STABILITY AND EQUILIBRIA OF LINEAR CONTROL SYSTEMS UNDER INPUT AND MEASUREMENT QUANTIZATION

By

HANZ RICHTER

Bachelor of Science
Pontificia Universidad Católica del Perú
Lima, Perú
1995

Master of Science
Oklahoma State University
Stillwater, Oklahoma
1997

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
DOCTOR OF PHILOSOPHY
December, 2001

STABILITY AND EQUILIBRIA OF LINEAR CONTROL SYSTEMS UNDER INPUT AND MEASUREMENT QUANTIZATION

Thesis Approved:

____
Thesis Adviser

____

____

____

____
Dean of the Graduate College

ACKNOWLEDGMENTS

This work has been made possible -in great part- due to the weariless support of a number of people. My graduate advisor, Dr. Eduardo A. Misawa has guided my work and encouraged me to pursue it, even during the less productive times. I sincerely thank his patience and vision. I also wish to thank the members of my advisory committee, Dr. Prabhakar Pagilla, Dr. Gary Yen, and, in particular, Dr. Don A. Lucca for motivating my interest in areas of Mechanical Engineering other than Systems and Control. I also wish to acknowledge Dr. Hongbing Lu for his past support to my academic development.

My sincere thanks also to the Oklahoma State University Distingushed Graduate Fellowship Program, whose financial support during three years of study have been invaluable and to Seagate Technology, the Oklahoma Center for the Advancement of Science and Technology, grant No. AR-992-016, and the National Science Foundation, grant CMS-9978748 for their sponsorship of the research projects which motivated this work.

Finally, I wish to especially thank my wife Francisca for her unqualified support of my work and my self, my parents, for being a model of endurance, and my father-in-law, Dr. Ramón García-Cobián, for his valuable remarks on some mathematical aspects of my work.

"If a man does his best, then what else is there?" - *General George Patton*

TABLE OF CONTENTS

Chapter Page

LIST OF TABLES

LIST OF FIGURES

Figure		Page

LIST OF SYMBOLS

$\mathbb{R}$	the set of real numbers
$\mathbb{Z}$	the set of integers
$\mathbb{Q}$	the set of rationals
Q	scalar quantization
$\mathcal{Q}$	vector quantization
sup	supremum (least upper bound)
inf	infimum (greatest lower bound)
gcd	greatest common divisor
lcm	greatest common multiple
$a\|b$	a divides b (the remainder of b/a is zero
Re	real part
$\|x\|$	absolute value of x
$\|\|F\|\|_p$	p-norm: $[\sum_{i=1}^{n} \|F_i\|^p]^{1/p}$
sgn(x)	1 if $x > 0$, -1 if $x < 0$, undefined if $x = 0$.
ceil(x)	least integer greater or equal to x.
$\bigcap$	set intersection
$\bigcup$	set reunion
$A \subset B$	set A is included in set B, with $A \neq B$
$A \subseteq B$	set A is included in or equal to set B

Chapter 1

Introduction

The study of quantization and finite precision effects in control has been motivated by the high accuracy requirements of modern systems. Despite the fact that the resolution of a data converter is improved by a factor of two when one more bit is used, commercial converters often have less than 14 bits. If one of such devices is used at the input and output of a digital controller, a limit cycle is likely to arise. The amplitude of quantization-induced limit cycles is reduced when the quantization step size is reduced. Therefore, a limitation on the number of bits in a data converter may result in a limit cycle whose amplitude is too large for a given performance requirement. In fact, an arbitrary state feedback gain which stabilizes the nominal system may result in a limit cycle, or the state trajectories may converge to an equilibrium point different than the origin. In the case of linear systems, it is possible to design feedback controllers which do not induce limit cycles, regardless of the quantization step size. In this research work, feedback controllers are derived such that the system remains stable, despite the presence of input and output quantizers. This is done for systems with either full state feedback, or with only output feedback and dynamic compensation. The system under study, in its more general form, is shown in Figure 1.1. A quantizer outputs a nonlinear function of its input. This function has, in addition, an infinite number of discontinuities, which creates difficulties in any analysis technique requiring differentiation. In this work, attention is centered on discrete time linear time-invariant (LTI) systems with quantized feedback. More precisely, the following closed-loop configurations are studied:

$$\begin{cases} x(k+1) = Ax(k) + Bu(k) \\ u = -\mathcal{Q}(Fx(k)) \end{cases} \tag{1.1}$$

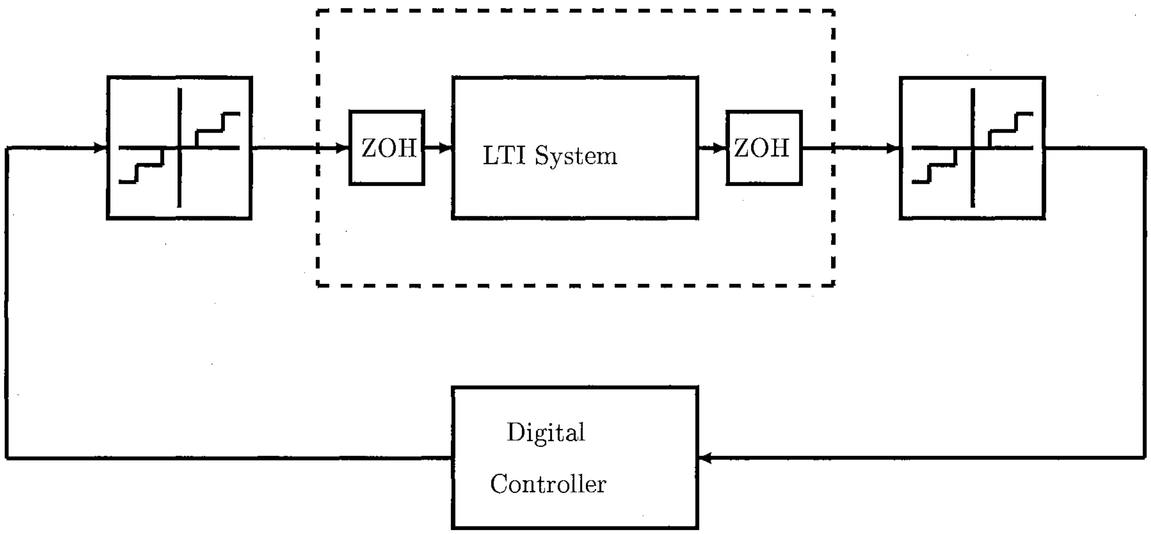

Figure 1.1: System configuration under study

and

$$\begin{cases} x(k+1) = Ax(k) + Bu(k) \\ u = -\mathcal{Q}(F\mathcal{Q}(x(k))) \end{cases} \tag{1.2}$$

where $\mathcal{Q}(.)$ denotes the quantization operator whose graph is shown in Figure 1.2. The quantization step size is denoted by q. The equations corresponding to the output feedback configurations are given in Chapter 4. The system in Eq.(1.1) corresponds to the practical case where only a digital-to-analog (D/A) converter is used, or, more realistically, when its resolution is much lower than that of the analog-to-digital converter (A/D). The system in Eq.(1.2) has more relevance in a practical context. It represents the common situation of a control computation based on quantized measurements which is quantized when it leaves the digital environment of the control computer and enters the plant. The number of bits of both input and output data converters is assumed to be equal.

Mathematically, the scalar and vector quantization functions are defined below:

Definition 1.1. *The scalar quantization function is a mapping $Q: \mathbb{R} \to \mathbb{R}$ with rule of correspondence*

$$\mathcal{Q}(y) = jq$$

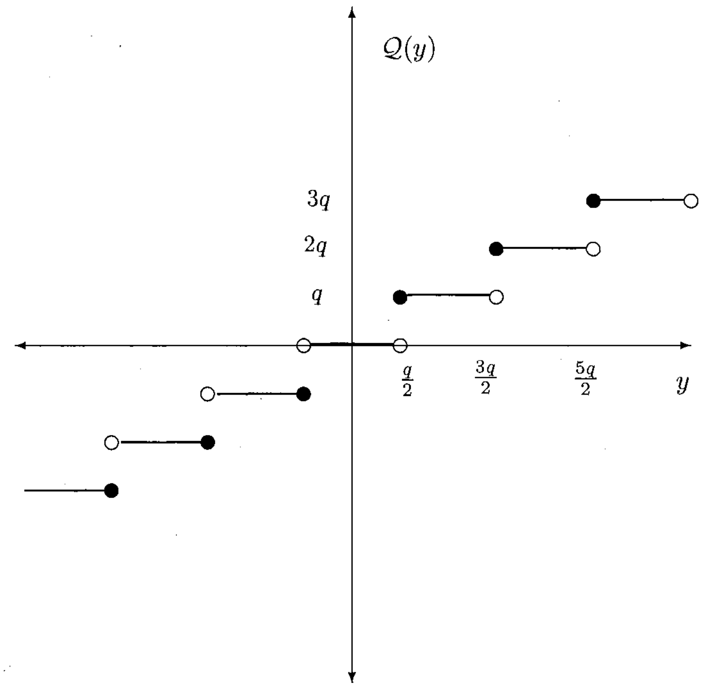

Figure 1.2: The scalar quantization operator

where j is an integer satisfying

$$\frac{(2j-1)q}{2} \leq y < \frac{(2j+1)q}{2} \quad , \; j \geq 1$$
$$\frac{(2j-1)q}{2} < y \leq \frac{(2j+1)q}{2} \quad , \; j \leq -1$$

The number $q \in \mathbb{R}^+$ is the quantization step size.

Definition 1.2. *The vector quantization function is a mapping $\mathcal{Q} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ with rule of correspondence*

$$\mathcal{Q}(x) = \xi \text{ , where}$$
$$\xi_i = Q(x_i)$$

As explained in Sections 1.1 and 1.2, the approach taken in this work is to utilize the well-established theory of Absolute Stability in the analysis. In this regard, the material of this work is presented for the first time, to the author's knowledge. This approach departs from the mainstream analysis technique which replaces the quantizer by a pass-through and an additive noise, or quantization error, which is bounded by $\frac{q}{2}$ in absolute value. The motivation for the use of Absolute Stability is that the scalar quantizer is a sector nonlinearity, and that some manipulations can be performed in the case of vector quantization (i.e.,when the state measurement is quantized) in order to analyze the system in terms of sectors.

1.1 Review of the Literature

The general topic of finite precision effects in control systems has received considerable attention. It is found that three broad problems predominate in the literature. Each problem has been approached by a different method. The next subsections are a brief overview, by no means exhaustive, of the representative works.

1.1.1 Optimal realizations of finite precision controllers

Control algorithms for mass-market devices such as automotive engines are often embedded in a digital signal processor, or DSP. Cost and size limitations often result in devices that must operate with a limited number of bits. It is known that the propagation of numerical errors due to finite precision arithmetic is sensitive to the state-space realization chosen for the controller. Therefore, the problem is to find state-space representations which minimize error accumulation. In particular, the optimal realization is sought. Works in this area are often associated with the names of Li and Gevers ([23], [24]), Istepanian ([19]), Collins [33] and others. In [19], a method for finding the optimal finite precision realization using numerical optimization is given. The method provides a realization that requires the minimum number of bits and gives the maximum stability bound. In [23] the optimal finite precision realization of an observer-controller combination is sought. The paper's main contribution is the derivation of an expression to calculate the sensitivity of the numerical errors with respect to the realization of a given transfer function. With this tool, the set of state-space realizations minimizing the sensitivity is computed. The paper by Chen *et.al.* [8] obtains results for the particular case of PID control.

Role of the delta operator

The delta operator is defined from the conventional forward-shift operator z as

$$\delta \triangleq \frac{z-1}{\Delta}$$

where Δ is an arbitrary positive number. For example

$$\delta x(k) = \frac{x(k+1) - x(k)}{\Delta}$$

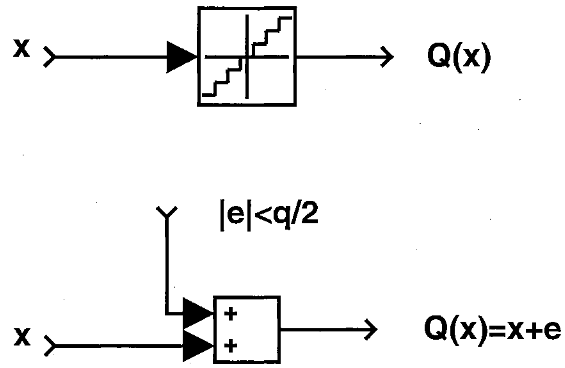

Figure 1.3: The additive model

When a transfer function for the controller is given, it may be written in terms of the delta operator, resulting in a set of coefficients different than the ones in the conventional z operator transfer function. Therefore, state space realizations are different. The works by [24], [27] and [33], among others, either claim or analyze the superiority of the delta operator parameterization. Arguably, a shortcoming of this family of approaches is that it lacks guarantees of stability, that is, the question of whether the optimal realization is stable is typically not addressed, under the rationale that it is sufficient to keep the errors small.

1.1.2 Statistical analysis and limit cycles in digital filters

The quantization error is always bounded in absolute value by one half of the quantization step size. This has motivated many researchers to replace quantization by an additive noise model, shown in Figure 1.3. This model has lead to a number of papers that address the existence of limit cycles in digital filters; for instance, Leclerc and Bauer [21] introduce a computer-aided test for existence of limit cycles; and Bose [5] analyzes the stability of second-order digital filters, providing stability regions for the filter coefficients. Several researchers have also focused on the spectral properties of the quantization noise. Based on this, controllers that are insensitive to noise in such spectrum can be designed. An example of this approach is the paper by Liu and Skelton [25], in which a design method is provided for LQG control, taking into account data converter quantization and roundoff error from the control computer. The paper, however does not guarantee that the designed gains will be stable. In Section 2.5 of this work it is shown how can a nominal LQR gain destabilize a system. In the seminal work by Widrow *et.al* [36], the quantizer is analyzed in depth in terms of its statistical properties. The probability distribution function (pdf) of the quantizer output is seen as a sampling of the input pdf. The characteristic function[1] of the output is used to derive reconstruction theorems with a surprising resemblance to the sampling reconstruction theorem due to Shannon. These theorems establish that if the characteristic function of the input is in some sense "bandlimited", then statistical properties of the output can be derived from those of the input, including the pdf and moments.

1.1.3 Exact approaches: Nonlinear systems analysis

In these approaches, quantization is analyzed in a deterministic setting, often arriving at stability conditions applicable to various kinds of control structures. Nonlinear analysis is used to arrive at such conditions, which are exact. Unlike the previous two, this kind of treatment has not been followed by many researchers. One of such approaches was published by Brockett and Liberzon [6]. In this paper, only quantized measurements are considered. The control signal is post-multiplied by a time-varying coefficient to make up for quantizer imperfections. The paper by Delchamps [12] is of great importance in the context of the present research work. It shows that open-loop unstable systems cannot be asymptotically stabilized in general when the measurements are quantized. However, it shows that, in certain cases, it is possible to bring the state arbitrarily close to zero and remain there indefinitely. A problem closely related to the subject of the present work is determining when a difference equation of the form

$$y(k) = Q[a_1 y(k-1)] + Q[a_2 y(k-2)] + .. + Q[a_n y(k-n)]$$

is BIBO stable. This problem is analyzed by Bauer and Leclerc in [3]. The results are in terms of the sum of the absolute values of coefficients a_i. This resembles the stability condition in the present work, which depends on the 1-norm of the feedback gain.

1.1.4 Consulted material on Absolute Stability

In this work, the theory of Absolute Stability is extensively used. This theory is well-documented for continuous time systems, and is part of standard textbooks on nonlinear control. For discrete time, however, there is less material available. This has motivated simple extensions to existing results. The background theory required has been extracted mainly from two papers by Hitz and B.D.0 Anderson ([17], [2]), and the textbooks by Hsu and Meyer [18] and Aizermann [1].

[1] The characteristic function of a random variable is the Fourier transform of the pdf.

1.2 Objectives, Scope and Methodology

The objectives of this work are to obtain a simple stability tests for discrete time systems of the form of Eqs. (1.1)and (1.2), and also for systems with output feedback and dynamic compensation. It is expected that the tests developed are useful only positively. That is, if a system passes the test, then it will be stable, but not necessarily viceversa. This is due to the fact that Absolute Stability typically provides only sufficient conditions. It is also a goal to characterize the equilibrium points of quantized control systems, in order to increase understanding of the subject. The methodology followed is constructive, starting from simple cases and building results progressively. Standard material has been placed at the appendix, while proofs of the contributed lemmas and theorems are in the chapters, with the exception of certain derivations which are sent to the appendix to avoid disrupting the logical flow.

1.2.1 Limitations

The following are general assumptions that pertain to the work as a whole. Particular assumptions are stated with the corresponding lemmas and theorems.

- The open-loop system is of single input and asymptotically stable, except in Section 2.4, where unstable systems are allowed.
- Control computations occur at infinite precision, or at a much higher precision than the resolution of the data converters.
- Both, input and output quantizers have the same quantization step size.
- Analysis is limited to regulation about the origin.

1.3 Contributions

To the best of the author's knowledge, this is the first time the problem of quantized feedback is addressed with the tools of Absolute Stability. The most relevant contributions are now listed:

1.3.1 Quantized Feedback with Precise State Measurememts (QI)

- Complete characterization of equilibria, including singular cases and criterion for uniqueness of equilibrium. Results for continuous and discrete systems.

- For continuous QI systems, proof of existence of a solution in the sense of Filippov.
- Stability of continuous QI systems. Chattering control for unstable systems.
- Stability of discrete QI systems: Parameterization of stabilizing gains.

1.3.2 Quantized Feedback and State Measurements (QIQM)

- Graphical method for equilibrium finding.
- Multiplicative perturbation theorem.
- Stability criterion for QIQM systems.
- Stabilization by gain scaling.
- Example of bifurcations in QIQM systems.

1.3.3 Quantized Feedback with Precise Output Measurement (QIO)

- Characterization of equilibria by reduction to QI case.
- Stability problem: Reduction to QI case.
- Relationship with output quantization case (IQO).

1.3.4 Quantized Input and Output Measurement (QIQO)

- Equilibrium problem: Sufficient condition for uniqueness of equilibrium and iterative method of solution.
- Stability analysis by multiplicative perturbation, discrete positive real, and small gain methods.

Perhaps the most important contribution of this work is the analysis of systems with quantized feedback based on quantized state measurements (QIQM). The result is a stability criterion restricting the location of the polar diagram of a system transfer function. The polar diagram must stay to the right of a vertical line located at a point which depends on the 1-norm of the feedback gain. The polar diagram itself depends on such norm. Due to this fact, the criterion is at this point, useful only for analysis; or for design by iterative process.

Chapter 2

Quantized Feedback with Precise State Measurements (QI)

2.1 Characterization of Equilibrium Points

2.1.1 Equilibrium analysis: Continuous time

For the analysis consider the following LTI system under quantized state feedback:

$$\begin{cases} \dot{x} = Ax + Bu \\ u = -\mathcal{Q}(Fx) \end{cases} \tag{2.1}$$

where $\mathcal{Q}(.)$ denotes the quantization operator. Existence of a solution to the above equations is guaranteed if a concept of solution of differential equations with discontinous right hand sides is used. In particular, the concept of solution introduced by Filippov is used in the appendix to show existence in this case. Assume A is nonsingular. Then the equilibrium points of the closed loop system are solutions of

$$x = A^{-1}B\mathcal{Q}(Fx) \tag{2.2}$$

One solution method is to use fixed-point iteration.

Example 2.1. Take $q = 1$ and

$$A = \begin{bmatrix} 1 & 2 \\ -3 & -2 \end{bmatrix}$$
$$B = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$
$$F = \begin{bmatrix} 1 & 1.8 \end{bmatrix}$$

Starting the iteration $x_{k+1} = A^{-1}BQ(Fx_k)$ with $x_0 = [5,\ 5]^T$ gives an equilibrium point at $x = [-1.5,\ 1.25]$.

The disadvantages of this method of solution are the lack of guaranteed convergence and the need for an initial guess. The method does not tell the number of equilibrium points or how to avoid multiple equilibrium points. This motivates the search for an analytical, closed-form solution of Eq.(2.2).

2.1.2 Solution to scalar case

As a starting point to constructing the solution, the following equation is studied

$$x = kQ(x) \tag{2.3}$$

where k is a constant, x is a scalar, and the quantization step is q. Graphically, solving the equation amounts to finding the intersections of $Q(x)$ with the line $y = \frac{1}{k}x$. Figure 2.1 shows the construction of the solution.

It is clear that a nonzero solution exists only when $0.5 \leq k < 1.5$. The solutions are just the pre-images of multiples of the quantization level through the straight line, that is

$$x_i = ikq\ ,\ i = 0, \pm 1, \pm 2... \tag{2.4}$$

Number of solutions

The number of solutions is obtained by the geometrical condition that the intersection point must be within $q/2$ from the center of the horizontal segments which comprise $Q(.)$. That is,

$$|ikq - iq| \leq \frac{q}{2}$$

so

$$|i| \leq \frac{1}{2|k-1|}$$

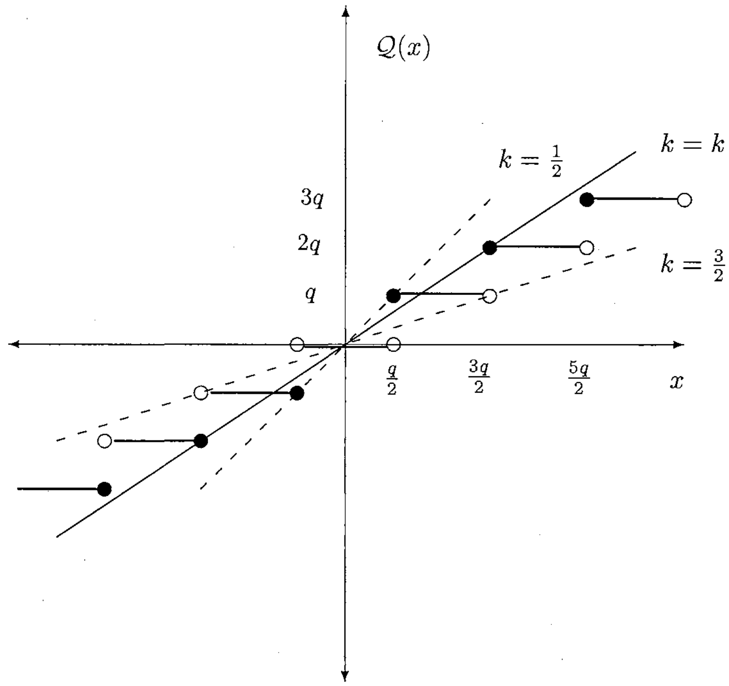

Figure 2.1: Solution to scalar quantization equation

More precisely, the solutions are given by $x_i = ikq$, where i is an integer such that:

$$
\begin{aligned}
|i| &\leq \frac{1}{2|k-1|}, \text{ when } k < 1 \\
|i| &< \frac{1}{2|k-1|}, \text{ when } k > 1
\end{aligned} \tag{2.5}
$$

If $k = 1$, the number of solutions is infinite, as seen graphically and from the above formula. Note that the number of solutions is independent of q.

Now consider the slightly more general scalar equation

$$x = kQ(\alpha x) \tag{2.6}$$

where k, $\alpha \neq 0$. Changing the variables to $x' = \alpha x$ results in the equation

$$\frac{x'}{\alpha k} = Q(x')$$

which reduces to the earlier case. Nonzero solutions exist when $0.5 \leq \alpha k < 1.5$. The solutions are of the form

$$x_i = ikq \ , \ i = 0, \pm 1, \pm 2 ... \tag{2.7}$$

and the range of index i is given by

$$\begin{aligned} |i| &\leq \frac{1}{2|\alpha k - 1|}, \text{ when } \alpha k < 1 \\ |i| &< \frac{1}{2|\alpha k - 1|}, \text{ when } \alpha k > 1 \end{aligned} \tag{2.8}$$

Note that the solutions given in Eq.(2.7) are independent of α.

2.1.3 Solution for the vector case

With the previous solution for the scalar case, the solution to the vector equation is straightforward. Consider the equation

$$x = GQ(Fx) \tag{2.9}$$

with $x \in R^n$, $G \neq 0$ a n-by-1 vector, and $F \neq 0$ a 1-by-n vector. Note that $Q(Fx)$ is a scalar. Write out the components:

$$x_j = g_j Q(Fx)$$

Suppose, without loss of generality, that $g_1 \neq 0$. If this were not true, any nonzero g_j can be used to construct the solution. All the components of the solution can be written in terms of the first one:

$$x_j = \frac{g_j}{g_1} x_1 \ , \ j = 2, 3, ...n$$

Substituting into the equation for x_1 gives

$$x_1 = g_1 Q(\frac{FG}{g_1} x_1)$$

which is a scalar equation of the form studied above. The condition for existence of nonzero solutions is $0.5 \leq FG < 1.5$. The solutions are given by $x_{1_i} = i g_1 q$, for $i = 0, \pm 1, \pm 2....$ The remaining components are obtained as

$$x_{j_i} = \frac{g_j}{g_1} x_{1_i}$$

The range of index i is given by

$$\begin{aligned} |i| &\leq \frac{1}{2|FG - 1|}, \text{ when } FG < 1 \\ |i| &< \frac{1}{2|FG - 1|}, \text{ when } FG > 1 \end{aligned} \tag{2.10}$$

Example 2.2. Consider the same A, B, and F values as in Example 2.1, along with $q = 1$. $A - BF$ is Hurwitz, however $FA^{-1}B = 0.75$, which violates the conditions for a single equilibrium point. Using the developed solution we have that $FG = FA^{-1}B < 1$, therefore $|i| \leq 2$. This means that there are 4 equilibrium points other than the origin. The points are: $(\mp 1.5q, \pm 1.25q)$, $(\mp 3q, \pm 2.5q)$.

2.1.4 Equilibrium analysis: Discrete time

Consider the following LTI discrete time system under quantized state feedback:

$$\begin{cases} x(k+1) = Ax(k) + Bu(k) \\ u(k) = -Q(Fx(k)) \end{cases} \tag{2.11}$$

Existence of a unique solution sequence is guaranteed, since all computations involved in calculating $x(k+1)$ from $x(k)$ are single-valued and well-defined, therefore the iteration can be continued indefinitely starting from any initial condition. Suppose $A - I$ is nonsingular. The equilibrium points are solutions to the vector equation

$$x = (A - I)^{-1} BQ(Fx) \tag{2.12}$$

The analysis used for the continuous time case applies replacing A by $A - I$. Note that the non-singularity requirement on $A - I$ means that A must not have eigenvalues at 1.

2.1.5 Equilibrium analysis of singular systems

Definition 2.1. *A LTI system is singular if its set of equilibrium points is not enumerable.*

Lemma 2.1. *The continuous time system 2.1 is singular if and only if A has at least one zero eigenvalue. Also, the discrete system 2.11 is singular if and only if A has at least one unity eigenvalue.*

Proof. Sufficiency, continuous time: Since the system matrix A is singular, equilibrium equation 2.2 has to be written in the form:

$$Ax = BQ(Fx) \tag{2.13}$$

First, apply elementary row operations to matrix A, performing the same changes in the corresponding rows of B. The row operations realize the process know as Gaussian elimination, which puts the matrix in echelon form. The quantity $Q(Fx)$ is scalar, therefore it will remain unchanged by the operations. Since A is singular, the process will reveal a number of zero rows, leaving the equation in the form

$$\begin{bmatrix} A' \\ \hline 0 \end{bmatrix} x = \begin{bmatrix} B_1' \\ \hline B_2' \end{bmatrix} Q(Fx) \tag{2.14}$$

This implies that $Q(Fx) = 0$. Substituting this back into Eq.(2.13) shows that the solutions must satisfy

$$\begin{cases} Ax = 0 \\ -\frac{q}{2} < Fx < \frac{q}{2} \end{cases} \tag{2.15}$$

Therefore, when A is singular, the set of solutions is dense, and lies at the intersection of the null space of A with the set of points for which Fx is rounded to zero. Mathematically, x is a solution if

$$x \in \text{span}(\text{null}(A)) \text{ and } |Fx| < \frac{q}{2}$$

Necessity: Suppose A is does not have a zero eigenvalue. Then the calculation of the number of equilibrium points given in 2.10 applies, and therefore the set of equilibrium points is enumerable, and possibly infinite. (In fact, index i provides a one-to-one mapping between the elements of equilibrium set and a countable set, the integers).

For the discrete time, the eigenvalues of $A - I$ are those of A minus one. Therefore $A - I$ is singular whenever A has an eigenvalue at one. The reasoning for continuous time applies if A is replaced by $A - I$. ■

Example 2.3. Consider a continuous system with singular A

$$\begin{aligned} A &= \begin{bmatrix} 1 & 2 \\ -2 & -4 \end{bmatrix} \\ B &= \begin{bmatrix} 1 \\ 3 \end{bmatrix} \\ F &= [2\ 1] \end{aligned} \tag{2.16}$$

The null space of A is of dimension 1 and is spanned by the vector $[-2\ 1]$. A parameterization of the equilibrium solutions is given by

$$x = \gamma[-2\ 1]^T$$

with $-1/6 < \gamma < 1/6$. Now consider a discrete system with singular $A - I$

$$A = \begin{bmatrix} 0 & -1 \\ \frac{1}{2} & \frac{3}{2} \end{bmatrix}$$
$$B = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{4} \end{bmatrix}$$
$$F = [\frac{1}{2}\ 1] \tag{2.17}$$

The null space of $A - I$ is of dimension 1 and is spanned by the vector $[-1\ 1]$. A parameterization of the equilibrium solutions is given by

$$x = \gamma[-1\ 1]^T$$

with $-1 < \gamma < 1$. Figure 2.2 shows the equilibrium line and some trajectories of the continuous time example. Similarly, Figure 2.3 illustrates the discrete time case. Note that discretization in time results in the state not achieving values in null$(A - I)$ for all initial conditions.In this example, none of the initial conditions simulated results in a trajectory intersecting null$(A - I)$

2.1.6 Conditions for uniqueness of equilibrium

The above findings can be summarized in the following lemmas:

Lemma 2.2. *The origin is the unique equilibrium point of continuous time system 2.1 if and only if A is nonsingular and $FA^{-1}B \geq 1.5$ or $FA^{-1}B < 0.5$.*

Lemma 2.3. *The origin is the unique equilibrium point of discrete time system 2.11 if and only if 1 is not an eigenvalue of A and $F(A - I)^{-1}B \geq 1.5$ or $F(A - I)^{-1}B < 0.5$.*

The necessity of the nonsingularity condition on A or $A - I$ is a restatement of Lemma 2.1.

2.2 Absolute Stability Analysis for Continuous Hurwitz Plants

In this section, it is shown how the theory of Absolute Stability can be directly applied to the problem at hand. The theory assumes that the system of differential equations has a solution. Existence of a solution to Eq. (2.1) is shown in the appendix for the definition of solution introduced by Filippov [13]. However, it is unclear if Filippov's concept of solution is compatible with the results

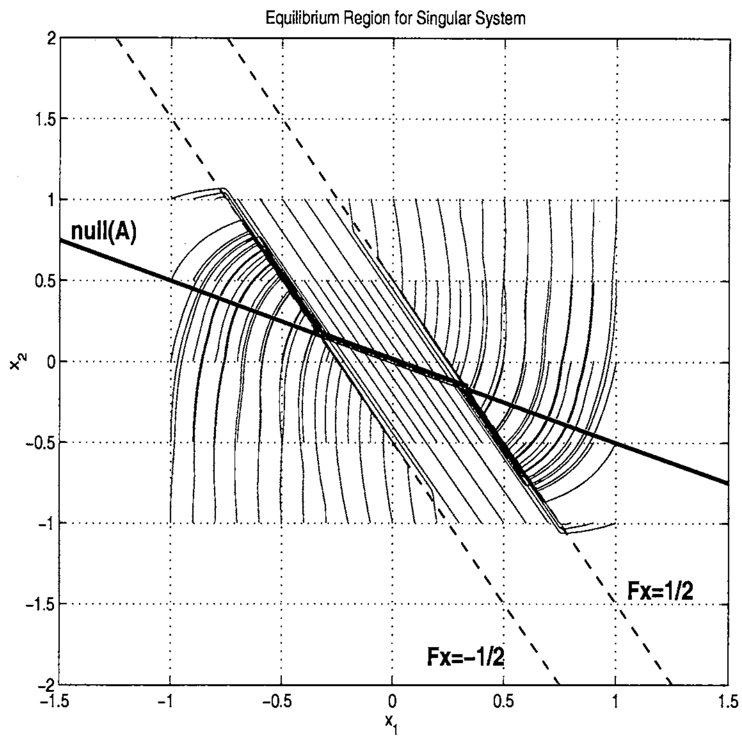

Figure 2.2: Equilibrium Line and Trajectories for Continuous Time Singular System

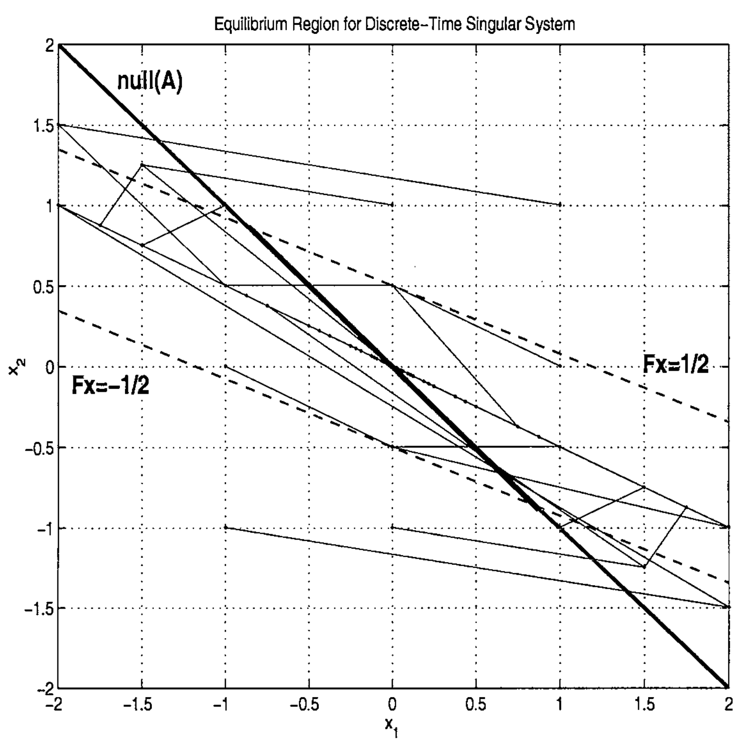

Figure 2.3: Equilibrium Line and Trajectories for Discrete Time Singular System

of Absolute Stability. In this work, it will be assumed that Eq. (2.1) has a solution that satisfies the requirements of the theory. Determining whether Filippov's -or other- solution definitions satisfy the requirements of Absolute Stability is out of the scope of this work. If a feedback gain F is given, the Popov frequency domain criterion can be applied to determine if a Hurwitz plant remains stable under quantized state feedback. Absolute Stability is a convenient method in the case of control quantization, because the nonlinearity is of the sector type. The Popov condition, circle criterion, and other Absolute Stability theories provide sufficient conditions for global asymptotic stability (G.A.S.) based on the sector bounds, regardless of the variation exhibited by the nonlinearity within the sector. Time-varying nonlinearity is allowed.

2.2.1 Continuous time analysis using the Popov condition

The following are formal definitions of sector nonlinearity and absolute stability.

Definition 2.2. *Let a class of functions be defined by*

$$\mathcal{S}(K_1, K_2) = \left\{ \sigma : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R} \;\middle|\; K_1 < \frac{\sigma(t,y)}{y} < K_2 \right\}$$

A function ϕ is said to be of the sector type with sector bounds $K_1 < K_2$ if $\phi \in \mathcal{S}(K_1, K_2)$. If equality is allowed in either side, we use the notations $\phi \in \mathcal{S}[K_1, K_2)$,$\phi \in \mathcal{S}(K_1, K_2]$, and $\phi \in \mathcal{S}[K_1, K_2]$

Definition 2.3. *The system*

$$\begin{cases} \dot{x} = Ax + Bu \\ y = Cx \\ u = -\mathcal{N}(y) \end{cases} \tag{2.18}$$

is absolutely stable in the sector $[K_1, K_2]$ if the origin is G.A.S. for all $\mathcal{N} \in \mathcal{S}[K_1, K_2]$.

The frequency domain condition first given by V.M. Popov applies for Hurwitz plants. For convenience and flow, the Popov theorem is reproduced below ([1])

Theorem 1 (Popov). *System 2.18 with A Hurwitz is absolutely stable in the sector $[0, K_2]$ if there exist a finite real number r such that the following inequality is satisfied for all $w \geq 0$*

$$Re[(1 + rwj)C(jwI - A)^{-1}B] + \frac{1}{K_2} > 0 \text{ for all } w \geq 0$$

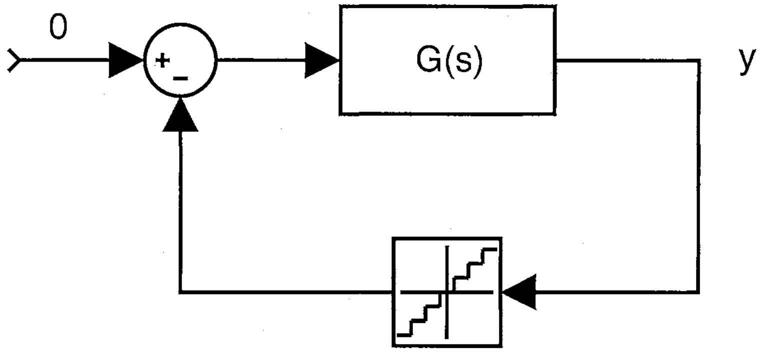

Figure 2.4: Stability of quantized system as a Luré problem

The system under study can be written as

$$\begin{cases} \dot{x} = Ax + Bu \\ y = Fx \\ u = -\mathcal{Q}(y) \end{cases} \tag{2.19}$$

The following transfer function can be defined

$$G(s) = F(sI - A)^{-1}B$$

and the closed loop system can be represented in the form of the Luré problem of absolute stability, as seen in Figure 2.4. The quantization nonlinearity $\mathcal{Q}$ is such that $\mathcal{Q} \in \mathcal{S}[0, 2]$, as can be seen graphically from Figure 1.2. Therefore the Popov theorem can be applied with $K_2 = 2$ and $C = F$.

Example 2.4. Consider the same A and B as in Example 2.1, but take $F = [1,\ 1.5]$. It can be verified that the Popov condition holds for $\alpha = 0$, implying G.A.S. for a quantization as coarse as desired. If we take $F = [1,\ 1.8]$, it can be verified that Popov's condition does not hold for any nonnegative α, however this alone does not imply that the origin is not G.A.S., since Popov's condition is only sufficient. Note that for the first case we have $FA^{-1}B = 0.375$, implying that the origin is the only equilibrium point, a necessary condition for G.A.S. In the second case, however, four nonzero equilibrium points were found earlier, making it impossible for the origin to be G.A.S.

Popov's condition can be used *a posteriori*, in order to check the suitability of a given F. It can be verified (at least by numerical counterexample) that uniqueness of equilibrium does not imply Popov condition. Therefore, for design purposes, it is desirable to parameterize all gains satisfying Popov criterion. This will be done in the next sections for discrete time systems.

2.3 Absolute Stability Analysis for Discrete Plants

The concept of discrete positive realness (DPR) is used in the derivation of absolute stability conditions. The theory introduced by Hitz and B.D.O. Anderson is presented at the appendix for completeness. To facilitate the discussion, two stability lemmas are stated below. The lemmas have been specialized to the SISO case. Let a LTI SISO discrete time system be described by

$$\begin{cases} x(k+1) = Ax(k) + Bu(k) \\ u(k) = -\phi(Cx(k)) \end{cases} \tag{2.20}$$

Let A have eigenvalues inside the open unit circle and $\phi \in \mathcal{S}(0, \overline{K})$. Define the transfer matrix

$$W(z) = \frac{1}{\overline{K}} + C(zI - A)^{-1}B$$

Theorem 2 (Discrete Popov Criterion). *If $W(z)$ is discrete positive real, then the system 2.20 is globally stable about the origin. $W(z)$ is DPR if*

$$\mathrm{Re}\left\{W(e^{jw}\right\} \geq -\frac{1}{\overline{K}} \quad \forall w \in \mathbb{R}$$

Theorem 3 (Anderson). *If there exist a real symmetric positive definite matrix P and real matrices L and U, such that*

$$\begin{aligned} A^T P A - P &= -LL^T \\ B^T P A &= C - U^T L^T \\ U^T U &= \frac{2}{\overline{K}} - B^T P B \end{aligned} \tag{2.21}$$

then the system 2.20 is globally stable about the origin.

Note that the dimensions of P are n-by-n, and L and U are of dimensions n-by-r and r-by-1 respectively, with r being any positive integer.

2.3.1 Frequency domain condition for stability of QI systems

Inclusion of interval ends

Theorem 2 yields asymptotic stability when $\phi \in (0, \overline{K})$. In the case of quantization, the nonlinearity belongs to a closed sector, namely $\mathcal{Q} \in [0, 2]$. Inclusion of the upper bound is handled by choosing $\overline{K} > 2$ to test or design a gain F. If the system is open-loop stable, zero can be included. This requires a detailed proof, which is postponed until Section 3.8, where it is carried out for the more general case of measurement quantization. The following Lemma gives a tool to analyze and design feedback gains:

Lemma 2.4. *Let $G(z) = F(zI-A)^{-1}B$ be the transfer function of the discrete system with quantized feedback, Eq.(2.11). The system is globally asymptotically stable about the origin if the graph of $G(e^{jw})$ lies to the right of a vertical line at $z = -1/2$ in the complex plane.*

2.3.2 Matrix Condition for Stability - Parameterization of stabilizing gains

In the case of quantized state feedback, the matrix equations of Eq.(B.4) must be used with $C = F$ and $\overline{K} > 2$. A gain F may not be found for some choices of L in Eq.(B.4). The limitation stems from the third equation, which can be rewritten as

$$||U||_2^2 = \frac{2}{\overline{K}} - B^T PB$$

For F to exist for a given L it is necessary and sufficient that $\frac{2}{\overline{K}} \geq B^T PB$. A scaling procedure can be used to characterize the set of gains in terms of three parameters: L, a vector V, and a scalar β. Multiply the first equation in Eq.(B.4) by a real positive number γ. Then, if P is a solution of the first equation using L, γP will be a solution using $\sqrt{\gamma}L$. Now, noting that $B^T PB > 0$, find the allowable range of γ from the last equation:

$$\gamma \leq \frac{2}{\overline{K}B^T PB}$$

Moreover, choose γ as follows:

$$\gamma = \beta\frac{2}{\overline{K}B^T PB}$$

where $\beta \in]0, 1]$. The procedure for computing a gain F is

1. Choose an n-by-r matrix $L \neq 0$, with any r, and solve for P from the discrete Lyapunov equation

2. Choose $\beta \in]0, 1]$ and compute

$$\gamma = \beta\frac{2}{\overline{K}B^T PB}$$

3. Choose an arbitrary r-by-1 vector U such that

$$||U||_2^2 = \frac{2}{\overline{K}} - \frac{2\beta\gamma}{\overline{K}}$$

4. Compute F from the middle equation, using $\sqrt{\gamma}L$, γP and V.

Varying β in $]0,\ 1]$ and V in the sphere of radius

$$\rho = \sqrt{\frac{2}{\overline{K}} - \frac{2\beta\gamma}{\overline{K}}}$$

should sweep all solutions for a given L. Therefore L and β, and V parameterize the space of solutions for F.

Example 2.5. Consider a discrete time system with

$$A = \begin{bmatrix} -\frac{1}{2} & -\frac{1}{4} \\ 1 & \frac{3}{4} \end{bmatrix}$$
$$B = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

For simplicity, take $r = 1$. Choose a starting $L = [1,\ 2]^T$. The discrete Lyapunov equation is solved using the Matlab command **dlyap(A',L*L')** yielding

$$P = \begin{bmatrix} 3.6247 & 4.2765 \\ 4.2765 & 5.9951 \end{bmatrix}$$

This gives $2/(\overline{K}B^TPB) = 0.224$. This means $\gamma \in]0,\ 0.224[$. Since $r = 1$, V is a scalar and is directly computed from the third of Eq.(B.4). Finally, a value of F is computed for each γ using V, $\sqrt{\gamma}L$ and γP from the middle equation. These calculations have been programmed in Matlab and the result is shown in Figure 2.5. Note that as $\beta \to 1$, the stability limit is approached

Parameterization is exhaustive for $r \geq n$

In the previous example, $n = 2$ and $r = 1$. One question that arises is whether the parameterization is onto the set of solutions regardless of r. The answer to this question is negative. This can be understood with the aid of the well-known inequality

$$\text{rank}(AB) \leq \min\{\text{rank}(A), \text{rank}(\text{B})\}$$

applied to this problem. To that effect, note that the range of values achieved by P increases with $\text{rank}(LL^T)$. Since $\text{rank}(L) \leq \min(n, r)$, the above inequality gives

$$\text{rank}(LL^T) \leq \min\{n, r\}$$

It follows that the parameterization is exhaustive for $r \geq n$.

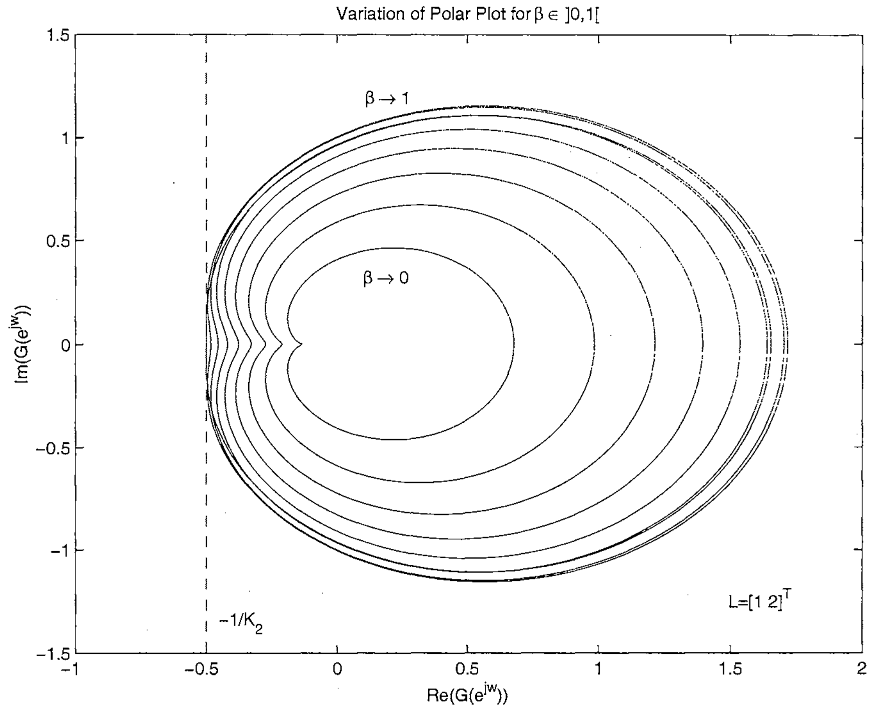

Figure 2.5: Variation of polar plot by parameter changes

2.4 Unstable Systems: Solution using Chattering Control

When the system under quantized feedback is open loop unstable, stabilization is not achievable [12]. This is readily seen by considering the following region of the state space

$$\Omega = \left\{ x \mid |Fx| < \frac{q}{2} \right\}$$

In this region, $\mathcal{Q}(Fx) = 0$, therefore the system evolves in open loop, with dynamics $\dot{x} = Ax$. When A contains unstable poles, the trajectories will move away from the origin, eventually reaching the boundary of Ω where attractiveness is recovered. The interplay of attractive and repulsive forces leads to a limit cycle. A simulation of the unstable system

$$A = \begin{bmatrix} 1 & 2 \\ -3 & 2 \end{bmatrix}$$

$$B = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

$$F = \begin{bmatrix} 2 & 2 \end{bmatrix}$$

shows the presence of a limit cycle. Note that $A - BF$ has stable eigenvalues at -1 and -2, and that $FG = -0.5$, indicating that the origin is the only equilibrium point. This example shows that

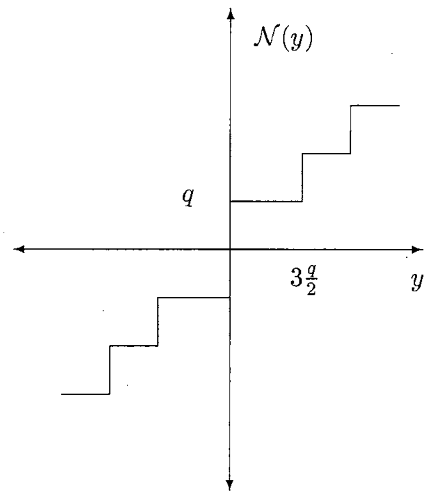

Figure 2.6: Modified nonlinear operator

uniqueness of equilibrium point does not imply G.A.S. when A is not Hurwitz.

2.4.1 Stabilization using local switching

One way to circumvent the problem of zero control in Ω is to modify the computed control law in the following way:

$$u_c = \begin{cases} Fx & \text{when } |Fx| \geq \frac{q}{2} \\ \text{sgn}(Fx)q & \text{when } |Fx| < \frac{q}{2} \end{cases} \tag{2.22}$$

Note that the effective control signal is

$$u = \mathcal{Q}(u_c) = -\mathcal{N}(y)$$

where $y = Fx$ and the new nonlinear operator $\mathcal{N}(y)$ is shown in Figure 2.6.

2.4.2 Absolute Stability Analysis

Consider that is in general not A not Hurwitz, and write the system using the new nonlinear operator as follows

$$\begin{cases} \dot{x} = Ax + Bu \\ y = Fx \\ u = -\mathcal{N}(y) \end{cases} \tag{2.23}$$

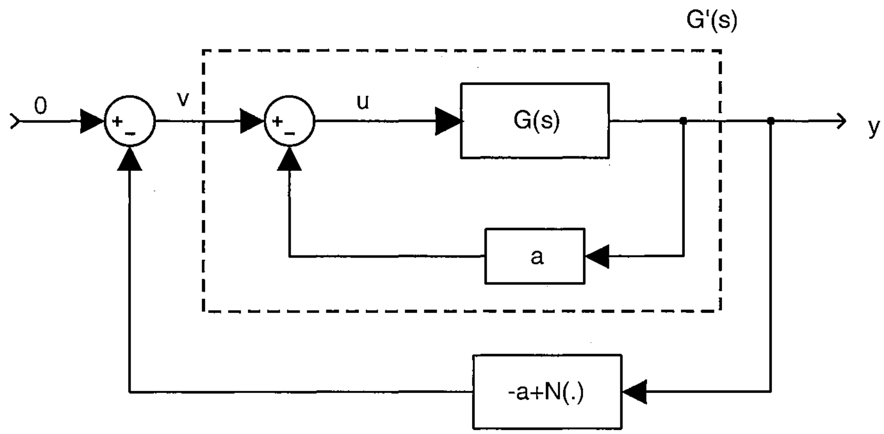

Figure 2.7: Pole shifting for stability analysis of quantized system as a Luré problem

It is assumed that a solution to Eq. (2.23) exists. The appendix shows that a solution in the sense of Filippov exists; however, as stated earlier, it will be assume that such solution is compatible with the standard results of Absolute Stability. Note that the nonlinearity $\mathcal{N} \in \mathcal{S}[\frac{2}{3}, \infty)$. Defining the transfer function as before, i.e.,

$$G(s) = F(sI - A)^{-1}B$$

the system can put in the form of the Luré problem provided we apply a fictitious feedback to compensate for A not being Hurwitz, in a procedure known as "pole shifting" [18]. The problem set up is shown in Figure 2.7. In state space form, $G'(s)$ can be represented as

$$\begin{cases} \dot{x} = Ax + Bu \\ u = v - aFx \\ y = Fx \end{cases}$$

or

$$\begin{cases} \dot{x} = (A - aBF)x + Bv \\ y = Fx \end{cases} \tag{2.24}$$

which leads to the representation

$$G'(s) = F(sI - (A - aBF)^{-1})B$$

Pole shifting and sector rotation

In order to keep the system unchanged after the application of pole-shifting feedback, the nonlinearity must be modified to become $\mathcal{N}'(y) = \mathcal{N}(y) - ay$, as illustrated in Figure 2.7. The sector bounds of $\mathcal{N}'$ are obtained below. $\mathcal{N}$ satisfies

$$\frac{2}{3} < \frac{\mathcal{N}(y)}{y} < \infty$$

subtracting a both sides:

$$\frac{2}{3} - a < \frac{\mathcal{N}(y) - ay}{y} < \infty$$

Therefore $\mathcal{N}' \in \mathcal{S}[\frac{2}{3} - a,\ \infty)$. The theory is directly applied when $K_1 = \frac{2}{3} - a \geq 0$ and a is such that $G'(s)$ is Hurwitz. Assuming such an a exists, it will suffice to obtain a gain F such that $G'(s)$ is absolutely stable in $[\frac{2}{3} - a,\ \infty]$. In order to simplify the problem, at the expense of restricting the allowable gains F to a set smaller than the set of all solutions, $G'(s)$ may be required to be absolutely stable in $[0,\ \infty]$. Graphically, the Nyquist diagram of $G'(s)$ must belong to the closed right half of the complex plane, or the phase shift must be within ± 90 degrees at all frecuencies. This requirement is in fact that $G'(s)$ be *strictly positive real*, or SPR for some $a \leq \frac{2}{3}$. For a definition of positive realness, refer to the appendix. The well-known lemma discovered by Kalman and Yakubovich [32] can be employed. This result provides a relationship between the output and input matrices of a stable system such that the transfer function is SPR.

Lemma 2.5 (Kalman-Yakubovich). *Let the LTI system*

$$\begin{cases} \dot{x} = Ax + Bu \\ y = Cx \end{cases} \tag{2.25}$$

be Hurwitz and controllable. Then, the transfer function

$$H(s) = C(sI - A)^{-1}B$$

is strictly positive real if and only if there exist symmetric positive-definite matrices P and Q such that

$$\begin{aligned} A^TP + PA &= -Q \\ C &= B^TP \end{aligned}$$

The above result is applied to the state space representation of $G'(s)$ given in Eq.(2.24). The Lyapunov equation becomes

$$(A^T - aF^TB^T)P + P(A - aBF) = -Q$$

Choosing $F = B^T P$ makes $G'(s)$ SPR. Substituting into the Lyapunov equation, we obtain the following algebraic Riccati equation in P:

$$A^T P + PA - P(2aBB^T)P + Q = 0 \tag{2.26}$$

This equation is readily solvable using Matlab. Note that a solution P automatically stabilizes $A - aBF = A - aBB^T P$ because we are actually solving a Lyapunov equation, so the Hurwitz condition is satisfied for $G'(s)$ and it is legal to apply the circle criterion. Note that a negative a may not be used for a solution to the Riccati equation to exist. However, the condition $a \leq 2/3$ is not too restrictive because F may be chosen large enough to stabilize $G'(s)$.

Example 2.6. Consider the unstable and controllable LTI system with

$$\begin{aligned} A &= \begin{bmatrix} 2 & 3 & 4 \\ 1 & 1 & 0 \\ -2 & 3 & 5 \end{bmatrix} \\ B &= \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix} \end{aligned} \tag{2.27}$$

Here the eigenvalues of A are 1 and $3.5 \pm 1.658j$. Choosing $Q = I$ and $0 < a = 0.5 < 2/3$ yields

$$P = \begin{bmatrix} 16.0354 & 21.4283 & 6.4161 \\ & 53.3461 & 12.4718 \\ & & 8.1934 \end{bmatrix}$$

and $F = B^T P = [12.8323 \;\; 24.9437 \;\; 16.3868]$.

The practical limitation of the above method is that, for unstable systems, the control law is of the switching type near the origin, which creates chatter.

2.5 Optimality issues: Failure of nominal LQR

The absolute stability properties of the linear quadratic regulator have been studied by several researchers ([10, 7]). In particular, the discrete linear quadratic regulator tolerates a gain reduction of 50 percent when arbitrary weights Q and R are used. This rules out the use of nominal LQR in quantized control, due to the fact that the quantization nonlinearity belongs to the sector $[0, 2]$.

Example 2.7. Consider again the system in Example 2.5. First take the following weights:

$$Q = \begin{bmatrix} 9.1 & 0.8 \\ 0.8 & 0.1 \end{bmatrix} > 0$$
$$R = 1$$

The resulting gain is $F = [-0.271914 \quad -0.111429]$. For this gain, $F(A-I)^1 B = 0.60438$, so there are several equilibrium points other than the origin. This gain violates the Popov condition. Now take

$$Q = \begin{bmatrix} 9.1 & -0.653939 \\ -0.653939 & 9.1 \end{bmatrix} > 0$$
$$R = 1$$

The resulting gain is $F = [0.412387 \quad 0.318871]$. In this case, $F(A-I)^{-1}B = -1.8758$, so the origin is the only equilibrium point. However, this gain violates the $-1/\overline{K}$ boundary. Simulation shows that a limit cycle arises.

2.5.1 Open problem: The true optimum gain

Results that restrict matrices Q and R for the LQR design to be absolutely stable in a given sector are available [22] . This can be directly applied to quantization, taking $\overline{K} = 2$. However, especially for coarse quantization, the true optimum constant feedback gain may differ from the one obtained with nominal LQR. The optimal control sequence, not necessarily constant gain feedback, can be obtained numerically using mathematical programming [29]. The problem of analytically finding the best static gain is still an open challenge.

Chapter 3

Quantized Input and State Measurements (QIQM)

3.1 Problem Statement

In Chapter 2 the situation where only the controller output is quantized is analyzed. The discrete time case has practical relevance if it is assumed that the analog-to-digital converter employed to introduce signals to the control computer has a quantization step size that is small relative to the ones used for measurement. Finite quantization at both plant input and output is a situation often encountered in practice. In this chapter, attention is focused in an important class of feedback controllers, i.e., linear full state feedback. For reasons of practical relevance, the analysis is limited to sampled linear plants which can be effectively treated as discrete-time systems, as customary. It is assumed that each state is measured through a sampling device followed by a quantizer, keeping with the general assumptions stated in Chapter 1, and then multiplied by the feedback gain F at a much higher numerical precision than the quantization step size of the input and output data converters. The calculated control $-F\mathcal{Q}(x)$ is quantized at the output of the control computer and applied to the plant's input hold device. Thus, the control input to the plant is, effectively

$$u(k) = -\mathcal{Q}(F\mathcal{Q}(x)) \tag{3.1}$$

This process is illustrated in Figure 3.1. Problems such as multiplicity of equilibria and instability -in the form of limit cycles or otherwise- appearing with a single quantizer are much aggravated with measurement quantization. Simulations indicate that some of the non-steady responses are also non-periodic, suggesting chaotic behavior [12]. Existence and uniqueness of a solution sequence

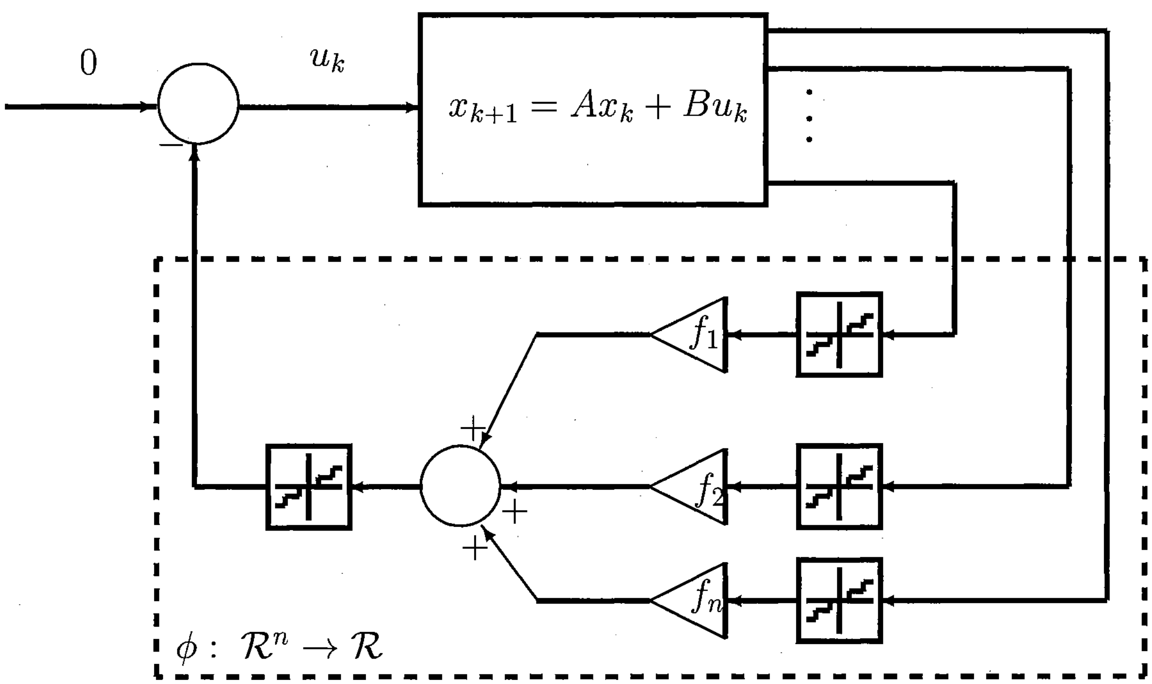

Figure 3.1: Problem set-up for QIQM systems

for each initial condition is guaranteed. This can be readily recognized by noting that the control signal (3.1) is a single-valued function, which together with the system matrices A and B uniquely determine the next value of the state. The goals of this chapter are to obtain a method for verifying if the origin is the only equilibrium point, and to derive a simple criterion that can be used to analyze the stability of a discrete LTI system with quantized measurements and quantized state feedback control. Unlike the QI case, where the sector condition sufficed, the analysis of the QIQM system requires an in-depth examination of the quantization nonlinearities.

3.2 Equilibrium Analysis of QIQM Systems

The complexity introduced by the nested quantization operators prevents one from obtaining a closed-form solution or criterion for single equilibrium, as done in Section 2.1.3. In this section, a graphical construction is described that can be used to predict the number of equilibrium points when a feedback gain F is known in advance. If $(A - I)$ is nonsingular, the equilibrium equation has the form

$$x = (A - I)^{-1} B \mathcal{Q}(F \mathcal{Q}(x)) \tag{3.2}$$

It is possible to obtain a sufficient criterion for the absence of nontrivial solutions for a scalar version of Eq. (3.2) [1], but the results do not carry over to the general vector case. To derive a graphical solution method, call $G = (A - I)^{-1} B$ and write out the components of a solution vector x as

$$x_j = g_j \mathcal{Q} \left[\sum_{k=1}^{n} f_k \mathcal{Q}(x_k) \right]$$

Noting that the outer quantization is scalar-valued, all components of the solution can be written in terms of the first one, assuming, without loss of generality, that the first component of G is nonzero. If this were not true, any nonzero component of G may be used.

$$x_k = \frac{g_k}{g_1} x_1 \ , \ k = 2, 3, ...n$$

Substituting into the equation for x_1 gives

$$x_1 = g_1 \mathcal{Q} \left[\sum_{k=1}^{n} f_k \mathcal{Q}(\frac{g_k}{g_1} x_1) \right] \tag{3.3}$$

Graphically, the solutions for the first component are found by intersecting the irregular staircase-shaped function with the straight line passing through the origin with slope $1/g_1$. Such method is readily applied to any set of matrices A, B, F and quantization step size q. The Matlab program **qiqm-equil** listed at the appendix performs the necessary computations and displays both graphs.

Example 3.1. Consider the following randomly-generated system matrices

$$\begin{aligned} A &= \begin{bmatrix} 0.6822 & 0.1509 & 0.8600 \\ 0.3028 & 0.6979 & 0.8537 \\ 0.5417 & 0.3784 & 0.5936 \end{bmatrix} \\ B &= \begin{bmatrix} 0.8462 \\ 0.5252 \\ 0.2026 \end{bmatrix} \\ F &= \begin{bmatrix} 0.6721 & 0.8381 & 0.0196 \end{bmatrix} \end{aligned}$$

Assume $q = 0.8$. The first element of $G = (A - I)^{-1} B$ is used in constructing the solution. Figure 3.2 shows the graphical construction used to find the first component of the solutions, according to

[1] See Section 4.3.1

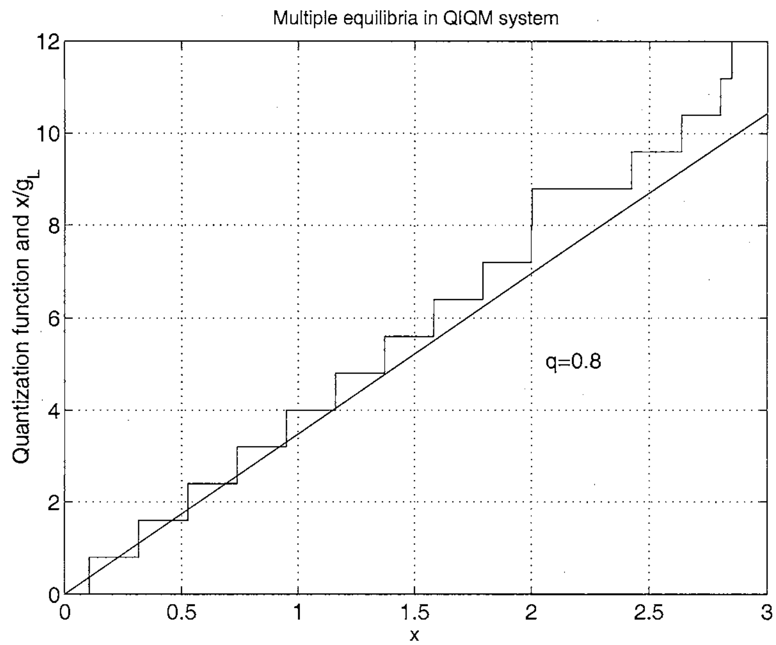

Figure 3.2: Multiple equilibria in QIQM system

Eq. (3.3). The complete solutions can be found using the values obtained graphically. Five non-trivial solutions are found, namely,

$$x_1 = \begin{bmatrix} 0.2299 & 0.8719 & 0.7192 \end{bmatrix}$$
$$x_2 = \begin{bmatrix} 0.4598 & 1.7439 & 1.4384 \end{bmatrix}$$
$$x_3 = \begin{bmatrix} 0.6896 & 2.6154 & 2.1574 \end{bmatrix}$$
$$x_4 = \begin{bmatrix} 0.9195 & 3.4874 & 2.8766 \end{bmatrix}$$
$$x_5 = \begin{bmatrix} 1.1494 & 4.3593 & 3.5958 \end{bmatrix}$$

Now consider another set of randomly-generated matrices:

$$A = \begin{bmatrix} 0.9501 & 0.4860 & 0.4565 \\ 0.2311 & 0.8913 & 0.0185 \\ 0.6068 & 0.7621 & 0.8214 \end{bmatrix}$$
$$B = \begin{bmatrix} 0.4447 \\ 0.6154 \\ 0.7919 \end{bmatrix}$$
$$F = \begin{bmatrix} 0.9218 & 0.7382 & 0.1763 \end{bmatrix}$$

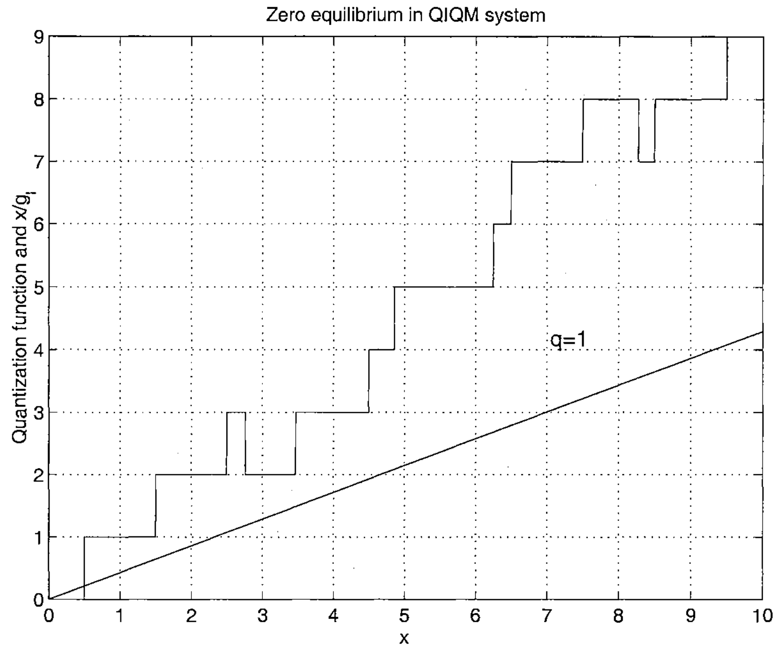

Figure 3.3: Zero equilibrium in QIQM system

Consider $q = 1$. The first element of $G = (A - I)^{-1}B$ is used in constructing the solution. Figure 3.3 shows that there is no non-trivial solution for the first component of the state vector, therefore the origin is the only equilibrium state.

3.3 Stability Analysis- Construction of Equivalent System

In the case where only a control quantizer is present, the quantizer represents a scalar sector-bounded nonlinearity which takes the linear system's defined output $y = Fx$ as its input. In the present QIQM case, such a system output cannot be defined directly, due to the fact that the overall nonlinear operation takes the *state* as its input and outputs a scalar. This is denoted in Figure 3.1 as a mapping ϕ from the state space $\mathbb{R}^n$ to the real line. The established results of Absolute Stability do not apply in this case. Since it is desired to exploit the results of this theory to the solution of the problem, the original system must be put in a form that contains a scalar sector nonlinearity, and an appropriate output must be defined on the linear part of the system. A perfect equivalence, however, is not possible to obtain if the nonlinearity is a function only of the system's defined output. A special theorem will be required to allow for a nonlinearity which multiplies the defined output by a bounded function of the state, henceforth called "multiplier".

3.3.1 Multiplier definition

The state equations of the original QIQM system, denoted Σ_0 are

$$\Sigma_0 : \begin{cases} x(k+1) = Ax(k) + Bu(k) \\ u(k) = -\mathcal{Q}(F\mathcal{Q}(x(k))) \end{cases} \tag{3.4}$$

Define the linear system $L_1 = (A, B, F, d)$ by Eq. 3.5

$$L_1 : \begin{cases} x(k+1) = Ax(k) + Bu'(k) \\ y(k) = Fx(k) + du'(k) \end{cases} \tag{3.5}$$

for some real number d. It is desired that L_1 have the same control input as Σ_0, and, at the same time, $u'(k)$ be defined in terms of the output of L_1. The following lemma provides a way to achieve this.

Lemma 3.1. *There exists a mapping $\alpha : \mathbb{R}^n \rightarrow \mathbb{R}$, called* **multiplier**, *such that*

$$u'(k) = -\mathcal{Q}[\alpha(x(k))y(k)] = u(k) \ \ \forall k$$

that is,

$$\mathcal{Q}[\alpha(x(k))(Fx(k) - d\mathcal{Q}(F\mathcal{Q}(x(k)))] = \mathcal{Q}[F\mathcal{Q}(x(k))]$$

The proof of this lemma is done by specifying a particular function $\alpha(x(k))$, and is presented in Section 3.7. Denote by Σ_1 the closed-loop system that results from applying the feedback $u = -\mathcal{Q}(\alpha(x(k))y(k))$ to system L_1. It will become clear that there exists a unique solution to system Σ_1 for any initial condition when the explicit formula for the multiplier is presented. The significance of Lemma 3.1 is that the input $u(k)$ in system Σ_1 can be seen as the quantization of the perturbed scalar output $y(k)$, where the perturbation factor is the scalar $\alpha(x(k))$. If α were constant, the results of Absolute Stability could be applied directly to the linear portion L_1, taking into account the sector to which the composition of α and the quantization nonlinearity would belong. Note that the Lyapunov stability of system Σ_1 implies that of Σ_0, since the state vectors are the same for both systems at all times. It is shown in Theorem 4 that even when $\alpha(x(k))$ is not constant, but bounded and nonnegative, the stability of Σ_1 -and therefore of Σ_0- can still be derived from a sector condition on the linear portion L_1. Figure 3.4 illustrates the structure of Σ_1, and Figure 3.5 shows the system cast in a form similar to the Luré problem.

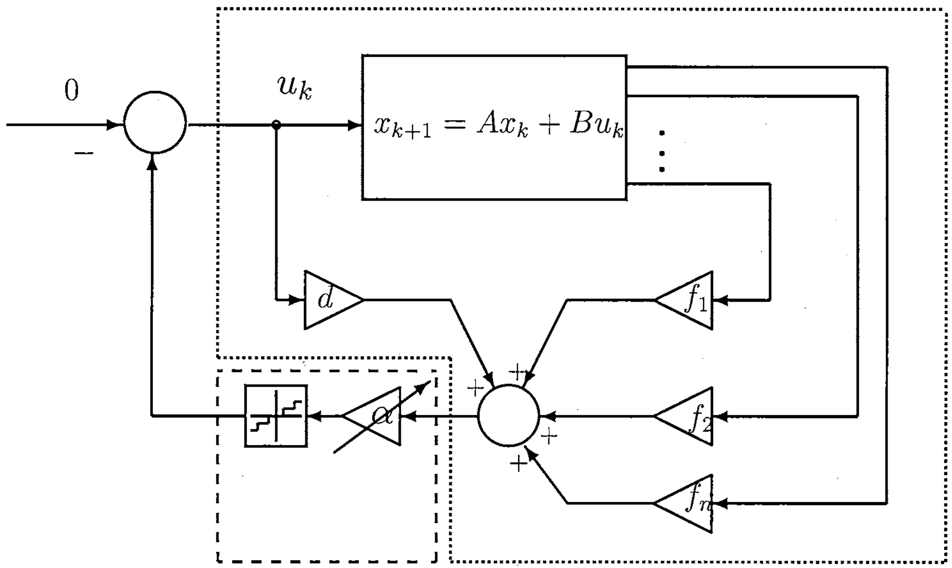

Figure 3.4: The equivalent system Σ_1

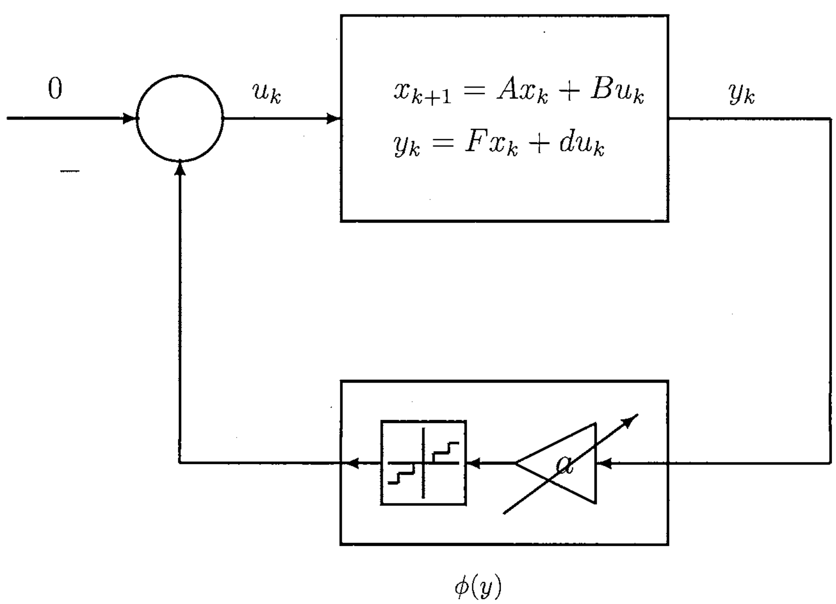

Figure 3.5: Almost a Luré problem

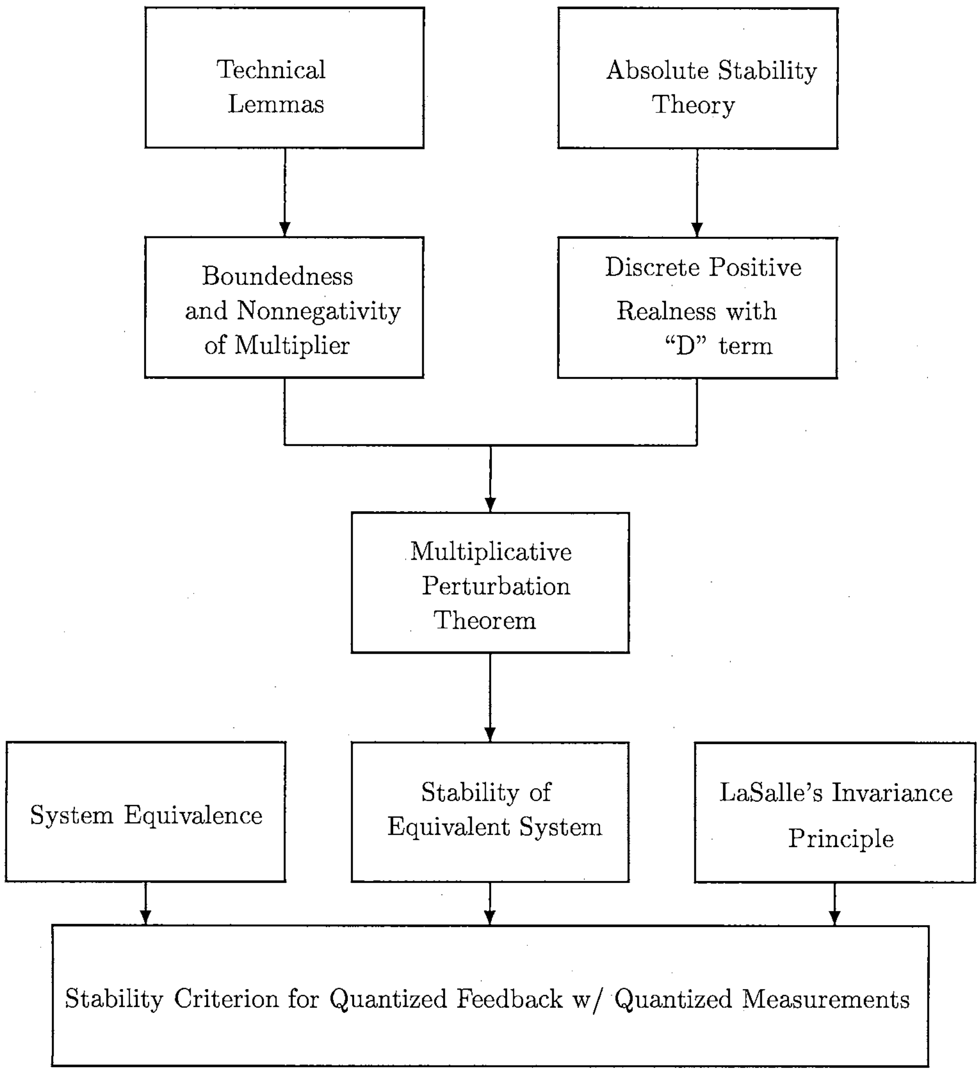

Figure 3.6: Logical dependencies for the derivation of the stability criterion

3.4 Overview of the Method

The derivation of the sufficient condition for stability of a QIQM system requires a number of results, some of which have been obtained previously. Several of the required results are derived here for the first time, to the author's knowledge. The key result required to derive the final stability condition is Theorem 4, which in turn requires an extension of Hitz and Anderson's [17] theorem on stability of DPR systems. This extension constitutes Lemma 3.2. Also, the application of Theorem 4 to the problem of QIQM requires existence, boundedness and non-negativity of the multiplier $\alpha(x(k))$. For the sake of clarity, a diagram with the logical dependencies involved is shown in Figure 3.6.

3.5 Stability of DPR Systems with Direct Transmission Term

This section is based on the original work by Hitz and B.D.O. Anderson [17]. In the cited paper, the main stability theorem is shown for transfer matrices without direct transmission term, i.e., without "D" term. An extension of the result for such cases is introduced in this work in the form of Lemma 3.2.

Lemma 3.2. *Let a discrete time system be represented by Eq.(3.5) Let $\mathcal{N}(y)$ be a scalar nonlinear function such that $\mathcal{N} \in \mathcal{S}(0,k)$. If the transfer function*

$$H(z) = C(zI - A)^{-1}B + D + \frac{1}{\overline{K}}$$

is DPR, then the closed-loop system obtained by applying the feedback

$$u(k) = -\mathcal{N}(y)$$

is globally stable about the origin.

Proof. By hypothesis, there exist matrices P,L and W satisfying the conditions in Eq.(B.2). Consider the Lyapunov function $V(x(k)) = x^T(k)Px(k)$. The change of this function along the system equations is

$$\Delta V(x(k)) = V(x(k+1)) - V(x(k)) = [x^TA^T - B^T\mathcal{N}(y)]P[Ax - B\mathcal{N}(y)] - x^TPx$$

Performing operations and incorporating the matrix equations, D cancels out and the following expression is obtained

$$\Delta V(x(k)) = -[L^Tx - W\mathcal{N}(y)]^T[L^Tx - W\mathcal{N}(y)] - 2\mathcal{N}(y)[y - \frac{\mathcal{N}(y)}{\overline{K}}]$$

The first term is negative semidefinite. The second term can be examined as follows. If $\mathcal{N}(y) > 0$ then $y > 0$. Since $\mathcal{N} \in \mathcal{S}(0, \overline{K})$,

$$y - \frac{\mathcal{N}(y)}{\overline{K}} > 0$$

and the second term is negative. If $\mathcal{N}(y) < 0$, it can be seen that the second term is also negative. If $\mathcal{N}(y) = 0$, the second term is zero. Therefore the change in $\Delta V(x(k))$ is negative semidefinite and the closed loop system is globally stable about the origin. If $\mathcal{N}(y) = 0$ only when $y = 0$ and the linear part is zero-state observable, asymptotic stability is obtained. Otherwise, more information about the nonlinearity is required in order to establish asymptotic stability. ∎

3.6 Absolute Stability with Multiplicative Perturbation of the Sector

In this section it will be shown that if the linear part of the closed loop system is absolutely stable in a sector that is large enough, then it will remain stable when the linear output y is multiplied by a bounded and nonnegative function of the state. This relates to the equivalent system presented in Section 3.3.

Theorem 4. *Let L_1 be a discrete time system represented by the equations*

$$\begin{cases} x(k+1) = Ax(k) + Bu(k) \\ y(k) = Cx(k) + Du(k) \end{cases} \tag{3.6}$$

Let $\alpha : \mathbb{R}^n \to \mathbb{R}$ be a mapping such that $\exists\ \overline{\alpha}$ finite satisfying $0 \leq \alpha(x) < \overline{\alpha}$ for all $x \in \mathbb{R}^n$. Let $\mathcal{N} : \mathbb{R} \to \mathbb{R}$ be a sector nonlinearity $\mathcal{N} \in \mathcal{S}[0, \bar{n}]$. Then if the transfer matrix

$$H(z) = F(zI - A)^{-1}B + D + \frac{1}{\overline{K}}$$

is DPR, and $\overline{\alpha n} < \overline{K}$, the closed-loop system formed by applying the feedback

$$u(k) = -\mathcal{N}[\alpha(x(k))y(k)]$$

is stable in the large.

Proof. By hypothesis, $H(z)$ is DPR.Then by Lemma B.2, there exists a real symmetric positive definite matrix P and real matrices L and W such that

$$\begin{aligned} A^T PA - P &= -LL^T \\ B^T PA &= F - W^T L^T \\ W^T W &= 2D + \frac{2}{\overline{K}} - B^T PB \end{aligned}$$

Consider the quadratic Lyapunov function $V(x(k)) = x^T(k)Px(k)$. The change of the function along the the equations of the closed-loop system is, dropping index k from the notation:

$$\Delta V(x(k)) = V(x(k+1)) - V(x(k)) = [x^T A^T - B^T \mathcal{N}(\alpha(x)y))]P[Ax - B\mathcal{N}(\alpha(x)y)] - x^T Px$$

Performing operations and incorporating the matrix equations, D cancels out and the following expression is obtained

$$\Delta V(x(k)) = -[L^T x - W\mathcal{N}(\alpha(x)y)]^T[L^T x - W\mathcal{N}(\alpha(x)y)] - 2\mathcal{N}(\alpha(x)y)[y - \frac{\mathcal{N}(\alpha(x)y)}{\overline{K}}] \tag{3.7}$$

The first term is clearly negative semidefinite. The second term can be examined as follows. If $\alpha(x)y > 0$ then by the sector condition on $\mathcal{N}$ and the sector inclusion inequality of the hypothesis it follows that

$$0 \leq \mathcal{N}(\alpha(x)y) \leq \bar{n}\alpha(x)y \leq \bar{n}\bar{\alpha}y < \overline{K}y$$

This implies

$$y - \frac{\mathcal{N}(\alpha(x)y)}{\overline{K}} > 0$$

so the term is negative or zero. If $\alpha(x)y < 0$ the above chain of inequalities is reversed, yielding

$$y - \frac{\mathcal{N}(\alpha(x)y)}{\overline{K}} < 0$$

Thus, the second term is negative semidefinite, being zero if $\mathcal{N} = 0$ or $y = 0$, thus Lyapunov stability follows. In order to prove asymptotic stability, further assumptions on the local behavior of $\mathcal{N}$ and system observability might be required. ∎

3.7 Multiplier Boundedness and Positivity

System equivalence and Theorem 4 can be used to analyze the original problem if a suitable multiplier $\alpha(x(k))$ can be found. In this section, a formula for such multiplier is provided, along with the derivation of its lowest upper bound, $\overline{\alpha}$. The functional form of the multiplier is simple, though the derivation of the lowest upper bound is involved and requires some results and developements from Number Theory. Part of the material of this section constitutes the proof of Lemma 3.1, postponed until now.

3.7.1 A formula for the multiplier

A few definitions are needed before the formula is introduced.

Definition 3.1. *A quantization node is an n-by-1 vector z such that $z_i = j_i q$, for $i = 1, 2, ..n$ and some integers j_i. When $q = 1$, z is an element of $\mathbb{Z}^n$*

Definition 3.2. *The quantization region around node z is defined as the set*

$$\Omega_z = \left\{x \in \mathbb{R}^n \;\middle|\; \mathcal{Q}(x) = z\right\}$$

Definition 3.3. *Define Ξ as the set*

$$\Xi = \left\{x \in \mathbb{R}^n \;\middle|\; |F\mathcal{Q}(x)| < \frac{q}{2}\right\}$$

Ξ is the set where the computed control is rounded to zero when passed through the D/A converter. Clearly, $\Omega_0 \subset \Xi$. Denote F_0 the set of x which satisfy $Fx = 0$. Define the set Λ as

$$\Lambda = \Omega_0 \bigcap F_0$$

Set Λ is also a subset of Ξ.

Definition 3.4. *Define $\alpha(x(k))$ for all $x(k)$ as*

$$\alpha(x(k)) = \begin{cases} \dfrac{Q(FQ(x(k)))}{Fx(k) - dQ(FQ(x(k)))} & , \textit{if } x \notin \Lambda \\ 0 \ , \textit{if } x \in \Lambda \end{cases} \tag{3.8}$$

It will be shown in the next sections that it is possible to choose d such that the denominator in the formula is zero only when $x \in \Lambda$. Since $\Lambda \subset \Xi$, the numerator will also be zero. Thus, an arbitrary value can be assigned to α and still maintain the identity of Lemma 3.1. In fact, if d is a suitable value and $x \notin \Lambda$, the denominator cancels the input $y(k) = Fx(k) + du(k) = Fx(k) - dQ(FQ(x(k)))$ and substitution of $\alpha(x(k))$ in Lemma 3.1 results in an identity, since $Q(Q(v)) = Q(v)$ for any v. If $x \in \Lambda$ then also $x \in \Xi$, therefore $Fx = 0$, and $-dQ(FQ(x(k))) = 0$, so the numerator, denominator and control $u(x)$ are all zero. Since the denominator is the defined output $y(k) = Fx(k) + du(k)$, the identity of Lemma 3.1 is satisfied regardless of the value chosen for $\alpha(x)$ when $x \in \Lambda$.

3.7.2 Multiplier behavior around node zero

Figure 3.7 shows a two-dimensional[2] depiction of the quantization region around the origin node. Inside Ω_0 the numerator of $\alpha(x(k))$ is zero, and its denominator is nonzero except for points that also satisfy $Fx = 0$. Therefore $\alpha(x(k)) = 0$ in Ω_0. The multiplier is continuous in Ω_0.

3.7.3 Existence of a suitable d

Now consider the quantization region around node $z \neq 0$. In this region, $FQ(x) = Fz$. It is desired to show that it is possible to choose d such that $Fx \neq dQ(Fz)$. Figure 3.8 is of great aid in understanding the problem. The set

$$F_z = \{x \in \mathbb{R}^n \mid Fx = Fz\}$$

is represented by the contour line passing through the node. The contour line corresponding to the quantized value of Fz has not been drawn, but could lie inside or outside the quantization region.

[2]Two-dimensional diagrams are intended to aid the mathematical developments, which are valid in any dimension.

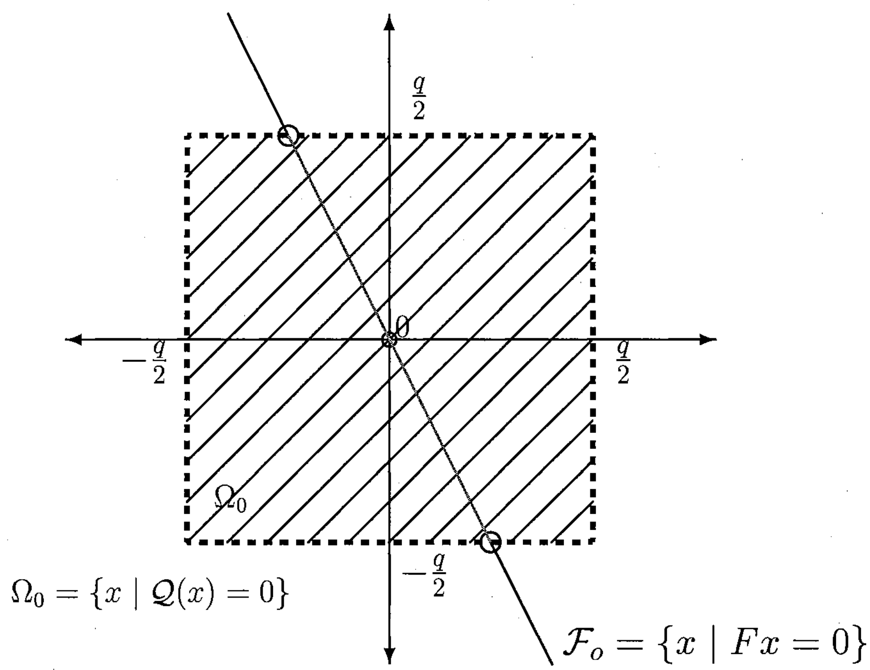

Figure 3.7: Quantization region around node 0

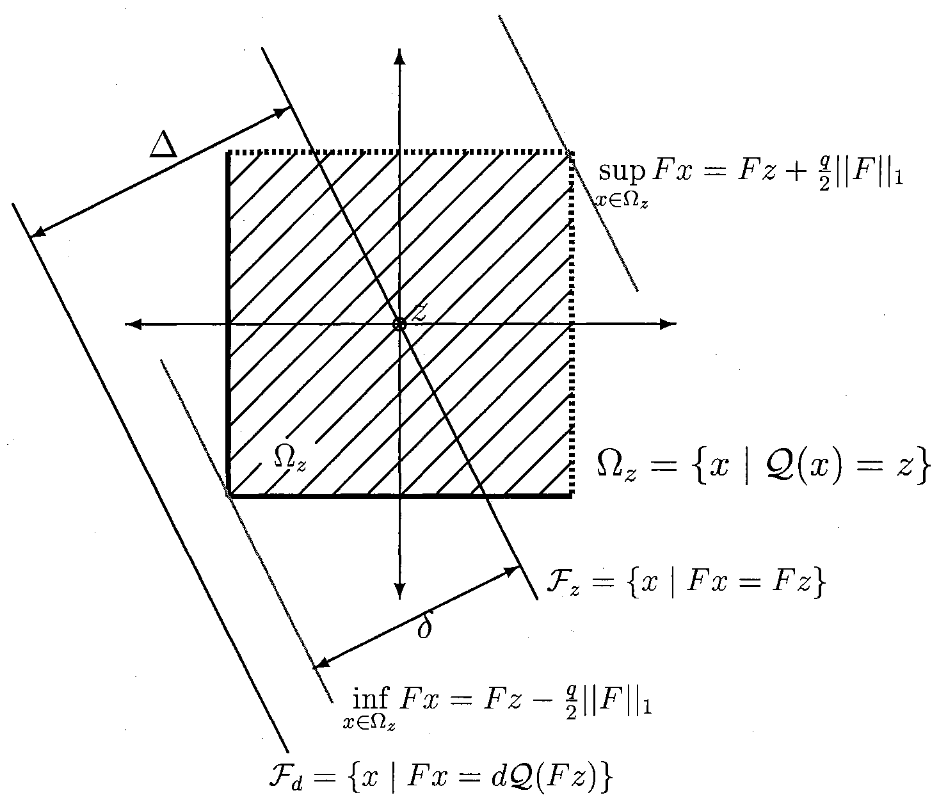

Figure 3.8: Quantization region around node z

The contour line corresponding to the value $dQ(Fz)$ is shown in the figure. To avoid a solution to $Fx = dQ(Fz)$, it is necessary and sufficient that this line not intersect Ω_z. In fact, suppose that there is a non empty intersection. Then, for points on the contour line we would have $Fx = dQ(FQ(x))$. Conversely, suppose the line is outside Ω_z. Then for a point of Ω_z the value of Fx will differ from $dQ(Fz)$ since contour lines of different values do not have points in common for a linear function Fx. Figure 3.8 also shows the contour lines with limiting values inside Ω_z. As shown in the appendix, the limiting values of Fx inside Ω_z are

$$\inf_{x\in\Omega_z} Fx = Fz - \frac{q}{2}||F||_1$$
$$\sup_{x\in\Omega_z} Fx = Fz + \frac{q}{2}||F||_1$$

Also derived in the appendix is a definition of distance between contour lines and the following formula for the distance δ_{12} between contour lines of values c_1 and c_2:

$$\delta_{12} = \frac{|c_1 - c_2|}{||F||_2}$$

Let Δ be the distance between the contour line of value Fz and the line of value $dQ(Fz)$. Let δ be the distance between the line Fz and either extreme line, of values $Fz \pm \frac{q}{2}||F||_1$. Then it is required that $\Delta > \delta$. Using the above formulas, this reduces to

$$|Fz - dQ(Fz)| > \frac{q}{2}||F||_1 \tag{3.9}$$

This condition must be satisfied only for z such that $|Fz| \geq \frac{q}{2}$. In fact, when $|Fz| < \frac{q}{2}$, the numerator of $\alpha(x(k))$ is zero. The denominator becomes just Fx, which is zero only when $x \in \Lambda$. In this case, $\alpha(x(k))$ attains its definition value of zero. The left-hand side of inequality 3.9 is now a function of a scalar variable, namely Fz. Figure 3.9 is a graph of the variation of such function when Fz is varied. Assuming a negative d, the minimum value attained by the function for $|Fz| \geq \frac{q}{2}$ is $|(Fz)_{min}| = -dq + \frac{q}{2}$ and occurs at $|Fz| = \frac{q}{2}$. The minimum must be greater than $\frac{q}{2}||F||_1$. Therefore, the following condition can be stated

Condition 1. *The multiplier $\alpha(x(k))$ in Eq.(3.8) is well-defined if*

$$d < \frac{1}{2}(1 - ||F||_1)$$

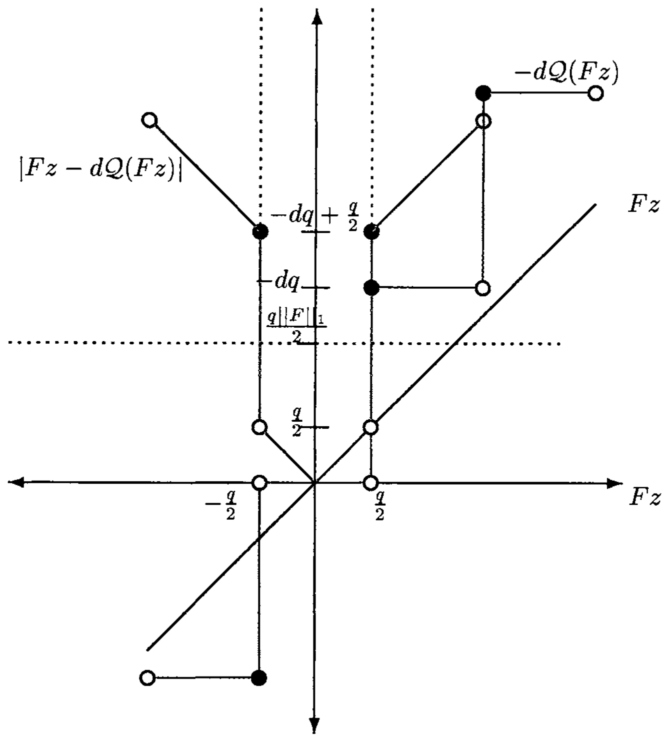

Figure 3.9: The choice of a suitable d

3.7.4 Limiting values of the multiplier inside a quantization region

In order to obtain the absolute least upper bound on $\alpha(x(k))$ over the state space, the supremum of the multiplier within a quantization region Ω_z is first sought. That is, it is desired to calculate

$$\sup_{x \in \Omega_z} \alpha(x)$$

Within a quantization region Ω_z with $z \neq 0$ the formula for $\alpha(x(k))$ becomes, dropping index k from the notation:

$$\alpha(x) = \frac{Q(Fz)}{Fx - dQ(Fz)} \tag{3.10}$$

Since z is constant within Ω_z, α is a function of the scalar quantity Fx. The function has a singularity at $Fx = dQ(Fz)$, but it is clear that such value of Fx does not occur inside Ω_z, if d is chosen following Condition 1. At either side of the singularity, the function $\alpha(x)$ is monotonic. Therefore the extreme values of the multiplier occur at the ends of the interval centered at Fz with radius $\delta = \frac{q}{2}||F||_1$. This is shown in Figure 3.10, where $Fz > 0$ and $d < 0$ have been assumed. The quantity of interest is the supremum of $\alpha(x)$ within the quantization region around node z. Before the value of the supremum is stated, it is convenient to show that $\alpha(x)$ is nonnegative for d which

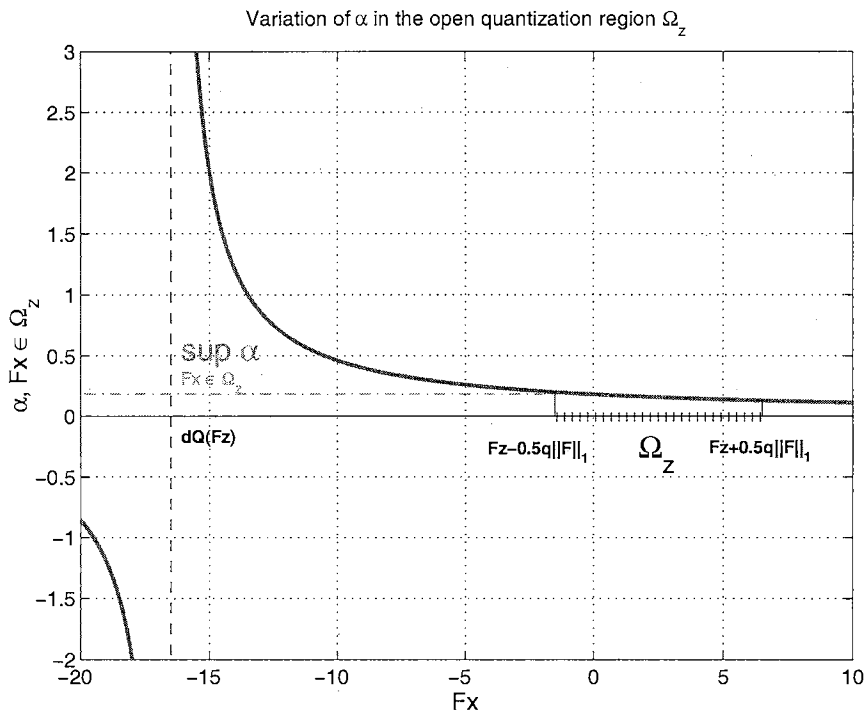

Figure 3.10: Variation of multiplier within a quantization region

also satisfies Condition 1.

Lemma 3.3. *Suppose d is chosen such that $d < \frac{1}{2}(1 - ||F||_1)$. Then $\alpha(x) \geq 0$ for all $x \in \mathbb{R}^n$.*

Proof. It is already known that if $d < \frac{1}{2}(1 - ||F||_1)$ then the variation of $\alpha(x)$ within a quantization region is monotonic. Also, the denominator of $\alpha(x)$ is constant within Ω_z.

Case i): $Fz \geq \frac{q}{2}$

Since $\alpha(x)$ varies monotonically and the denominator is constant within Ω_z, the limiting values of $\alpha(x)$ are obtained by evaluating it at the ends of the interval which define Ω_z. That is, $Fx = Fz \pm \frac{q}{2}||F||_1$ need to be considered. First, take the plus sign. The denominator is

$$\text{den}_{max} = Fz + \frac{q}{2}||F||_1 - dQ(Fz)$$

The following stems from Condition 1:

$$-dQ(Fz) > -\frac{Q(Fz)}{2}(1 - ||F||_1)$$

Combining the two previous inequalities gives

$$\text{den}_{max} > Fz - \frac{Q(Fz)}{2} + (\frac{Q(Fz)}{2} + \frac{q}{2})||F||_1$$

which is nonnegative, since $2Fz \geq Q(Fz) > 0$ for $Fz \geq \frac{q}{2}$. Now take the minus sign. The denominator is

$$\text{den}_{min} = Fz - \frac{q}{2}||F||_1 - dQ(Fz)$$

Following the same steps, it is obtained that

$$\text{den}_{min} > Fz - \frac{Q(Fz)}{2} + (\frac{Q(Fz)}{2} - \frac{q}{2})||F||_1$$

which is nonnegative, since $2Fz \geq Q(Fz) \geq q$ for $Fz \geq \frac{q}{2}$. Therefore, both the numerator is nonnegative and the denominator is positive, resulting in a nonnegative $\alpha(x)$.

Case ii): $Fz \leq -\frac{q}{2}$

Similarly, first take the plus sign. The denominator is

$$\text{den}_{max} = Fz + \frac{q}{2}||F||_1 - dQ(Fz)$$

Again using Condition 1 gives:

$$\text{den}_{max} < Fz - \frac{Q(Fz)}{2} + (\frac{Q(Fz)}{2} + \frac{q}{2})||F||_1$$

which is nonpositive, since $2Fz \leq Q(Fz) \leq -q$ for $Fz \leq -\frac{q}{2}$. Now take the minus sign. The denominator is

$$\text{den}_{min} = Fz - \frac{q}{2}||F||_1 - dQ(Fz)$$

Following the same steps, it is obtained that

$$\text{den}_{min} < Fz - \frac{Q(Fz)}{2} + (\frac{Q(Fz)}{2} - \frac{q}{2})||F||_1$$

which is nonpositive, since $2Fz \leq Q(Fz) \leq -q$ for $Fz \leq -\frac{q}{2}$. Therefore, both the numerator is nonpositive and the denominator is negative, resulting in a nonnegative $\alpha(x)$.

Case iii): Suppose $-\frac{q}{2} < Fz < \frac{2}{2}$

In this case $\alpha(x) = 0$. It is thus proved that Condition 1 not only guarantees boundedness of $\alpha(x)$, but also that it is never negative. The multiplier can now be interpreted as a bounded variation which keeps a sector in the first and third quadrants. ∎

A direct consequence of Lemma 3.3 is that the extreme values of α are the same for all pairs of values Fz_1, Fz_2 outside Ω_0 such that $Fz_1 = -Fz_2$. Then, the upcoming search for the absolute least upper bound is performed only over positive Fz. Now, the supremum of α over a quantization region is summarized in the following Lemma:

Lemma 3.4. *If d is chosen to satisfy $d < \frac{1}{2}(1 - ||F||_1)$, then the supremum of the set of values attained by the multiplier $\alpha(x)$ for $x \in \Omega_z$ is given by*

$$\sup_{x \in \Omega_z} \alpha(x) = \frac{Q(Fz)}{(Fz - \frac{q}{2}||F||_1) - dQ(Fz)} \tag{3.11}$$

3.7.5 The supremum of the multiplier over the state space

In this section the final result required for the application of Theorem 4 is provided. The following quantity is sought:

$$\sup_z \sup_{x \in \Omega_z} \alpha(x)$$

or

$$\bar{\alpha} = \sup_z \frac{Q(Fz)}{(Fz - \frac{q}{2}||F||_1) - dQ(Fz)}$$

The supremum within Ω_z is a function of the scalar variable Fz. Figure 3.11 shows a graph of the numerator $Q(Fz)$, the denominator $Fz - \frac{q}{2}||F||_1 - dQ(Fz)$, and the ratio $\overline{\alpha}$ as a function of Fz. The numerator is zero for $|Fz| < \frac{q}{2}$. Then, the maximum must be sought for $|Fz| \geq \frac{q}{2}$; moreover, only $Fz \geq \frac{q}{2}$ needs to be considered, in view of the consequence of Lemma 3.3. As seen in Figure 3.11, the ratio decreases as Fz is increased. Therefore, the absolute supremum, or maximum, occurs for $Fz = \frac{q}{2}$, if such a z exists.

3.7.6 Number-theoretical issues

The set of quantization nodes together with the operations of addition and multiplication constitutes a mathematical entity known as a group. When the gain F belongs to the set of quantization nodes, the operation $h = Fz$ is closed in this set. This means that the value $Fz = \frac{q}{2}$ is never attained. In particular, when $q = 1$ and the elements of F are integers, the closest positive value attained is 1. When the gain has arbitrary components, it constitutes a difficult problem to determine which is the attained value which is larger or equal than and is the closest to $\frac{q}{2}$. This problem is illustrated in Figure 3.12.

It is desired to find the node for which the contour line approaches the value $\frac{q}{2}$ as much as possible from above. Mathematically, define the set

$$\tilde{F} = \left\{ Fz \mid z \in \mathbb{Z}_q,\ Fz \geq \frac{q}{2} \right\} \tag{3.12}$$

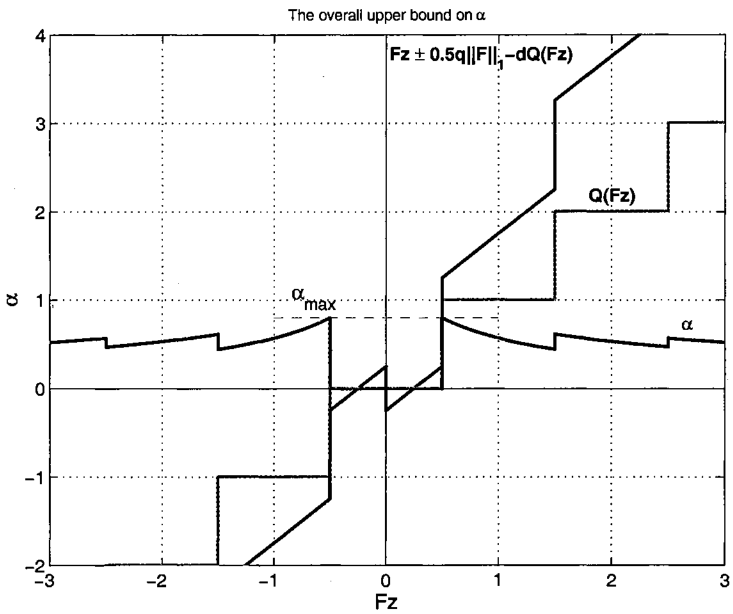

Figure 3.11: Overall variation of multiplier

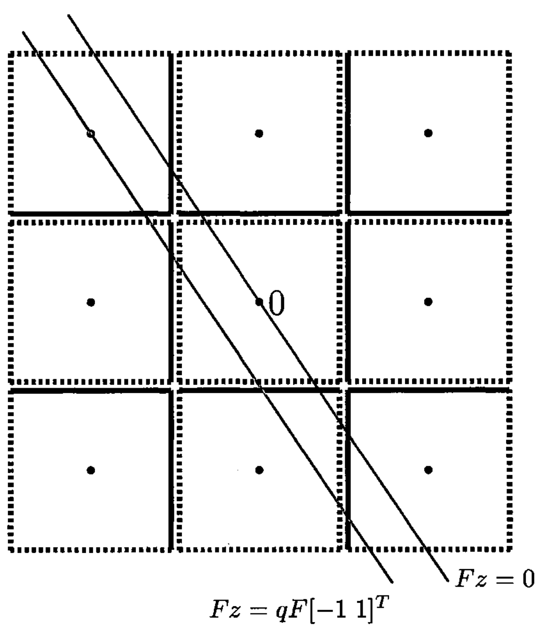

Figure 3.12: Finding a contour line approaching $Fz = \frac{q}{2}$ from above.

where $\mathbb{Z}_q$ is the set of all quantization nodes. The quantity of interest is, in the general case,

$$F_q = \inf \tilde{F} \tag{3.13}$$

When this quantity is known, the multiplier bound is expressed as

$$\bar{\alpha} = \frac{\mathcal{Q}(F_q)}{F_q - d\mathcal{Q}(F_q) - \frac{q}{2}||F||_1} \tag{3.14}$$

3.7.7 Calculation of the infimum

The problem is equivalent to computing a function

$$\Gamma(F) = \inf\left\{Fz | z \in \mathbb{Z}^n,\ Fz \geq \frac{1}{2}\right\}$$

since $\tilde{F} = q\Gamma(F)$. Consider the set

$$C_F = \{Fz | z \in \mathbb{Z}^n,\ Fz \geq 0\}$$

The following statement holds: [16, 4]

Proposition 3.1. *If F cannot be decomposed as $F = \sigma F_Q$, where σ is a real number and F_Q is a vector in $\mathbb{Q}^n$, then* $\inf C_F = 0$ *and C_F is dense. Therefore there exists a z for which Fz is arbitrarily close to $\frac{1}{2}$. On the contrary, if F can be written as $F = \sigma F_Q$, where σ is a real number and F_Q is a vector in $\mathbb{Q}^n$, then* $\inf C_F = c > 0$ *and C_F is the set of all positive integer multiples of c.*

Note that if F cannot be decomposed as rational vector multiplied by a real scalar, the sought infimum F_q will equal $\frac{q}{2}$. In the opposite case, one needs to find the smallest multiple of c which is greater than or equal to $\frac{1}{2}$. The infimum in this case could also be $\frac{q}{2}$, being actually a minimum. A sufficient condition for F not being the product of a rational vector and a real scalar is that F has irrational elements which are not all integer multiples of each other. Clearly, only F vectors with rational elements are of interest in an engineering application. For example, the finite number of bits used in digital computers automatically involves rational stored quantities. Therefore, a standing assumption on F is that it is a vector of $\mathbb{Q}^n$. Therefore, attention is focused on how to find the infimum of C_F for such case. For this purpose, an elementary result from the theory of Diophantine equations is now presented.

Diophantine equations

The equation with integer coefficients a, b and v

$$az_1 + bz_2 = v \tag{3.15}$$

where integer solutions for z_1 and z_2 are sought, is called a linear Diophantine equation. Equations with more than two variables or of higher degree are still collectively known as Diophantine, being the equation in Fermat's last theorem a notable example. It can be shown [28] that Eq. 3.15 has a solution if and only if the greatest common divisor of a and b is a divisor of v, that is, if $\gcd(a,b)|v$. This can be generalized to the following

Proposition 3.2. *The Diophantine equation*

$$\sum_{i=1}^{n} a_i z_i = v \tag{3.16}$$

has a solution for z in $\mathbb{Z}$ if and only if $gcd(a_1, a_2, ...a_n)|v$.

Finding the infimum of C_F -in this case a minimum- corresponds to finding the smallest positive value attained by the quantity Fz over all vectors z of $\mathbb{Q}^n$. The following lemma is required

Lemma 3.5. *Let F constant and z be vectors of $\mathbb{Z}^n$. The smallest positive value attained by Fz when z is varied over $\mathbb{Z}^n$ is given by the greatest common divisor of the absolute values of the components of F.*

Proof. Suppose, without loss of generality, that all components of F are nonnegative. Then, by Proposition 3.16, a vector z such that $Fz = v$ exists for a given integer v if and only of $\gcd(F_1,\ F_2,\ ....F_n)$. If $v = 0$ there is a trivial solution, however v is not positive. The integer v must be increased until a solution to the Diophantine equation is found. This happens when $v = \gcd(F_1,\ F_2,\ ....F_n)$, giving a quotient of one. If some components of F are negative, one just needs to change the sign of the corresponding components of z which attained the minimum when F had nonnegative components. ∎

To account for F having rational components, a common denominator must be factored out. Let the components of F be represented in fractional form as

$$F_i = \frac{n_i}{d_i}$$

where n_i and d_i are integers, with $d_i \neq 0$ for $i = 1, 2, ..n$. The quantity F can be expressed as

$$F = \frac{1}{\mathrm{lcm}(d_1, d_2, ..d_n)} \left[\frac{n_1}{d_1}\mathrm{lcm}(d_1, d_2..d_n) \quad \frac{n_2}{d_2}\mathrm{lcm}(d_1, d_2..d_n) \quad .... \frac{n_n}{d_n}\mathrm{lcm}(d_1, d_2..d_n) \right]$$

where lcm denotes the least common multiple. The vector components are now integer, therefore Lemma 3.5 applies. The minimum positive value attained by Fz is now readily expressed as

$$c = \min C_F = \frac{1}{\mathrm{lcm}(d_1, d_2, ..d_n)} \gcd \left[\left| \frac{n_1}{d_1}\mathrm{lcm}(d_1, d_2..d_n) \right|, \left| \frac{n_2}{d_2}\mathrm{lcm}(d_1, d_2..d_n) \right| .... \left| \frac{n_n}{d_n}\mathrm{lcm}(d_1, d_2..d_n) \right| \right]$$

Lemma 3.6. *The supremum of the multiplier $\alpha(x)$ over the state space is given by*

$$\bar{\alpha} = \frac{\mathcal{Q}(F_q)}{F_q - d\mathcal{Q}(F_q) - \frac{q}{2}||F||_1}$$

where $F_q = \mathbf{ceil}(\frac{1}{2c})qc$ *and*

$$c = \frac{1}{lcm(d_1, d_2, ..d_n)} gcd\left[\left|\frac{n_1}{d_1} lcm(d_1, d_2..d_n)\right|, \left|\frac{n_2}{d_2} lcm(d_1, d_2..d_n)\right| .... \left|\frac{n_n}{d_n} lcm(d_1, d_2..d_n)\right|\right]$$

$\mathbf{ceil}(x)$ *is the integer nearest to x which is greater or equal to x.*

Proof. The expression for $\bar{\alpha}$ in terms of F_q was derived in Section 3.7.6. The expression for F_q is a direct consequence of the previous calculations. In fact, the minimum positive element of C_F is c. This element needs to be multiplied by the least positive integer k such that $kc \geq \frac{1}{2}$. This can be computed using the "ceiling" function of the Lemma. ■

An algorithm has been implemented in the Matlab program **infval** listed at the appendix. The algorithm takes an arbitrary vector F and quantization stepsize, and returns the value of F_q. The number of decimals in F must be limited, for the conversion to rational fractions can create large denominators which cause overflows when computing their least common multiple.

Example 3.2. Take $F = [0.3243,\ -0.2120,\ 9.1245]$. A rational representation is $F = [\frac{359}{1107},\ -\frac{53}{250},\ \frac{2272}{249}]$. The least common multiple of the denominators is 22970250 (note the large magnitude for just four decimals and three F components). Therefore F can be written as

$$F = \frac{1}{22970250}[7449250,\ -4869693,\ 209592000]$$

The greatest common divisor of the absolute values of the integer components is 1. Then $c = \frac{1}{22970250}$. The integer number by which c has to be multiplied in order for the product to be greater or equal than $\frac{1}{2}$ is given by $k = \text{ceil}(\frac{1}{2c}) = 11485125$. Multiplying k by c results in a value of 0.5, to a precision of 16 digits. Thefore the value of F_q is simply

$$F_q = q \times \text{ceil}(\frac{1}{2c}) \times c = \frac{q}{2}$$

Now consider $F = [6.25\ \ -12.5]$. A rational representation is $F = [\frac{25}{4},\ -\frac{25}{2}]$. The least common multiple of the denominators is 4. F can be written as

$$F = \frac{1}{4}[25,\ -50]$$

The greatest common divisor of the absolute values of the integer components is 25. Then $c = \frac{25}{4}$. The integer number by which c has to be multiplied for the product to be greater than $\frac{1}{2}$ is 1. The value of F_q is in this case $\frac{25q}{4}$. Note that if it is desired to find the value of z at which Fz equals F_q, one has to solve a Diophantine equation. In this example, the equation

$$25z_1 - 50z_2 = 25$$

yield the values of z_1 and z_2 which produce c. Multiplying those values by q gives F_q. The expression $z_1 = 1 + 2z_2$ generates all integer solutions.

3.7.8 Summary

The multiplier $\alpha(x)$ is given in Eq.(3.8). If $d < \frac{1}{2}(1 - ||F||_1)$, then $\alpha(x(k))$ is well-defined and satisfies Lemma 3.1. Moreover, $\alpha(x)$ satisfies

$$0 \leq \alpha(x) \leq \overline{\alpha} \;\; \forall x \in \mathbb{R}^n$$

where

$$\bar{\alpha} = \frac{\mathcal{Q}(F_q)}{F_q - d\mathcal{Q}(F_q) - \frac{q}{2}||F||_1}$$

3.8 Stability Theorem for QIQM systems

In this section, the objective set for the present work is accomplished. A theorem is proposed, along with its proof. An numerical example is also provided.

Theorem 5. *Let a LTI discrete time system under quantized feedback with quantized state measurements (QIQM) be described by the equations*

$$\begin{cases} x(k+1) = Ax(k) + Bu(k) \\ u(k) = -\mathcal{Q}(F\mathcal{Q}(x(k))) \end{cases} \tag{3.17}$$

Suppose A has eigenvalues inside the open unit circle. Define the transfer function

$$G(z) = F(zI - A)^{-1}B$$

Then, the closed-loop system is globally asymptotically stable about the origin if

$$\inf_{w \in \mathbb{R}} Re\{G(e^{jw})\} > \varepsilon \tag{3.18}$$

where

$$\varepsilon = \frac{||F||_1}{4}\left\{\frac{q}{\mathcal{Q}(F_q)} + 1\right\} - \frac{F_q}{2\mathcal{Q}(F_q)} - \frac{1}{4}$$

Proof. First put the system in the equivalent form

$$\begin{cases} x(k+1) = Ax(k) + Bu(k) \\ y(k) = Fx(k) + du(k) \\ u(k) = -\mathcal{Q}(\alpha(x(k))y(k)) \end{cases} \tag{3.19}$$

where $\alpha(x(k))$ is the multiplier defined in Eq.(3.8). By Lemma 3.1, the control signal is effectively $u(k) = -\mathcal{Q}(F\mathcal{Q}(x))$, therefore the state of this system is identical to that of the original system. Suppose one of the allowed values of d is used in the equivalent system. Then $\alpha(x)$ is nonnegative and upper-bounded by $\overline{\alpha}$ discussed in Section 3.7.5. In order to prove Lyapunov stability, Theorem 4 is invoked. In this context, the quantizer is the nonlinear function $\mathcal{N}$ of the Theorem, with $\mathcal{N} \in \mathcal{S}[0,2]$, that is, $\overline{n} = 2$. Global stability is obtained if the transfer function

$$H(z) = F(zI - A)^{-1}B + d + \frac{1}{\overline{K}}$$

is DPR, with $2\bar{\alpha} < \overline{K}$, that is, if

$$\operatorname{Re}\left\{G(e^{jw})\right\} \geq -d - \frac{1}{\overline{K}}$$

for all frequencies w. The maximum sector allowed by the linear system is given by the bound

$$\overline{K}_{crit} = \frac{1}{-d - \inf \operatorname{Re}\left\{G(e^{jw})\right\}}$$

The critical sector bound must be greater than $\bar{\alpha}\overline{n} = 2\bar{\alpha}$. Enforcement of this condition and substitution of the bound

$$\bar{\alpha} = \frac{\mathcal{Q}(F_q)}{F_q - d\mathcal{Q}(F_q) - \frac{q}{2}||F||_1}$$

results in the inequality

$$d > \frac{q||F||_1}{2\mathcal{Q}(F_q)} - 2\inf \operatorname{Re}\left\{G(e^{jw})\right\} - \frac{F_q}{2}\mathcal{Q}(F_q)$$

The inequality of Condition 1 must also be considered, as well as the requirement that the critical sector $\bar{K}_{crit}$ be positive. This last requirement is automatically satisfied, since [3].

$$\inf_{w \in \mathbb{R}} \operatorname{Re}\left\{G(e^{jw})\right\} \leq 0$$

The following is obtained

$$\frac{1}{2}(1 - ||F||_1) > d > -2\inf \operatorname{Re}\left\{G(e^{jw})\right\} + \frac{q||F||_1 - 2F_q}{2\mathcal{Q}(F_q)}$$

[3]This will be the case for any plant with D=0.

The inequality of the Theorem follows directly from the above. Note that enforcing the above inequality guarantees that a d exists such that $\alpha(x)$ is well-defined, nonnegative, and bounded by $\bar{\alpha}$. However d dissapears naturally from the formulation, as it is just an artifact in the construction of the stability result. The above inequality can be rewritten to take the form of the inequality of the theorem. In order to prove asymptotic stability, steps beyond the proof of Theorem 4 are required. The second term of Eq.(3.7) has to be zero for $\Delta V(x)$ to be zero. The proof of Theorem 4 shows that the second term is zero only if $\mathcal{N}(\alpha(x)y) = 0$ or $y = 0$. When either equality is satisfied, the first term is reduced to $-LL^T x$, which must also be zero for $\Delta V(x)$ to be zero. Note also that in our case

$$\mathcal{N}(\alpha(x)y) = \mathcal{Q}(F\mathcal{Q}(x))$$

Suppose $y = 0$. This means

$$y = Fx - d\mathcal{Q}(F\mathcal{Q}(x)) = 0$$

but, from Section 3.7.3, this only happens in $\Lambda \subset \Omega_0$ defined in Section 3.7.1, where $\mathcal{Q}(x) = 0$. These findings lead to a description of the set where $\Delta V(x) = 0$:

$$R = \{x \mid \Delta V(x) = 0\} = \left[\left\{x \mid |F\mathcal{Q}(x)| < \frac{q}{2}\right\} \bigcup \{x \mid \mathcal{Q}(x) = 0\}\right] \bigcap \text{null}(LL^T) \tag{3.20}$$

Clearly, the intersection is contained in the left set i.e., the one that is written as the union of two sets. Then only one of the sets participating in the intersection needs to be considered. No information is available about LL^T, therefore the left set is taken. The shape of this set is illustrated for a two-dimensional example in Figure 3.13. To complete the proof of asymptotic stability, the discrete version of LaSalle's Invariance Principle [20] is invoked. The principle states that if the difference in the Lyapunov function is negative or zero in a bounded set Ω, and zero only in a set R, the trajectories will converge to M, the largest invariant set contained in R. The relevant sets for this case will now be identified. Let Ω be an arbitrary bounded region of the state space. For instance define Ω by the set of points x such that $V(x) < l$, for some positive l. The change in the Lyapunov function is negative or zero in this set, that is $\Delta V(x) \leq 0$ in Ω. From the above reasonings, the set R corresponds to Eq.(3.20). The largest invariant set $M \subseteq R$ is R itself, therefore the trajectories must stay within R. Note that the dynamic equation of the system inside R is simply

$$Ax(k+1) = Ax(k)$$

Since A does not have an eigenvalue equal to one, the trajectories must converge asymptotically to the origin. It is concluded that all trajectories starting in Ω converge asymptotically to the

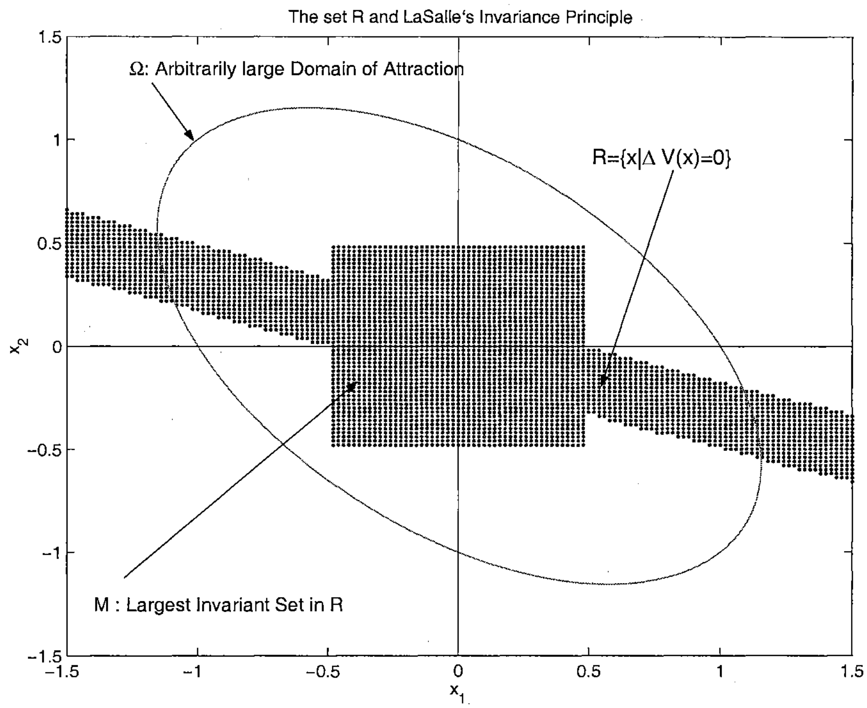

Figure 3.13: The various sets in LaSalle's Invariance Principle

origin. Since the argument is valid for Ω bounded, but arbitrarily large, global asymptotic stability is proven. ■

A Matlab program called **qiqm-check** which automates the stability test is listed at the appendix.

3.9 Numerical Example

To illustrate the concepts developed in this chapter, consider again the system of Example 2.5. A has eigenvalues inside the unit circle, therefore the stability criterion of Theorem 5 applies. A series of simulations were performed with 11 nominally stabilizing gains. The simulation diagram is shown in Figure 3.14. The sampling time was set to 0.1, but this quantity is irrelevant for stability analysis of the discrete-time system. Initial conditions were set to a large value, since stabilizing gains are small and it is desired to have nonzero control for relevant simulations. The initial condition is $x_0 = [50.45\ 40.55]$. Table 3.1 summarizes the results of the simulations. From the data in the table, it is verified that satisfaction of the criterion implies stability. The study is also useful to gain insight about how conservative the criterion could be. Of a total of 7 occurrences where the criterion was not satisfied, the system converged to the origin in 2. However, there is no guarantee that the system will be asymptotically stable for all initial conditions. In fact, reversing the sign of the initial condition

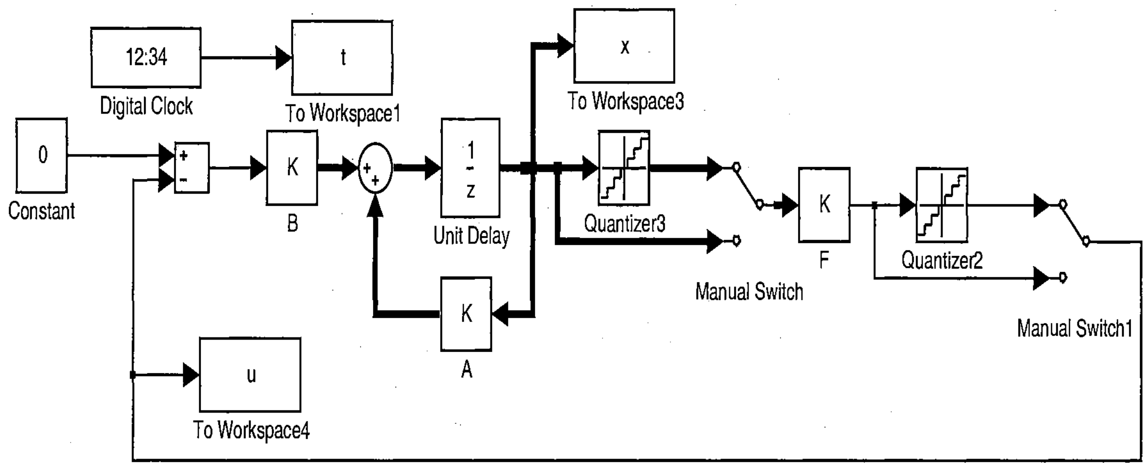

Figure 3.14: Simulink diagram for simulation study

Gain	Criterion Satisfied	Asymptotic Convergence to Zero
[0.4 0]	No	No
[0 0.4]	No	Yes
[0.4 0.4]	No	Yes
[-0.1 -0.1]	No	No
[-0.1 0]	Yes	Yes
[-0.1 0.1]	Yes	Yes
[0 -0.1]	No	No
[0 0.1]	Yes	Yes
[0.1 -0.1]	No	No
[0.1 0]	Yes	Yes
[0.1 0.1]	Yes	Yes

Table 3.1: Summary of simulation results

on x_2 results in a limit cycle when any of the two gains for which the theorem was inconclusive but were apparently stable for the original initial condition. Also, with the first gain in the table a limit cycle is obtained. If the initial condition is changed to $x_0 = [10 \;\; -40]$, asymptotic convergence is obtained. This example suggests that the stability test may not necessarily be conservative. Also, the equilibrium test developed in Section 3.2 can be used to reduce uncertainty, for if the test shows multiple equilibria, global asymptotic stability is ruled out. Figure 3.15 shows four typical behaviors found in systems with quantization, three of which correspond to the system in example. The fourth behavior was obtained by simulating the a marginally stable system found in Franklin, Powell and Workman, [15] page 729. The system matrices and gain used are

$$
\begin{aligned}
A &= \begin{bmatrix} 1.000 & 5.3316 \\ 0 & 0.9993 \end{bmatrix} \\
B &= \begin{bmatrix} 0.0133 \\ 0.0050 \end{bmatrix} \\
F &= \begin{bmatrix} 10.2851 & 147.9711 \end{bmatrix}
\end{aligned}
$$

The sampling rate used in the simulation was $T = 6.6 \times 10^{-6}$. A non-periodic permanent chatter is observed in one of the states, while the other undergoes a limit cycle.

3.10 Synthesis Issues

The stability test developed in the previous section, although sufficient, provides a means to analyze a system with a given feedback gain. The dual problem of finding gains which stabilize the QIQM system is harder, and only one method is considered here. Performance is not addressed. The method, however, is exhaustive; that is, all gains that satisfy the stability test are spanned by the introduced scaling. In future research, the space of nominally stabilizing gains along with scaling factors can be searched for a combination that optimizes a performance criterion. The key idea is to recognize that any nominally stabilizing gain will ultimately pass the test when scaled by a constant factor, since the system is open-loop stable. That is, as the scaling factor approaches zero, the system approaches its open-loop properties, including stability.

3.10.1 Scaling Procedure

Recall that the stability criterion is given by Eq. (3.18). Let F_0 be an arbitrary nominally stabilizing gain. Consider that the scaled gain $F = \sigma F_0$ is actually used. The extremal point in the Nyquist

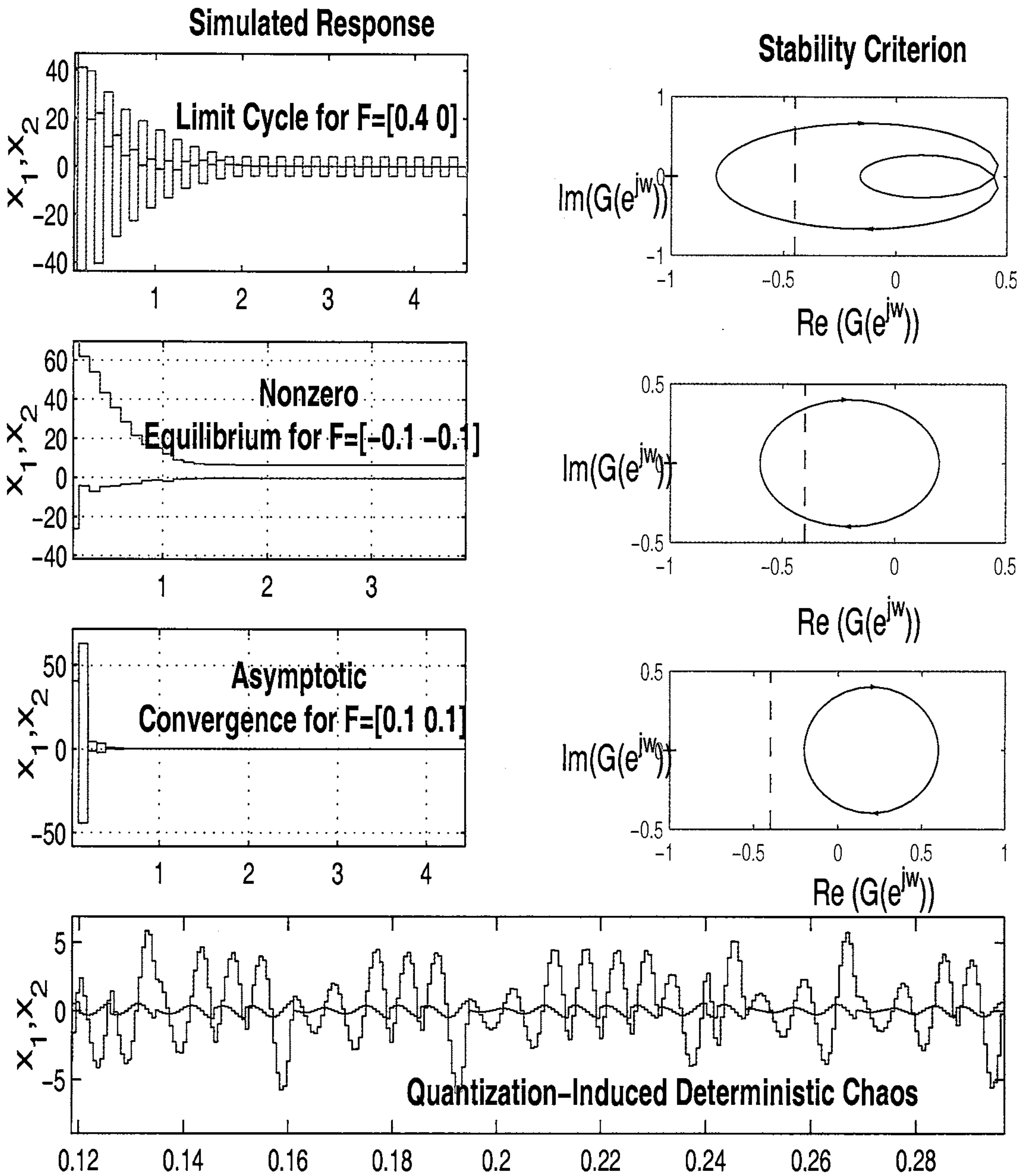

Figure 3.15: Qualitative responses in QIQM systems

diagram is scaled linearly with σ, in fact

$$\inf_{w\in\mathbb{R}} \text{Re}\left\{(\sigma F_0)(Ie^{jw} - A)^{-1})B\right\} = \sigma \inf_{w\in\mathbb{R}} \text{Re}\left\{F_0(Ie^{jw} - A)^{-1})B\right\} \tag{3.21}$$

The stability limit ε however, is scaled differently. To see this, consider the case when $F_q = \frac{q}{2}$. Following the results of Section 3.7.6, it is seen that the expression for the stability limit is simplified. In this case, it becomes

$$\varepsilon = \frac{\sigma||F_0||_1}{2} - \frac{1}{2} \tag{3.22}$$

The graphs of the scaled extremal point and the stability limit against the scaling factor σ for any fixed F_0 are straight lines. The lines have, in general, different slopes and intercepts. This implies the existence of a crossover point with abcissa σ_c that divides the σ axis into to open intervals. The set of scaling factors which generate a stable gain is one of these intervals. In the general case of F_q, the scaling is not linear, however the existence of a crossover point is understood, and the same method applies.

3.10.2 An expression for the critical scaling factor

The critical scaling factor is the point where the scaled gain is in the stability boundary. It can be found by solving the equation

$$\sigma_c \inf_{w\in\mathbb{R}} \text{Re}\left\{F_0(Ie^{jw} - A)^{-1})B\right\} = \frac{\sigma||F_0||_1}{2} - \frac{1}{2}$$

The solution is

$$\sigma_c = \frac{1}{||F_0||_1 - 2\inf_{w\in\mathbb{R}} \text{Re}\left\{F_0(Ie^{jw} - A)^{-1})B\right\}} \tag{3.23}$$

Note that if it is found that $\sigma_c = 1$, this indicates that the original gain is in the stability boundary. In order to render a gain stable by scaling, σ_c needs to be found. This can be done by evaluating another scaling factor or by explicit plotting. A major disadvantage of the scaling approach is that the critical scaling factor may be very far from one, implying that any performance objective that was intended to be met with the original nominal gain will be lost. Therefore another direction for the gain must be tried. A Matlab program called **scale-F** that automates the method is listed at the appendix and is used in the next example.

Example 3.3. Take the second order system of Example 2.5. Figure 3.16 shows the results of running **scale-F** with $Fo = [0.4,\ 0]$, which was shown earlier not to pass the stability test. The line passing through the origin is the locus of the extremal point of the Nyquist diagram as the scaling

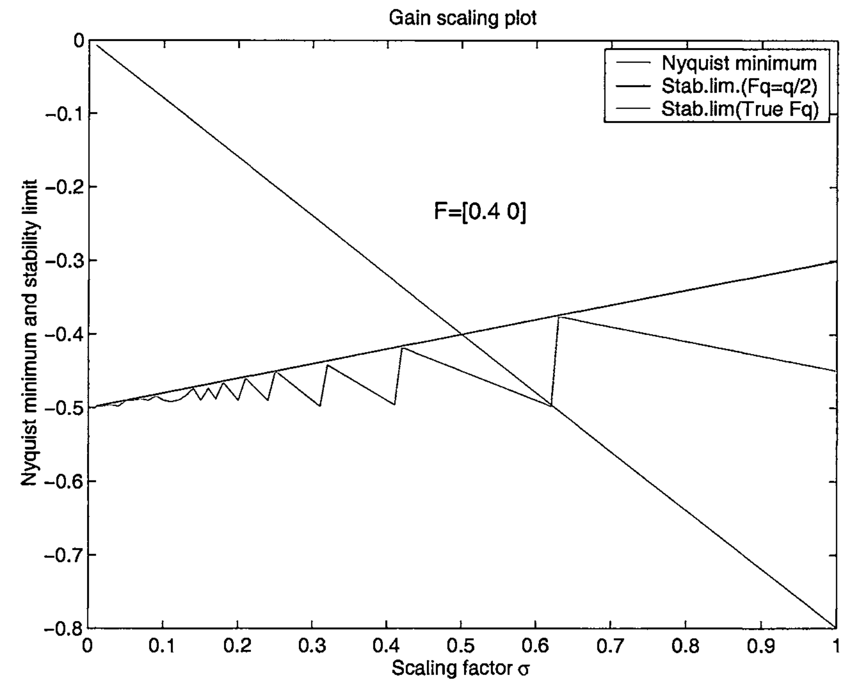

Figure 3.16: Scaling procedure in QIQM systems

factor is varied. The jagged line is the stability limit computed exactly using the theory presented earlier. The other straight line is the stability limit obtained assuming $F_q = \frac{q}{2}$. The critical scaling computed with the simplified stability limit is 0.5. The true critical scaling is roughly 0.62. The initial gain needs to be multiplied by a number between 0 and 0.62 to satisfy the stability test. Now consider $Fo = [0.1,\ 0.1]$. This gain was earlier shown to satisfy the stability test. Figure 3.17 shows that the critical scaling factor is 1.66. Therefore the original gain can be either reduced, or increased by a factor less than 1.66.

3.10.3 Bifurcations and Catastrophe in QIQM Systems

If the resolution is high enough, one can see that the stability limit plot has discontinuities, i.e., vertical segments. This implies that there could be a drastic change in system behavior if one of such discontinuities is in the neighborhood of the critical scaling factor, as it happens in the previous example with $F = [0.4,\ 0]$. A bifurcation diagram where the parameter is the scaling factor is depicted in Figure 3.18. The vertical axis represents the values of state x_1 when enough time has been allowed for the trajectories to set into either a equilibrium value or a limit cycle. At $\sigma = 0.624$ there is an abrupt change in the behavior of the system, commonly known as *catastrophe.* More bifurcations can be observed, where the amplitude of the limit cycle is suddenly increased.

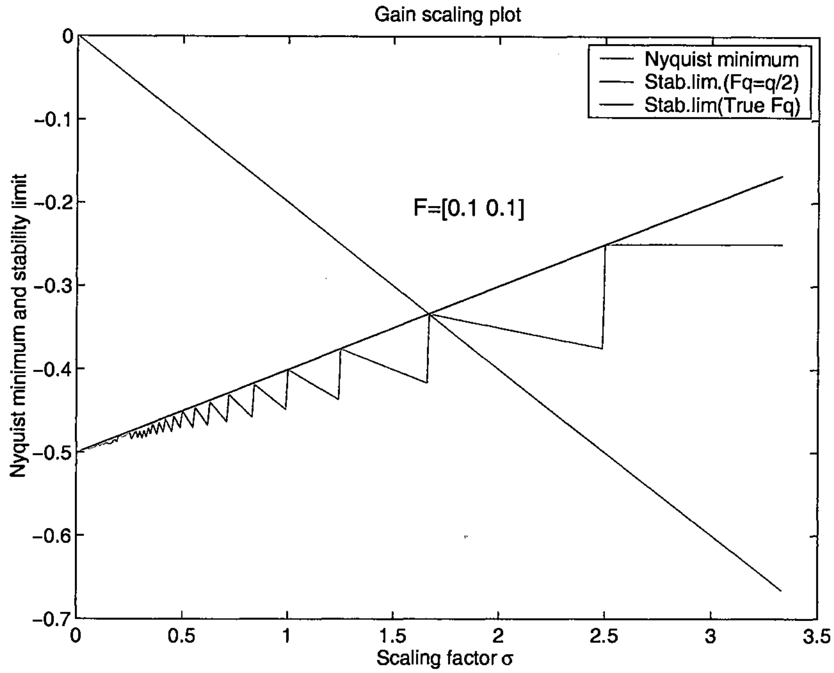

Figure 3.17: Scaling procedure in QIQM systems

Note that the stability limit obtained experimentally, i.e., the first bifurcation point, corresponds to a scaling factor of 0.624. This is in excellent agreement with the stability theorem, which gives a value of 0.620 for the critical factor. The program that illustrates these concepts is called **bifurc-quant** and is listed at the appendix.

3.10.4 Unsolvability Issues

The problem of synthesizing a gain which not only satisfies the stability criterion, but also performance requirements could be better approached if a closed-form expression were available for

$$\inf_{w \in \mathbb{R}} \operatorname{Re}\left\{F(Ie^{jw} - A)^{-1})B\right\}$$

When the infimum is actually a minimum, this calculation involves the differentiation of a rational function with respect to w and then finding the roots of a polynomial. Because of the "unsolvability of the quintic" limitation, a formula for the minimum point is not possible to obtain for arbitrary order systems. For systems between 2nd and 4th order, the formula is so complicated that it also prevents a direct algebraic treatment.

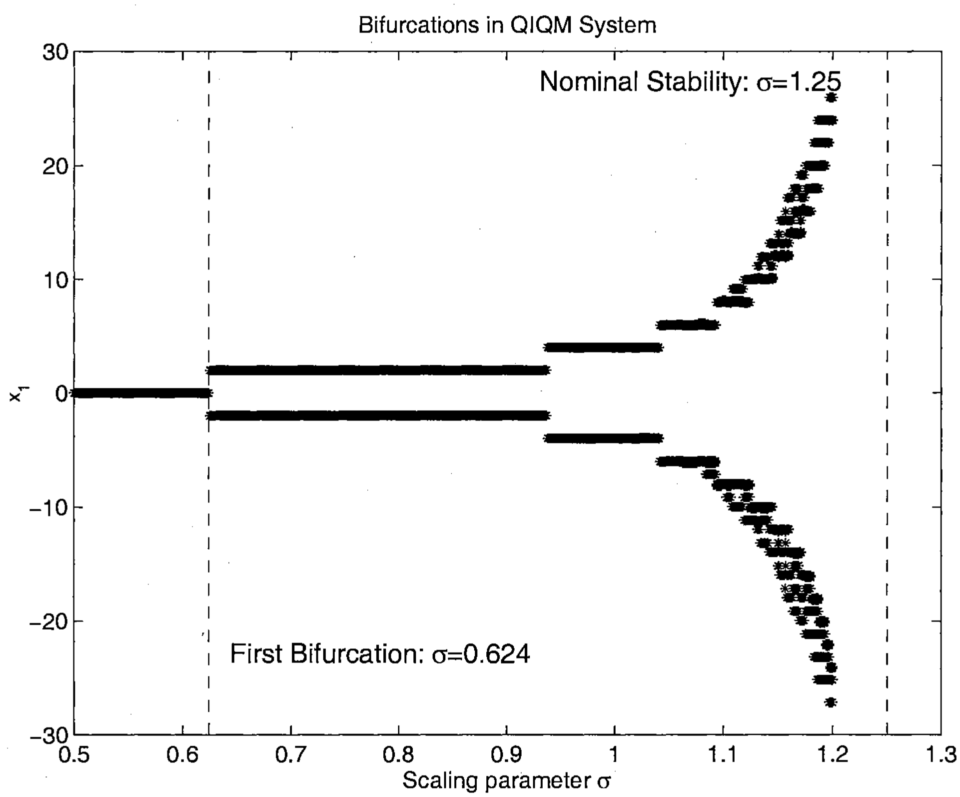

Figure 3.18: Bifurcations in QIQM Systems

Chapter 4

Quantized Input and Output Measurement (QIQO)

4.1 Problem Statement

The problem of evaluating the equilibrium points and stability of SISO systems with output feedback under quantization is analyzed in this chapter. In Section 4.2, the case where significant quantization is found only at system input is analyzed (QIO case). In Section 4.3, quantization of both, input and output is considered (QIQO) case. Section 4.4 offers remarks about the case where only an output measurement quantizer is present (IQO case). It is shown how this case reduces to problems already studied in other sections. While the QIO case can be directly analyzed using Absolute Stability, the technique has to be substantially modified in the QIQO case, and still provides with a stability condition which is unacceptably conservative. The multiplicative perturbation method used in Chapter 3 for QIQM systems is also considered. Although a MIMO version of the multiplicative perturbation theorem of Section 3.6 is proven, the method has not been applied due to difficulties in finding bounded and nonnegative multipliers. This motivates the consideration of another method; namely, the Small Gain Theorem. This method yields a simpler and less conservative stability test, and allows to show that the controller can "cooperate" with the plant in fighting quantization by means of keeping the loop gain sufficiently small. The equilibrium equations for QIQO systems have not been explicitly solved; however, a sufficient condition for the origin to be the only equilibrium point is provided, along with an iterative algorithm that can find all solutions numerically.

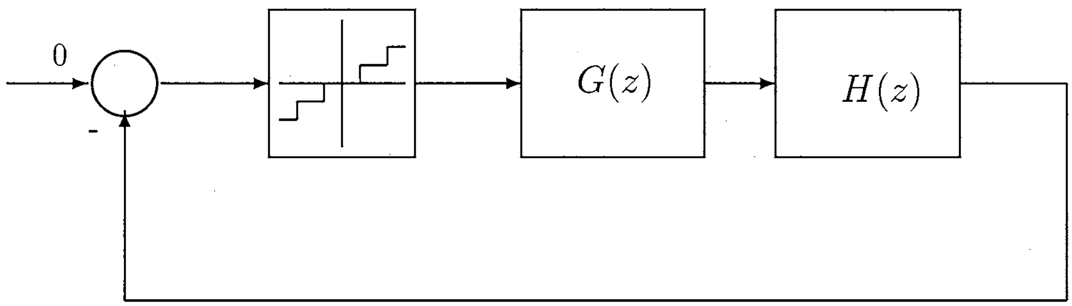

Figure 4.1: QIO System Setting

4.2 Quantized Input with Precise Output Measurement (QIO)

4.2.1 System Definition and Assumptions

Consider the system configuration depicted in Figure 4.1. Let $G(z)$ and $H(z)$ represent the SISO transfer functions of plant and controller, respectively. A standing assumption will be that the plant is strictly proper, that is the "D" term of a state space realization is zero. The controller transfer function, however, is only required to be proper. The reason for this assumption is that the existence of a unique solution to the difference equations in the feedback configuration is not guaranteed when both plant and controller are not strictly proper. A choice has to be made as to which transfer function is allowed to have a nonzero "D". Most physical systems are strictly proper, and for added flexibility in designing the controller, $H(z)$ is the one allowed to be proper, and $G(z)$ is required to be strictly proper.

4.2.2 Equilibrium Problem in QIO Systems

Let a minimal state space realization for the plant be $G(z) = (A_p, B_p, C_p, 0)$, and consider also $H(z) = (A_c, B_c, C_c, D_c)$. The equations that describe the closed-loop system are:

$$\begin{cases} x_p(k+1) = A_p x_p(k) - B_p \mathcal{Q}(y_c(k)) \\ x_c(k+1) = A_c x_c(k) + B_c y_p(k) \\ y_p(k) = C_p x_p(k) \\ y_c(k) = C_c x_c(k) + D_c y_p(k) \end{cases} \tag{4.1}$$

Note that existence of a unique solution is guaranteed, since all computations are single-valued and there are no algebraic constraints. Suppose that neither A_p or A_c have eigenvalues at $z = 1$. Denote by I_p and I_c the identity matrices with sizes equal to the orders of the plant and controller, respectively. Denote by $\bar{x}_p$ and $\bar{x}_c$ the equilibrium states of plant and controller. The equilibrium equations can be written as

$$\begin{cases} \bar{x}_p = -(I_p - A_p)^{-1} B_p \mathcal{Q}(C_c \bar{x}_c + D_c C_p \bar{x}_p) \\ \bar{x}_c = (I_c - A_c)^{-1} B_c C_p \bar{x}_p \end{cases}$$

Substituting the second equation into the first one above gives

$$\bar{x}_p = -(I_p - A_p)^{-1} B_p \mathcal{Q}\left[(C_c(I_c - A_c)^{-1} B_c + D_c) C_p \bar{x}_p\right]$$

Multiplying by C_p on both sides and recognizing the well-known matrix formula for a transfer function yields the following quantization equation:

$$\bar{y}_p = -G(1)\mathcal{Q}[H(1)\bar{y}_p] \tag{4.2}$$

This equation has the form of Eq. (2.6) studied earlier. Nothe that if zero is the only solution to Eq. (4.2) then zero is the only solution for $\bar{y}_c$, since

$$\bar{y}_c = (C_c(I_c - A_c)^{-1} B_c + D_c)\bar{y}_p = H(1)\bar{y}_p$$

Moreover, if zero is the only solution for the outputs, it is readily seen that the controller and plant states also admit only trivial solution. Therefore, using the results of Section 2.1.2, a condition for the origin to be the only equilibrium point can be derived considering Eq. (4.2) only. The condition is summarized in the following

Lemma 4.1. *The origin is the only equilibrium point of system (4.1) if and only if* A_p *and* A_c *do not have eigenvalues at 1 and* $G(1)H(1) \leq -1.5$ *or* $G(1)H(1) > -0.5$.

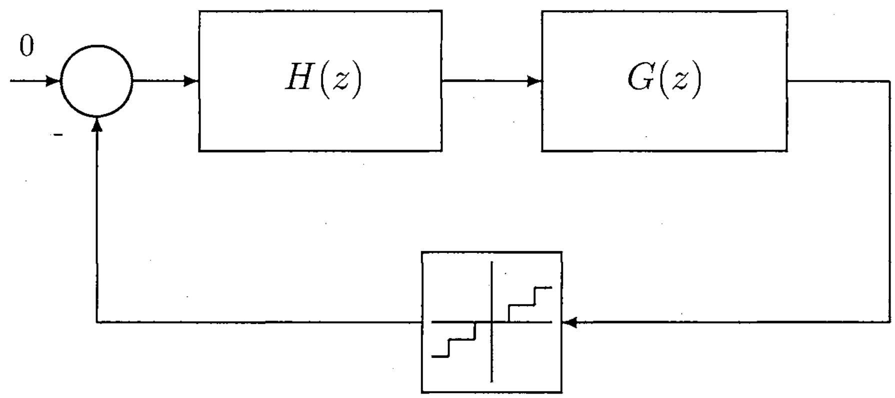

Figure 4.2: Rearrangement of QIO System

Note that for nontrivial equilibria to exist, the DC gain of the loop is required to be negative. This implies that the loop is effectively of positive feedback, suggesting that only unstable systems might have nontrivial equilibria.

4.2.3 Stability of QIO systems

Reduction to QI case

The closed-loop system can be rearranged as shown in Figure 4.2, since the scalar quantization operator is an odd function. The system representation in Eq. (4.1) can be written in a convenient fashion by realizing the cascade connection of $H(z)$ and $G(z)$:

$$\begin{cases} \left[\begin{array}{c} x_p(k+1) \\ \hline x_c(k+1) \end{array}\right] = \left[\begin{array}{c|c} A_p & 0 \\ \hline B_c C_p & A_c \end{array}\right] \left[\begin{array}{c} x_p(k) \\ \hline x_c(k) \end{array}\right] + \left[\begin{array}{c} B_p \\ \hline 0 \end{array}\right] u_p(k) \\ y_c(k) = [D_c C_p \mid C_c] \left[\begin{array}{c} x_p(k) \\ \hline x_c(k) \end{array}\right] \\ u_p(k) = -\mathcal{Q}(y_c(k)) \end{cases} \tag{4.3}$$

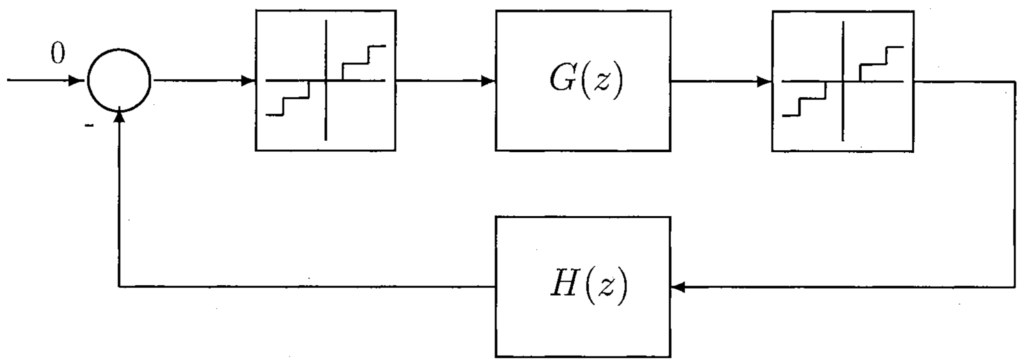

Figure 4.3: QIQO System Setting

It is now clear that this system can be viewed as having quantized state feedback with precise state measurements (QI), a case studied in Chapter 2. In fact, the above system can be rewritten by defining an augmented state $x(k) = [x_p(k)\ x_c(k)]^T$:

$$\begin{cases} x(k+1) = Ax(k) + Bu(k) \\ u(k) = -\mathcal{Q}(Fx(k)) \end{cases}$$

where the feedback gain is $F = [D_cC_p \mid C_c]$, and the augmented system matrix definitions are also obvious. Note that the equilibrium problem may also be approached from this state space representation. This can be seen by recognizing that $F(A-I)^{-1}B$ equals $-G(1)H(1)$.

4.3 Quantized Input and Output Measurement (QIQO)

The case of compensated systems with quantization at plant input and output has a high practical relevance. The system configuration is shown in Figure 4.3. Although it would be desirable to obtain an exact equilibrium point characterization and a stability test along the lines of the previous work, it will be seen that those tasks prove difficult. The equilibrium analysis will be reduced to a simpler system of equations involving rounding. The equations may be solved numerically by an iterative procedure. A useful result is still possible to obtain, in the form of a sufficient condition for the absence of nonzero equilibria. With regards to stability, the discrete positive real approach used throughout this Thesis needs to be modified to allow for a nondiagonal matrix of sector nonlinearities present in the feedback path of a MIMO system. This method yields a stability condition which is needlessly restrictive. The multiplicative perturbation method of Chapter 3 used QIQM systems is also considered, but the method has not been applied due to difficulties in finding bounded and nonnegative multipliers. This motivates the consideration of the Small Gain Theorem as a simpler and alternative method, which yields less conservative results.

4.3.1 Equilibrium Analysis of QIQO Systems

The assumptions that $G(z)$ must be strictly proper and $H(z)$ proper still hold in these developments. Let a minimal state space realization for the plant be $G(z) = (A_p, B_p, C_p, 0)$, and consider also $H(z) = (A_c, B_c, C_c, D_c)$. The equations that describe the closed-loop system are:

$$\begin{cases} x_p(k+1) = A_p x_p(k) - B_p \mathcal{Q}(y_c(k)) \\ x_c(k+1) = A_c x_c(k) + B_c \mathcal{Q}(y_p(k)) \\ y_p(k) = C_p x_p(k) \\ y_c(k) = C_c x_c(k) + D_c \mathcal{Q}(y_p(k)) \end{cases} \tag{4.4}$$

Note that existence of a unique solution is guaranteed, since all computations are single-valued and there are no algebraic constraints. Suppose that neither A_p or A_c have eigenvalues at $z = 1$. Denote by I_p and I_c the identity matrices with sizes equal to the orders of the plant and controller, respectively. Denote by $\bar{x}_p$ and $\bar{x}_c$ the equilibrium states of plant and controller. The equilibrium equations can be written as

$$\begin{cases} \bar{x}_p = -(I_p - A_p)^{-1} B_p \mathcal{Q}\left[C_c \bar{x}_c + D_c \mathcal{Q}(C_p \bar{x}_p)\right] \\ \bar{x}_c = (I_c - A_c)^{-1} B_c \mathcal{Q}(C_p \bar{x}_p) \end{cases}$$

Substituting the second equation into the first one above gives

$$\bar{x}_p = -(I_p - A_p)^{-1} B_p \mathcal{Q}\left[(C_c (I_c - A_c)^{-1} B_c + D_c) \mathcal{Q}(C_p \bar{x}_p)\right]$$

Multiplying by C_p on both sides and recognizing the well-known matrix formula for a transfer function yields the following quantization equation:

$$\bar{y}_p = -G(1) \mathcal{Q}[H(1) \mathcal{Q}(\bar{y}_p)] \tag{4.5}$$

This equation, unlike Eq. (2.6) studied before, contains nested quantization, which introduces unexpected complexity in the solution. In the following developments, a prototype scalar equation is studied.

The nested quantization equation

Consider the equation

$$x = g\mathcal{Q}(f\mathcal{Q}(x)) \tag{4.6}$$

where x is the unknown, all quantities are real and $fg \neq 0$.

Proposition 4.1. *If $fg < 0$ then $x = 0$ is the only solution to Eq. (4.6).*

Proof. Suppose $x_o \neq 0$ is a solution with $fg < 0$. Then

$$\frac{x_o}{g} = \mathcal{Q}(f\mathcal{Q}(x_o))$$

therefore

$$\frac{x_o^2}{g^2} = \frac{x}{g}\mathcal{Q}(f\mathcal{Q}(x)) > 0$$

Since $\mathcal{Q}(f\mathcal{Q}(x))$ has the same sign as $f\mathcal{Q}(x)$ and, moreover, the same sign as fx, if follows that

$$\frac{x}{g}xf > 0$$

that is

$$x^2\frac{f}{g} > 0$$

from which the contradiction that f/g and thus fg are positive arises. ∎

Proposition 4.2. *If x is a solution to Eq. (4.6), then $-x$ is also a solution.*

This follows directly from the odd character of the quantization operator. Because of this fact, it is only necessary to look for positive solutions to the quantization equation. Also, only the case $f > 0$, $g > 0$ needs to be considered. In fact, suppose $f < 0$ and $g < 0$. Define $y = -x$ and write the equation as

$$y = -g\mathcal{Q}((-f)\mathcal{Q}(y))$$

It is then sufficient to look for positive solutions to the above quantization equation with $f' = -f > 0$ and $g' = -g > 0$.

Reduction to a system of equations with rounding

Let $q > 0$ be the quantization stepsize and x be a positive solution to Eq. (4.6). Then the following must hold:

$$\mathcal{Q}(x) = qi \tag{4.7}$$

$$\frac{x}{g} = jq \tag{4.8}$$

for some positive integers i and j. From Eq. (4.7) write

$$qi - \frac{q}{2} \leq x < qi + \frac{q}{2} \tag{4.9}$$

Write Eq. (4.6) as

$$\frac{x}{g} = \mathcal{Q}(fqi)$$

This occurs if and only if

$$fqi - \frac{q}{2} < \frac{x}{g} \leq fqi + \frac{q}{2} \tag{4.10}$$

Divide Eq. (4.9) by $g > 0$ and use Eq. (4.8). Similarly, use Eq. (4.8) in Eq. (4.10) to obtain the following system of inequalities, where q has cancelled out:

$$\begin{cases} \frac{i}{g} - \frac{1}{2g} \leq j < \frac{i}{g} + \frac{1}{2g} \\ fi - \frac{1}{2} < j \leq fi + \frac{1}{2} \end{cases} \tag{4.11}$$

Any pair of integers $i \geq 1$ and $j \geq 1$ satisfying the above system will generate a solution. Multiplying the first inequality in the above system by $g > 0$ shows that it is equivalent to a system of equations:

$$\begin{cases} \text{round}(fi) = j \\ \text{round}(gj) = i \end{cases} \tag{4.12}$$

The striking symmetry in the equations is deceiving, since a general, closed-form solution seems impossible to obtain, as it will be seen next. The system of inequalities in Eq. (4.11) has a solution if several conditions are met. First, the intersection of the intervals specified for j must be non-empty. Second, the intersection must contain at least an integer. This second condition can satisfied only if one of the intervals contains itself an integer. Those conditions need also to be stated when the inequalities are written as intervals for i. The only tractable condition is the first one, for there is no general test to determine if an interval contains an integer. That is, if $a < b$ are interval ends, numerical computation is required to decide if the interval contains an integer. A sufficient condition, however, is that the interval length be greater than one. The approach in the present work will be to obtain a sufficient condition for the absence of nonzero solutions and to establish a numerical algorithm to find all solutions.

Sufficient condition for the absence of nonzero solutions

The j intervals in the inequalities of Eq. (4.11) have non-empty intersection over the real numbers if

$$\begin{cases} \frac{i}{2} + \frac{1}{2g} > fi - \frac{1}{2}, \text{ and} \\ fi + \frac{1}{2} \geq \frac{i}{g} - \frac{1}{2g} \end{cases}$$

Only $i \geq 1$ needs to be considered. First suppose $f > 1/g$. Then the inequalities become

$$\frac{-\frac{1}{2g} - \frac{1}{2}}{f - \frac{1}{g}} \leq i < \frac{\frac{1}{2g} + \frac{1}{2}}{f - \frac{1}{g}}$$

The left term is negative, and therefore less than one, so the inequalities reduce to

$$1 \leq i < \frac{\frac{1}{2g} + \frac{1}{2}}{f - \frac{1}{g}} \tag{4.13}$$

Now suppose $f < 1/g$. Similarly, the inequalities reduce to

$$1 \leq i \leq \frac{\frac{1}{2g} + \frac{1}{2}}{-(f - \frac{1}{g})} \tag{4.14}$$

The inequalities in Eq. (4.11) need to be rewritten as intervals for i:

$$\begin{cases} \frac{j}{f} - \frac{1}{2f} \leq i < \frac{j}{f} + \frac{1}{2f} \\ gj - \frac{1}{2} < i \leq gj + \frac{1}{2} \end{cases} \tag{4.15}$$

The above system has exactly the same form as Eq. (4.11) if f and g are exchanged and i and j are exchanged. Taking advantage of this symmetry, the following inequalities can be written directly:

$$1 \leq j < \frac{\frac{1}{2f} + \frac{1}{2}}{g - \frac{1}{f}}, \text{ when } g > 1/f \tag{4.16}$$

and

$$1 \leq j \leq \frac{\frac{1}{2f} + \frac{1}{2}}{-(g - \frac{1}{f})}, \text{ when } g < 1/f \tag{4.17}$$

Now let k denote the number of solutions to Eq. (4.6). It is clear that inequalities Eq. (4.13, 4.14, 4.16, 4.17) provide bounds for k. That is,

$$\begin{cases} k < \min\left\{\frac{\frac{1}{2f}+\frac{1}{2}}{g-\frac{1}{f}}, \frac{\frac{1}{2g}+\frac{1}{2}}{f-\frac{1}{g}}\right\} & \text{, when } f > 1/g \\ k \leq \min\left\{\frac{\frac{1}{2f}+\frac{1}{2}}{-(g-\frac{1}{f})}, \frac{\frac{1}{2g}+\frac{1}{2}}{-(f-\frac{1}{g})}\right\} & \text{, when } f < 1/g \end{cases}$$

which can be compactly expressed in the following

Lemma 4.2. *Let k be the number of positive solutions to Eq. (4.6). Then*

$$\begin{cases} k < \frac{1}{fg-1}[1+\min(f,g)] & \text{, when } f > 1/g \\ k \leq \frac{1}{1-fg}[1+\min(f,g)] & \text{, when } f < 1/g \end{cases} \tag{4.18}$$

Note that when $f = 1/g$, it can be shown that it is necessary that $f > -1$ for a nontrivial solution to exist. Since it is assumed that $f > 0$ and the conditions developed are only necessary, there could be any number of nontrivial solutions, ranging from zero to infinity.

The main purpose of the above calculations is to derive a sufficient condition for zero to be the only solution. This can be done by requiring that k, i.e., the number of nontrivial solutions, be less than one. Upon doing this, the following lemma can be stated:

Lemma 4.3. *Zero is the only solution of Eq. (4.6) if $fg \geq \frac{3}{2} + \frac{\min(f,g)}{2}$ or $fg < \frac{1}{2} - \frac{\min(f,g)}{2}$.*

Application to equilibrium of QIQO systems

The above results can be applied to analyze multiplicity of equilibria in QIQO systems by taking $f = H(1)$ and $g = -G(1)$. The following lemma summarizes the findings:

Lemma 4.4. *The origin is the only equilibrium point of the QIQO system of Eq. (4.19) if*

$$-G(1)H(1) \geq \frac{3}{2} + \frac{\min(H(1), -G(1))}{2}$$

or

$$-G(1)H(1) < \frac{1}{2} - \frac{\min(H(1), G(1))}{2}$$

Example 4.1. Consider the following strictly proper and proper transfer functions for plant and controller, respectively:

$$G(z) = -\frac{1}{z - 0.5}$$

$$H(z) = \frac{z - 0.6}{z - 0.1}$$

Consider a quantization stepsize of $q = 1$. Although unnecessary for the equilibrium analysis, it can be verified that the closed-loop system is nominally stable. This fact, however, facilitates a simulation that shows the system outputs indeed converging to the nonzero equilibrium values predicted by the theory. Noting that $f = H(1) = \frac{4}{9}$ and $g = -G(1) = 2$, the case $f < 1/g$ holds. Then, using the second of Eq. (4.18) gives a bound for the number of positive solutions:

$$k \leq \frac{1}{1 - fg}[1 + \min(f, g)] = 13$$

Therefore, up to 13 positive equilibrium points for the plant and controller outputs are possible. The actual solutions have to be found by numerically iterating Eq. (4.12) for 13 iterations. The Matlab program **solve-qiqo** listed at the appendix automates these tasks. The program output is

```
>> sols=solve_qiqo(4/9,2,13)


sols =


     2     4     6     8
```

The returned values are those of i for which there is positive solution to Eq. (4.12). The corresponding values of j can be found through the same equations, and are $j = 1, 2, 3, 4$. According to Eq. (4.8), the equilibrium values of the plant output are just

$$\bar{y}_p = gjq = 2j$$

The equilibrium values of the controller output are

$$\bar{y}_c = H(1)Q(\bar{y}_p)$$

Figure 4.4 shows system outputs converging to the equilibrium value $(\bar{y}_p,\ \bar{y}_c) = (2,\ 8/9)$, which is predicted by the above equations. Note that, given state space realizations for plant and controller, it is also possible to calculate the equilibrium values of the whole states.

4.3.2 Stability Analysis of QIQO Systems

In this section, it will be assumed that both, plant and controllers are asymptotically stable. This assumption stands in addition to the properness requirements considered earlier. Three methods have been considered to analyze the stability of QIQO systems. First, an equivalent system consisting of a purely linear MIMO system with decoupled nonlinearities in the feedback path was considered. The nonlinearities are quantizers multiplicatively perturbed by functions of the state. This is to be done in the spirit of the method used for QIQM systems in Chapter 3. The method fails because of the difficulty in finding bounded and nonnegative multipliers. However, the generalization of the multiplicative perturbation theorem of Section 3.6 holds. The method is described in Section 4.3.2, and the MIMO multiplicative perturbation theorem is given in the appendix. The second method

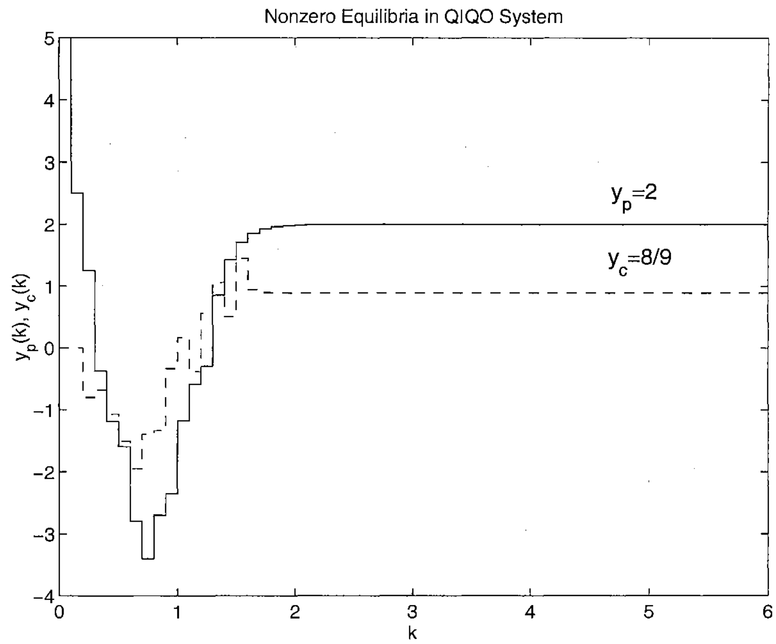

Figure 4.4: Nonzero equilibrium in QIQO System

attempts to generalize the Absolute Stability theory for MIMO systems with output-crossing nonlinearities. This is done by requiring that certain MIMO transfer function be DPR. The method does result in a stability criterion; however its conservativeness is unacceptable. The method is described in Section 4.3.2. The third method is much simpler and gives better results. It directly uses the Small Gain theorem, formulated for discrete systems. The required definitions and results of l_p stability theory are given in Section 4.3.2, along with the application to the QIQO case and a numerical example.

Multiplicative Perturbation Method

This method attempts to generalize the techniques and results used for QIQM systems in Chapter 3. Let Σ_0 denote a state space realization of the original QIQO system:

$$\Sigma_0 : \begin{cases} x_p(k+1) = A_p x_p(k) - B_p \mathcal{Q}(y_c(k)) \\ x_c(k+1) = A_c x_c(k) + B_c \mathcal{Q}(y_p(k)) \\ y_p(k) = C_p x_p(k) \\ y_c(k) = C_c x_c(k) + D_c \mathcal{Q}(y_p(k)) \end{cases} \tag{4.19}$$

Define the MIMO linear system L_1 by

$$L_1 : \begin{cases} \left[\begin{array}{c} x_p(k+1) \\ \hline x_c(k+1) \end{array}\right] = \left[\begin{array}{c|c} A_p & 0 \\ \hline 0 & A_c \end{array}\right] \left[\begin{array}{c} x_p(k) \\ \hline x_c(k) \end{array}\right] + \left[\begin{array}{c|c} B_p & 0 \\ \hline 0 & B_c \end{array}\right] \left[\begin{array}{c} u'_p(k) \\ \hline u'_c(k) \end{array}\right] \\ \\ \left[\begin{array}{c} y'_p(k) \\ \hline y'_c(k) \end{array}\right] = \left[\begin{array}{c|c} C_p & 0 \\ \hline 0 & C_c \end{array}\right] \left[\begin{array}{c} x_p(k) \\ \hline x_c(k) \end{array}\right] + \left[\begin{array}{c|c} d_p & 0 \\ \hline 0 & d_c \end{array}\right] \left[\begin{array}{c} u'_p(k) \\ \hline u'_c(k) \end{array}\right] \end{cases} \tag{4.20}$$

for some real numbers d_p and d_c. If suitable multipliers α_p and α_c could be found such that

$$u'_p(k) = -\mathcal{Q}\left[\alpha_p(x_p(k), x_c(k))y'_p(k)\right] = -\mathcal{Q}(y_c(k)) = u_p(k) \tag{4.21}$$

$$u'_c(k) = -\mathcal{Q}\left[\alpha_c(x_p(k), x_c(k))y'_c(k)\right] = \mathcal{Q}(y_p(k)) = u_c(k) \tag{4.22}$$

$$\tag{4.23}$$

then closing the loop on L_1 with controls u'_p and u'_c would achieve two desired objectives: denote by Σ_1 the closed loop; then, Σ_0 and Σ_1 will have the same values of state at all instants, and therefore their internal stability will be equivalent. With equal importance, L_1 would be a linear system with decoupled nonlinearities in the feedback path, which are multiplicatively perturbed by a function of the state. The required results regarding stability of DPR MIMO systems with "D" term and the multiplicative perturbation theorem hold, and are presented at the appendix. The method falls short of being applicable because appropriate multipliers cannot be found. In fact, consider the candidate definitions, dropping index k from the notation:

$$\alpha_p(x_p, x_c) = \frac{\mathcal{Q}(y_c)}{y'_p} = \frac{\mathcal{Q}(y_c)}{y_p - d_p\mathcal{Q}(y_c)} \tag{4.24}$$

$$\alpha_c(x_p, x_c) = -\frac{\mathcal{Q}(y_p)}{y'_c} = \frac{\mathcal{Q}(y_p)}{y_c + (d_c - D_c)\mathcal{Q}(y_p)} \tag{4.25}$$

The multipliers are functions of two variables, therefore their boundedness and nonnegativity cannot be established if the variables are considered to vary independently. The variations of y_p and y_c are actually coupled through the system equations, which makes the determination of existence of suitable d_p and d_p a very difficult task. This motivates the search for another approach.

DPR Method

The method consists in viewing the closed loop as a two-input, two-output linear system with sector nonlinearities in the feedback path, as it is done is standard Absolute Stability. Traditional methods cannot be applied, however, since the plant outputs cross when entering the nonlinearities, resulting

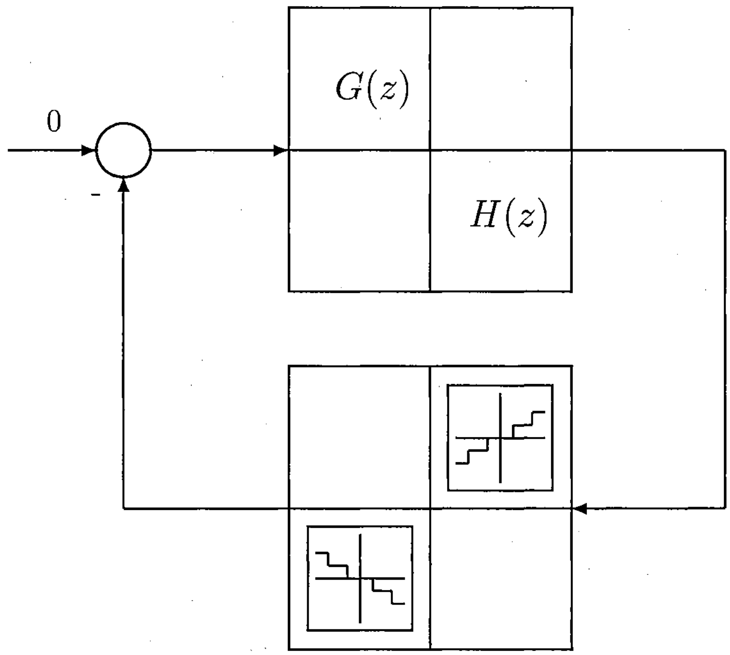

Figure 4.5: QIQO Problem in MIMO Form

in a non-diagonal matrix of nonlinearities at the feedback path. Figure 4.5 shows the problem from this point of view. The method requires certain transfer function to be DPR, but it removes the notion of nonlinearities belonging to specific sectors. Instead, it asks for a special algebraic property that must be satisfied by the combination of nonlinearities. Rather than stating results followed by their proofs, a constructive approach will be followed to facilitate understanding. Let $G(z)$ and $H(z)$ be the plant and controller transfer functions of a QIQO system, respectively, with $G(z)$ strictly proper and $H(z)$ proper. The condition sought is that the system be globally asymptotically stable if certain transfer function is DPR. A choice of DPR transfer function that resembles the standard theory is the following

$$M(z) = \left[\begin{array}{c|c} G(z) & 0 \\ \hline 0 & H(z) \end{array}\right] + F \tag{4.26}$$

The role of the diagonal matrix containing the sector bounds is taken by F in this case, which is only required to be symmetric. The question is, what conditions must F satisfy in order for the system to be G.A.S. when $M(z)$ is DPR. In order to determine this, suppose that $M(z)$ is indeed DPR and use the Lyapunov approach employed throughout this work. The system can be written in MIMO form as follows:

$$\left\{\begin{array}{l}\left[\begin{array}{c}x_p(k+1)\\ \hline x_c(k+1)\end{array}\right]=\left[\begin{array}{c|c}A_p & 0\\ \hline 0 & A_c\end{array}\right]\left[\begin{array}{c}x_p(k)\\ \hline x_c(k)\end{array}\right]+\left[\begin{array}{c|c}B_p & 0\\ \hline 0 & B_c\end{array}\right]\left[\begin{array}{c}u_p(k)\\ \hline u_c(k)\end{array}\right]\\ \\ \left[\begin{array}{c}y_p(k)\\ \hline y_c(k)\end{array}\right]=\left[\begin{array}{c|c}C_p & 0\\ \hline 0 & C_c\end{array}\right]\left[\begin{array}{c}x_p(k)\\ \hline x_c(k)\end{array}\right]+\left[\begin{array}{c|c}0 & 0\\ \hline 0 & D_c\end{array}\right]\left[\begin{array}{c}u_p(k)\\ \hline u_c(k)\end{array}\right]\end{array}\right. \tag{4.27}$$

Let A, B, C and D denote the MIMO system matrices in obvious correspondence with the above equations and let $x = [x_p \ x_c]^T$ be the augmented state. The original QIQO system is reproduced by the above equations if the following controls are used:

$$u(k) = -\Phi(x(k)) = \begin{bmatrix} u_p(k) \\ u_c(k) \end{bmatrix} = \begin{bmatrix} -\mathcal{Q}(y_c(k)) \\ \mathcal{Q}(y_p(k)) \end{bmatrix} \tag{4.28}$$

Since $M(z)$ is DPR, there exist matrices P,L and W which satisfy Lemma B.2, that is

$$\begin{aligned} A^T PA - P &= -LL^T \\ B^T PA &= C - W^T L^T \\ W^T W &= D + D^T + 2F - B^T PB \end{aligned} \tag{4.29}$$

Choosing the Lyapunov function $V(x(k)) = x^T(k)Px(k)$ and finding its change along system trajectories results in the cancellation of D and the following expression:

$$\Delta V = -[L^T x - W\Phi]^T [L^T x - W\Phi] - 2[\Phi^T y - \Phi^T F \Phi]$$

Being the first term negative semidefinite, it is desired to find conditions on matrix F such that the second term is also negative semidefinite. This reduces to analyzing the properties of the nonlinearity, in this case the quantizer. To simplify the analysis, it is supposed that $F = \text{diag}(f_1,\ f_2)$. Moreover, it will be seen later that off-diagonal terms in F do not increase the generality or reduce the conservativeness of the method. The expression inside brackets in the second term is expanded as

$$\Psi(y_p, y_c) = -f_1 \mathcal{Q}^2(y_c) - f_2 \mathcal{Q}^2(y_p) + \mathcal{Q}(y_c) y_p - \mathcal{Q}(y_p) y_c$$

Defining

$$\overline{\Psi}(y_p, y_c) = -\Psi(y_p, y_c) \tag{4.30}$$

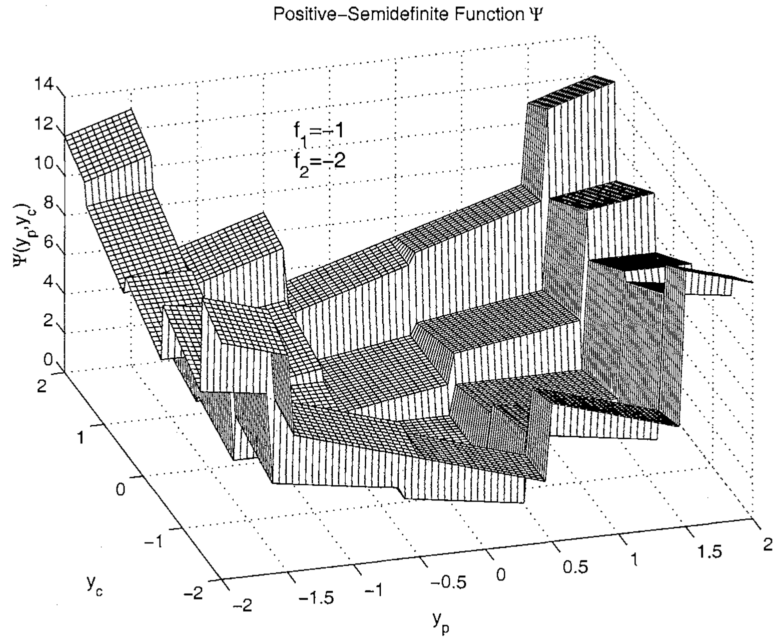

Figure 4.6: Positive-Semidefinite Function Ψ

it is seen that $\overline{\Psi}$ has to be negative semidefinite to obtain global stability. The shape of the faceted function Ψ is shown in Figure 4.6. It is shown in the appendix that this occurs if and only if $f_1 < -1/2$ and $f_2 < -1/2$, and moreover, $\overline{\Psi} = 0$ only for $\mathcal{Q}(y_p) = \mathcal{Q}(y_c) = 0$. This last fact helps in proving asymptotic stability, for it implies that both controls are zero and each stable transfer function must decay asypmtotically to zero. A rigurous proof requires the discrete version of LaSalle's theorem, as it was done for QIQM systems. Therefore, the closed loop system is G.A.S. if $M(z)$ is DPR, with F diagonal such that $f_{ii} < -\frac{1}{2}$. Since $M(z)$ is itself diagonal, the requirement is imposed on each transfer function. While it is possible to find a $H(z)$ which is DPR by a margin of more than 1/2, this never happens for strictly proper plants, since

$$\inf \text{Re}\, G(e^{jw}) \leq 0$$

for a strictly proper G. The result suggests that no strictly proper plant can pass the stability test and that the controller cannot influence the plant's ability to resist quantization-induced limit cycles. The results are thus unusable. Allowing a more general F will again result in a matrix of the form $F + F^T$ in the definition of $\overline{\Psi}$, which is again symmetric. The off-diagonal terms only add complication, for the DPR condition is still applied to the diagonal elements of $M(z)$. Thus, more restrictions are added to those already introduced by a diagonal F matrix. A possible alternative is to require that a different transfer function be DPR, but this would create difficulties when trying to apply the matrix conditions for DPR given in Lemma B.2.

Small Gain Approach

The Small Gain theorem, commonly known in its continuous-time formulation also holds in difference systems. The result formalizes the intuitive concept that the overall gain in a closed loop must be less than one for system outputs to converge asymptotically to zero. This can be verified in a straightforward manner for static systems (i.e., gains coupled in a negative feedback loop). In order to formulate the teorem for dynamic systems, certain definitions are required. In what follows, only the basic definitions and results will be given. For further details, refer to [35].

Definitions and Results

Denote by $\mathcal{S}$ the space of sequences x_i for $i \geq 0$. Define the set

$$l_p = \left\{ x \in \mathcal{S} \;\middle|\; \sum_{i=0}^{\infty} |x_i|^p < \infty \right\}$$

Note that, unlike the corresponding case of continuous-time signals, there is no need to define an "extended" set composed of sequences whose truncations at $i = I$ are in l_p for all $I \geq 0$. This is due to the fact that all sequences of $\mathcal{S}$ have that property. In other words, there is no "finite escape time". Therefore, the set $\mathcal{S}$ acts as an extension of l_p. Define the norms

$$||x||_p = \left[\sum_{i=0}^{\infty} |x_i|^p \right]^{\frac{1}{p}}$$

$$||x||_\infty = \sup_i |x|_i$$

Definition 4.1. *A mapping $R : \mathcal{S} \to \mathcal{S}$ is said to be l_p-stable if $Rx \in l_p$ for all $x \in l_p$. Also, R l_p-stable with finite gain and zero bias if R is l_p-stable and $\exists\, \gamma < \infty$ such that*

$$||Rx||_p \leq \gamma ||x||_p$$

In the context of the present work, the mapping R is a discrete transfer function.

Theorem 6 (Small Gain). *Consider the system in Figure 4.7. Suppose that G_1 and G_2 are causal and l_p-stable with finite gains γ_1 and γ_2. Then $y_1 \in l_p$ and $y_2 \in l_p$ if $\gamma_1 \gamma_2 < 1$ and $\psi_1 \in l_p$ and $\psi_2 \in l_p$.*

Application to QIQO Systems

In order to apply the Small Gain theorem to systems with quantization, it is necessary to combine the quantizer and transfer function to form a single operator and show that the combination is l_p

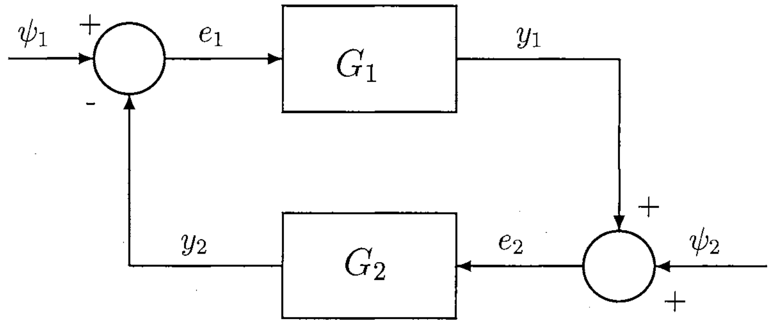

Figure 4.7: Setting for Small Gain Theorem

stable if the transfer function is itself stable. Also, it is necessary to compute the finite gain of the combination. Suppose $G(z)$ is a proper transfer function with all poles inside the unit circle. Then $G(z)$ is l_p-stable with finite gain and zero bias, since [11]

$$||Gx||_p \leq ||G||_p||x||_p$$

The most common p-norms of a transfer function are given below in terms of the impulse response sequence g_i:

$$\begin{aligned} ||G||_1 &= \sum_{i=0}^{\infty} |g_i| \\ ||G||_2 &= \sqrt{\sum_{i=0}^{\infty} |g_i|^2} \\ ||G||_\infty &= \sup_i |g_i| \end{aligned}$$

Therefore, the l_p gain of $G(z)$ is simply its l_p norm. The quantizer is also l_p stable. In fact consider an input sequence $x_i \in l_p$, with norm $||x||_p$. The worst-case scenario occurs when x is the sequence of transition points, that is

$$x = \{.... - 3q/2, -q/2, q/2, 3q/2....\}$$

In this case, the output sequence is twice the input, and it is straightforward to verify that its norm is $2||x||_p$ for any p. Therefore

$$||\mathcal{Q}(x)||_p \leq 2||x||_p$$

so the l_p gain of a quantizer is 2. Now consider the combination of quantizer and transfer function depicted in Figure 4.8. The overall l_p gain is less or equal to the product of the individual ones. In fact:

$$||y|| \leq ||G||_p||x_q||_p \leq ||G||(2||x||)$$

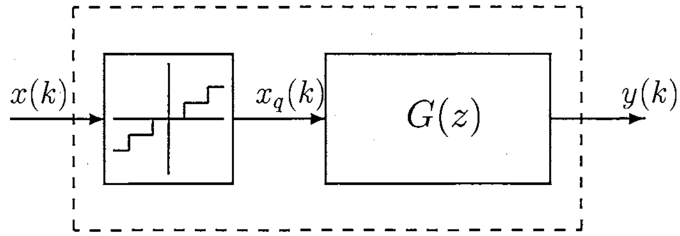

Figure 4.8: Combined Quantizer-Transfer Operator

p	$\|\|G\|\|_p$	$\|\|H\|\|_p$	Loop Gain
1	2	1.55	3.1
2	1.15	1.12	1.29
∞	2	1.45	2.9

Table 4.1: l_p Gains in QIQO system

Therefore the combined operator is l_p-stable with zero bias and finite gain $\gamma_p = 2||G||_p$. Application of the Small-Gain theorem is straighforward and results in the following

Lemma 4.5. *The QIQO system of Eq. (4.19) is globally asymptotically stable about the origin if*

$$||G||_p||H||_p < \frac{1}{4}$$

The proof follows directly from the above considerations by combining the plant and controller with their input quantizers and noting that $\psi_1 = \psi_2 = 0 \in l_p$ for regulation about the origin. The choice of p is done in a case by case basis, using the value that gives less conservative results.

Example 4.2. Consider the transfer functions of Example 4.1. Table 4.1 shows the values of the gains for different values of p. As seen in the table, the loop gain is higher than $\frac{1}{4}$. The Small Gain Theorem is only sufficient, and conservative because it is a worst-case approach. However, it is known from Example 4.1 that this system has multiple equilibrium points, and therefore cannot be G.A.S. about the origin. In order for the system to satisfy the Small Gain theorem, the gain of the loop must be reduced to be less than 0.25. Taking the less conservative case of $p = 2$, a factor of $0.25/1.29 = 0.19$ must be introduced in the system. Including the factor in the plant, the new transfer function is

$$G'(z) = \frac{-0.19}{z - 0.5}$$

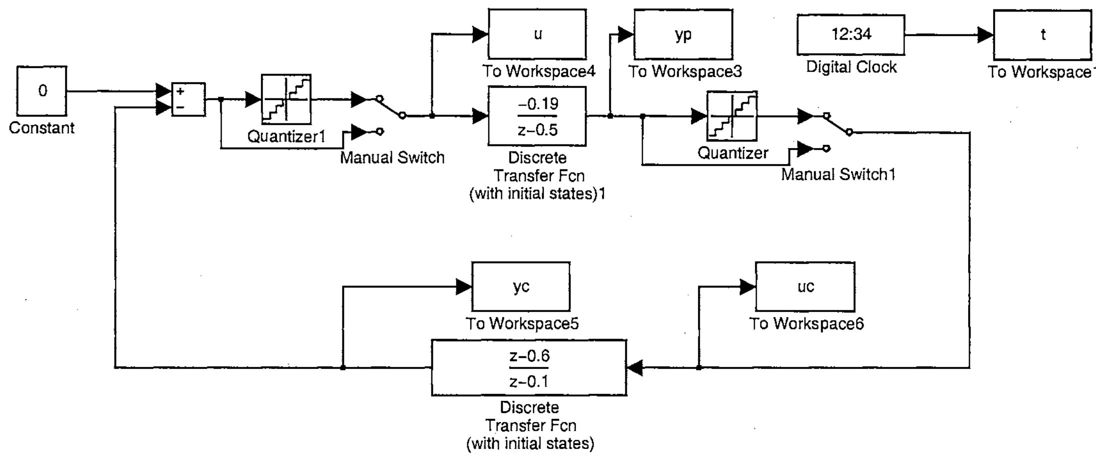

Figure 4.9: Simulation of QIQO System

The equilibrium analysis can be carried out again with $f = H(1) = \frac{4}{9}$ and $g = -G(1) = 0.38$, where the case $f < 1/g$ holds. Then, using the second of Eq. (4.18) gives a bound for the number of positive solutions:

$$k \leq \frac{1}{1 - fg}\left[1 + \min(f, g)\right] = 1.66$$

Therefore at most one nonzero solution can happen. To find the actual solutions the Matlab program **solve-qiqo** is executed with one iteration:

```
>> solve_qiqo(f,g,1)

ans =

     []
```

This shows that the origin is the only equilibrium point. A simulation, illustrated in Figures 4.9 and 4.10 confirms the results.

4.4 A Note on Plants with only Output Quantization (IQO)

Results have been obtained for the QIQO and QIO cases. It is now shown that the analysis of systems with only measurement quantization reduces to the QIO case.

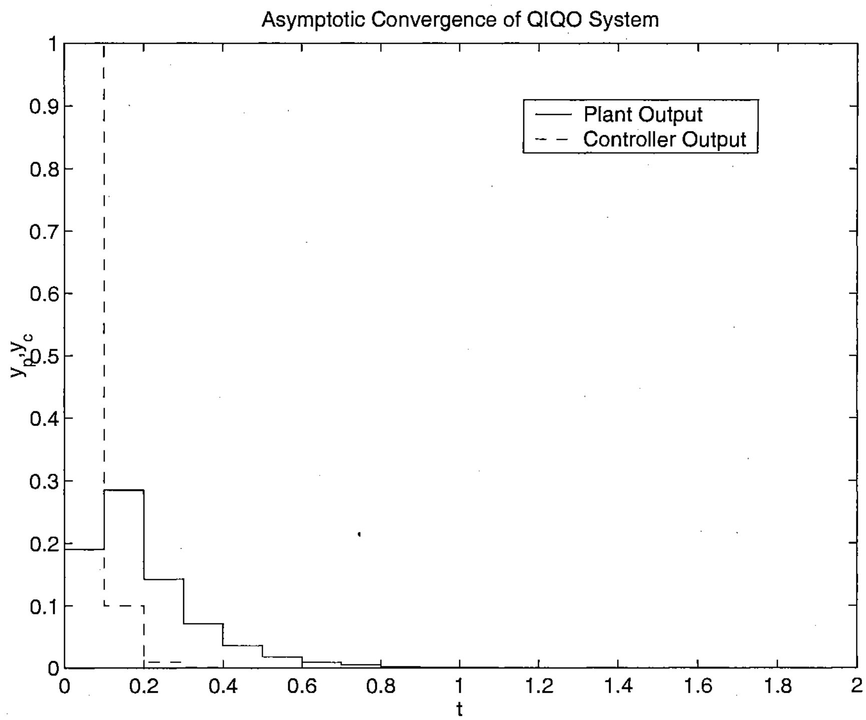

Figure 4.10: Asymptotic Convergence of QIQO System

4.4.1 Equilibrium Problem

The controller transfer function and the quantizer can be moved around the loop and placed at the plant input, since the quantizer is an odd function. This is depicted in Figure 4.11. The equilibrium problem reduces to a QIO case with the roles of G and H exchanged. Note that the QI0 condition for uniqueness of equilibrium is symmetric with respect to G and H.

4.4.2 Stability Problem

The closed-loop an be represented in state-space form as

$$\begin{cases} \left[\begin{array}{c} x_p(k+1) \\ \hline x_c(k+1) \end{array}\right] = \left[\begin{array}{c|c} A_p & B_pC_c \\ \hline 0 & A_c \end{array}\right] \left[\begin{array}{c} x_p(k) \\ \hline x_c(k) \end{array}\right] + \left[\begin{array}{c} B_pD_c \\ \hline B_c \end{array}\right] u_c(k) \\ y_p(k) = [C_p \mid 0] \left[\begin{array}{c} x_p(k) \\ \hline x_c(k) \end{array}\right] \\ u_c(k) = -\mathcal{Q}(y_p(k)) \end{cases} \tag{4.31}$$

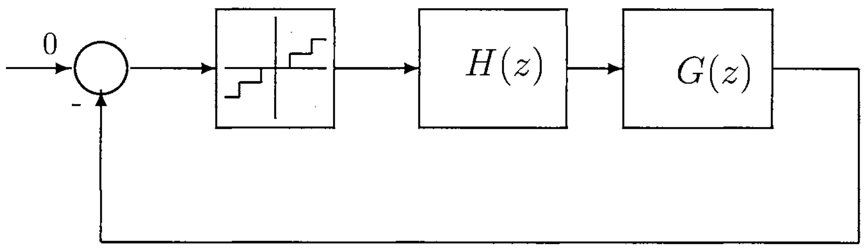

Figure 4.11: Reduction of IQO System to QIO Case

Therefore it reduces to the QIO case, which in turn reduced to the QIQM case, as described in Section 4.2.3.

Chapter 5

Conclusions and Topics for Future Research

5.1 Conclusions

Absolute Stability proves useful for analysis and design of systems with quantized feedback. The stability of full-state quantized feedback is straightforward to analyze when the states are precisely measured (QI systems). The concept of discrete positive-realness (DPR) is closely connected to the Absolute Stability analysis of discrete-time systems. In summary, a system with a sector-bounded nonlinearity on the feedback path is stable if an associated transfer function is DPR. For the design problem in QI systems, a parameterization of all gains which render the associated transfer function DPR was obtained. Using such parameterization, stable gains can be generated by choice of one matrix, one scalar which varies in $[0,1]$, and one vector parameter of fixed norm. For open-loop unstable systems, it is known that stabilization by state feedback is impossible due to the existence of a finite region of zero control surrounding the origin. A chattering solution to the problem is introduced by modifying the feedback law. The new feedback law effectively behaves as a relay in the neighborhood of the origin. For this case, the gains can be obtained from the solution of an algebraic Riccati equation. This solution lacks practical relevance due to its reliance on an infinitely fast switching of the control signal, similar to the one in mathematical sliding mode control.

The problem of quantized state feedback based on quantized measurements is much harder. The equilibrium equations cannot be solved in closed form. However, a graphical method that gives all solutions is developed, and is applicable to systems of any order. A constructive process is followed in order to arrive at a stability criterion. The criterion provides a sufficient condition for a system with a given feedback gain and quantization step size to be globally asymptotically stable. The stability criterion resembles the well-known Circle Criterion, but with a bound on the location of the polar plot which depends on the gain itself. More exactly, it depends on the 1-norm of the feedback gain. Since the location of the polar plot also depends on the gain, a practical design procedure may need to be iterative. It is shown that any nominally stabilizing gain can be scaled down until the system passes the stability test. However, this cannot be taken as a final design procedure, since any benefits attained by a nominal design will be lost after scaling of the gain. The parametric behavior of the system when the gain is scaled displays bifurcations. The sudden transition from asymptotic convergence to limit cycle, and from one limit cycle amplitude to another when the scaling factor is changed is evidenced by an example. Usage of the QIQM stability theorem is simple and direct, and is not necessarily conservative. An example shows that the critical scaling factor predicted by the theorem and the one obtained by simulation are in excellent agreement. A method for the computation of the least upper bound of t he multiplier has been completely specified and borrows concepts from Number Theory, in particular Diophantine equations. This bound is necessary to complete the stability evaluation of a given system using the theorem.

Systems with output feedback and quantization are analyzed in different ways, depending on whether quantization is present at the plant input, output or both. Systems with quantization only at plant input (QIO) reduce to the QI case in terms of equilibrium and stability analysis. Similarly, systems with quantization only at plant output (IQO) reduce to the QI case. Thus, a complete characterization of equiliria is possible, and the stability treatment is as general as the one given for the QI case.

Systems with quantization at both, plant input and output (QIQO) are more difficult to analyze. The equilibrium equations cannot be solved explicitly, but are reduced to a simpler system of 2 equations with rounding. A bound for the number of nontrivial solutions has been obtained, which, in turn, is used to provide a sufficient condition for the origin to be the only equilibrium point. The stability problem in QIQO systems was approached from three different angles. A multiplicative perturbation method analogous to the one used in QIQM systems fails due to difficulties in finding adequate multipliers. However, the MIMO version of the multiplicative perturbation theorem holds, and is proven in the appendix. A DPR method was attempted by realizing the closed-loop as a 2-input, 2-output system with sector nonlinearities in the feedback path. Since the outputs cross at the nonlinearities, the standard theories are not applicable. Removing the notion of sector and requiring more general properties of the feedback nonlinearities resulted in a very conservative criterion. The last method considered was the Small Gain Theorem, which gives less conservative results and is simple to apply.

Several Matlab programs have been written to automate the tasks performed in the various equilibrium and stability tests.

The author believes that the results obtained are a first step in developing a fairly general set of tools for the analysis of design of control systems with quantization.

5.2 Topics for Future Research

5.2.1 Systems with Quantization in the State Measurements

The closed-loop system

$$\begin{cases} x(k+1) = Ax + Bu \\ u(k) = -FQ(x) \end{cases}$$

has not been studied in this work. It is anticipated that the equilibrium problem does not pose difficulties. The stability problem, however, requires a more careful examination. Here, as in the QIQM case, the nonlinearity is many-to-one, therefore the standard Absolute Stability does not apply.

5.2.2 Design for Performance in QIQM Systems

A possible way to include performance constraints in QIQM systems is to study the relationship between the matrix conditions for DPR and the Riccati equations that arise in linear quadratic control. One known difficulty is that the problem is implicit, for the sector bound to be included in the matrix conditions depends on the feedback gain, more specifically, on its 1-norm. A search method may need to be considered.

5.2.3 Galois Field Theory and Number-Theoretical Transforms

Richalet ([30],[31]), Melikov [26], Vidal [34] and other researchers have explored the applicability of Galois field theory- in particular by defining number-theoretical transforms- to the analysis of discrete-time systems where the variables belong to a finite field and the operations of sum and multiplication take special definitions. For instance, in [30], an operational calculus is constructed for sequences that take binary values and difference equations whose sums and multiplications are those of Boolean algebra. A generalization of this calculus and the definition of the Laplace-Galois transform is presnted in [31], where the sequences take values in a more general Galois field. The starting point in defining a transform is to show how a sequence and a rational function with coefficients in the field can be put in one-to-one correspondence. Some familiar properties of the Z-transform, such as the conversion between powers of z and delays and the final and initial value theorem hold for the number-theoretical transform. Number-theoretical transforms defined for sequences and systems over finite fields is a natural setting for studying systems with quantization and over/underflow. In fact, contrary to the assumption that internal control computations happen at infinite precision and only the inputs and outputs are quantized, digital signal processors operate ultimately with elements of a finite set of integer numbers. Also, the common experience of a positive overflow quantity showing up as a negative underflow is reproduced by the addition and multiplication operations over a finite field. For example, in Boolean algebra, one has that $1 + 1 = 0$, which is interpreted as an overflow quantity showing up as the least element of the finite field. Application of these concepts to the analysis of digital control systems would require to model quantization and overflow in intermediate calculations. If a discrete system is modeled to include overflowing operations and quantization of intermediate results, the number-theoretical transform can be directly applied and yield, for example, the solution sequence when the input and the initial conditions are known. Note that in the finite field of definition, and using the appropriate concepts of sum and multiplication, systems having quantization and overflow appear as "linear".

Appendix A

Filippov Solutions of Continuous-Time QI Systems: Existence and Uniqueness

A.1 Preliminary Results

Some definitions and results from [13] and [14] are now presented. Let $f(t, x_1, x_2, ..x_n)$ be a vector function $f : \mathcal{R}^{n+1} \to \mathcal{R}^{n+1}$. Consider the system of equations

$$\frac{dx}{dt} = f(t, x) \tag{A.1}$$

Definition A.1. *A vector function $x(t)$ defined on the interval (t_1, t_2) is called a solution to Eq. (A.1) if:*

- *$x(t)$ is absolutely continuous*
- *For every neighborhood N_δ of $x(t)$ with radius δ, the vector $\frac{dx(t)}{dt}$ belongs to the smallest convex closed set of values achieved by $f(t, x')$, where x' varies over all of N_δ except a set of measure zero (i.e., the points of N_δ where f is discontinuous or undefined).*

Definition A.2 ([9]). *The function $Y : X \to \mathcal{R}$, with $X \subseteq \mathcal{R}$ is said to be summable if*

$$\sup_{A \in \mathcal{F}(X)} \sum_{t \in A} |Y(t)| < \infty$$

where $\mathcal{F}(X)$ denotes the set of all finite subsets of X. Observe that any bounded function is summable in the set where it is bounded.

Condition 2 (B in [13]). *The system of equations Eq. (A.1) is said to satisfy Condition B in a region Q of $\mathcal{R}^n$ if $f(x)$ is defined almost everywhere in Q* [1]*, is measurable, and for any closed bounded region $D \subseteq Q$ there exists a summable function $B(t)$ such that almost everywhere in D the following bound holds*

$$||f(t,x)||_2 \leq B(t)$$

Definition A.3. *A solution to Eq. (A.1) is said to be unique from the right in a region G if for any point $(t_o, x_o) \in G$ there is no other solution in G for $t \geq t_o$ satisfying the initial condition $x(t_o) = x_o$.*

Theorem 7 (Theorem 4 in [13]). *Let the right hand side of Eq. (A.1) be measurable in a region G and satisfy Condition B. Then for arbitrary initial conditions $x(t_o) = a$, where $(t_o, a) \in G$, a solution to Eq. (A.1) exists satisfying the initial conditions, and defined on the interval $[t_o - d,\ t_o + d]$ where d is such that the $(n+1)$ dimensional "cylinder"*

$$|t - t_o| \leq, \quad |x - a| \leq \left| \int_{t_o}^{t_o \pm d} B(t)dt \right|$$

is situated entirely inside G.

Lemma A.1 (Corollary of Theorem 10 in [13]). *A right-unique solution of Eq. (A.1) exists if i) Condition B is satisfied and, ii) for almost all (t,x) and (t,z) where $||x - z|| < \epsilon$, the following holds for some K constant and any $\epsilon > 0$:*

$$(x - z)^T (f(t,x) - f(t,z)) \leq K||x - z||_2^2$$

A.2 Existence and Uniqueness of Solutions of Continuous-Time QI Systems

A.2.1 Proof of Existence

Showing that a solution exists consists in identifying the right hand side f and showing that it satisfies Condition B. The system of differential equations is written as

$$\frac{dx}{dt} = Ax - B\overline{\mathcal{Q}}(Fx) = f(x)$$

where $\overline{\mathcal{Q}}$ denotes either the quantizer or the modified nonlinear operator $\mathcal{N}$ defined in Section (2.4), which have similar properties, as far the above results are concerned. Condition B is satisfied by

[1] i.e., in all of Q except a subset of Q of measure zero.

$f(x)$. In fact, $\mathcal{N}$ and $\mathcal{Q}$ are defined in all $\mathcal{R}^n$. Also, let $D \subseteq \mathcal{R}^n$ be a closed bounded region. Then there exists $m \geq 0$ such that $||x||_2 \leq m$ for all x in D. Using the induced 2-norm of A and the triangle inequality it follows that

$$f(x) = ||Ax - B\overline{\mathcal{Q}}(Fx)||_2 \leq ||Ax||_2 + ||B||_2|\overline{\mathcal{Q}}(Fx)| \leq ||A||_2 m + ||B||_2|\overline{\mathcal{Q}}(Fx)|$$

Note that

$$|\overline{\mathcal{Q}}(Fx)| = |Fx + \gamma|$$

with $|\gamma| \leq \frac{q}{2}$. Also, by Hölder's inequality,

$$|Fx| \leq ||F||_2||x||_2$$

Therefore it follows that

$$f(x) \leq ||A||_2 m + ||B||_2(||F||_2 m + \frac{q}{2}) = B$$

The function $B(t)$ is actually constant, and therefore summable. The measurability assumption is verified by the fact that $f(x)$ is integrable.

A.2.2 Considerations of Uniqueness

The inequality of Lemma (A.1) is now examined. Substituting the expression for f in the left-hand side and expanding:

$$\begin{aligned}(x-z)^T(f(x)-f(z)) = (x-z)^T A(x-z) - (x-z)^T B(\overline{\mathcal{Q}}(Fx) - \overline{\mathcal{Q}}(Fz)) &= \\ |(x-z)^T A||(x-z)| - (x-z)^T B(Fx + \gamma_1 - Fz - \gamma_2) &\leq \\ ||(x-z)||_2^2(||A||_2 - ||BF||_2) + |(x-z)^T B|q &\leq \\ ||(x-z)||_2^2(||A||_2 - ||BF||_2) + ||(x-z)||_2||B||_2 q &\end{aligned}$$

Therefore the following inequality needs to be satisfied

$$||(x-z)||_2^2(||A||_2 - ||BF||_2 - K) + ||(x-z)||_2||B||_2 q \leq 0$$

for some constant K, with $||x - z|| < \epsilon$. Clearly, the above inequality cannot be satisfied since it contains a term linear in $||(x - z)||_2$. Therefore more than one solution may exist for a given initial condition.

Appendix B

Discrete Positive-Real Transfer Functions and Discrete Popov's Criterion

B.1 Positive-Real Transfer Matrices

The background theory presented in this appendix is entirely based on B.D.O. Anderson's papers [17],[2]. Theorem 8 is extended in Section 3.5 to allow for systems with a direct transmission term D. Let a continuous-time LTI system be represented by the state equations

$$\begin{cases} \dot{x} = Ax + Bu \\ y = Cx + Du \end{cases}$$

The transfer matrix

$$G(s) = C(sI - A)^{-1}B + D$$

is said to be *positive-real* (p.r.) if it satisfies the following conditions:

- The elements of $G(s)$ are analytic in $\mathrm{Re}(s) > 0$
- $G(s) + G^*(s) \geq 0$ in $\mathrm{Re}(s) > 0$

The above conditions are equivalent to the following three conditions together:

- $G(s)$ has simple poles on the imaginary axis

- For any pole of $G(s)$ with zero real part, the residue matrix is positive semidefinite Hermitian
- $G(jw) + G^T(-jw) \geq 0$ for all $w \in \mathcal{R}$ for which jw is not a pole of $G(s)$

Furthermore, the following lemma provides a necessary and sufficient condition for a transfer matrix to be p.r.:

Lemma B.1. *A square transfer matrix $G(s)$ with elements analytic in $Re(s) > 0$, simple poles on the imaginary axis and finite $G(\infty)$ is positive real if, and only if, there exists a real symmetric positive definite matrix P and real matrices L and W such that*

$$\begin{aligned} PA + A^T P &= -LL^T \\ B^T P &= C - W^T L^T \\ W^T W &= D + D^T \end{aligned} \tag{B.1}$$

Note that the analycity requirement on G forbids RHP poles, but allows a pole at the origin.

B.2 Discrete Positive-Real Transfer Matrices

Let a discrete-time LTI system be represented by the state equations

$$\begin{cases} x(k) = Ax(k) + Buk \\ y = Cx(k) + Du(k) \end{cases}$$

The transfer matrix

$$G(z) = C(zI - A)^{-1}B + D$$

is said to be *discrete positive-real* (d.p.r.) if it satisfies the following conditions:

- The elements of $G(z)$ are analytic in $|z| > 1$
- $G(z) + G^*(z) \geq 0$ in $|z| > 1$

The above conditions are equivalent to the following three conditions together:

- $G(z)$ has simple poles on the unit circle
- For any pole z_0 of $G(z)$ on the unit circle, the matrix $Q = \bar{z}_0 K$ is positive semidefinite Hermitian, where K is the residue matrix of $G(z)$ at $z = z_0$.

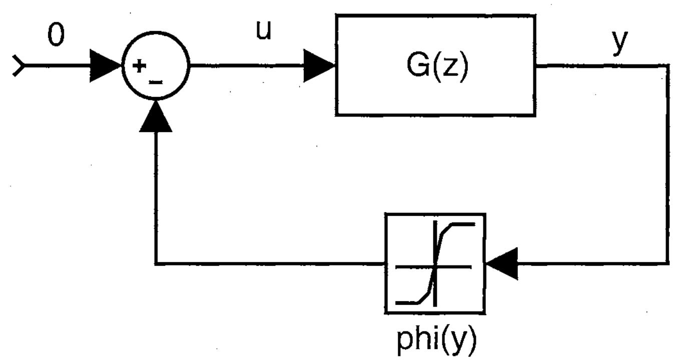

Figure B.1: Discrete Luré problem

- $G(e^{jw}) + G^T(e^{-jw}) \geq 0$ for all $w \in \mathcal{R}$ for which e^{jw} is not a pole of $G(z)$

As in the continuous case, a lemma providing a necessary and sufficient condition for a transfer matrix to be d.p.r. is given:

Lemma B.2. *The square transfer matrix $G(z)$ with no poles in $|z| > 1$ and simple poles unit circle is positive real if, and only if, there exists a real symmetric positive definite matrix P and real matrices L and W such that*

$$\begin{aligned} A^T PA - P &= -LL^T \\ B^T PA &= C - W^T L^T \\ W^T W &= D + D^T - B^T PB \end{aligned} \tag{B.2}$$

B.3 Discrete Popov's Criterion

Consider a discrete version of the Luré problem, represented in Figure B.1 The system can be described by the following equations

$$\begin{cases} x(k+1) = Ax(k) + Bu(k) \\ y(k) = Cx(k) \\ u(k) = -\phi(y(k)) \end{cases} \tag{B.3}$$

The transfer function of the linear part of the system is

$$G(z) = C(zI - A)^{-1}B$$

For the remaining presentation, let us restrict the attention to the SISO case. Also, note that the theory presented is applicable when A has no poles on or outside the unit circle, and when $G(z)$

contains no zero-pole cancellations. Let the nonlinear function $\phi(.)$ be of the "sector" type, that is, restricted to belong to the first and third quadrants. The nonlinearity can be time-varying, but it is such that $\phi \in [0,\, k)$. Mathematically this is stated as

$$0 \leq \phi(y)y < \overline{K}y^2$$

The following theorem provides a sufficient condition for the closed-loop system to be globally uniformly stable or uniformly asymptotically in the large (UASIL):

Theorem 8. *Define $H(z) = C(zI - A)^{-1}B + \frac{1}{\overline{K}}$. Then, if $H(z)$ is discrete positive real, the system in Eq.(B.3) is uniformly stable in the large. If the sector nonlinearity is in addition time-invariant, $\phi(y) = 0$ only for $y = 0$ and the system is observable, UASIL is obtained.*

The DPR property of $H(z)$ can be tested by

$$\mathrm{Re}\left\{G(e^{jw})\right\} + \frac{1}{\overline{K}} \geq 0,\ \forall w \in \mathcal{R}$$

Graphically, the polar plot of $G(z)$ must lie to the right of a vertical line that passes through (-1/$\overline{K}$,0), resembling the well-known Circle Criterion.

B.3.1 Discrete Strict Positive-Realness

The familiar concept of SPR transfer functions has a discrete counterpart, called DSPR.
A square transfer matrix $G(z)$ is called *discrete strictly positive-real* if $G(z)$ is asymptotically stable and

$$G(e^{jw}) + G^T(e^{-jw}) > 0, w \in [0,\, 2\pi]$$

In the SISO case, this reduces to [10]

$$\mathrm{Re}(G(e^{jw})) > 0,\ \forall w \in \mathcal{R}$$

Graphically, the polar plot of $G(z)$ must lie in the right half of the complex plane, without touching the imaginary axis.

B.3.2 Matrix conditions for Discrete Positive Realness

Theorem 8 states that if the transfer function $H(z)$ is DPR, the closed loop system is stable. The constant $\frac{1}{\overline{K}}$ plays the role of matrix D in the realization (A, B, C, D) of $G(z)$ of Lemma B.2. Therefore, the matrix equations given earlier in (B.2) provide sufficient conditions for stability of the closed loop system. The closed-loop system is UASIL if $\phi(.)$ is time-invariant and there exist matrices $P = P^T > 0$, L and W such that

$$\begin{aligned} A^T P A - P &= -LL^T \\ B^T P A &= C - W^T L^T \\ W^T W &= \frac{2}{\overline{K}} - B^T P B \end{aligned} \tag{B.4}$$

Appendix C

Derivation of the Distance between Contour Lines

Let the sets S_1 and S_2 be defined by

$$\begin{aligned} S_1 &= \left\{ x \in \mathcal{R}^n \,\middle|\, Fx = c_1 \right\} \\ S_2 &= \left\{ x \in \mathcal{R}^n \,\middle|\, Fx = c_2 \right\} \end{aligned}$$

Define the distance between contour lines by

$$\Delta = \min_{x_2 \in S_2} ||x_2 - x_1||_2 \qquad \text{for some } x_1 \in S_1$$

Although not formally proved here, it is clear that the distance is independent of x_1 for any pair of contour sets, since they are hyperplanes. Multiply the definition of Δ by $||F||_2$ on both sides. Then,

$$||F||_2 \Delta = ||F||_2 \min_{x_2 \in S_2} ||x_2 - x_1||_2 = ||F||_2 ||x_2^* - x_1||_2$$

where x_2^* is the element of S_2 for which the minimum is achieved. Now, by Hölder's inequality:

$$||F||_2 ||x_2^* - x_1||_2 \geq |F(x_2^* - x_1)|$$

that is,

$$||F||_2 ||x_2^* - x_1||_2 \geq |c_2 - c_1|$$

If we take

$$||F||_2 \Delta = |c_2 - c_1|$$

then

$$\|F\|_2\Delta \leq \|F\|_2\|x_2^* - x_1\|_2 = \|F\|_2 \min_{x_2 \in S_2} \|x_2 - x_1\|_2$$

It follows that

$$\Delta = \frac{|c_2 - c_1|}{\|F\|_2}$$

Appendix D

Derivation of the Extremal Values of Fx in Ω_z

The problem is to determine

$$F_z^+ = \sup_{x \in \Omega_z} Fx$$

and

$$F_z^- = \inf_{x \in \Omega_z} Fx$$

where

$$\Omega_z = \{x \in \mathcal{R}^n \mid \mathcal{Q}(x) = z\}$$

The components of x in Ω satisfy the fundamental quantization inequality

$$z_i - \frac{q}{2} < x_i < z_i + \frac{q}{2}$$

where, depending on the sign of z_i, some of the inequality signs might allow equality. Since a supremum is sought, it is not necessary to make distinctions. Since Fx is a linear function of x, the extreme values occur at the boundary of Ω_z. In fact, to maximize Fx we choose $x_i = z_i + \frac{q}{2}$ when $f_i \geq 0$ and $x_i = z_i - \frac{q}{2}$ when $f_i < 0$. That is, the maximizing x is

$$x_i^* = z_i + \frac{q}{2}\text{sgn}(f_i)$$

Now,

$$F_z^+ = \sum_{i=1}^{n} f_i(z_i + \text{sgn}(f_i)\frac{q}{2})$$

using the definition of 1-norm:

$$||F||_1 = \sum_{i=1}^{n} |f_i|$$

and the identity

$$|a| = a \operatorname{sgn}(a)$$

it follows that

$$F_z^+ = Fz + \frac{q}{2}||F||_1$$

Similar arguments lead to the formula for the infimum:

$$F_z^- = Fz - \frac{q}{2}||F||_1$$

Appendix E

Stability of DPR MIMO Systems with Direct Transmission Term

The following is a generalization of Lemma 3.2 for square MIMO systems. Let a discrete time system be represented by the state equations

$$x(k+1) = Ax(k) + Bu(k) \tag{E.1}$$

$$y(k) = Cx(k) + Du(k) \tag{E.2}$$

Here $x(k)$, $y(k)$ and $u(k)$ are n, m and m-dimensional column vectors, respectively; and matrices A,B,C and $D = D^T$ have appropriate dimensions. Let the control vector be defined by

$$u(k) = \begin{bmatrix} -\Phi_1(y_1(k)) \\ -\Phi_2(y_2(k)) \\ . \\ . \\ -\Phi_m(y_m(k)) \end{bmatrix} \tag{E.3}$$

where the nonlinear functions satisfy

$$\Phi_i \in \mathcal{S}(0, k_i),\ i = 1, 2..m \tag{E.4}$$

Define the m-by-m matrix K as

$$K = \text{diag}(k_1, k_2.., k_m)$$

Lemma E.1. *The discrete system described by Eq. (E.2) with* $D = D^T$ *is globally stable about the origin when the controls (E.3) are applied if the transfer function*

$$J(z) = C(zI - A)^{-1}B + D + K^{-1}$$

is discrete positive real.

Proof. Consider the Lyapunov function

$$V(k) = x^T(k)Px(k)$$

Following a procedure similar to the proof of Lemma 3.2, D cancels out and the change of the Lyapunov function along system trajectories is found to be

$$\Delta V = -[L^T x - W\Phi]^T[L^T x - W\Phi] - 2[\Phi^T y - \Phi^T K^{-1}\Phi]$$

where $\Phi = -u$. The first term is clearly negative semidefinite. Negative semidefiniteness of the second term follows directly from the sector conditions (E.4). Therefore the change in $V(k)$ is negative semidefinite and the closed loop is globally stable about the origin. Asymptotic stability requires further assumptions on the behavior of the nonlinearity. ■

Appendix F

Multiplicative Perturbation Theorem for DPR MIMO Systems

Theorem 9. *Let a discrete system be represented by Eq. (E.2) with* $D = D^T$. *Let* $\alpha_i : \mathcal{R}^n \rightarrow \mathcal{R}$ *for* $i = 1, 2, ..m$ *be mappings such that* $\exists\ \overline{\alpha}_i$ *finite satisfying* $0 \leq \alpha_i(x) < \overline{\alpha}_i$ *for all* $x \in \mathcal{R}^n$. *Let* $\mathcal{N}_i : \mathcal{R} \rightarrow \mathcal{R}$ *be sector nonlinearities* $\mathcal{N}_i \in \mathcal{S}[0, \bar{n}_i]$. *Let the m-by-m matrix K be defined as*

$$K = diag(k_1, k_2.., k_m)$$

Then if the transfer matrix

$$J(z) = C(zI - A)^{-1}B + D + K^{-1}$$

is DPR, and $\overline{\alpha}_i\overline{n}_i < \overline{k}_i$, *al this for* $i = 1, 2, ..m$, *then the closed-loop system formed by applying the feedback*

$$\begin{aligned} u_1(k) &= -\mathcal{N}_1[\alpha_1(x(k))y_1(k)] \\ u_2(k) &= -\mathcal{N}_2[\alpha_2(x(k))y_2(k)] \\ &. \qquad\qquad (F.1) \\ &. \qquad\qquad (F.2) \\ u_m(k) &= -\mathcal{N}_m[\alpha_m(x(k))y_m(k)] \end{aligned}$$

is stable in the large.

Proof. By hypothesis, $J(z)$ is DPR.Then by Lemma B.2, there exists a real symmetric positive definite matrix P and real matrices L and W such that

$$\begin{aligned} A^T PA - P &= -LL^T \\ B^T PA &= C - W^T L^T \\ W^T W &= 2D + 2K^{-1} - B^T PB \end{aligned}$$

Consider the quadratic Lyapunov function $V(x(k)) = x^T(k)Px(k)$. The change of the function along the the equations of the closed-loop system is, dropping index k from the notation:

$$\Delta V(x(k)) = V(x(k+1)) - V(x(k)) = [x^T A^T - B^T \mathcal{N}(\alpha(x)y))]P[Ax - B\mathcal{N}(\alpha(x)y)] - x^T Px$$

where $\mathcal{N}(\alpha(x)y)$ is the column vector with components $\mathcal{N}_i(\alpha_i y_i)$. Performing operations and incorporating the matrix equations, D cancels out and the following expression is obtained

$$\Delta V(x(k)) = -[L^T x - W\mathcal{N}(\alpha(x)y)]^T[L^T x - W\mathcal{N}(\alpha(x)y)] - 2[\mathcal{N}^T(\alpha(x)y)y - \mathcal{N}^T(\alpha(x)y)K^{-1}\mathcal{N}(\alpha(x)y)] \tag{F.3}$$

The first term is clearly negative semidefinite. The second term can be examined as follows. If $\alpha_i(x)y_i > 0$ then by the sector condition on $\mathcal{N}_i$ and the sector inclusion inequality of the hypothesis it follows that

$$0 \leq \mathcal{N}_i(\alpha_i(x)y_i) \leq \bar{n}_i\alpha_i(x)y_i \leq \bar{n}_i\bar{\alpha}_i y_i < \overline{k}_i y_i$$

This implies

$$\mathcal{N}(\alpha_i(x)y_i)y_i > \mathcal{N}^2(\alpha_i(x)y_i)k_i^{-1} \tag{F.4}$$

The second term is thus negative semidefinite, since

$$\mathcal{N}^T(\alpha(x)y)y - \mathcal{N}^T(\alpha(x)y)K^{-1}\mathcal{N}(\alpha(x)y) = \sum_{i=1}^{m} \mathcal{N}(\alpha_i(x)y_i)y_i - \mathcal{N}^2(\alpha_i(x)y_i)k_i^{-1}$$

and each term of the sum is negative semidefinite, from Eq. (F.4). If $\alpha(x)y < 0$ the above chain of inequalities is reversed, so the second term is negative semidefinite, being zero if $\mathcal{N} = 0$ or $y = 0$, thus Lyapunov stability follows. In order to prove asymptotic stability, further assumptions on the local behavior of $\mathcal{N}$ and system observability might be required. ■

Appendix G

Negative Semidefinite Conditions for $\overline{\Psi}$

Let

$$\overline{\Psi}(y_p, y_c) = f_2 \mathcal{Q}^2(y_p) + f_1 \mathcal{Q}^2(y_c) + \mathcal{Q}(y_p) y_c - \mathcal{Q}(y_c) y_p$$

Divide the y_p y_c plane into square quantization regions. Within a quantization region, $\overline{\Psi}$ is linear, so the supremum of its values occurs at the boundaries of the region. Consider a generic quantization region centered at (iq, jq), where q is the quantization stepsize and i, j are integers. Then

$$\overline{\Psi}(i,j) = f_2 i^2 q^2 + f_1 j^2 q^2 + iqy_c - jqy_p$$

1) Suppose $i, j \neq 0$

Then the supremum occurs for $y_c = jq + q/2$ and $y_p = iq - q/2$. Substituting:

$$\frac{1}{q^2} \overline{\Psi}_{max}(i,j) = f_2(i^2 + \frac{i}{2}) + f_1(j^2 + \frac{j}{2})$$

Clearly, the above function is negative semidefinite for $f_1 < 0$ and $f_2 < 0$, being zero only when $i = j = 0$.

2) Suppose $i = 0$, $j \neq 0$

Then $\overline{\Psi}(i,j) = f_1 j^2 q^2 - jqy_p$, and the supremum occurs at $y_p = -q/2$. Substituting:

$$\frac{1}{q^2} \overline{\Psi}_{max}(j) = f_1 j^2 + \frac{j}{2}$$

For this to be negative for all integers $j \neq 0$, with $f_1 < 0$ it is required that $f_1 < -1/2$. This also guarantees that $\overline{\Psi}$ is not zero in those regions.

3) Suppose $j = 0$, $i \neq 0$

Then $\overline{\Psi}(i,j) = f_2 i^2 q^2 + jqy_c$, and the supremum occurs at $y_c = q/2$. Substituting:

$$\frac{1}{q^2}\overline{\Psi}_{max}(i) = f_2 i^2 + \frac{i}{2}$$

For this to be negative for all integers $i \neq 0$, with $f_2 < 0$ it is required that $f_2 < -1/2$. This also guarantees that $\overline{\Psi}$ is not zero in those regions.

4) Suppose $i = j = 0$

This is the only case for which the function is zero when $f_1 < -1/2$ and $f_2 < -1/2$.

Summarizing, $\overline{\Psi}$ is negative semidefinite for $f_1 < -1/2$ and $f_2 < -1/2$, being zero only when $|y_c| < q/2$ and $|y_p| < q/2$.

Appendix H

Matlab Programs

H.1 GCD of vector components

```
%vgcd.m
%This calculates the G.C.D. of a vector
function num=vgcd(F)
num=F(1);
for i=1:max(size(F)),
   num=gcd(num,F(i));
end;
```

H.2 LCM of vector components

```
%vlcm.m
%This calculates the l.c.m. of a vector

function num=vlcm(F)
num=F(1);
for i=1:max(size(F)),
   num=lcm(num,F(i));
end;
```

H.3 Infimum for multiplier bound calculation

```
%infval.m
%Calculates Fq, the infimum of the set
%
%           Gamma(F)={Fz|z in Z_q^n and Fz>=q/2}
%
%given vector F and quantization stepsize q.
%Calls vgcd (greatest common divisor of vector elements) and vlcm
%(least common multiple of vector components)

%Hanz Richter,  Fall 2001

function Fq=infval(F,q);

%Limit the precision of F to a few decimals
%to prevent overflow problems when calculating the least
%common multiple of the denominator of the fractional
%representation of the elements of F.

decimals=4;
F=round(F*10^decimals)/10^decimals;

%Convert F to the form: real*integer_vector

[num,den]=rat(F);

sigm=1/vlcm(den); %real and positive(rat sticks the sign
%in the numerator)

Fint=abs(num)*vlcm(den)./den;  %integer_vector

alph=sigm*vgcd(Fint);
```

```
n=ceil(0.5/alph);

Fq=q*n*alph;
```

H.4 QIQM stability check

```
%qiqm_check.m
%Tests is a given discrete-time system passes the sufficient
%condition for global asymptotic stability when the loop is
%closed using quantized input and state measurements.

%Syntax: m=qiqm_check(A,B,F,q)
%A,B are state space matrices (A must have eigenvalues in unit circle)
%F is the state feedback gain, to be used in the control law
%u=-Q(FQ(x)), where Q is the quantization operator with step size q.
%The stability margin (distance between min(Re(G(e^jw)) and the
%qiqm stability limit) is returned in m.
%If invoked without left hand argument, the command plots the
%Nyquist diagram and stability limit.

%Hanz Richter, 2001

function [m]=qiqm_check(A,B,F,q)

if max(abs(eig(A-B*F)))>=1
   disp('Closed-loop system is not nominally stable');
end;

[re,im]=dnyquist(A,B,F,0,1);
minpoint=min(re);
Fq=infval(F,q);
limit=0.25*norm(F,1)*(q/quant(Fq,q)+1)-Fq/(2*quant(Fq,q))-0.25;
```

```
if nargout==1
   m=minpoint-limit;
   if m>0
      disp('Test passed. Closed loop system globally asymptotically stable');
   else
      disp('Test failed. Closed loop system may not be globally asymptotically stable');
   end;
else
   maxim=max(abs(im)); %for plot scaling
   plot(re,im,'b',re,-im,'b');
   hold on;
   plot([limit limit],[-maxim maxim],'r--');
   xlabel('Re(G(e^jw))');
   ylabel('Im(G(e^jw))');
   hold off
end;
```

H.5 QIQM equilibrium check

```
%qiqm_equil.m
%Performs a graphical test for multiplicity of equilibrium points
%in systems with input and state measurement quantization.
%Syntax: qiqm_equil(A,B,F,q,X)
%A,B are discrete system matrices, F is the state feedback gain used
%in the control law u=-Q(F(Q(x)), where Q is the quantization operator
%with stepsize q.
%X is the horizontal range used for plotting in the format
%[Xstart:Xstep:Xend]
%The index of the G=(A-I)^{-1} used in the solution is returned in
%base_index. This can be used to find all solutions to the vector
%equation.

%Hanz Richter, 2001
```

```
function [base_index]=qiqm_equil(A,B,F,q,X)

n=max(size(A));
if det(A-eye(n))==0
   disp('A-I is singular. Equilibrium set is dense');
else
   G=inv(A-eye(n))*B;

   %choose the first nonzero element of G
   base_g=G(1);
   i=1;
   while (base_g==0) & (i<=n-1)
      i=i+1;
      base_g=G(i);
   end;

   base_index=i;

   for i=1:max(size(X)),
      staircase(i)=quant(F*quant(X(i)*G/base_g,q),q);
   end;

   plot(X,staircase,X,X/base_g,'r');
   xlabel('x');
   ylabel('Quantization function and x/g_1')
end;
```

H.6 Gain scaling in QIQM

```
%scale_F.m

%This function helps in synthesizing stable gains for QIQM
```

```
%Syntax: sigmac=scale_f(A,B,Fo,q)
%A,B are discrete system matrices. Fo is the initial gain, which must be
%nominally stabilizing. q is the quantization stepsize. The critical scaling
%factor is returned in sigmac. This is done assuming Fq=q/2.
%If the function is invoked without left-hand argument, the loci of the
%minimum of the real part of the Nyquist plot and the actual stability limit
%(using the full expression for Fq computed by infval) are plotted.

%Hanz Richter, August 22 2001

function sigmac = scale_F(A,B,Fo,q)

%Check validity of initial gain:
if max(abs(eig(A-B*Fo)))>=1
   disp('Initial gain not nominally stabilizing')
elseif nargout==1
   [re,im]=dnyquist(A,B,Fo,0,1);
   %Compute critical scaling
   sigmac=1/(norm(Fo,1)-2*min(re));
else
   [re,im]=dnyquist(A,B,Fo,0,1);
   %Compute critical factor for plot scaling purposes
   sigmac=1/(norm(Fo,1)-2*min(re));
   %Show up to 200 percent of the critical
   sigmas=[0.01:0.01:sigmac*2];
   %Line corresponding to Fq=q/2:
   line1=sigmas*norm(Fo,1)/2-1/2;
   %Curve corresponding to actual Fq:
   for i=1:max(size(sigmas)),
      inf=infval(sigmas(i)*Fo,q);
      line2(i)=.25*norm(sigmas(i)*Fo,1)*(q/quant(inf,q)+1)-0.5*inf/quant(inf,q)-0.25;
   end;
   %Plot:
```

```
   plot(sigmas,sigmas*min(re),'b',sigmas,line1,'k',sigmas,line2,'r');
   xlabel('Scaling factor \sigma');
   ylabel('Nyquist minimum and stability limit')
   title('Gain scaling plot');
   legend('Nyquist minimum','Stab.lim.(Fq=q/2)','Stab.lim(True Fq)')
end;
```

H.7 QIQM Example - Stability

```
%endlich.m

%Sampling rate

T=0.1;

%Hurwitz controllable system:

A=[-1/2 -1/4;1 3/4];
B=[1;2];

%Set of nominally stabilizing gains

Fset={[0.4 0],[0 0.4],[0.4 0.4],[-0.1 -0.1],[-0.1 0],[-0.1 0.1],[0 -0.1],[0 0.1],
[0.1 -0.1],[0.1 0],[0.1 0.1]};

Fq={0.8 0.8 0.8 0.5,0.5,0.5,0.5,0.5,0.5,0.5,0.5};

for i=1:11,
    i
    F=Fset{i};
    limit(i)=0.25*norm(F,1)*(1/round(Fq{i})+1)-Fq{i}/(2*round(Fq{i}))-0.25;
    subplot(2,1,1)
    plot([limit(i) limit(i)],[-1 1],'r--');
```

```
    hold on
    dnyquist(A,B,F,0,1);
    hold on
    pause


%RUN SIMULATIONS HERE

    subplot(2,1,2);
    stairs(t,x(:,1));
    hold on
    stairs(t,x(:,2));
    hold off
    clear x
    clear t
    pause
end;
******************************************************************************
*****************CHAOTIC RESPONSE EXAMPLE************************************

%chaos.m

%A chaotic sytem
%HD model from Franklin, Powell & Workman, p. 729, 1992.

A=[1 5.3316;0 0.9993];
B=[0.0133;0.0050];

T=66e-5;
p=[-1000*0.7+1000*sqrt(1-0.49)*j -1000*0.7-1000*sqrt(1-0.49)*j];
zp=exp(p*T);

F=place(A,B,zp);
```

```
abs(eig(A-B*F))
```

H.8 Bifurcations in QIQM

```
%bifurc_quant.m
%This plots a bifurcation diagram for a QIQM system
%                x(k+1)=Ax(k)-BQ(FQ(x))
%Hanz Richter, 2001

A=[-1/2 -1/4;1 3/4];
B=[1;2];
F=[0.4 0];
q=1;

plotiter=20;
skipiter=20;

for sigma=0.5:0.001:1.2,
   x=[40;50];
   for k=1:skipiter,
      x=A*x-B*quant(sigma*F*quant(x,q),q);
   end;
   for k=1:plotiter,
      x=A*x-B*quant(sigma*F*quant(x,q),q);
      plot(sigma,x(1),'*','markersize',5);
      hold on;
   end;
end;
```

H.9 Equilibrium finding in QIQO

```
%solve_qiqo
%This function solves the system of scalar quantization equations
```

```
%
%         round(fi)=j
%          round(gj)=i
%
%for j,i>=1 and f,g>0. The theoretical maximum number of solutions
%must be provided and is used as a search stop criterion.
%Syntax:
%[sols]=solve_qiqo(f,g,N)
%A vector with the solutions for i is returned in sols.
%Hanz Richter, 2001


function [sols]=solve_qiqo(f,g,N)
sols=[];
k=0;
if (f<=0 | g<=0)
   disp('Please rewrite the equation to have positive constants');
else
   for i=1:N,
      if round(g*round(f*i))==i
         k=k+1;
         sols(k)=i;
      end;
   end;
end;
```

Bibliography

[1] M.A. Aizerman and F.R. Gantmacher. *Absolute Stability of Regulator Systems.* Holden-Day, 1964.

[2] B.D.O. Anderson. A system theory criterion for positive real matrices. SIAM Journal on Control, 5(2):171–182, 1967.

[3] P.H. Bauer and L.J. Leclerc. Bounded-input-bounded-output properties of nonlinear discrete difference equations - applications to digital filters. IEEE Transactions on Automatic Control, 39(8):662–666, 1992.

[4] V. Birkhoff and D. MacLane. *A survey of modern algebra.* Macmillan, 1965.

[5] T. Bose. Asymptotic stability of two-dimensional digital filters under quantization. IEEE Transactions on Signal Processing, 42(5):1172–1177, 1994.

[6] R. Brockett and D. Liberzon. Quantized feedback stabilization of linear systems. IEEE Transactions on Automatic Control, 45(7):1279–1289, 2000.

[7] V. Chellaboina and W. Haddad. Stability margins of discrete-time nonlinear nonquadratic optimal regulators. In *Proc. 37th IEEE Conference on Decision and Control,* volume TA06, pages 1786–1791, 1998.

[8] S. Chen and J. Wu. Optimizing stability bounds of finite-precision pid controllers. IEEE Transactions on Automatic Control, 44(11):2149–2153, 1999.

[9] C. Constantinescu, W. Filter, K. Weber, and A. Sontag. *Advanced Integration Theory.* Kluwer Academic Publishers, 1998.

[10] J.J. Da Cruz and J.C. Geromel. Frequency-domain approach to the absolute stability analysis of discrete-time linear-quadratic regulators. IEE Proceedings, 137D(2):104–106, 1990.

[11] M. Dahleh and I. Diaz-Bobillo. *Control of Uncertain Systems: A Linear Programming Approach.* Prentice Hall, 1994.

[12] D. Delchamps. Stabilizing a linear system with quantized state feedback. IEEE Transactions on Automatic Control, 35(8):916–924, 1990.

[13] A.F. Filippov. Differential equations with discontinuous right hand side. American Mathematical Society Translations, 2(42):199–231, 1964.

[14] A.F. Filippov. *Differential Equations with Discontinuous Righthand Sides.* Kluwer Academic Publishers, 1988.

[15] G. Franklin, D. Powell, and M. Workman. *Digital Control of Dynamic Systems.* Addison-Wesley, 1992.

[16] R. García-Cobián. Private communication. 2001.

[17] B.E Hitz and B.D.O. Anderson. Discrete positive-real functions and their application to system stability. IEE Proceedings, 116(1):153–155, 1969.

[18] J. Hsu and A. Meyer. *Modern Control : Principles and Applications.* Mc.Graw-Hill, 1968.

[19] R.H. Istepanian and S. Chen. Optimal finite-precision controller realization of sampled-data systems. Int. Journal of Systems Science, 31(4):429–438, 2000.

[20] V. Lakshmikantham and D. Trigiante. *Theory of Difference Equations: Numerical Methods and Applications*, volume 181 of *Mathematics in Science and Engineering.* Academic Press, Inc., Boston, 1988.

[21] L.J. Leclerc and P.H. Bauer. A computer-aided test for global asymptotic stability of digital filters. In *International Conference on Acoustics, Speech, and Signal Processing*, volume 3, pages 1881–1884, 1991.

[22] T. Lee and S. Lee. Iee synthesis of control laws with specified gain and phase margin and poles region. IEE Proceedings, 135D(6):474–478, 1988.

[23] G. Li and M. Gevers. Optimal finite precision implementation of a state-estimate feedback controller. IEEE Transactions on Circuits and Systems, 37(12):1487–1498, 1990.

[24] G. Li and M. Gevers. Comparative study of finite wordlength effects in shift and delta operator parameterizations. IEEE Transactions on Automatic Control, 38(5):803–807, 1993.

[25] K. Liu and R. Skelton. Optimal controllers for finite wordlength implementation. IEEE Transactions on Automatic Control, 37(9):1294–1304, 1992.

[26] G.Y. Melikov and R.G. Faradzhev. An algrithm for calculation of the inverse laplace-galois transform. Automation and Remote Control, (1):74–80, 1984.

[27] R. Middleton and G. Goodwin. Improved finite word length charachteristics in digital control using delta operators. IEEE Transactions on Automatic Control, 31(11):1015–1021, 1986.

[28] R.A. Mollin. *Fundamantal Number Theory with Applications.* CRC Press, 1998.

[29] T. Munakata and K. Samsak. A method of minimal-time control to a quantized linear discrete system with rational coefficients. Journal of the Japan Society of Mechanical Engineering, 42C(2):323–327, 1998.

[30] J. Richalet. Calcul operationnel booleen. L'Onde Electrique, (439):1003–1012, 1963.

[31] J. Richalet. Operational calculus for finite rings. IEEE Transactions on Circuit Theory, CT-12(4):558–570, 1965.

[32] J.J. Slotine and W. Li. *Applied Nonlinear Control.* Prentice-Hall, 1991.

[33] T. Song and E. Collins. Improved closed-loop stability for fixed-point controller realizations using the delta operator. Int. Journal of Robust and Nonlinear Control, 11(1):41–57, 1999.

[34] P. Vidal. *Non-linear Sampled-Data Systems.* Gordon and Breach, 1969.

[35] M. Vidyasagar. *Nonlinear Systems Analysis.* Prentice Hall, 1993.

[36] B. Widrow and I. Kollar. Statistical theory of quantization. IEEE Transactions on Automatic Control, 45(2):353–361, 1996.

2

VITA

Hanz Richter

Candidate for the Degree of

Doctor of Philosophy

Thesis: STABILITY AND EQUILIBRIA OF LINEAR CONTROL SYSTEMS UNDER INPUT AND MEASUREMENT QUANTIZATION

Major Field: Mechanical Engineering

Biographical:

Personal Information: Born in Providence, Rhode Island, on December 19, 1969, the son of Alfonso and Ana Maria Richter.

Education: Hans Christian Andersen School, Lima, Peru, 1986; Pontificia Universidad Catolica del Peru, Bachelor of Science in Mechanical Engineering, 1995; Oklahoma State University, Stillwater, Oklahoma, Master of Science in Mechanical Engineering, specialized in Dynamic Systems and Control. Completed the requirements for the Doctor of Philosophy degree with a major in Mechanical Engineering at Oklahoma State University in December, 2001.

Experience: Raised in Lima, Peru; employed as an engineering intern in the Peruvian Air Force, High Thrust Engines Test Bench, 1992; engineering intern in the San Vicente mining company, Junin, Peru, 1994; employed as Design Engineer in Marine Construction and Design Co., Callao, Peru, 1994; employed as a Research Assistant, Teaching Assistant and Lecturer, Oklahoma State University, Department of Mechanical and Aerospace Engineering; employed as an engineering intern in Seagate Technology, Servo Engineering, 2001; awarded a National Research Council Associateship at NASA Stennis Space Center, 2002.

Professional Memberships: American Society of Mechanical Engineers.