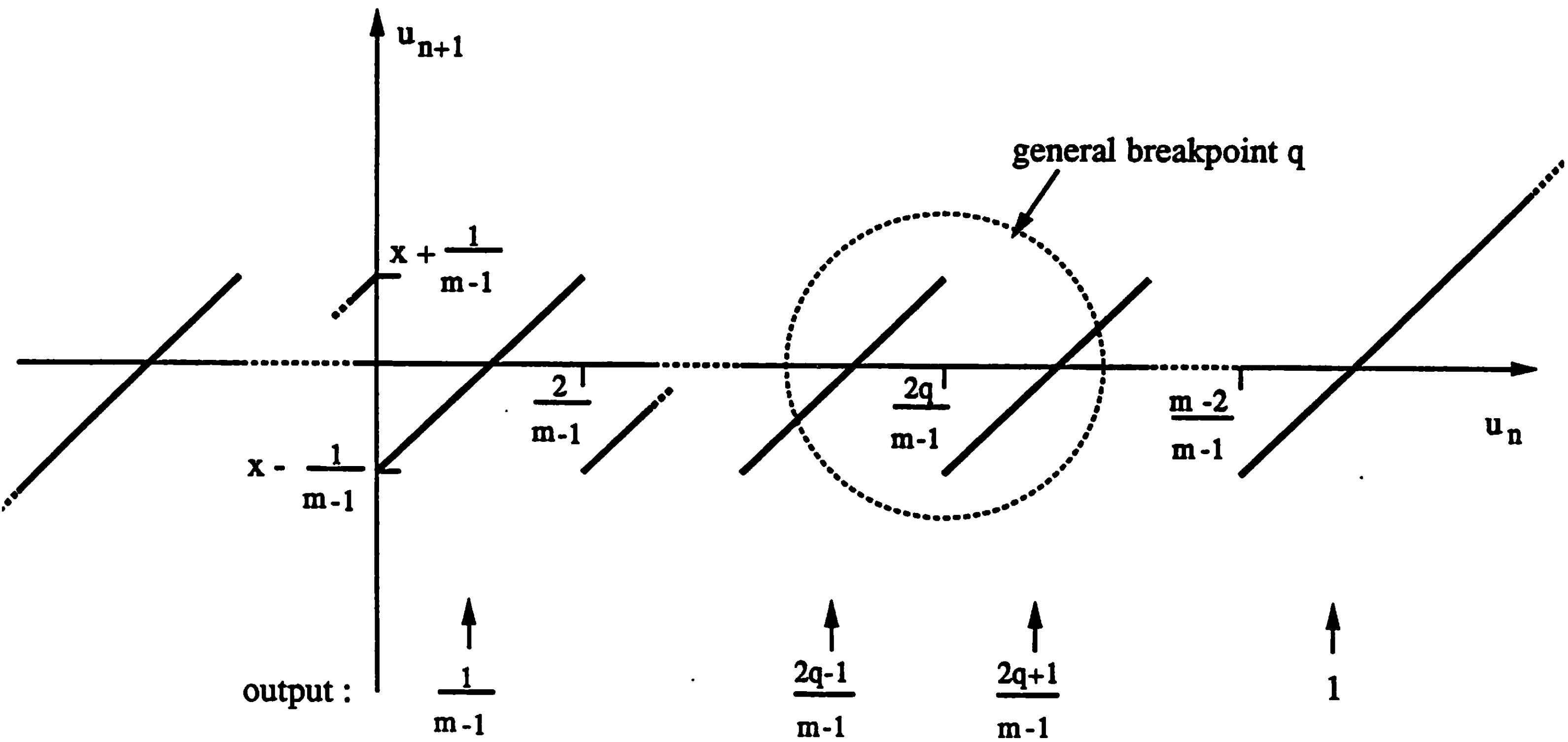
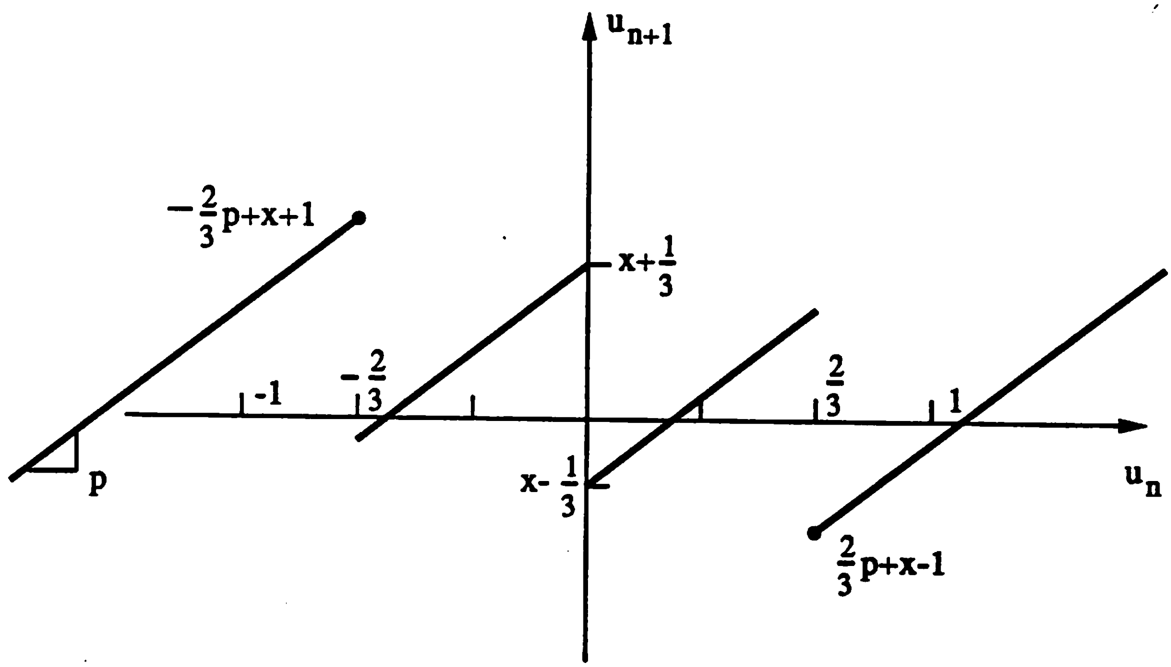

u_{n+1}
$\frac{2}{3}$
$\frac{2}{3}$
u_n
(i)
u_{n+1}
$x+\frac{1}{3}$
$x-\frac{1}{3}$
$x+\frac{1}{3}$
u_n
$x-\frac{1}{3}$
(ii)
u_{n+1}
$x+\frac{1}{3}$
$x-\frac{1}{3}$
$x-\frac{1}{3}$
$\frac{2}{3}$
$x+\frac{1}{3}$
u_n
(iii)

Figure 5



u_{n+1}
general breakpoint q
$x + \frac{1}{m-1}$
$x - \frac{1}{m-1}$
$\frac{2}{m-1}$
$\frac{2q}{m-1}$
$\frac{m-2}{m-1}$
u_n
output :
$\frac{1}{m-1}$
$\frac{2q-1}{m-1}$
$\frac{2q+1}{m-1}$
1

Figure 6



u_{n+1}
$-\frac{2}{3}p+x+1$
$x+\frac{1}{3}$
$-\frac{2}{3}$
-1
$\frac{2}{3}$
1
p
$x-\frac{1}{3}$
u_n
$\frac{2}{3}p+x-1$

Figure 7

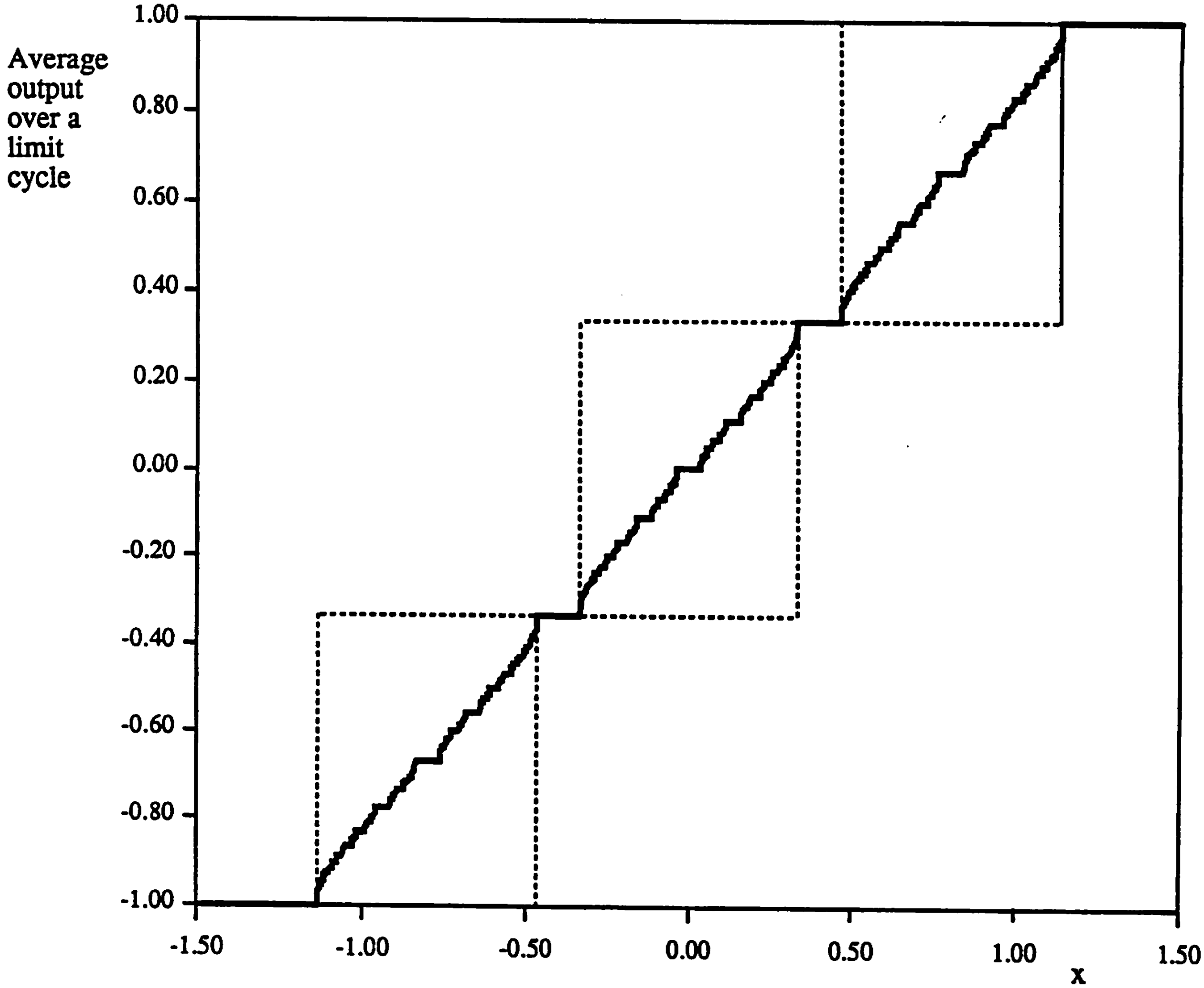

Average output over a limit cycle
1.00
0.80
0.60
0.40
0.20
0.00
-0.20
-0.40
-0.60
-0.80
-1.00
-1.50
-1.00
-0.50
0.00
0.50
1.00
1.50
x

Figure 8

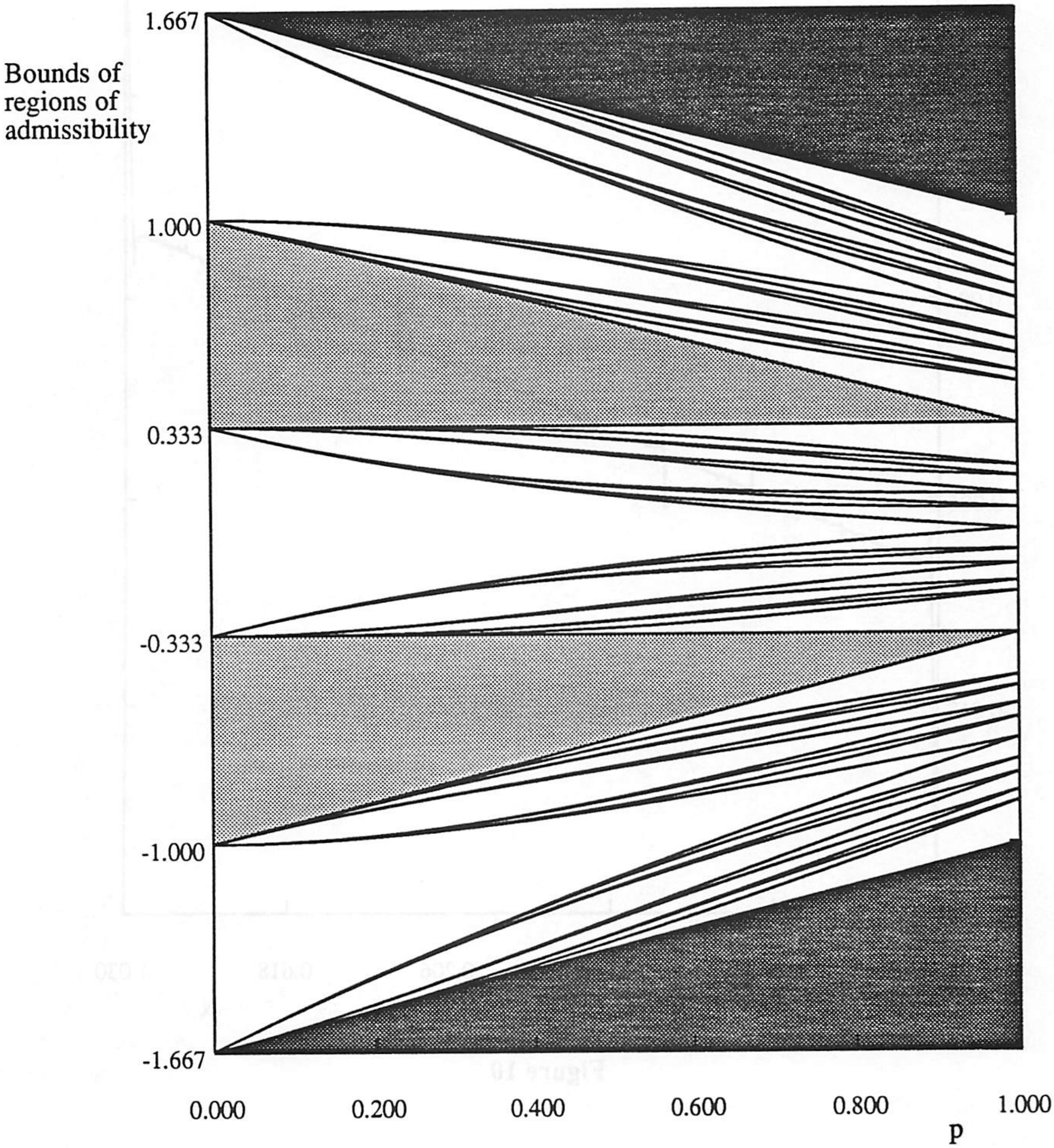

Bounds of regions of admissibility
1.667
1.000
0.333
-0.333
-1.000
-1.667
0.000
0.200
0.400
0.600
0.800
1.000
p

Figure 9

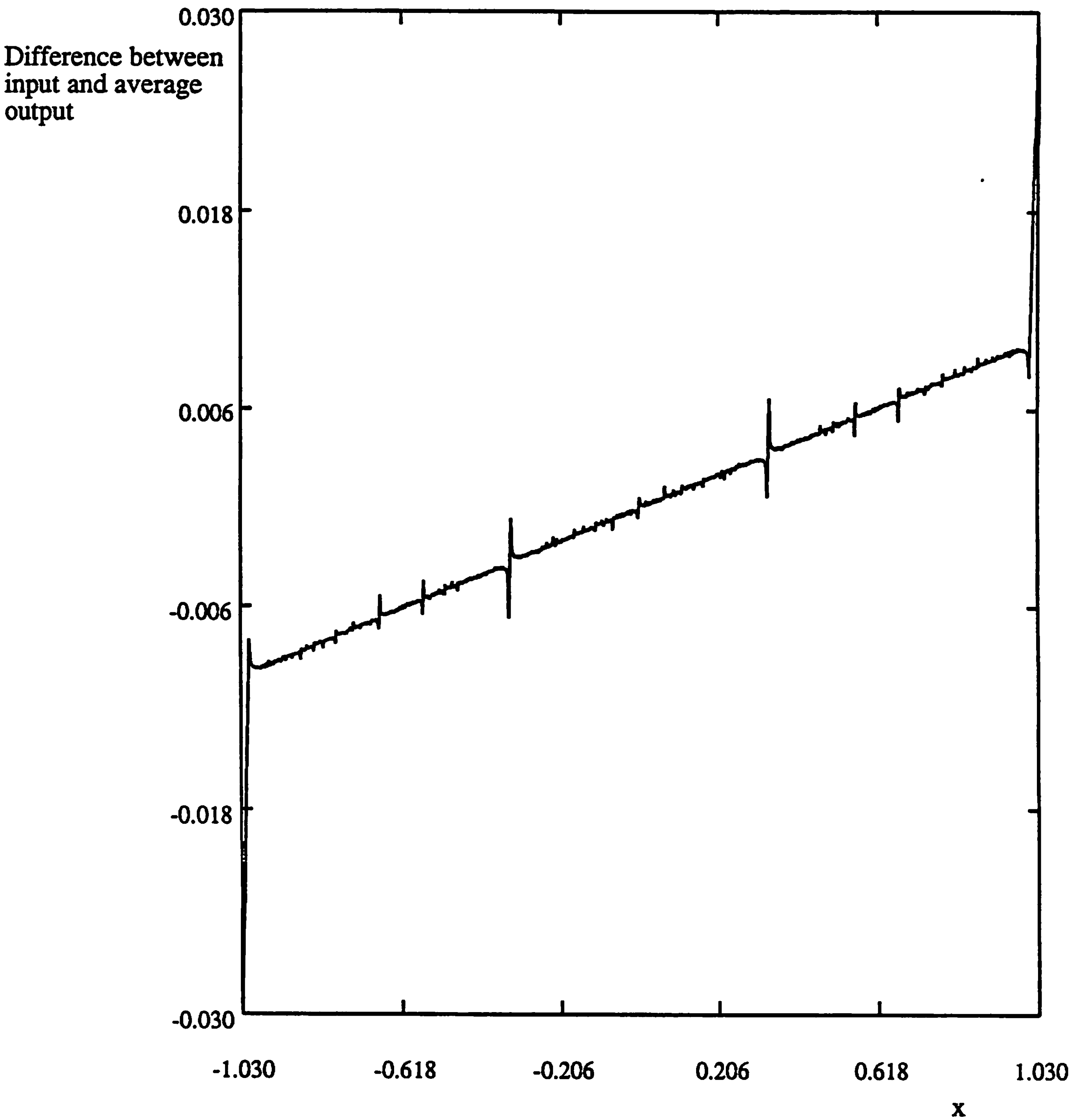

0.030
Difference between
input and average
output
0.018
0.006
-0.006
-0.018
-0.030
-1.030
-0.618
-0.206
0.206
0.618
1.030
x

Figure 10

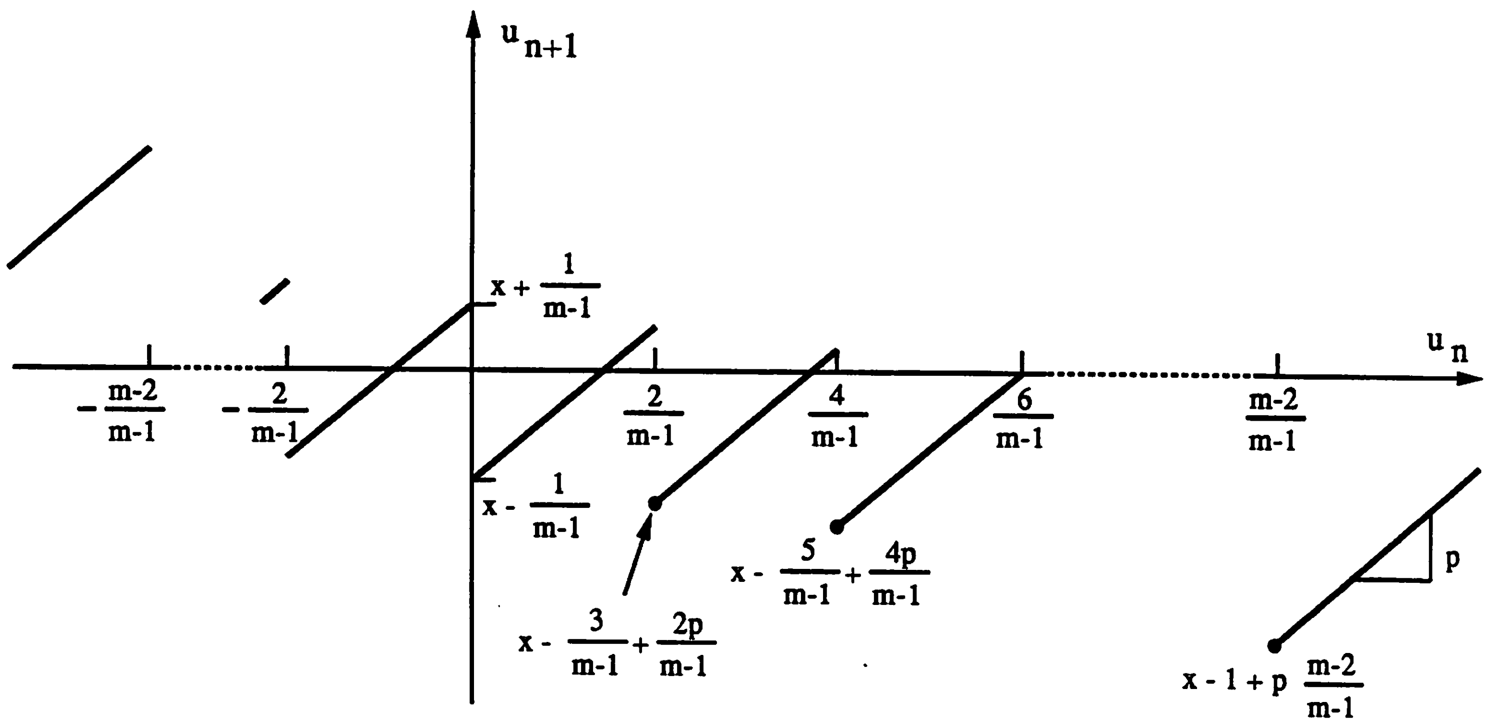

u_{n+1}
u_n
$x + \frac{1}{m-1}$
$x - \frac{1}{m-1}$
$-\frac{m-2}{m-1}$
$-\frac{2}{m-1}$
$\frac{2}{m-1}$
$\frac{4}{m-1}$
$\frac{6}{m-1}$
$\frac{m-2}{m-1}$
$x - \frac{3}{m-1} + \frac{2p}{m-1}$
$x - \frac{5}{m-1} + \frac{4p}{m-1}$
$x - 1 + p\,\frac{m-2}{m-1}$
p

Figure 11

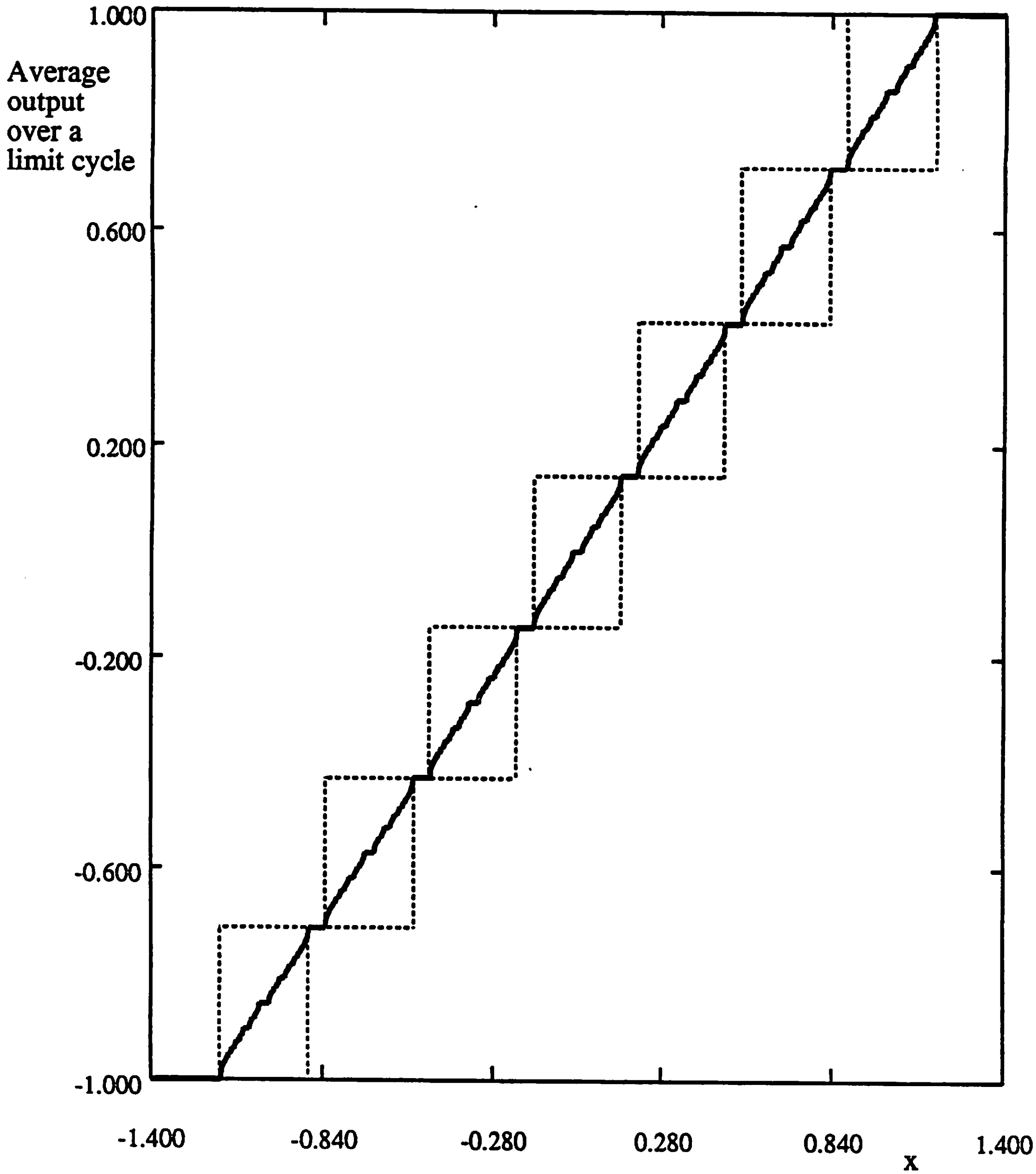
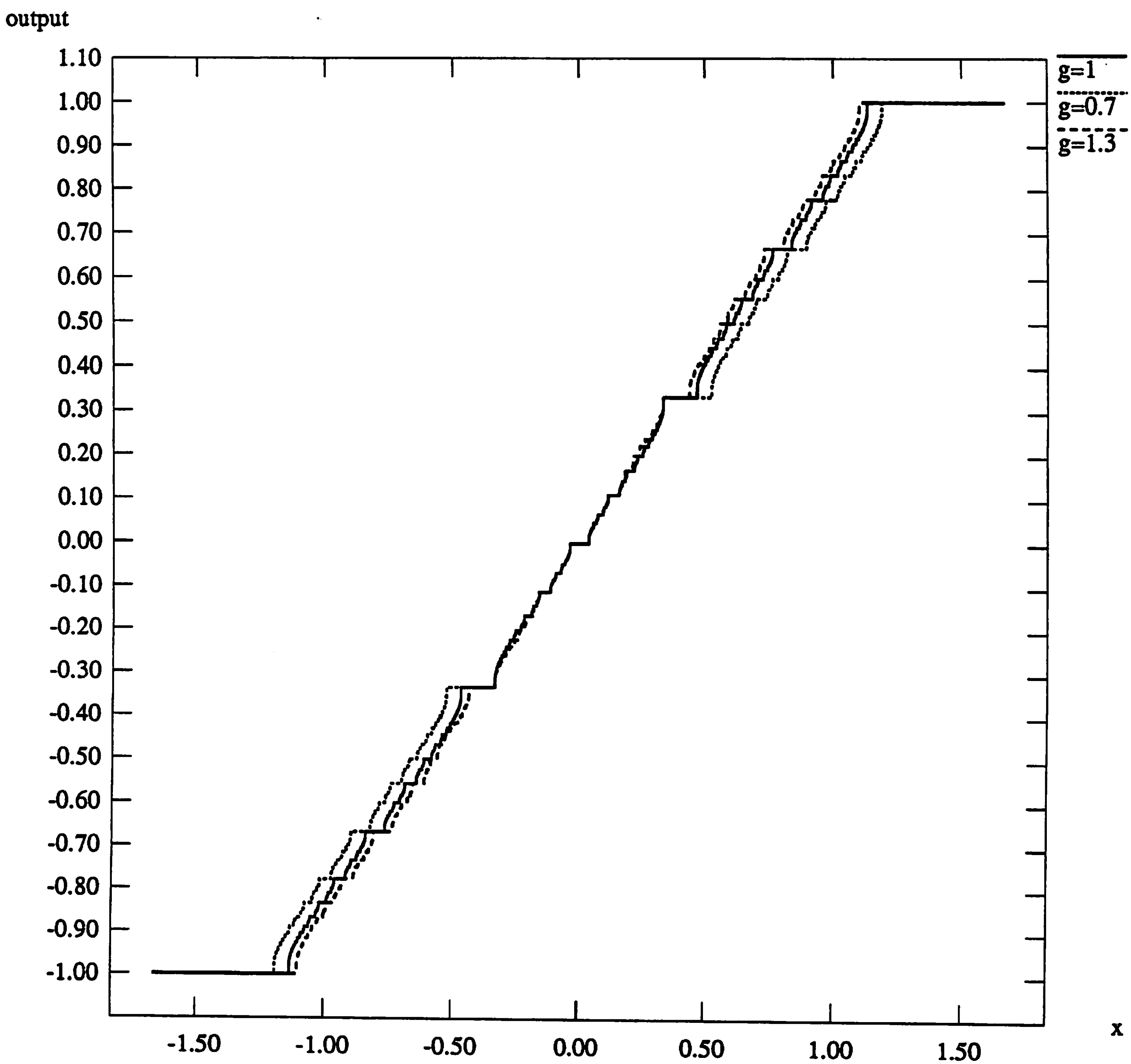
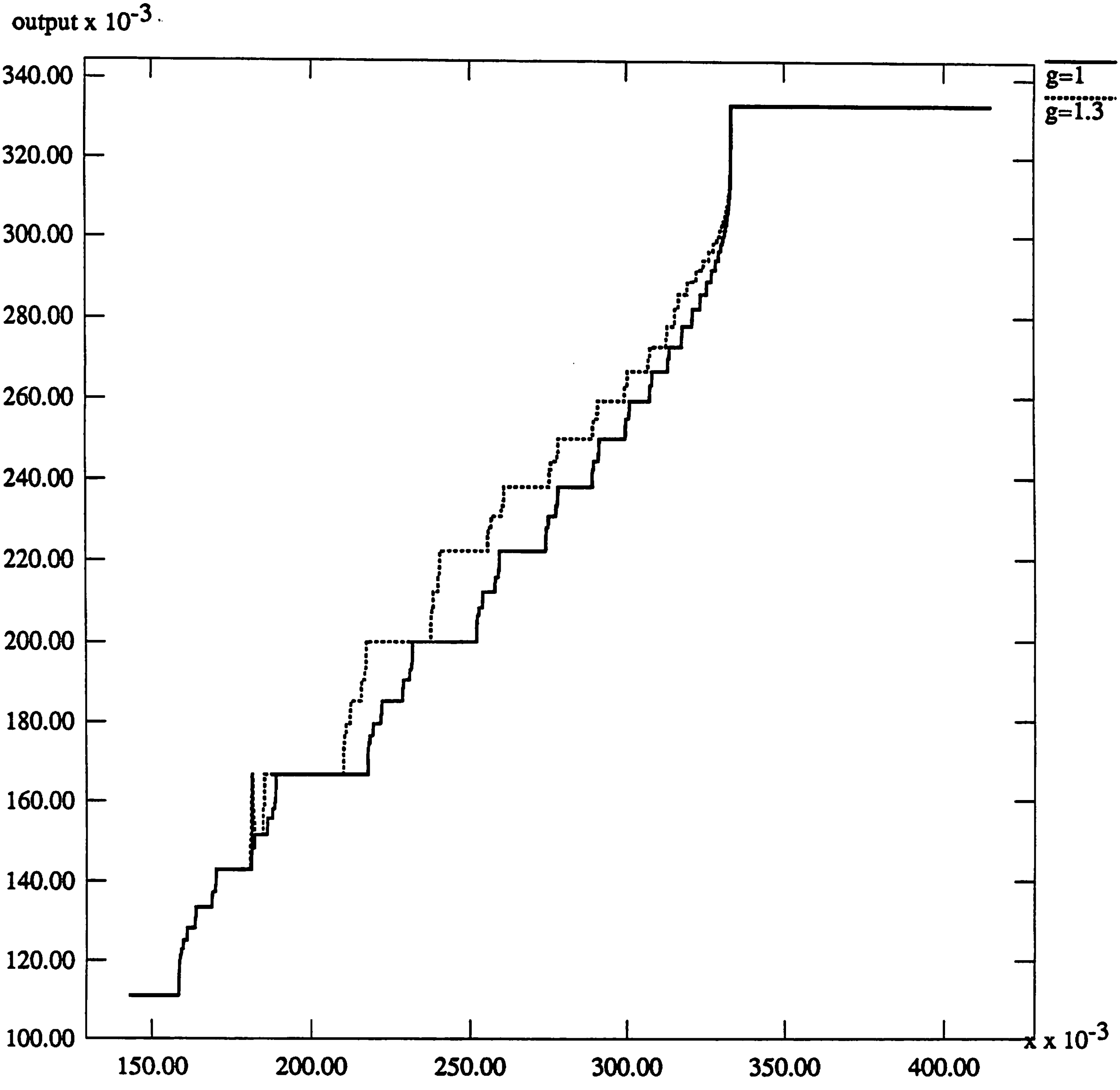

1.000
Average
output
over a
limit cycle
0.600
0.200
-0.200
-0.600
-1.000
-1.400
-0.840
-0.280
0.280
0.840
x
1.400

Figure 12



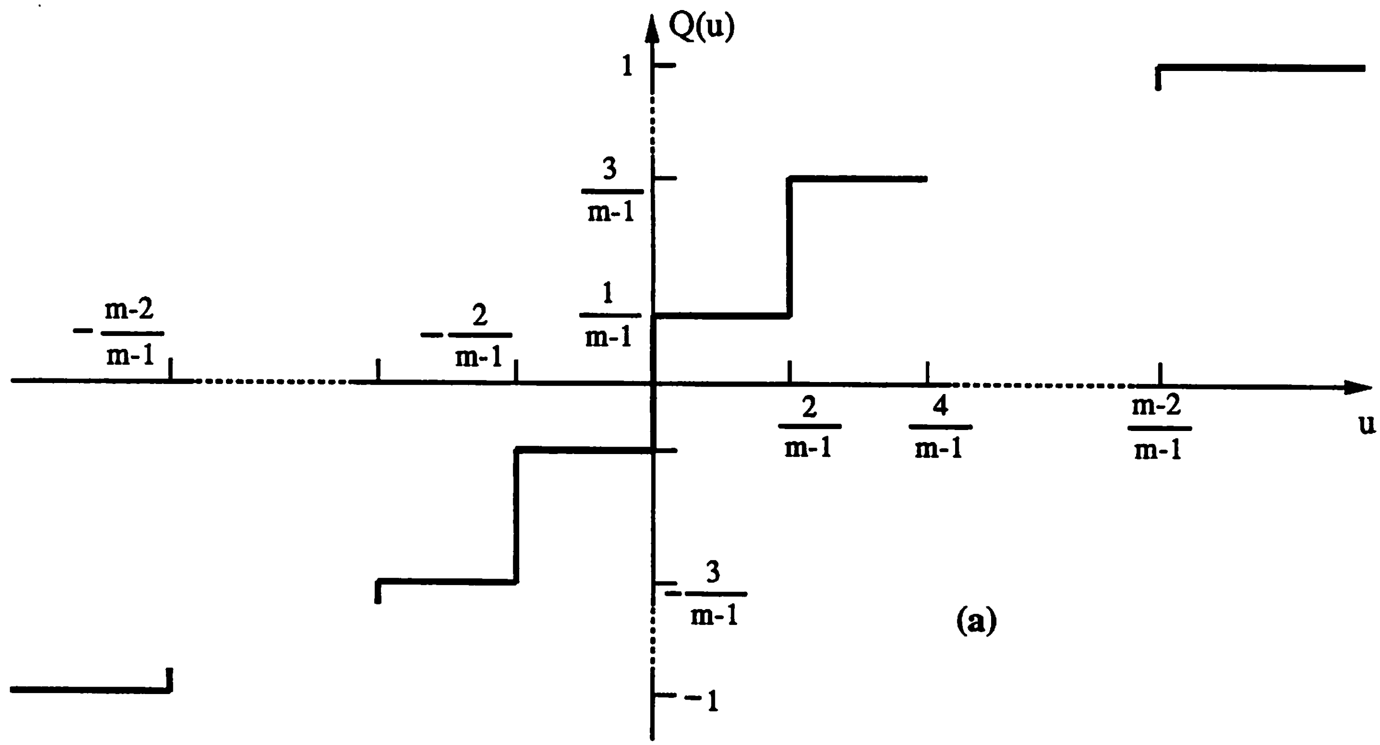
output
1.10
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
-0.10
-0.20
-0.30
-0.40
-0.50
-0.60
-0.70
-0.80
-0.90
-1.00
-1.50
-1.00
-0.50
0.00
0.50
1.00
1.50
x
g=1
g=0.7
g=1.3

Figure 13

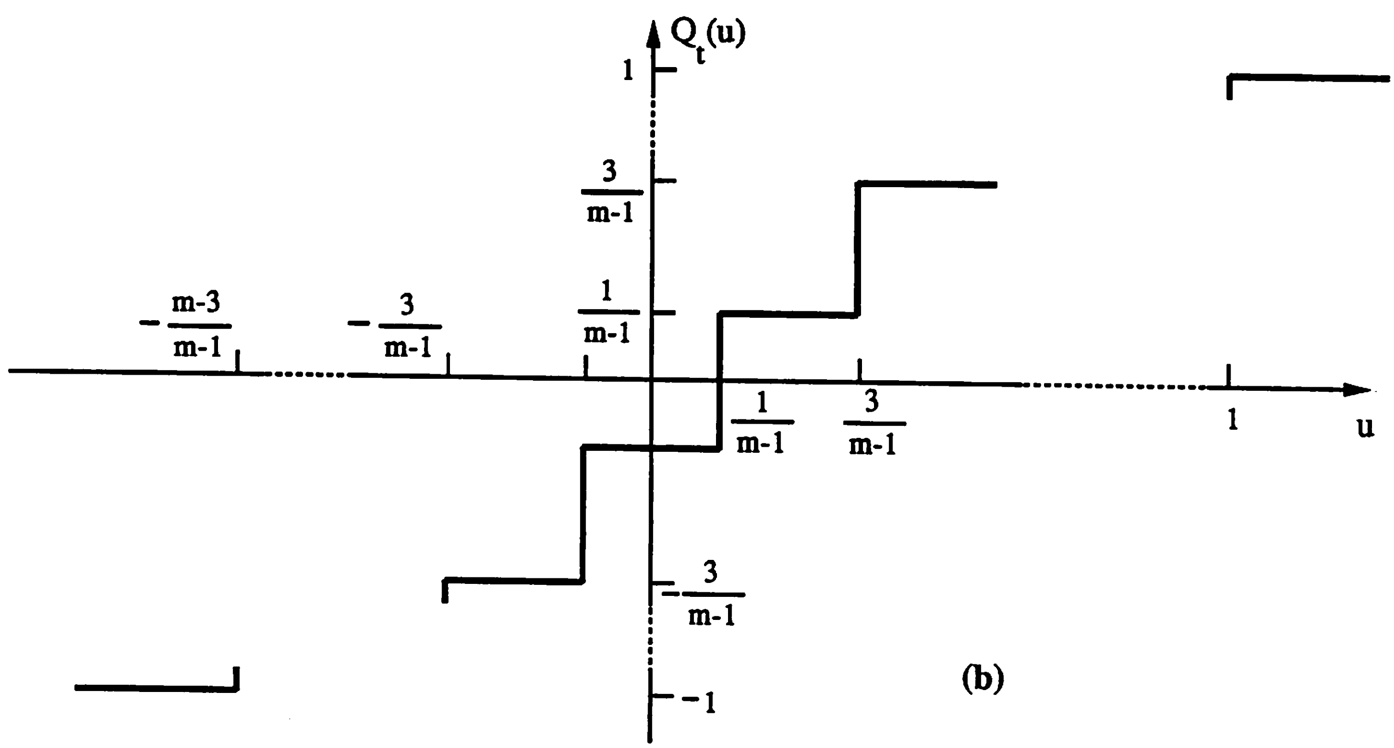


output x 10^{-3}
340.00
320.00
300.00
280.00
260.00
240.00
220.00
200.00
180.00
160.00
140.00
120.00
100.00
150.00
200.00
250.00
300.00
350.00
400.00
x x 10^{-3}
g=1
g=1.3

Figure 14

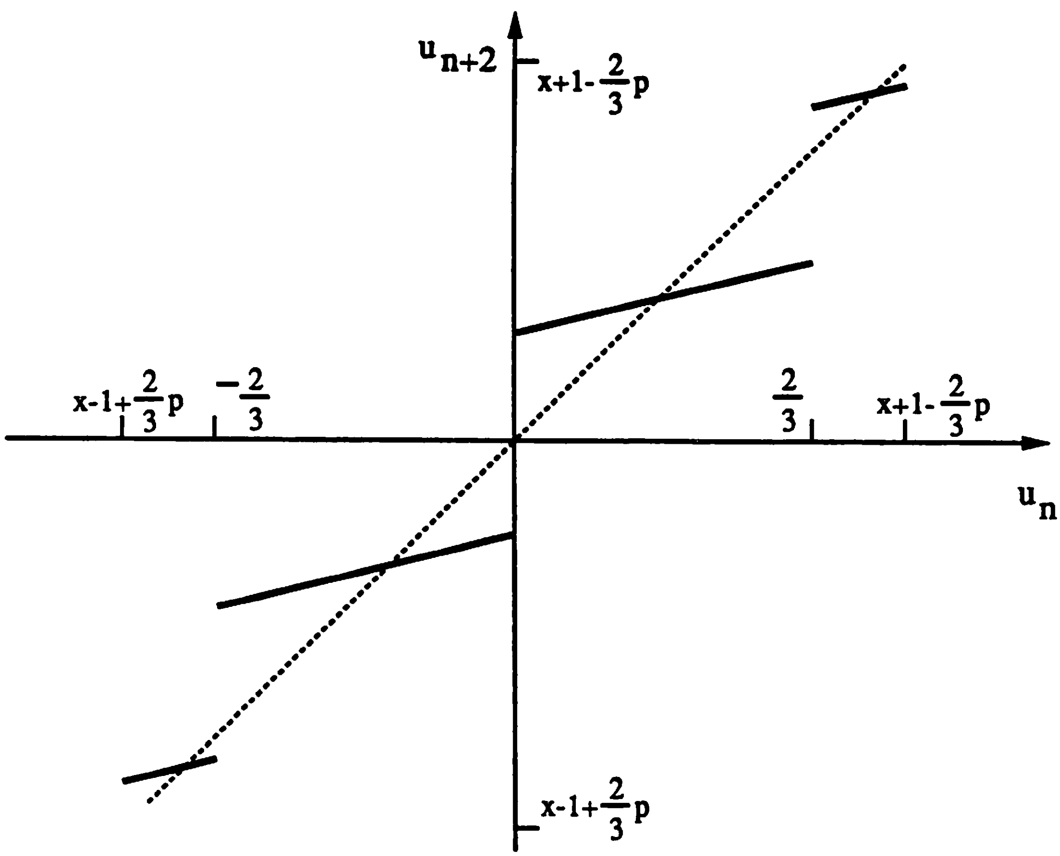


Q(u)
1
3
m-1
1
m-1
- m-2
m-1
- 2
m-1
2
m-1
4
m-1
m-2
m-1
u
- 3
m-1
-1
(a)



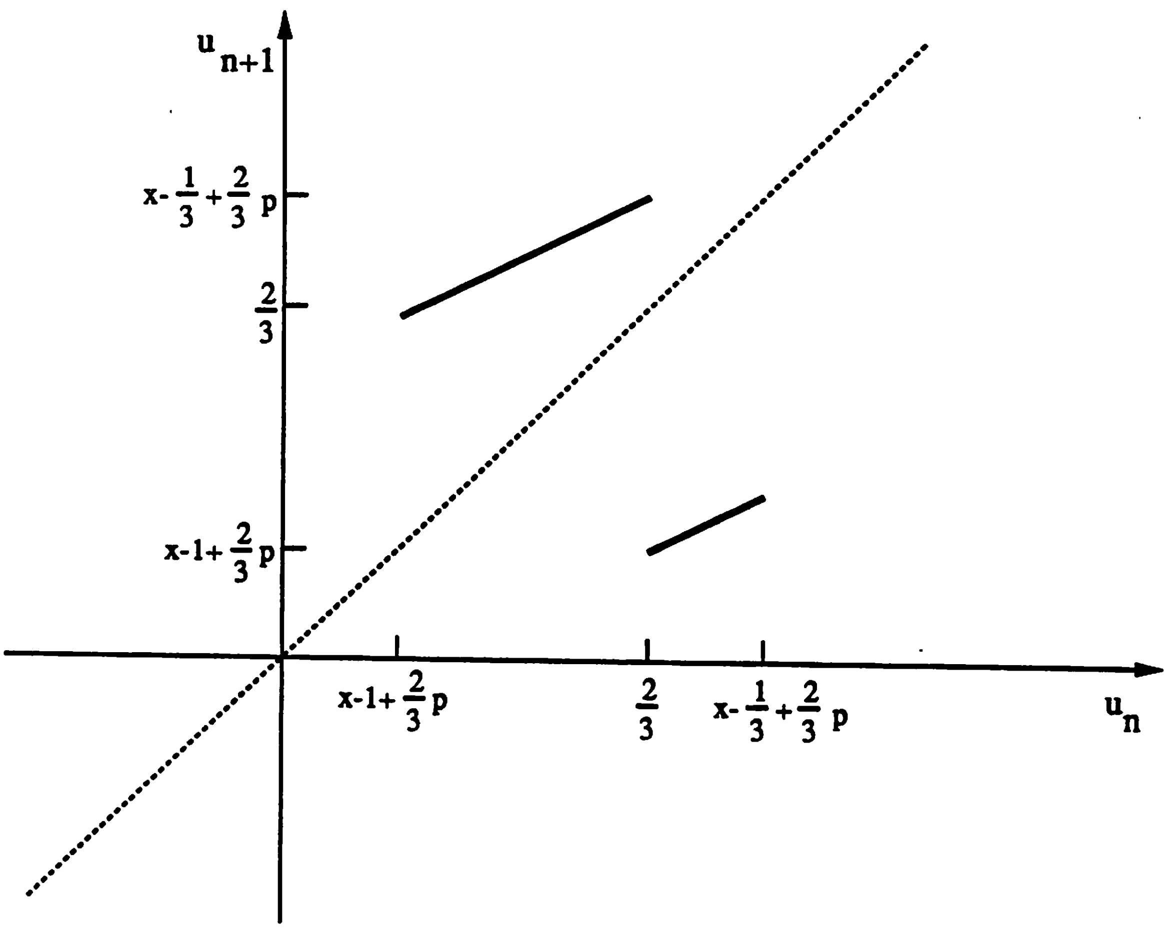
$Q_t(u)$
1
3
m-1
1
m-1
- m-3
m-1
- 3
m-1
1
m-1
3
m-1
1
u
- 3
m-1
-1
(b)

Figure 15



u_{n+2}
$x+1-\frac{2}{3}p$
$x-1+\frac{2}{3}p$
$-\frac{2}{3}$
$\frac{2}{3}$
$x+1-\frac{2}{3}p$
u_n
$x-1+\frac{2}{3}p$

Figure A1



u_{n+1}
$x-\frac{1}{3}+\frac{2}{3}p$
$\frac{2}{3}$
$x-1+\frac{2}{3}p$
$x-1+\frac{2}{3}p$
$\frac{2}{3}$
$x-\frac{1}{3}+\frac{2}{3}p$
u_n

Figure A2

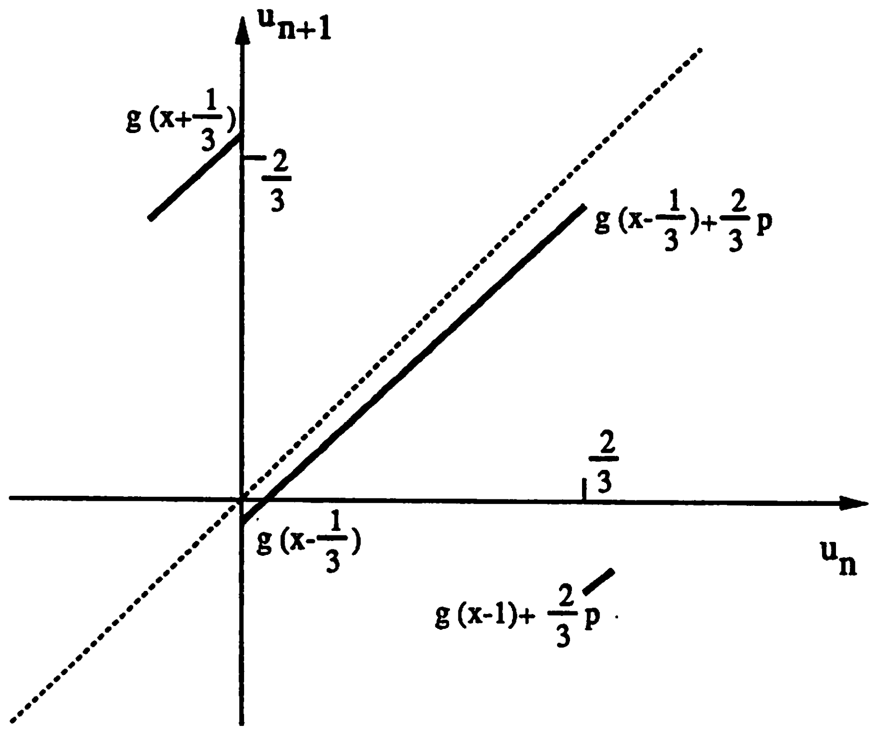

u_{n+1}
g (x+$\frac{1}{3}$)
$\frac{2}{3}$
g (x-$\frac{1}{3}$)+$\frac{2}{3}$p
$\frac{2}{3}$
g (x-$\frac{1}{3}$)
u_n
g (x-1)+ $\frac{2}{3}$p

Figure C1

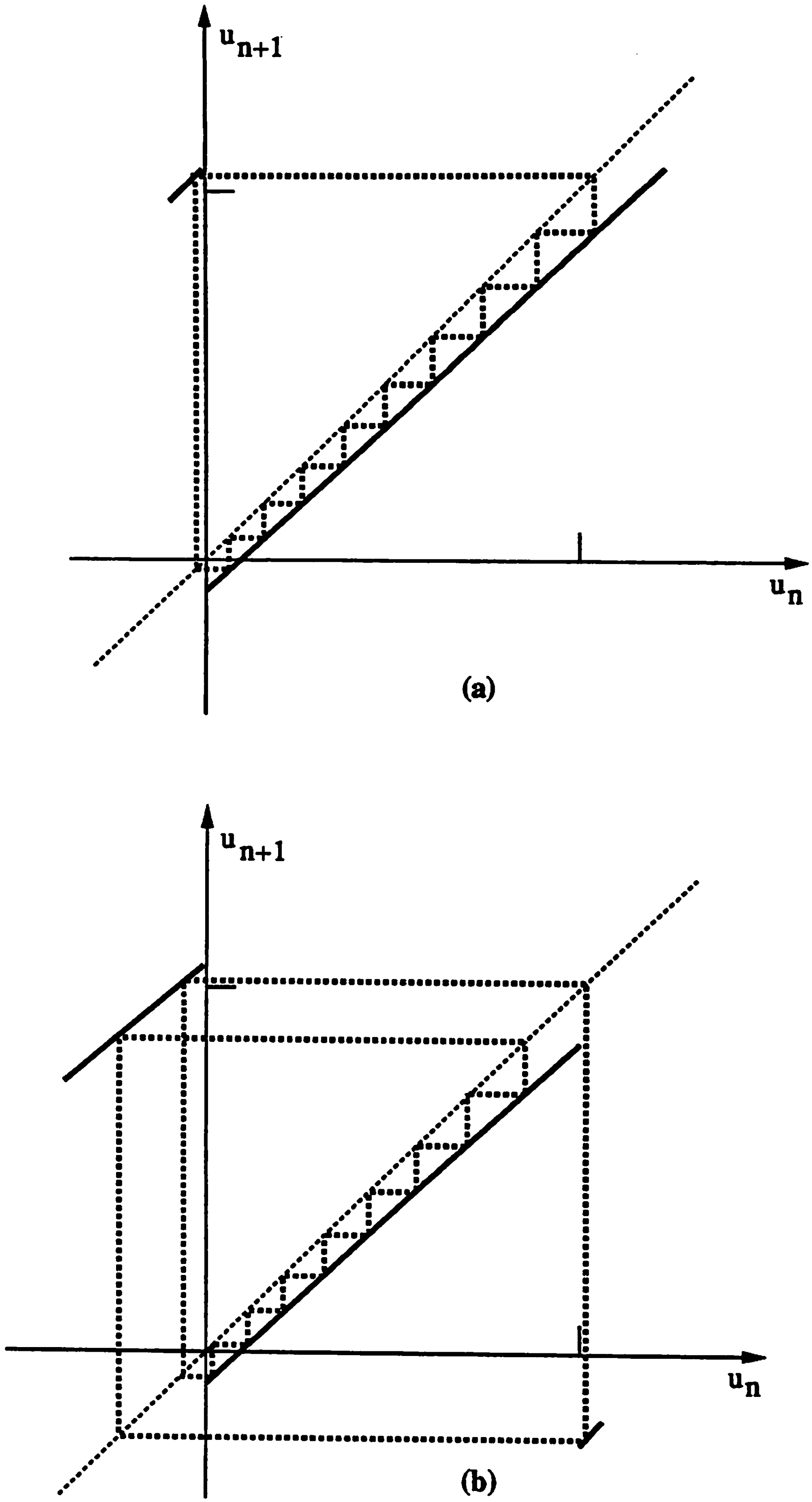


Figure C2