Cross directional control of basis weight on paper machines using Gram polynomials

by Kristinn Kristinsson


A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE STUDIES
DEPARTMENT OF ELECTRICAL ENGINEERING

We accept this thesis as conforming
to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA
January 1994

© Kristinn Kristinsson, 1994
In presenting this thesis in partial fulfillment of the requirements for an advanced degree at the University of British Columbia, I agree that the library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

(Signature)

Department of Electrical Engineering

The University of British Columbia
Vancouver, Canada
Date: January 1994
Abstract

In this thesis a control algorithm for cross-directional basis weight control based on orthogonal polynomials is presented. Instead of controlling the cross-directional values at each point, the parameters of the orthogonal polynomials are controlled, taking into account the maximum bending moment of the slice lip and the relative stiffness of the slice with respect to the actuators. Problems with slice lip bending are dealt with in two ways. Implicitly, by modelling it as a polynomial such that the slice is not bent more than the highest degree of the polynomial, which is chosen beforehand. Explicitly, by calculating the bending moment and then use quadratic programming to limit the bending moment. By modeling the measurements and the actuators positions by polynomials the mapping between the actuators and the sensors, which is the key element of the current control systems, is effectively eliminated. Computational requirements are greatly reduced by avoiding inverses of big matrices.
Table of Contents

Abstract ................................................................. ii
List of Tables ............................................................ vi
List of Figures ........................................................... vii
Acknowledgments ......................................................... xiv
Epigraph ................................................................. xv

1 Introduction ......................................................... 1

1.1 Literature overview ................................................. 4
1.2 Motivation ............................................................. 8
1.3 Contribution of this work ........................................... 10
1.4 Outline of the thesis ................................................ 11

2 Discrete orthogonal polynomial approximation using Gram polynomials 12

2.1 Orthogonal polynomials ............................................. 12
  2.1.1 Creating Gram polynomials .................................... 13
  2.1.2 Gram goes to Legendre ......................................... 20

2.2 Approximations by Gram polynomials ............................ 22
  2.2.1 Autocorrelation .................................................. 26
  2.2.2 Akaike’s Information Criterion ................................. 27

2.3 Frequency characteristics of Gram polynomials ................. 30

2.4 Summary ............................................................. 32
List of Tables

4.1 Location of the PRBS inputs and the corresponding output measurements deduced from Figure 4.39. ................................................................. 63

4.2 Number of data boxes between the maximum values of the cross correlations between the 11 responses in Figure 4.43. ................................................................. 67

5.3 The points of bifurcation and the ratio of the differences between points of bifurcation. . . . 112
List of Figures

1.1 A simplified paper machine ........................................... 1
1.2 CD profile for industrial data and the average for 9 sections of the profile. .............. 4
2.3 Gram polynomials of first to sixth order according to equations 2.12–2.14 evaluated at the 100 points shown with the fine grid on the top x-axis ........................................... 15
2.4 The world of Gram polynomials (first to third order) according to Equation 2.17, plotted for N=100 .......................................................... 16
2.5 Gram polynomials according to Equation 2.22 for 100 points. Plotted for the first to sixth order .......................................................... 17
2.6 Plot of the first 30 Gram polynomials for 100 data points ........................................... 21
2.7 Relationship between the number of data points and the order at which the maximum absolute value of the polynomial is greater than 2, 10 and 100. The circles in the lower left hand corner denote a fourth and an eight order polynomial for 20 data points .......... 22
2.8 Gram goes to Legendre. The smooth solid line is the Legendre polynomial in both plots ........................................... 23
2.9 Approximation of an industrial basis weight CD profile with Gram polynomials to the 10th order. The small inserted graph shows the 10 parameters of the Gram polynomial approximation ........................................... 24
2.10 Measured data and 4, 10, 20 and 40th order approximation using Gram polynomials ........................................... 25
2.11 Parameters of the Gram polynomials in Figure 2.10 and two times the standard deviation of the approximation error ........................................... 26
2.12 Measured data and the measured data with 40th order approximation subtracted .......... 27
2.13 Autocorrelation of a profile minus the Gram polynomial approximation for 4, 10, 20 and 40th order approximation. The 95% confidence interval is shown with the dotted lines .......... 28
2.14 Akaike’s Information Criterion for the profile from Figure 2.9. The solid line shows the AIC, the other two lines show the AIC when it is modified to greater penalize model complexity. The circles in the plot denote the minimum for the curves .......... 29
2.15 Left half of 100th order Gram polynomial for 4096 points. It is shown with the polynomial partitioned into 8 zones, i.e. the left half is partitioned into 4 zones, separated with different line types in the graph. .............................................. 31

2.16 Spectrum of 100th order polynomial. ................................................................. 31

2.17 Spectra of 100th order poly, calculated for the four sections of the left half of the polynomial shown in Figure 2.15. ................................................................. 32

2.18 The original profile (dotted line) and the residual profile (solid line) when a 20th order Gram polynomial is removed. Power spectrum for the original profile (dotted line) and the profile when 20th order Gram polynomial is removed (solid line). ................................. 33

2.19 The original profile (dotted line) and the residual profile (solid line) when a 40th order Gram polynomial is removed. Power spectrum for the original profile (dotted line) and the profile when 40th order Gram polynomial is removed (solid line). ................................. 33

3.20 Paper Machine. ................................................................................................. 36

3.21 O-frame. ........................................................................................................... 36

3.22 Response to a slice lip movement at one actuator position. The slice lip deflection is denoted by • with O denoting the deflection at a slice lip actuator. The + marks what the basis weight might look like. ................................................................. 38

3.23 CD responses, according to Laughlin's interpretation from the literature, see also Equation 3.45. ................................................................................................. 41

3.24 Reciprocal condition numbers of $G$, where it is created according to equation 3.43 with $p_0 = 1$, $p_1 = 2r$ and $p_2 = -r$, for various sizes of the matrix. When $G$ is badly conditioned, the reciprocal condition number is close to 0. .............................................. 43

3.25 Reciprocal condition number for a 105 x 105 matrix with $p_0 = 1$, $p_1 = 2r$ and $p_2 = -r$. Basis weight response for a machine with a singular interaction matrix, reciprocal condition number equal to zero. .............................................. 43

3.26 Dry weight measurements and the input to the corresponding slice lip location. ................. 44

3.27 Plot of the weight response to a straight line slice lip profile. The upper half is using $G$ matrix from Equation 3.50 and the lower half is if the edges of the matrix are made smooth. ................................................................. 45
3.28 Geometry when the number of actuators is not the same as the number of measurements.

It is assumed that the first measurement is half a data box width from the edge.

3.29 Meshed surface of the $G_x$ matrix for an interaction matrix of size 105x105 and $p_0 = 1$,

$p_1 = 0.4$ and $p_2 = 0.2$.

3.30 Every second diagonals of $G_x$ from Figure 3.29. The zeroth (main) diagonal is not shown.

3.31 Reciprocal condition number of $G_x$ for a fixed number of Gram polynomials but varying the sizes of the interaction matrix, $G$, and its off diagonal elements which are $p_1 = 2r$ and $p_2 = -r$.

3.32 Reciprocal condition numbers of $G_x$ for $G$ of size 105 x 105 with $p_0 = 1$, $p_1 = 2r$, $p_2 = -r$ and varying the number of Gram polynomials.

3.33 Plot of singular values of $G_x$ for an interaction matrix of dimension 105 x 105 with $p_0 = 1$, $p_1 = 2r$, $p_2 = -r$ and with $r = 0.3$.

3.34 Model of slice lip.

3.35 Deflection of the beam in Figure 3.34 for different stiffnesses, when an actuator at $x = 1$ is moved one unit.

3.36 Deflection of a beam when a actuator at $x = 1$ is moved to one from the initial position of zero (shown for the actuator as a dotted line with the endpoints denoted with small circles).

3.37 Slice lip.

4.38 An experiment to identify the interaction matrix, $G$.

4.39 Slice lip actuators and basis weight measurements from an industrial data set. To get a good contrast in the picture the mean value of every CD point is subtracted from the raw data.

4.40 Slice lip set point and measured basis weight as a function of time, from an industrial data set.

4.41 The cross correlation between the input and the output for the 11 PRBS signals from an industrial data set.
4.42 Estimates of steady state gain, $p_j$, between the slice lip actuators and the corresponding measurements from Table 4.1 for an industrial data set. .............................................. 66

4.43 The estimated response across the sheet, for 11 actuators from an industrial data set. The plus signs mark the final estimates from Figure 4.42. ............................... 66

4.44 The 11 estimated response shapes from Figure 4.43 shifted according to the first line in Table 4.2 such that they overlay. ...................................................... 67

4.45 Average estimated response of the weight profile by moving a single actuator. ............. 68

4.46 Basis weight response to a flat slice lip (at 10$\mu m$) with 105 actuators for two different models of a paper machine. ................................................................. 70

4.47 Input, $u_c$, and output, $y_c$, used for the estimation of $G_x$, when there are no limits on the slice lip bending. ................................................................. 72

4.48 Overlay of 59 slice lip profiles, measurement profiles and bending moments corresponding to the inputs and outputs in Figure 4.47. ............................... 73

4.49 Estimation of each row of $G_x$ when no limit is put on the slice lip bending. The true parameters for this simulation, are shown with dotted lines. ........................................... 74

4.50 Estimate of $G_x$ after 59 scans when no limits are put on the slice lip bending. The estimated matrix is denoted with * and dotted lines between its elements. The true matrix used in the simulation is denoted with o and solid lines. ........................................... 75

4.51 Input, $u_c$, and output, $y_c$, used to estimate $G_x$, when there are limits on the slice lip bending. ................................................................. 76

4.52 Overlay of 59 slice lip profiles, measurement profiles and bending moments corresponding to the inputs and outputs in Figure 4.51. ............................... 77

4.53 Estimation of each row of $G_x$ when limit is put on the slice lip bending. The true parameters for this simulation, are shown with dotted lines. ........................................... 78

4.54 Estimate of $G_x$ after 59 scans when limits are put on the slice lip bending. The estimated matrix is denoted with * and dotted lines between its elements. The true matrix used in the simulation is denoted with o and solid lines. ........................................... 79
4.55 Input and output used to estimate the $G_x$ matrix when the dynamics is estimated as well.

4.56 Estimation of the parameterized interaction matrix $G_x$ when the dynamics is estimated as well. The dotted lines show the true values.

4.57 Estimation of $a$, for the 10 rows in the previous figure.

4.58 Estimation of each row of $K_0$, the inverse of $G_x$. The dotted lines show the true values.

4.59 The estimate of $K_0$ after 59 scans. The estimated matrix is denoted with * and dotted lines between its elements. The true matrix used in the simulation is denoted with o and solid lines.

5.60 Dahlin controller for a plant.

5.61 Dahlin controller with control action taken every 3 scans with the closed loop pole, $p = 0.05$ and the open loop pole is, $f_p = 0.1$. Dahlin controller with control action taken every scan with the closed loop pole, $p = 0.37$ and the same open loop pole. The * denotes the output at every third scan.

5.62 The signals in a closed loop control with a Dahlin controller; with and without limits on the control action, and with and without integrator antiwindup. The plant pole is $f_p = 0.3$ and the desired closed loop pole is $p = 0.05$.

5.63 Controller with antiwindup.

5.64 Block diagram of basis weight control strategy using Gram polynomials.

5.65 Block diagram of basis weight control strategy using Gram polynomials with an integrator antiwindup. The controller, $C$, from Figure 5.64 is $\frac{B_s(q^{-1})}{A_s(q^{-1})}$.

5.66 Input, $u_c$, and output, $y_c$. The system is initially in open loop but the loop is closed at scan 40.

5.67 Estimation of each row of $G_x$. The true parameters for this simulation, are shown with dotted lines. The system is initially in open loop but the loop is closed at scan 40.

5.68 Slice lip profile, bending moments and output at scan 100.

5.69 Slice lip profile, bending moments and output at scan 100, when no limits are used for the slice lip.
5.89 Meshed surface of the $G_z$ matrix for uniform $G$ using the Gram polynomials from Figure 2.3. ................................................................. 110

5.90 A bifurcation diagram and pseudo phase plane plots for different values of $p$. ........... 111

5.91 Response of the system with $p = 0.55$, for 1% difference in initial condition on the desired value of the CD value at location 10. ........................................ 112
I would like to use this opportunity to thank my research supervisor and colleague, Prof. Guy A. Dumont, for his advice and guidance during the course of this work. I would also like to thank Prof. Michael Davies and other past and present members of the Paper Machine Group at the Pulp and Paper Centre for their valuable discussion.

Thanks to Dr. Oliver Lee, Leoncio Estevez-Reyes, Navid Mohsenian, Dr. Ye Fu and Jahan Ghofraniha for reading parts of the thesis. Also thanks to the students and the staff at the Pulp and Paper Centre at the University of British Columbia for all their help and assistance, and for making my stay here very enjoyable.

Acknowledgment is made to Paprican's Frank Leslie Mitchell memorial fellowship, for financial support, and to Jóns Thórarinssonar memorial fellowship.

Special thanks to friends and family in Iceland that have helped in one way or another during the course of my study.

Last but not the least my heartfelt thanks to Anna Dóra for her encouragement when I was working on my thesis.
I know why there are so many people who love chopping wood. In this activity one immediately sees the results.

Albert Einstein
Chapter 1

Introduction

The three R's (reduce, reuse and recycle) have been catching on recently as a means of reducing the consumption of raw materials. Paper, which in Canada is manufactured out of wood chips, is one of the prime materials being targeted by the three R’s. One way of reducing usage of raw materials is in the manufacturing of paper, where a single percent reduction in fiber usage translates to preservation of some 800 km$^2$ of forest per year [79], which is about 2.5% of the area of Vancouver Island. Not only can we reduce the amount of paper we consume, but we can also reduce the amount of fibers needed to manufacture a sheet of paper.

The weight of paper per unit area, or the basis weight (BW), is a fundamental physical property of paper [30]. Other important properties are moisture and caliper. These three properties are interdependent; changing one will affect the rest, but it has been observed that the oven dry basis weight profile is dominant [69]. When basis weight is expressed in g/m$^2$ or mass per unit area, it is referred to as grammage. Grammage variation is synonymous with substance variation and basis weight variation [32]. Basis weight is the most commonly used term and will therefore be used throughout this thesis when referring to mass or weight per unit area of manufactured paper.

![Figure 1.1: A simplified paper machine.](image)

1
Chapter 1: Introduction

The paper machine (see Figure 1.1), invented almost 200 years ago [68], has been constantly improved over the years. It has become wider and faster, and modern paper machines make hundreds of tons of paper per day. With the increase in their width and speed, the problem of controlling them has become more difficult. The goal of control is to make the paper as uniform as possible. Variations will undoubtedly exist in the manufactured paper sheet, because of variations in raw materials and the complicated pulping process prior to the paper machine. To better understand the source of these variations, and ultimately to remove them, it has become necessary to classify variations in the paper sheet based on their source or how they might be reduced. Variations are generally divided into 3 components, based on their source, following Burkhard’s and Wrist’s paper in 1954 [16]:

— Machine Direction (MD) part. Variations occurring along the sheet travel on the paper machine and affecting the whole width of the machine.

— Cross Direction (CD) component. Variations occurring perpendicular to the sheet travel on the paper machine. The CD variations generally change much slower than the MD variations and can be assumed to be nearly time-invariant or stable with time. The CD variations will therefore persist in successive profiles.

— Random (R) element. Part of the variations that occurs neither along nor across the machine. This part is also referred to as the residual part, because variations that cannot be attributed to either CD or MD, are assumed residual or random.

An alternative classification scheme has appeared in recent times as measurement suppliers have started categorizing basis weight variations according to how they are controlled. Variations that can be controlled by CD actuators are called CD, while those controllable by MD control loops are called MD [30]. The different definitions can be attributed to the fact that variations that appear in the machine direction can be better controlled by actuators in the cross direction — they have faster response time and shorter time lag since they are closer to the measurement point. Researchers have also recently started looking at ways to represent the paper sheet as a two-dimensional process in which the CD and MD are two mutually-dependent dynamical processes [123].

The goal of control is to increase uniformity of the manufactured product. A reduction in variability influences nearly every factor which tends to increase production and improve paper quality [16]. Many benefits can be gained from reduced cross direction fluctuation, e.g. the average basis weight can be set lower and the average moisture can be set higher, resulting in a reduction in raw material usage. A more uniform profile also produces a stronger web and a more uniform reel.
leading to a reduction in the number of sheet breaks on the paper machine and in the converting process. It has been argued that variations in CD are more important than variations in MD or random variations [30, 31]. Since the manufactured paper is wound into rolls the MD and random variations will average out for each roll. The paper roll is usually cut into few smaller rolls, the CD variations will therefore be particular to a specific roll. This can be observed from Figure 1.2 where a CD profile (dotted line) along with its average for 9 sections of the profile (solid line), are shown. Assuming that the MD and random variations average to zero for the production of a roll of paper, and that the CD profile will be present in the whole roll, the 9 rolls will all have different averages.

It is stated in [26]: “it is a fact that changes across the machine are always greater than they are in the machine direction”. That statement is supported in several references [17, 125], however, it may be due to the fact in most cases the MD is already controlled [50]. Care must be exercised when using variance to judge the performance of a CD controller. The reduction might, in part, be due to low profile resolution [95]. The measured profile mapped to the actuators might even show zero variance whereas the unmapped profile has some streaks [103]. It can, for example, be shown that the expected value of the standard deviation for normally distributed random numbers reduces by $1/\sqrt{2}$ if the random numbers are averaged over two data boxes. A better technique for variability analysis makes use of profile auto-spectra [96], or power spectra. Fourier analysis and power spectra can be used [74] to show that the part of the variance with wavelength bigger than or equal to twice the size of the actuator spacing can be reduced, whereas it is not possible to reduce shorter wavelengths. However, it has been reported for an industrial application that wavelengths slightly shorter than two actuator spacings can be reduced when the controller makes use of full resolution measurements [63].

One of the major problems with current CD control systems is to correctly align actuators and measurements. Inaccurate mapping is mentioned in [4] as one source of poor controller performance. Misalignment of even a few inches can result in sawtooth profiles, since actuators are adjusted to compensate for errors outside their control zone [89]. For optimal control results, measurement zones should align with control zones. When the measured zone on the dry end, corresponds to the correct actuator zone on the headbox, performance is maximized [17]. If they do not align, the data must be transformed, with some loss in profile resolution [89]. As we shall see, the problem of misalignment between actuators and measurements can be overcome through the use of orthogonal polynomials.
1.1 Literature overview

There have been several review articles on the control of pulp and paper processes [76, 117], as well as several surveys on the use of CD control in the industry [108–110]. In [36] a survey is given on the use of advanced control methods in the process industry in general, and [35, 37] give a survey on the use of advanced control in the pulp and paper industry and adaptive control in papermaking in particular.

The three factors that have governed the industrial implementation of CD control systems are: sensors, actuators and process control computers [121]. With the introduction of the microcomputer in the early 1980’s, interest in CD profile control increased, indeed the 1980’s have been called the decade of cross direction profile control [106]. With faster computers, the profile could be measured in smaller and smaller data boxes, which consequently called for smaller and smaller actuator spacings. Now it is common to have three to five measurements for each actuator for the control of basis weight. On the other hand, one early investigation on CD basis weight control assumed that there were more actuators than measurements when automatically controlling the basis weight profile [116], but that is clearly inadequate, as it is not possible to control variations that are not visible in the measurements. Traditionally the measured profile was mapped to the actuators before calculating the control action, but recently, use has been made of the high resolution profile when calculating the control action.

Figure 1.2: CD profile for industrial data and the average for 9 sections of the profile.
Chapter I: Introduction

[63]. The idea here is that the controller will control high frequency variations better if it knows they are present. The controller did reduce wavelengths slightly shorter than two actuators spacing.

As mentioned above, Burkhard and Wrist [16] were the first to classify basis weight variations into three components. Initial efforts at reducing variations concentrated on the control of MD variations [5], mainly because CD sensors were not available and more computer power was required for control of cross directional variations. CD control first became possible with the invention of sensors to measure the paper across the sheet. Initial efforts were mainly focused on troubleshooting machine faults rather than on control. In 1957, Cuffey [28] obtained a picture of basis weight profiles with an optical basis weight scanner. He used a bank of lamps over the moving paper sheet to illuminate the sheet; a rotating mirror reflected the light to a lens and then to a photosensor. With this setup it was observed that a movement of a single slice lip actuator caused the basis weight to change not only at that location but also at the adjacent locations. It was also observed that the slice lip opening changes at adjacent actuators when a single actuator is altered. In 1967, Åström published one of the early papers on computer control of paper machines [5]. This paper was on machine direction control of basis weight and summarized a project, dating back a few years, to implement computer control in industrial settings. A minimum variance controller was designed by use of linear stochastic control theory. The beta-ray sensor utilized was able to traverse the paper web but it stayed at a fixed position. An application of adaptive control to moisture and basis weight in the machine direction was reported in 1975 by Cegrell and Hedqvist [19]. In 1969, about 10 years after Cuffey’s paper, Dahlin presented an algorithm based on exponential filtering, to extract the MD and CD component from raw data, and gave a description of the sensors used for basis weight and moisture measurements [33]. Around the same time Beecher and Bareiss reported on their efforts to use a closed-loop analog controller on the slice lip for basis weight control; the inverse of an interaction matrix was used to describe how the measurements and actuator settings were related [8]. They had 16 motorized slice actuators on a 2.3 m wide machine capable of speeds between 0.76 m/s and 3.81 m/s. Standard deviation from the mean of the basis weight profile was reduced. Better results were obtained a couple of years later when the system was made digital [44]. With installation of a titanium slice, further reduction in basis weight variability was observed [26].

In 1975, Carey and co-workers [17] at Kodak reported a reduction in fibre usage following installation of a CD control system, which used the interaction matrix approach. Two years later Boyle [13] showed that the interaction matrix approach could result in excessive bending and/or displacement of the slice lip and that the solution may well be infeasible, calling for actuators to
be more than full open or less than fully closed [14]. Two methods emerged to limit bending and
displacement of the lip. The first method, quadratic programming, minimized variations subject to
hard limits on the actuators movements; however, it was found to be computationally intensive. The
second method, minimization of quadratic objective function, penalized actuators moves or bending
of the slice. In 1977, Boyle [13] described a quadratic programming formulation for simulation and
noted that computational considerations limit the applicability of the method to only a few actuators.
The following year, the same author [14] revealed a simulation study, which used a quadratic objective
function for a system with unequal number of actuators and sensors. A drawback with this method
is that a matrix of size $\text{number of actuators} \times \text{number of measurements}$ has to be inverted or its
pseudoinverse found, although this task can be somewhat simplified by taking advantage of the
symmetry of the matrix.

A particular problem with stiff slice lips according to Garnett [52] is that the lip cannot make the
moves necessary to control some profiles. This concern was repeated a year later by Boyle [15]; the
author suggested that tests should be carried out on slice actuators, prior to purchase of a CD weight
control system to evaluate the flexibility of the slice lip. Cutshall and Hamel [31] give an overview
in 1988 of some factors that need to be evaluated when slice lip automation is considered.

In Wilkinson and Hering [127], a linear quadratic optimal design using on line identification of
the interaction matrix is used in industrial settings. This work, which was partly based on earlier
work by Richards [100] and Wilhelm [124] from 1982, gave between 44–55% reduction in basis-
weight variations on several installations. The same year, Wilhelm and Fjeld [126] reviewed some
control algorithms for cross directional control. They formulated control methods for both the inverse
interaction matrix approach and for quadratic optimal design. They noted that strong coupling in
the interaction matrix could result in the matrix being singular and therefore not invertible. They
suggested that strong coupling simply represented a poor choice of actuator spacing. The problem
could be avoided by controlling groups of actuators together, or computing the control based on
a wide spacing of actuators and then using interpolation to calculate the settings of the remaining
 actuators. This work was extended three years later by Chen and Wilhelm [24] and Chen, Snyder
and Wilhelm [25]. They looked at both quadratic programming (QP) and quadratic penalty function
(QPF) methods; the latter is a form of quadratic optimal design. The QP gives an optimal solution,
however, time and memory requirements for the QP method prohibit real-time control. Therefore,
they turned to QPF, which gives a suboptimal solution. Generalized least squares was used to identify
a spatial non-causal model of the process and more than 50% reduction in standard deviation of BW
was achieved with an adaptive control algorithm based on QPF [25]. This algorithm was shown to be equally appropriate to manage the wide response model for heavy weight machines [21]. Later, it was used as a basis for controlling a single CD profile with multiple actuators and for using the CD actuators to assist in MD control [57]. Interactions in CD have been a cause for great concern in basis weight control. The interactions depend on the operating point of the paper machines, from which the need to identify the interactions follows. One such identification scheme is described in a patent granted in 1991 [22] for determining cross direction responses and mappings of measurements onto actuators used for control. The responses were determined by means of PRBS signals and by calculating the covariances between the input and the outputs. The perturbing signals were defined by PRBSs with multiple signals being selected to be statistically independent [22].

In the early 1980's researchers with STFI in Sweden developed and installed a CD control system in a mill. Karlsson et al. [72, 73] observed a 20% decrease in standard deviation for basis weight using an undisclosed control method. Some of the highlights of the control package were given by Elliot in 1984 [45, 46]. The algorithm was used in several installations to reduce CD variations [95, 96, 105]. Hansson et al. [59] have also reported on the use of the STFI control package, where a slice lip was used to control moisture on a kraft machine. The paper machine had 48 slice actuators and a width of 7.1 m, with slice opening at 50 mm.

In 1987 Bergh and MacGregor [9] employed LQG (linear quadratic gaussian) control theory to jointly control MD and CD moisture variations using spatially distributed actuators. They modeled the disturbances for a sheet forming process as a multivariate time series. A multivariate LQG control with recursive identification to sequentially adjust slice lip actuators with a traversing robot is given in Halousková et al. [58].

None of the control strategies mentioned so far have addressed the issue of error in the interaction model. An exact model will never be possible because of wave propagation on the wire. In 1988, Laughlin [79, 81] came up with a MIMO (multiple input multiple output) robust control scheme based on internal model control (IMC) [80]. He was able to guarantee both robust stability\(^1\) and robust performance\(^2\), by assuming knowledge of the parameter uncertainty. A major drawback with this scheme is that the gain matrix describing the relationship between actuators and sensors has to

\(^1\) A closed loop control system is said to be robustly stable over a set of plants if and only if it is stable for all plants in the set.

\(^2\) A closed loop control system is said to exhibit robust performance if for all frequencies the sum of weighted sensitivity and complementary sensitivity functions is less than one.
be positive definite. In [79], a table of reported matrices from the literature, shows that five out of thirteen matrices are not positive definite [79]. Furthermore the author does not take into account the constraints of the slice lip, so the control action from this scheme may not be realizable.

There have also been reports about reduction in cross directional variation, an increase in production and/or reduction in raw material usage with the installation of slice lip control. A few of them can be found in references [10, 12, 47, 53, 34, 60, 82, 91, 119, 128].

Matsuda observed in 1990 [88] that increasing the number of measurements to twice the number of actuators improved control because narrow streaks could be eliminated. A year later Yamamoto [129] reported an application of fuzzy logic for profile control of a paper machine. The author reported using only 4 rules to eliminate the sawtooth form from the CD profile. A patent describing the use of a fuzzy logic controller on a slice lip was granted in 1992 [104]. In the patent an example is given of a paper machine with 360 measurements and 37 actuators. The fuzzy rules are derived from manual manipulations by a skilled human operator.

Because of the complexity in controlling the cross direction dry weight Amyot, Nuyan, and Evans designed an expert system to help operators of a paper machine troubleshoot the cross direction basis weight control system [4, 48].

1.2 Motivation

The objective of this thesis is to develop a strategy for control of CD variations on paper machines which can address the shortcomings of current control methods, namely problems with interactions in the cross direction, and the need to map the measurements to the actuators. Since the basis weight is a fundamental property of paper, cross-directional basis weight control has been chosen for this study. The sampling rate on the CD basis weight loop is slow (around 30 seconds), so this is essentially a steady-state control problem where the purpose is to solve the decoupling problem between the numerous actuators [36].

Most current techniques for cross-directional basis weight control manipulate the cross directional values at each point separately from each other. Therefore, a large interaction matrix from the slice lip to the measurements has to be identified and its inverse found. The inverse matrix or mapping between measurements and actuators is the key element of current control methods and, if not correctly done, a sawtooth profile will result. Filtering is used in order to protect the slice lip from excessive bending. The filtering may take the form of sending the slice lip set points to a low pass
filter; of penalizing bending when calculating the slice lip profile; or of placing some hard limits on the movement of actuators. Another way to deal with excessive bending is to compute control settings of a few actuators and the remaining actuators settings are then found by interpolation using a polynomial of a certain order [126]. The actuators can also be treated in groups to simulate a change in the spacing of the actuators, since some paper grades might not be able to take advantage of close spacings [125]. Current industrial trend is towards an increased number of CD points. Measurement data boxes as small as 1 cm are reported by Taylor in 1991 [115]. For a 10 m wide paper machine this implies 1000 data boxes. To make full use of this increased measurement resolution, the number of actuators in the cross direction has increased as well. However, this gives rise to large interactions between adjacent CD points so that an actuator movement in one CD position affects many of its neighbors. Such interactions cannot be ignored when calculating a new control move and adjacent actuators must work not to exceed the bending limits of the slice lip they are controlling. Another problem associated with high resolution profiles is that in order for a CD control system to operate effectively, it is necessary to accurately determine which positions in the CD profiles will correspond to the locations of the slice screws. If this is not done, self-oscillation will result in the profile when a slice lip actuator tries to compensate for an error outside its area [95].

Recently, work has been devoted to the extraction of the cross direction (CD) and machine direction (MD) profiles from scanned measurements [20, 23, 39, 40, 70, 92, 122]. The main reason for extracting profiles from measurements is to exploit the information obtained for control. In this thesis, knowledge of the CD profile will be used for control.

In this work, the CD profile or its shape is approximated by discrete orthogonal polynomials and then the parameters of the polynomials are controlled rather than the cross directional values at each point. The slice lip has certain limits on bending so its profile is a smooth curve that is naturally described by a polynomial of a given order. By modelling the slice lip with a polynomial, problems with slice lip bending are dealt with implicitly because the slice lip will not bend more than the highest degree of the chosen polynomial. It will also be shown in this thesis how bending of the slice lip can easily be limited by modifying the parameters of the polynomials. Fewer parameters are controlled and fewer parameters need to be identified because the degree of the polynomial is much smaller than the number of CD points. Computational requirements are greatly reduced by avoiding inversion of big matrices and inverses are found for a diagonal matrix. Finally, by modelling the measurements and actuators positions with polynomials, mapping between actuators and measurements, which is the key element of current control systems, is effectively eliminated.
Modelling profiles by polynomials has already been used in the pulp and paper industry to control wet end moisture profiles on a paper machine through application of pressure to the ends and center of the press rolls [75]. The moisture profile was approximated by a set of polynomials and then the coefficients of the polynomials were controlled. Since only three actuators were used, second order polynomials were considered and consequently not much could be gained by use of orthogonal polynomials. Modelling profiles with discrete orthogonal polynomials has previously been used in Sendzimir steel rolling mills [41–43, 56, 101, 102]. Models for these plants can be derived from first physical principles and, therefore, analytical models exist. Only about 10 actuators and up to 31 measurements were used for cross directional control, so that only low order polynomials were required. Paper machines, in contrast, have complex interactions in CD that are difficult to model. Furthermore, the paper sheet is made up of small fibers that tend to flocculate, and are thus difficult to control. The number of measurements on paper machines is about 20 times that on steel mills and the number of actuators is about 10 times. Consequently, higher order polynomials are needed to describe the shape profile of manufactured paper.

1.3 Contribution of this work

1. Use of Gram polynomials to describe profiles on paper machines. Several forms of the Gram polynomials are investigated and it is shown that the scaling of the polynomials is important. Some of the frequency characteristics of the Gram polynomials are evaluated.

2. Use of Gram polynomials to describe slice lip deflection and bending moments. Problems with slice lip bending are implicitly dealt with by modelling it as a polynomial series, since the control method will not bend the slice more than the highest degree of the series. Slice lip bending is related to the parameters of the orthogonal polynomials describing the slice lip, thereby allowing the bending moments to be restricted, directly from the parameters of the orthogonal polynomial of the slice lip deflection.

3. Profile control design in the Gram space. Instead of controlling the CD values at each point, the parameters of the Gram polynomials are controlled, taking into account the maximum permissible slice bending. Because the number of polynomials is much smaller than the number of CD points (10 polynomials seem sufficient to describe most profiles in our industrial data), this technique greatly reduces computational requirements by avoiding inversion of large matrices and by using the fact that the polynomials are orthogonal. Furthermore, we effectively
eliminate the need for mapping between the actuators and the data boxes, a crucial difficulty of current control systems.

1.4 Outline of the thesis

The discrete orthogonal polynomials used in this work and how they are used to approximate a profile or vector are discussed in chapter 2. A model is developed in chapter 3 to relate i) slice lip deflection to measured basis weight profile, ii) actuator setpoints to slice lip deflection, and iii) parameters of the discrete orthogonal polynomial approximation of the measured output to parameters of the discrete orthogonal polynomials of the slice lip. In chapter 4 identification of the interaction and the parameterized matrices is described. Chapter 5 presents a simulation of the result of applying the method described in this thesis to control a paper machine via a Dahlin controller. Finally, in chapter 6 some conclusions are drawn, and a future work is discussed.
Chapter 2

Discrete orthogonal polynomial approximation using Gram polynomials

Polynomial approximation is a technique used to obtain a smooth approximation to noisy measurements [99]. Instead of including measurement errors in the approximation, we seek to find an approximation which gives the general trend of the data without being too concerned with local fluctuations [90].

The use of orthogonal polynomials achieves certain computational simplifications. There are many different orthogonal polynomials that can be employed. For discrete measurements it is natural to choose discrete orthogonal polynomials. Some discrete orthogonal polynomials are appropriate for uniformly-spaced data points. Others can be used when data points are to be weighted.

2.1 Orthogonal polynomials

Let \( y(x) \) be a sequence of measurements observed at \( x = \{x(i); \ i = 0, \ldots, N - 1\} \), where \( N \) is the number of data points. Let \( P_j(x) \) be a sequence of functions defined for every \( x(i) \) and \( j = 0, 1, \ldots, m \), where \( m \) is the highest degree of the polynomials. The plan is to write a function, \( y(x) \), as a linear combination of \( P_j(x) \):

\[
y(x) = y_c(0)P_0(x) + y_c(1)P_1(x) + \cdots + y_c(m)P_m(x)
\]

or

\[
y(x) = [P_0(x) \ P_1(x) \ \cdots \ P_m(x)] \begin{bmatrix} y_c(0) \\ y_c(1) \\ \vdots \\ y_c(m) \end{bmatrix}
\]

with the parameters \( y_c(j) \) to be determined such that the weighted square of the difference between the true (or measured) function and the approximated one is a minimum. In an equation form this is expressed as a least-squares problem:

\[
\min_{y_c} \left\{ \sum_{i=0}^{N-1} w(x(i)) \left[ y(x(i)) - \sum_{j=0}^{m} P_j(x(i))y_c(j) \right]^2 \right\}
\]

or

\[
\min_{y_c} \left\{ \sum_{i=0}^{N-1} w(x(i)) \left[ y(x(i)) - \sum_{j=0}^{m} P_j(x(i))y_c(j) \right]^2 \right\}
\]

12
Chapter 2: Discrete orthogonal polynomial approximation using Gram polynomials

To minimize the above expression, its partial derivatives with respect to $y_c(j)$ can be found and the result set equal to zero. This gives:

$$
\sum_{i=0}^{N-1} w(x(i)) \left[ y(x(i)) - \sum_{j=0}^{m} P_j(x(i)) y_c(j) \right] P_k(x(i)) = 0 \quad (2.4)
$$

or with a little rearrangement:

$$
\sum_{i=0}^{N-1} w(x(i)) y(x(i)) P_k(x(i)) = \sum_{j=0}^{m} y_c(j) \sum_{i=0}^{N-1} w(x(i)) P_j(x(i)) P_k(x(i)) \quad (2.5)
$$

Calculation of $y_c(j)$ from the above equation is greatly simplified if the set of functions $\{P_j(x)\}$ is orthogonal over the set of points $x(i)$. By definition a set of functions $\{P_j(x)\}$ is orthogonal over the set of points $x(i)$ with respect to the weight function $w(x)$ if

$$
\sum_{i=0}^{N-1} w(x(i)) P_j(x(i)) P_k(x(i)) = \begin{cases} 0 & \text{if } j \neq k \\ \lambda_i & \text{if } j = k \end{cases} \quad \text{for } j, k = 0, 1, \ldots m \quad (2.6)
$$

Furthermore, for equally spaced data points and the particular case when $w(x(i)) = 1$, the discrete orthogonal polynomials are called Gram polynomials [99 p.259]. They are sometimes referred to as discrete Legendre orthogonal polynomials (DLOP’s) because it is observed that the numeric coefficients of the DLOP’s are, apart from the sign and possibly a scaling factor, precisely the same as those of the shifted Legendre polynomials on the interval $[0,1]$ [93]. In [66 p. 290] it is mentioned that the polynomials are sometimes referred to as Chebyshev polynomials, although that name is usually reserved for continuous polynomials with $w(x) = 1/\sqrt{1-x^2}$.

2.1.1 Creating Gram polynomials

A recurrence relationship to create discrete orthogonal polynomials is given in [18, 51, 67, 99] and can be derived from the Gram-Schmidt orthogonalization of the powers $1, x, x^2, \ldots x^m$ [78]. The recursive formula in [67, 99] is:

$$
P_{m+1}(x) = (x - \alpha_{m+1}) P_m(x) - \beta_m P_{m-1}(x) \quad (2.7)
$$

where

$$
\alpha_{m+1} = \frac{\sum_{i=0}^{N-1} w(x(i)) x(i) \{P_m(x(i))\}^2}{\sum_{i=0}^{N-1} w(x(i)) \{P_m(x(i))\}^2} \quad (2.8)
$$
and
\[
\beta_m = \frac{\sum_{i=0}^{N-1} w(x(i))x(i)P_{m-1}(x(i))P_m(x(i))}{\sum_{i=0}^{N-1} w(x(i))(P_{m-1}(x(i)))^2} = \frac{\sum_{i=0}^{N-1} w(x(i))(P_m(x(i)))^2}{\sum_{i=0}^{N-1} w(x(i))(P_{m-1}(x(i)))^2}
\]  
(2.9)

and the initial conditions for the recursive equation is given by
\[
P_0(x) = 1 \quad \text{and} \quad P_{-1}(x) = 0
\]  
(2.10)

For an equally spaced \( x \) on the interval \([-a, a]\) and for the particular case \( w(x(i)) = 1 \) for all \( i \), \( \alpha_{m+1} = 0 \) so the recursive relationship becomes
\[
P_{m+1}(x) = xP_m(x) - \beta_m P_{m-1}(x)
\]  
(2.11)

or
\[
P_m(x) = xP_{m-1}(x) - \beta_{m-1}P_{m-2}(x)
\]  
(2.12)

If the interval the polynomials are defined over is chosen as \([-1, 1]\), the polynomials will be evaluated at the equally spaced points:
\[
x(i) = 2\frac{i}{N-1} - 1 \quad \text{for} \quad i \in [0, N-1]
\]  
(2.13)

and the parameter \( \beta_{m-1} \) defined as
\[
\beta_{m-1} = \frac{\sum_{i=0}^{N-1} (P_{m-1}(x(i)))^2}{\sum_{i=0}^{N-1} (P_{m-2}(x(i)))^2} = \frac{(m-1)^2(N^2-(m-1)^2)}{4(m-1)^2-1} \begin{pmatrix} N-1 \end{pmatrix}^2
\]  
(2.14)

The first few polynomials are
\[
P_1(x(i)) = 2\frac{i}{N-1} - 1
\]
\[
P_2(x(i)) = 4\frac{i^2}{(N-1)^2} - 4\frac{i}{N-1} + 2\frac{N-2}{3(N-1)}
\]
\[
P_3(x(i)) = 8\frac{i^3}{(N-1)^3} - 12\frac{i^2}{(N-1)^2} + 4\frac{i}{N-1} \frac{6N^2-15N+11}{5(N-1)^2} - 2\frac{N-2}{5(N-1)^2} \frac{(N-3)}{(N-1)^2}
\]  
(2.15)

Figure 2.3 shows the Gram polynomials calculated using Equations 2.12–2.14 for 100 points (N=100). For clarity, these polynomials are shown as continuous lines, even though they are defined only at the discrete abscissas indicated with tick marks at the top of the figure. The maximum value of
Chapter 2: Discrete orthogonal polynomial approximation using Gram polynomials

Figure 2.3: Gram polynomials of first to sixth order according to equations 2.12–2.14 evaluated at the 100 points shown with the fine grid on the top x-axis.

the polynomials gets smaller and smaller as the degree of the polynomials gets higher and higher. Accordingly, the parameters for the higher order polynomials can become very large.

Another recursive scheme for calculating orthogonal polynomials for \(N\) evenly spaced measurements is \([7]\):

\[
P_{m+1}(i) = P_m(i)P_1(i) - \frac{m^2(N^2 - m^2)}{4(m^2 - 1)}P_{m-1}(i)
\]

\(i = 0, 1, \ldots, N - 1\) and \(m = 1, 2, \ldots, N - 2\)  

The starting conditions are \(P_0(i) = 1\) and \(P_1(i) = x(i) = i - \frac{N-1}{2}\). With the same initial conditions, the recursion can also be written as:

\[
P_m(i) = P_{m-1}(i)P_1(i) - \frac{(m-1)^2(N^2 - (m-1)^2)}{4(4(m-1)^2 - 1)}P_{m-2}(i)
\]

\(i = 0, 1, \ldots, N - 1\) and \(m = 2, 3, \ldots, N - 1\)

The first three polynomials are then written as:

\[
P_1(i) = i - \frac{N-1}{2}
\]

\[
P_2(i) = i^2 - i(N-1) + \frac{(N-1)(N-2)}{6}
\]

\[
P_3(i) = i^3 - i^2\frac{3(N-1)}{2} + i\frac{6N^2 - 15N + 11}{10} - \frac{(N-1)(N-2)(N-3)}{20}
\]
Chapter 2: Discrete orthogonal polynomial approximation using Gram polynomials

Figure 2.4 shows the Gram polynomials calculated using Equation 2.17 for 100 points, and plotted as a function of $x$. Here, the maximum value of the polynomials gets larger and larger as the degree gets higher and higher. Due to the poor scaling of these polynomials, the scaling of the parameters here becomes even worse than for the first set of polynomials (see Figure 2.3). For example $P_3(0) \approx -N^3/20 = -100^3/20 = -50000$, whereas $P_1(0) \approx -N/2 = -100/2 = -50$. Poor scaling of the polynomials is apparent in Figure 2.4 where the first and the second order polynomials are barely visible because the third order polynomial has much larger values.

![Figure 2.4](image_url)

**Figure 2.4:** The world of Gram polynomials (first to third order) according to Equation 2.17, plotted for $N=100$.

Yet another scheme for creating discrete orthogonal polynomials is derived in [90]. For $N$ equally spaced points, the polynomials are written as:

$$P_m(i) = \sum_{j=0}^{m} (-1)^j \binom{m}{j} \binom{m+j}{j} \frac{i^{(j)}}{(N-1)^{(j)}}$$

(2.19)

The product $i^{(j)} = i(i-1)(i-2) \cdots (i-j+1)$, where $j$ is a positive integer, is called the factorial $j$th power of $i$ [66], the $j$th fading factorial [93] or the backward factorial function of order $j$ [1]. Remembering that the binomial coefficients are written as $\binom{m}{j} = \frac{m!}{j!(m-j)!}$, Equation 2.19 can also be written as:

$$P_m(i) = \sum_{j=0}^{m} (-1)^j \binom{m+j}{j}^{(2j)} \frac{i^{(j)}}{(N-1)^{(j)}}.$$

(2.20)
Chapter 2: Discrete orthogonal polynomial approximation using Gram polynomials

The orthogonality of the polynomials is not affected by multiplication of the polynomials with a scalar. The general form for the polynomials can therefore be written as [66 p. 289]:

\[
P_m(i) = c_{mN} \sum_{j=0}^{m} \frac{(-i)^j (m+j)^{(2j)}}{(j!)^2} \frac{i^{(j)}}{(N-1)^{(j)}}
\]

(2.21)

The \(N\) points, \(i\), are defined over the interval \([0, N-1]\), and \(m\) is the polynomial order. The polynomials, with \(c_{mN} = 1\), from degrees one to four, inclusive, are:

\[
P_1(i) = 1 - 2\frac{i}{N-1}
\]

\[
P_2(i) = 1 - 6\frac{i}{N-1} + 6\frac{i(i-1)}{(N-1)(N-2)}
\]

\[
P_3(i) = 1 - 12\frac{i}{N-1} + 30\frac{i(i-1)}{(N-1)(N-2)} - 20\frac{i(i-1)(i-2)}{(N-1)(N-2)(N-3)}
\]

(2.22)

\[
P_4(i) = 1 - 20\frac{i}{N-1} + 90\frac{i(i-1)}{(N-1)(N-2)} - 140\frac{i(i-1)(i-2)}{(N-1)(N-2)(N-3)} + 70\frac{i^{(4)}}{(N-1)^{(4)}}
\]

Those polynomials are plotted in Figure 2.5 for 100 points and up to the sixth order, as a function of \(x\). It can be seen that the six polynomials shown in the graph, all have the same maximum value and all the odd polynomials have the same minimum value. These polynomials are much better scaled than the previous two.

![Gram polynomials](image)

**Figure 2.5:** Gram polynomials according to Equation 2.22 for 100 points. Plotted for the first to sixth order.
It can be shown that Equations 2.12–2.14 (the first form of the polynomials given) can be written as Equation 2.21 with

\[ c_{mN} = \frac{(-1)^m m! (N-2)^{(m-1)} 2^{m-1}}{(N-1)^{m-1} (2m-1)^{(m)}} \] (2.23)

and Equation 2.17 (the second form of the polynomials) can be transformed into Equation 2.21 by means of

\[ c_{mN} = \frac{(-1)^m m! (N-1)^{(m)}}{2(2m-1)^{(m)}} \] (2.24)

with its respective initial conditions.

A recursive form for the creation of Gram polynomials is given in [99] for \( N \) points, \( k \), in the interval \(- (N-1)/2 \) to \( (N-1)/2 \) as:

\[ \frac{1}{\epsilon_{m+1}} P_{m+1}(k) = \frac{k}{\epsilon_m} P_m(k) - \frac{\beta_m}{\epsilon_{m-1}} P_{m-1}(k) \]
\[ \beta_m = \frac{m^2 (N^2 - m^2)}{4(4m^2 - 1)} \]
\[ \epsilon_m = \frac{(2m)!}{(m!)^2 (N-1)^{(m)}} \] (2.25)

This can be transformed into the form [54]

\[ P_m(k) = \frac{(N-1)(2m-1)}{m(N-m)} \frac{2k}{N-1} P_{m-1}(k) - \frac{(m-1)(N-1+m)}{m(N-m)} P_{m-2}(k) \] (2.26)

for \( m \geq 1 \), with \( P_0(k) = 1 \) and \( P_{-1}(k) = 0 \). The nonrecursive form is given in [66, 99] as:

\[ P_m(k) = \sum_{j=0}^{m} \frac{(-1)^{j+m} (j+m)^{(2j)} (N-1)_j^2 + k^j}{(j!)^2 (N-1)^{(j)}} \] (2.27)

The first few polynomials, calculated with the recursive or the nonrecursive equations, are:

\[ P_1(k) = \frac{2k}{N-1} \]
\[ P_2(k) = 1 - \frac{N-1}{N-1} + k + 6 \frac{(N-1)_2 + k}{(N-1)(N-2)} \] (2.28)
\[ P_3(k) = -1 + 12 \frac{N-1}{N-1} + k - 30 \frac{(N-1)_2 + k}{(N-1)(N-2)} + 20 \frac{(N-1)_3 + k}{(N-1)^{(3)}} \]

\(^3\) It is sometimes easier to assume that \( N-1 \) is an even number, such that \( (N-1)/2 \) will be an integer, but it is not required in any of the analysis that follows.
Using the same index as usual, i.e. let $i$ run from 0 to $N - 1$, the recursive Equation 2.26 can be written as

$$P_m(i) = \frac{(N - 1)(2m - 1)}{m(N - m)} \left(1 - \frac{2i}{N - 1}\right) P_{m-1}(i) - \frac{(m - 1)(N - 1 + m)}{m(N - m)} P_{m-2}(i)$$

which gives the same polynomials as the general form (Equation 2.22) with $c_{mN} = 1$. A number of properties of DLOP’s are listed in [93] and it also lists an algorithm for constructing the polynomials with a minimum number of operations. The algorithm is as follows:

**Initialization:**

\[n = N - 1\]

\[c_0 = 0\]

\[c_1 = n - i - i\]

\[c_2 = n\]

\[\delta_0 = n\]

\[\delta_1 = c_1 + c_1\]

\[\delta_2 = n\]

\[P_0 = 1\]

\[P_1 = c_1/n\] (2.30)

**Iterations for $m = 2, 3, \ldots$**

\[\delta_0 = \delta_0 + 2\]

\[\delta_2 = \delta_2 - 2\]

\[c_0 = c_0 + \delta_0\]

\[c_1 = c_1 + \delta_1\]

\[c_2 = c_2 + \delta_2\]

\[P_m = (c_1 P_{m-1} - c_0 P_{m-2})/c_2\] (2.31)

Since $m$, $N$, and $i$ are integers, the coefficients $c_0, c_1, c_2$ and $\delta_0, \delta_1, \delta_2$ will also be integers. The total number of operations required for this scheme to calculate a single point, $i$, is:

\[m\] real divisions

\[2m - 2\] real multiplications

\[m - 1\] real subtractions

\[4m - 3\] integer additions

\[m + 1\] integer subtractions

19
The other $N - 1$ points would need the same number of real operations, however if the integer additions and subtractions are stored, the number of integer subtractions would only be two for each point and the number of integer additions would be $m$. This scheme does not require any integer multiplications like Equation 2.29, and does not require $m^2$ operations as in Equation 2.19. There is also an algorithm to compute the multiple of the binomial coefficients in Equation 2.19, \( \binom{m}{j} \binom{m+j}{j} \), by forming a simple matrix using addition only, similar to the formation of Pascal's triangle, and then multiplying two elements of the matrix to compute the coefficients [1].

As a last resort, tables of orthogonal polynomials can be used. Fisher and Yates [49] give tables of integer orthogonal polynomials for equally spaced points for order 1 to 5 and number of data points from 3 to 75. They mention that more extensive tables are available for up to 104 data points and also for order up to 9 but only for 52 data points. With today's computers it is quite easy to generate all necessary integer polynomials, or scaled such that the first or the last value of every polynomial is equal to 1. Care must be taken when dealing with limited word lengths to represent large numbers since round-off errors might become a problem.

Unless stated otherwise, in this thesis, Equation 2.29 is used in order to create the Gram polynomials, since it will produce polynomials that are all of comparable magnitude as long as the number of data points is large enough compared to the polynomial order. Figure 2.6 shows a plot of the first 30 polynomials for 100 data points. It can be seen that the higher order polynomials become larger and larger as the polynomial order gets close to the number of discrete data points. Figure 2.7 shows at what order the maximum of the absolute value of the polynomials is greater than 2, 10, and 100 for number of data points between two and 1000 points. It will be shown in chapter 5 how this large absolute value for a polynomial will adversely affect the controller output.

2.1.2 Gram goes to Legendre

It has already been mentioned that a certain form of the Gram polynomials has exactly the same parameters as the shifted Legendre polynomials. The Gram polynomials with $c_{m,N} = 1$ will also converge to the continuous Legendre polynomials for infinite number of data points [66]. It is
observed that the Gram polynomials (Equation (2.22)) can be written as:

\[
\begin{align*}
P_1(i) &= x(i) = 1 - \frac{i}{N - 1} \\
P_2(i) &= \frac{1}{2} \left\{3(P_1(i))^2 \frac{N - 1}{N - 2} - \frac{N + 1}{N - 2}\right\} \\
P_3(i) &= \frac{1}{2} \left\{5(P_1(i))^3 \frac{(N - 1)^2}{(N - 2)^2} - 3(P_1(i)) \frac{N^2 - \frac{7}{3}}{(N - 2)^2}\right\} \\
P_4(i) &= \frac{1}{8} \left\{35(P_1(i))^4 \frac{(N - 1)^3}{(N - 2)^3} - 30(P_1(i))^2 \frac{(N - 1)(N^2 - \frac{13}{2})}{(N - 2)^3} + 3 \frac{(N + 3)(N + 1)}{(N - 4)(N - 2)}\right\}
\end{align*}
\]

In the limit when \(N\) goes to infinity the fractions involving \(N\) in the above equations go to one and the above equations can be written as:

\[
\begin{align*}
P_1(x) &= x \\
P_2(x) &= \frac{1}{2} \{3x^2 - 1\} \\
P_3(x) &= \frac{1}{2} \{5x^3 - 3x\} \\
P_4(x) &= \frac{1}{8} \{35x^4 - 30x^2 + 3\}
\end{align*}
\]
Chapter 2: Discrete orthogonal polynomial approximation using Gram polynomials

Figure 2.7: Relationship between the number of data points and the order at which the maximum absolute value of the polynomial is greater than 2, 10 and 100. The circles in the lower left hand corner denote a fourth and an eighth order polynomial for 20 data points.

These equations are exactly the same as the Legendre equations on the interval \([-1, 1]\) [11]. Figure 2.8 shows the fourth and eighth order Gram polynomial for 20, 30, 50, and 100 points as well as the continuous Legendre polynomial. It can be seen that, as the number of data points is increased, the Gram polynomial approaches the continuous Legendre polynomial. Furthermore, the 20 point Gram polynomial is much closer to the fourth order Legendre polynomial than it is to the eighth order Legendre polynomial, because fourth order is small in comparison to 20 points whereas eighth order is not, as can be see from the circles in the lower left hand corner in Figure 2.7. The requirement for Equation 2.32 to be written as Equation 2.33, i.e. for Gram to go to Legendre, is then \( N \gg m \).

2.2 Approximations by Gram polynomials

By approximating the CD profile with polynomials, the shape of the profile can be controlled, and the number of parameters to control can be greatly reduced. This amounts to controlling the low frequency components of the profile. This idea has been used in Sendzimir steel rolling mill during research done at the University of Strathclyde [101, 102] and the University of Sheffield [41, 56].

Now define a matrix, \( x_0 \), consisting of the first \( m \) orthogonal polynomials (it is assumed that the mean value of the profile to estimate is zero, i.e. \( P_0 \) does not need to be included) evaluated
Chapter 2: Discrete orthogonal polynomial approximation using Gram polynomials

Figure 2.8: Gram goes to Legendre. The smooth solid line is the Legendre polynomial in both plots.

at \( N \) points as (see Equation 2.2):

\[
x_0 = [P_1 \, P_2 \, \cdots \, P_m]^T
\]

The relationship between the profile, \( y \), and the parameters of the orthogonal polynomials, \( y_c \), is given by

\[
y(t) = x_0^T y_c(t) + \epsilon(t)
\]

where \( \epsilon \) consists of terms of order \( m + 1 \) and higher (more specifically, order \( m + 1 \) to \( N - 1 \), because a polynomial of order \( N - 1 \) will exactly match the measurements at the \( N \) data points). The least squares estimate for the \( y_c \) vector is:

\[
\hat{y}_c(t) = (x_0 x_0^T)^{-1} x_0 y(t)
\]

The matrix \( x_0 x_0^T \) is a diagonal matrix because it consists of orthogonal vectors, so its inverse is easily calculated. In the case of orthonormal polynomials, the matrix simply becomes the identity
Chapter 2: Discrete orthogonal polynomial approximation using Gram polynomials

The least squares estimate (Equation 2.36) can be written as:

\[
\hat{y}_c(t) = \begin{bmatrix}
\frac{1}{P_1^T P_1} & 0 & \cdots & 0 \\
0 & \frac{1}{P_2^T P_2} & \cdots & \vdots \\
\vdots & \ddots & \ddots & \vdots \\
0 & \cdots & 0 & \frac{1}{P_m^T P_m}
\end{bmatrix}
\begin{bmatrix}
P_1 \\
P_2 \\
\vdots \\
P_m
\end{bmatrix}^T y
\]

(2.37)

\[
= \begin{bmatrix}
\frac{1}{P_1^T P_1} P_1^T y \\
\frac{1}{P_2^T P_2} P_2^T y \\
\vdots \\
\frac{1}{P_m^T P_m} P_m^T y
\end{bmatrix}
\]

It is clear from the above equation that the elements of the \( y_c \) vector can be calculated independently of each other. If more parameters are needed the next parameter can be calculated from:

\[
\hat{y}_c(t, m + 1) = \frac{1}{P_{m+1}^T P_{m+1}} P_{m+1}^T y
\]

(2.38)

where the second variable in \( \hat{y}_c(t, m + 1) \) denotes the \( m + 1 \) element in the parameter vector. Hence, more parameters can be added without affecting the ones already calculated.

Figure 2.9: Approximation of an industrial basis weight CD profile with Gram polynomials to the 10th order. The small inserted graph shows the 10 parameters of the Gram polynomial approximation.
Chapter 2: Discrete orthogonal polynomial approximation using Gram polynomials

Figure 2.10: Measured data and 4, 10, 20 and 40th order approximation using Gram polynomials.

An example of approximating an industrial basis weight CD profile, with the polynomials defined by Equation 2.21 and $c_mN = 1$, is shown in Figure 2.9. The measured profile, $y$, is drawn with a solid line connecting the measured points. The inset bar graph in the upper left corner shows the values of the estimated parameters, $y_e$. The estimated profile, $\hat{y}$, is drawn in the big graph with a dashed line. Here, the order of polynomial approximation is chosen arbitrarily, but there exist methods that could be employed to estimate the optimal polynomial order for a given profile. Some of these methods will be discussed below. The particular profile in Figure 2.9 has been approximated with different polynomial orders and the results are shown in Figure 2.10. The parameters needed for approximation are shown in Figure 2.11. From this plot, it can be seen that little improvement in approximation is obtained in using polynomials from the 22nd order to 40th order because most of these parameters are small in comparison with those for the 21st and 22nd order. Furthermore, the plot of $2\sigma$, twice the standard deviation of the approximation error, versus order, shows that the largest drops are indeed where the parameters are the biggest. The figure shows that the standard deviation gets smaller and smaller with increasing polynomial order, and it will in fact be zero when the order of the polynomials is equal to one less than the number of data points. In practice, the
measured data has noise and it is not desirable to approximate noise. In the words of Hildebrand [66]:

"it is foolish to attempt to determine a polynomial of high degree which fits the vagaries of such data exactly and hence, in all probability, is represented by a curve which oscillates violently about the curve which represents the true function."

The measured profile, together with the approximation error for a 40th order approximation, is shown in Figure 2.12. It can be seen that the low frequency components of the measured profile have disappeared and the peak to peak value is lower. Therefore, if the lower order polynomials are removed from the profile, the profile becomes more uniform, and the problem with different averages for different sections of the profile — a major concern in CD control — is eliminated.

2.2.1 Autocorrelation

Assuming the true profile is an orthogonal polynomial and the measured profile is the true profile mixed with white noise, the “whiteness” of the residues could be used to calculate the optimum order of polynomials to use. If the residues are white the autocorrelation function would have a value of one at zero lag and a value of zero for all other lags. For finite length sequences the value of the
autocorrelation sequence for non zero lag will not be exactly zero. Therefore, a confidence interval or error bound is calculated which the autocorrelation should fall within. The confidence interval depends on the sample size used. A widely used measure for autocorrelation functions is to use 95% confidence interval, for number of data points greater than 60 it is $2/\sqrt{N}$, where $N$ is the number of data points. If the autocorrelation function is white, it can be said that one out of 20 points in the autocorrelation function can lie outside the 95% confidence interval by pure chance. A part of the autocorrelation function of the four approximations from Figure 2.10 is shown in Figure 2.13. The percentage of the autocorrelation of the residues for the 4th, 10th, 20th, and 40th order approximation, that is inside the 95% confidence bounds, are 88%, 91%, 90%, and 89% respectively. None of the approximation errors can be considered white and the whiteness does not seems to be related to the number of polynomials used. The whiteness of the residues is therefore not a good measure of fit unless something is known about the noise.

2.2.2 Akaike’s Information Criterion

A perfect fit can be obtained if a polynomial order is chosen as one less than the number of data points. Such a model would be overly complex and would approximate the noise. A balance should
therefore be found between the complexity (parsimony) of the model and the quality of the fit. A model should be as simple as possible, yet sufficiently flexible to cover all aspects of the process.

The most natural way to search for a suitable model structure is simply to test a number of different ones and compute the resulting models. One formalized way has been pursued by Akaike (see for example [3]). For model errors that are Gaussian with unknown variance, the Akaike’s information theoretic criterion (AIC) is written as

$$AIC(m) = \log \left( \frac{1}{N} \sum_{i=1}^{N} (y - \hat{y}(m))^2 \right) + \frac{2m}{N} \quad (2.39)$$

where $N$ is the number of data points, $m$ is the number of parameters and $\hat{y}(m)$ is the profile approximated with an $m$th order polynomial. The first term in the above equation measures approximation error and will decrease with increasing model complexity; the second term accounts for model complexity and increases sequentially. The idea is that as the approximation error becomes smaller with increasing complexity, a minimum is reached where reduction in prediction error is more than offset with the increased model complexity. Figure 2.14 (the solid line) shows a plot
Chapter 2: Discrete orthogonal polynomial approximation using Gram polynomials

Figure 2.14: Akaike’s Information Criterion for the profile from Figure 2.9. The solid line shows the AIC, the other two lines show the AIC when it is modified to greater penalize model complexity. The circles in the plot denote the minimum for the curves.

of the AIC for the profile from Figure 2.9. It is found that the optimum polynomial order for this particular profile is 86. If the order is increased further the increased complexity does not warrant the reduction in the approximation error. A deficiency in the AIC is that it only penalizes the addition of parameters in calculating the approximation error. For a general ARMA type model an increase in the order of the $A$ and/or the $B$ polynomial is usually required when increasing model complexity. This case only requires the input and output to be stored for one extra time step. On the other hand, when a polynomial order is added to a model using orthogonal polynomials, it is also necessary to calculate the new polynomial using a recursive equation. If we look at the complexity of a model as a cost associated with computing, it is necessary to factor this complexity into the AIC. If it is assumed that the cost of calculating the new polynomial is the same as the cost of increasing model parameter vector, the dashed line in Figure 2.14 is obtained. This criterion is used in [83] to reduce the probability of selecting a model with one more parameter than the true parameter vector. It can be seen that the optimum polynomial order is now 22. If the cost of creating an extra polynomial is assumed to be double that of increasing the parameter vector, the dotted line is Figure 2.14 is obtained, which has a minimum at polynomial order of 10.

To guard against approximating the noise in the profile, a polynomial of high degree should
not be used [66]. It will be shown in the next chapter how the condition number of a transfer matrix worsens when the order of the polynomial approximation is increased. Clearly the AIC in its original form gives too many parameters. The form suggested in [83] gives a good balance between complexity and fit.

2.3 Frequency characteristics of Gram polynomials

The frequency characteristics of the Gram polynomials were investigated using Fourier analysis. It was found that, unlike the amplitude of the polynomials, the frequency characteristics do not depend on the number of data points used because the number of zero-crossings remains the same. For this analysis, a 100th order polynomial is investigated and $2^{12} = 4096$ data points were chosen because it is convenient to calculate fast Fourier transform (FFT) for sequences that have numbers of elements that are powers of two. If the zeros of the polynomial are equidistant, the polynomial would be expected to have its major frequency around $(100-1)/2 = 49.5$ cycles per profile.

Figure 2.15 shows the left half of the 100th order Gram polynomial defined for 4096 points (the right half is identical because even order polynomials are symmetric). The left half is then further divided into 4 quarters and a frequency distribution for each part was calculated. Figure 2.16 shows the spectrum of the 100th order Gram polynomial defined over 4096 points in the interval $[-1,1]$. A 100th order polynomial has $\frac{100-1}{2} = 49.5$ periods of oscillations. It is seen in Figure 2.16 that the polynomial has most of its frequencies distributed slightly below 49.5 periods over its entire profile. The reason that it is lower than 49.5 periods is that the "edges" of the polynomials have slightly higher frequencies, i.e. more cycles per sheet width. This is evident in the leftmost quarter in Figure 2.15. Therefore a large portion of the 49.5 cycles are used for the first quarter. The remaining three quarters will have fewer cycles as can be seen in Figure 2.17, where the spectra is calculated for the four different 512 points sections of the left half of the profile. The edge quarter from Figure 2.15 is shown to have frequency component corresponding to about 65 periods over the sheet width, whereas the other three quarters have frequencies corresponding to between 30 and 40 cycles per sheet width.

To show how the Gram polynomials can be used to remove some frequencies, a 1000-point profile is created which consists of 5 sinusoidal waves with different frequencies (1, 4, 10, 30, and 100 cycles), added together. The dotted line in Figure 2.18 shows the profile created. The solid line shows the profile when the first 20 Gram polynomials have been removed from the profile. The bottom figure shows the frequencies in the original profile (the dotted line) and the profile with the
Chapter 2: Discrete orthogonal polynomial approximation using Gram polynomials

Figure 2.15: Left half of 100th order Gram polynomial for 4096 points. It is shown with the polynomial partitioned into 8 zones, i.e. the left half is partitioned into 4 zones, separated with different line types in the graph.

Figure 2.16: Spectrum of 100th order polynomial.
2.4 Summary

In this chapter discrete orthogonal polynomials with equal weight given to every data point, Gram polynomials, have been described. It is shown how a set of data, a profile for example, can be approximated with Gram polynomials. Several methods are given to create the Gram polynomials, using both sequentially and recursive algorithms. A unified sequential formula is given for all the methods. It is observed that the Gram polynomials will approximate the Legendre polynomials when the order of the polynomials is much smaller than the number of data points which gives the Gram polynomials another name, discrete Legendre orthogonal polynomials. Some methods are discussed
Chapter 2: Discrete orthogonal polynomial approximation using Gram polynomials

Figure 2.18: The original profile (dotted line) and the residual profile (solid line) when a 20th order Gram polynomial is removed. Power spectrum for the original profile (dotted line) and the profile when 20th order Gram polynomial is removed (solid line).

Figure 2.19: The original profile (dotted line) and the residual profile (solid line) when a 40th order Gram polynomial is removed. Power spectrum for the original profile (dotted line) and the profile when 40th order Gram polynomial is removed (solid line).
for finding the optimum number of polynomials to approximate a profile. It is found that a modified AIC gives a good balance between complexity and fit. Finally the frequency characteristics of the Gram polynomials are investigated and they are found to have their main frequency component slightly lower than half their order.
Chapter 3

Model for cross-directional basis weight response

A basis weight control system is made up of three main parts: i) the sensor, a gauge which scans the paper sheet, ii) the process control computer, and iii) the slice lip actuators [106]. The process control computer uses the information obtained by the scanning sensors to calculate the new slice lip actuator settings.

3.1 Paper machine

A paper machine, on which stock is drained and dried, is used to manufacture a sheet of paper of infinite length and up to 10 m wide. Thick stock, a water/fibre mixture with a fibre concentration around 2%, comes from the machine chest. The thick stock is diluted with white stock and sent to the headbox where the concentration is about 0.5%. The purpose of the headbox is to take the inlet stock flow and uniformly distribute the proper amount of fiber suspension across the paper machine moving forming wire at the right speed and in a uniform way, through an opening which is between 50 – 100 mm [98]. The product coming out of the paper machine will almost reverse the ratio of water to fibre, since typical newsprint paper contains 5% moisture. There are two ways of forming the sheet on the moving wire, either by a single wire (fourdrinier) or by a twin-wire former, which has wires on both sides of the forming sheet. The sketch of a simplified paper machine in Figure 3.20 shows a Fourdrinier machine. On the wire the fibres form into a sheet after water drainage using gravity and suction, and in the case of twin wire, also by pressing. On the fourdrinier wire there is a visible line of demarcation corresponding to the point where a glossy layer of water on the top of the stock suddenly fades away [112]; this line is called the dry line. The sections of the profile with high basis weight will require a longer time to dry than parts with low basis weight. Therefore, if the dry line is measured it could be used to predict the basis weight profile. In the dryer section the fibre bonds develop and more of the remaining water is evaporated by heating. In the calender section the sheet is pressed between metal rolls to reduce thickness and smooth the surface. Finally the paper is wound onto a reel [111].

Moisture, defined as percent water by weight, and basis weight (BW), defined as weight per unit area, are two of the most important characteristics of paper. BW includes both the weight of the
water and of the fibres, therefore, in order to measure the amount of fibres in the paper the weight of the water has to be subtracted to obtain the dry weight.

Burkhard and Wrist [16] proposed, a now widely accepted, classification of basis weight variations in three components:

1. The machine direction (MD) component which is time varying and independent of the location across the machine
2. The cross direction (CD) component which describes the profile and changes much slower than the MD.
3. The remaining signal is the random (R) component and varies in both time and CD position.

Lately, some researchers have also started looking at variations in the vertical or thickness direction (ZD) [29]. The variations in ZD are primarily caused by movement of fines, the small fibres, and fillers, used for creating a smooth surface.

The paper characteristics are measured with sensors mounted on an O-frame, i.e. a guiding rail on which they travel back and forth across the machine (see Figure 3.21). Due to the MD movement of the paper and the CD movement of the sensor, the measurements form a zigzag pattern on the paper.
sheet as shown in Figure 3.21. The measurements are thus a composite of the CD and MD variations, from which the CD and MD profiles must be extracted. A simple way is through an exponential filter. Let \( x^n(t) \) denote the raw measurement obtained at the measurement box \( n \) during scan \( t \), then

\[
y^n(t) = (1 - a)y^n(t - 1) + ax^n(t)
\] (3.40)

where \( y^n(t) \) is the estimated profile for box \( n \) and scan \( t \), and \( 0 < a < 1 \) is the exponential filter weighting factor [33, 40]. The MD profile can be calculated as the average of \( y^n(t) \) over one scan and the CD profile would then be the deviation of \( y^n(t) \) from the average. This method of separating the MD and CD is very slow because a small \( a \) is used to separate the MD and CD. Consequently, control action is taken every 3-5 scans so that the result of the last control action is clearly seen before a decision is made to do another control. Over the last few years, methods have been proposed in order to estimate the moisture CD profile [20, 23, 39, 40, 70, 92] and the basis weight CD profile [122] more quickly. Vendors have increased the speed of the scanner and sample the profile in smaller and smaller data boxes in their quest to reduce the variability of the paper and to better separate the CD and MD [115]. Using analysis of variance, some have come up with a method to alternatively decrease or increase the speed of the scanner in order to get the best measurements [2]. Others have used analysis of variance to quantify what can and cannot be done to reduce the variations [114]. Recently, MD variations have been estimated within a scan, and hence faster MD variations are reduced, resulting in overall reduction of MD variations [118, 120]. The algorithm in [70] consists of a modified least-squares parameter identifier for estimating cross direction profile deviations, a Kalman filter for estimating machine direction disturbances, and takes into account the nonlinear relationship between the MD and CD observed by Lindeborg [84, 85]. It is observed for basis weight [122] that there is no nonlinear relationship between the MD and CD, and that the model for MD variations is a second order model instead of a first order model, as is the case for the moisture model.

Basis weight on a paper machine is routinely controlled in the cross direction by bending the slice lip of the headbox. Because of coupling, movement of a single actuator will cause the slice to be deflected between several actuators, even though the actuators settings stay the same for the neighbouring actuators (see Figure 3.22). Hence, moving one actuator of the slice lip will in general affect more than one measurement point in the cross direction. Another source of coupling is waves on the fourdrinier caused by the slice lip opening. Disturbances can also be caused by the wake effect within the headbox [61, 62]. Therefore, two steps are needed to model the response of an actuator movement on the basis weight. The first stages involves modelling the slice lip and its deflection to
Chapter 3: Model for cross-directional basis weight response

A movement of a single actuator. The slice lip deflection is then calculated to the same resolution as the basis weight is measured. Slice lip deflection will cause a wave to form on the forming fabric and this wave will travel in the CD direction and spread further (Figure 3.22). At present, no models exist to describe that wave and consequently, identification of the actuator response to the measured basis weight is needed. Therefore, the paper machine is usually represented by a block diagonal matrix with number of rows equal to the number of basis weight measurements, and number of columns equal to the number of slice actuators. The block diagonal has to be identified. Usually, several responses across the machine are identified. Then it is assumed that the response to a slice lip movement is the average of those several responses, and is therefore the same, regardless of where it is on the machine (with the possible exceptions of the edges). It will be shown in this chapter that it is not possible to have the same response unless the number of measurements is an integer multiple of the number of actuators. The interaction matrix can have eigenvalues close to zero; therefore, its inverse, which could be used in decoupling the actuators, is not easily calculated.

![Figure 3.22: Response to a slice lip movement at one actuator position. The slice lip deflection is denoted by • with ○ denoting the deflection at a slice lip actuator. The + marks what the basis weight might look like.](image)

3.2 Basis Weight

It is claimed in [96] that the dominant constituent of the response is the result of the water waves created on the fourdrinier wire originating from the slice lip deflection, and setting somewhere before the dry line. As said before, no models currently exist in modelling the waves that appear on the
wire in the paper machine. For that reason the response model from the slice lip deflection to the basis weight has to be identified from the input-output data.

It was noted early on by researchers looking into CD control problems that the process response is not the same in the center of the paper machine as it is at the quarter points or at the edges [106]. The main reason is the asymmetric deflection of the slice lip at the edges. Part of the edge phenomena also relates to waves being bounced back from the deckle board, the board that is used to contain the stock on the wire. A method to reduce or eliminate the deckle waves is reported in [98].

One distinctive characteristic of all paper machine cross directional models is their large dimension. To measure basis weight, a beta ray gauge traverses the paper, measuring at $N$ points across the sheet. Typically $N$ can be between 50 and 500 although recent systems can provide as much as 2000–3000 points. Those measurements are then used to calculate settings for up to 100 actuators located upstream at the slice lip of the headbox (see Figure 3.20). The calculation of the actuators settings is based on CD response models, which have three common elements: actuators dynamics, time-delay, and interactions.

It is commonly assumed that all actuators can be modeled by the same first-order lag dynamics

$$\frac{k_a}{1 - p_a q^{-1}}$$

where $k_a$ is the gain and $p_a$ is the time constant. The sampling method currently employed on paper machines uses a sampling period that is much longer than this time constant. Hence, the actuator dynamics can be ignored in most cases, simply because it is not possible to observe it. Since the basis weight measurements are taken at the finishing stage of the paper machine, and the slice lip actuators are located on the head box, some time will pass before changes made at the slice will affect the measurements. Given that the transportation lag is written as $q^{-d}$, the transfer function from an actuator to its corresponding measurement would be:

$$g_a(q^{-1}) = \frac{k_a}{1 - p_a q^{-1}} q^{-d}$$

Control action by one actuator generates water waves on the wire which influences the basis-weight at several locations around the actuator. The number of locations affected depends on the type of paper being manufactured — heavier paper produces wider interactions [72]. Assuming the same
spatial response for all actuators, the response matrix, $G$, will therefore be banded diagonal:

$$G = \begin{bmatrix}
    p_0 & p_1 & \cdots & p_k & 0 & \cdots & 0 \\
p_1 & p_0 & p_1 & \cdots & \cdots & \cdots & \vdots \\
\vdots & p_1 & \cdots & \cdots & \cdots & \cdots & 0 \\
p_k & \cdots & \cdots & \cdots & \cdots & \cdots & p_k \\
0 & \cdots & \cdots & \cdots & p_1 & \cdots & \cdots \\
\vdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
0 & \cdots & 0 & p_k & \cdots & p_1 & p_0 & p_1 & p_0
\end{bmatrix}$$  \hspace{1cm} (3.43)

Here, it is also assumed that the response is symmetrical such that the elements above and below the diagonal are the same, and the width of the CD response is $2k + 1$ measurement data boxes. It is desirable to emphasize the fact that the matrix need not be symmetrical. The elements below the diagonal would then be subscripted with a minus. The input output relationship is:

$$y(t) = G \frac{k_a}{1 - p_0 q^{-1}} q^{-d} u(t)$$  \hspace{1cm} (3.44)

Moving an actuator flexes the slice lip and puts waves on the wire. The slice lip deflection can be modeled and will contribute to part of the structure of the interaction matrix. The deflection can be measured directly on the slice or observed in the measurements if there are more measurements than the number of actuators. The measurements will, however, contain more than just the deflection of the slice. They will also contain the waves produced on the wire by slice lip deflection. Work is under way at the Pulp and Paper Centre to model the wave process. It should be emphasized that the interaction matrix does not only describe the slice lip deflection but also the waveform that travels in the cross direction as a result of the slice lip deflection.

Several interaction matrices have been published or can be read from low resolution graphs in the literature [14, 24, 21, 28, 64, 72, 100, 116, 119, 126, 127]. Thirteen such matrices are reported in [79] for a wide range of paper machines, from light newsprint to heavy cardboard.

<table>
<thead>
<tr>
<th>$p_0$</th>
<th>$p_1$</th>
<th>$p_2$</th>
<th>$p_3$</th>
<th>$p_4$</th>
<th>$p_5$</th>
<th>$p_6$</th>
<th>$p_7$</th>
<th>$p_8$</th>
<th>$p_9$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0</td>
<td>0.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.0</td>
<td>0.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.0</td>
<td>0.5</td>
<td>-0.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.0</td>
<td>0.1</td>
<td>-0.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.0</td>
<td>0.45</td>
<td>-0.55</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.0</td>
<td>0.4</td>
<td>-0.5</td>
<td>0.05</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.0</td>
<td>-0.15</td>
<td>0.03</td>
<td>-0.01</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.0</td>
<td>0.2</td>
<td>-0.1</td>
<td></td>
<td>-0.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.0</td>
<td>0.4</td>
<td>-0.2</td>
<td>-0.4</td>
<td>-0.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.0</td>
<td>1.2</td>
<td>0.6</td>
<td>-0.4</td>
<td>-0.9</td>
<td>-0.2</td>
<td>-0.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.0</td>
<td>1.3</td>
<td>0.8</td>
<td>-0.6</td>
<td>-0.3</td>
<td>0</td>
<td>-0.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.0</td>
<td>0.9</td>
<td>0.7</td>
<td>0.8</td>
<td>1.0</td>
<td>0.6</td>
<td>-0.5</td>
<td>-0.4</td>
<td>-0.2</td>
<td>-0.2</td>
</tr>
</tbody>
</table>
Figure 3.23: CD responses, according to Laughlin’s interpretation from the literature [79], see also Equation 3.45.

The corresponding responses to a slice lip actuator movement are shown in Figure 3.23 for the first nine responses (left figure), and the last three responses (right figure). It can be seen that the lighter grades have a negative side lobe in most cases but the response is quite narrow compared to the heavy weight paper where the center of response is not necessarily the largest response [21].

One of the major problems with banded diagonal matrices is that they can easily approach singularity — their condition number can grow indefinitely. The condition number of a matrix is the ratio between its largest and smallest singular value. Singular values are calculated for real valued matrices using:

$$\sqrt{\text{eig}(G^T G)}$$  \hspace{1cm} (3.46)

where \(\text{eig}(\cdot)\) is the eigenvalue function. For example, the determinant of a \(5 \times 5\) matrix with \(p_0 = 1\), \(p_1 = 2r\), and \(p_2 = -r\), where \(r\) is a parameter used to vary the magnitude of the side lobe, is:

$$\begin{vmatrix}
1 & 2r & -r & 0 & 0 \\
2r & 1 & 2r & -r & 0 \\
-r & 2r & 1 & 2r & -r \\
0 & -r & 2r & 1 & 2r \\
0 & 0 & -r & 2r & 1
\end{vmatrix} = 56r^5 + 42r^4 - 24r^3 - 19r^2 + 1 = 56(r - 0.6404)(r + 0.7985)(r + 0.4163)(r + 0.3909)(r - 0.2149)$$  \hspace{1cm} (3.47)
If \( r \) is equal to any of the roots of the determinant then this matrix will be singular and its condition number is infinite. Reciprocal condition numbers (condition numbers between zero and one, where zero indicates singularity) can be plotted for a matrix with \( p_0 = 1, p_1 = 2r, \) and \( p_2 = -r \) where \( r \) goes from 0 to 0.5 for different sizes of the \( G \) matrix (see Figure 3.24). It can be seen when there is any significant cross coupling it is very likely that the matrix will be ill conditioned, i.e. reciprocal condition number close to zero. For a similar plot of condition numbers of \( G \) matrix with \( p_{0,1,2} = [1, r, -r] \) see [79]. Processes with high condition number can be difficult to control, especially in the presence of uncertainty in the interaction matrix. Because of the uncertainty, the direction of a large input may not correspond exactly to the low plant-gain direction, and the amplification of these large input signals may be much larger than expected from the model. It can be shown that disturbances entering the process in the direction of the minimum singular vector for a process with bad condition number can lead to a poor performance, because the control action can be taken in the wrong direction [107]. In the limit as the \( G \) matrix becomes singular, the plant is uncontrollable, as the actuator settings cannot be determined based on the measurements [79].

Singularity is not restricted to the two forms of the interaction previously mentioned, it is also of concern in matrices of the form \( p_{0,1} = [1, r], p_{0,1} = [1, -r], \) and \( p_{0,1,2} = [1, 2r, r] \). In later chapters, a paper machine with 105 actuators will be employed, it is of interest therefore to look at a 105 \( \times \) 105 matrix with \( p_0 = 1, p_1 = 2r, \) and \( p_2 = -r \). Its reciprocal condition number for \( r \) between 0.15 and 1.0 is plotted in the upper half of Figure 3.25. The corresponding CD response for cases where the matrix is singular is plotted in the lower half of Figure 3.25. These responses are typical of those reported in the literature for paper machines. Therefore it can be concluded that singularity of the interaction matrix can quite easily happen on paper machines.

It can be seen from industrial data that the model of the slice lip movement to the dry weight measurements is simply a gain and a delay. This is shown in Figure 3.26 where the mean values are removed from the input and the output. The solid line represents the output or the measured basis weight in one CD box, and the dotted line represents the input or the slice lip adjustment for one actuator. The model for this paper machine at its operating conditions can therefore be written as:

\[
y = p_0 q^{-d} u
\]  

(3.48)

If the interactions of adjacent CD locations are included in the model it becomes:

\[
y = G q^{-d} u
\]  

(3.49)
Figure 3.24: Reciprocal condition numbers of $G$, where it is created according to equation 3.43 with $p_0 = 1$, $p_1 = 2r$ and $p_2 = -r$, for various sizes of the matrix. When $G$ is badly conditioned, the reciprocal condition number is close to 0.

Figure 3.25: Reciprocal condition number for a $105 \times 105$ matrix with $p_0 = 1$, $p_1 = 2r$ and $p_2 = -r$. Basis weight response for a machine with a singular interaction matrix, reciprocal condition number equal to zero.
where $G$ is a block diagonal matrix. The $G$ matrix will be estimated in chapter 4. Now we will look at some of the forms the $G$ matrix will take.

3.2.1 Square interaction matrix

When the number of actuators is equal to the number of measurements, the $G$ matrix will be square. There are two main problems associated with this matrix form that are also prevalent for a nonsquare $G$ matrix. The most serious problem is mapping, that is, determining which actuator belongs to which data box. Another problem is how to model the edges. If the edges are not accounted for, the model might show some edge effects that are not present in the real process. As an example, assume that $G$, defined by Equation 3.43, is a $20 \times 20$ matrix with $k = 3$ or:

$$
G = \begin{bmatrix}
  p_0 & p_1 & p_2 & p_3 & 0 & 0 \\
  p_1 & p_0 & p_2 & p_3 & 0 & 0 \\
  p_2 & p_1 & p_0 & p_3 & 0 & 0 \\
  p_3 & p_2 & p_1 & p_0 & \cdots & \cdots \\
  0 & p_3 & p_2 & p_1 & p_0 & p_3 \\
  \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
  0 & 0 & p_3 & p_2 & p_1 & p_0 \\
  0 & 0 & 0 & p_3 & p_2 & p_1 \\
  0 & 0 & 0 & 0 & p_3 & p_2 \\
  0 & 0 & 0 & 0 & 0 & p_3 \\
\end{bmatrix}
$$

(3.50)

where the edge elements of the interaction matrix that might need adjustments are bold faced. If the actuators are set such that the slice is a straight line from $-1$ to $1$, the resulting basis weight profile,
Chapter 3: Model for cross-directional basis weight response

Figure 3.27: Plot of the weight response to a straight line slice lip profile. The upper half is using $G$ matrix from Equation 3.50 and the lower half is if the edges of the matrix are made smooth.

with $p_0 = 1$, $p_1 = 0.6$, $p_2 = 0.2$, and $p_3 = -0.2$, chosen arbitrarily, will have ripples at the ends (upper half of Figure 3.27). The ripples are artifacts which arise because the sum of the middle rows of the $G$ matrix is equal to $p_0 + 2(p_1 + p_2 + p_3) = 2.2$ whereas the first three rows have column sum equal to 1.6, 2.2 and 2.4, respectively. If the $G$ matrix is fixed such that the first three elements in the first column (bold faced in Equation 3.50) are set to $p_0 + p_1 + p_2 + p_3$, $p_1 + p_2 + p_3$, and $p_2 + p_3$ (and the same thing is done for the other edge), the effects of the edges of the matrix are reduced somewhat (lower half of Figure 3.27). However, the ripples are not completely removed from the output profile. To do so, it would require tuning of the $G$ matrix based on the input signal.

3.2.2 Nonsquare interaction matrix

When the number of actuators does not match the number of measurements the interaction matrix will be nonsquare for the unmapped measurements. Again, the edges of the interaction matrix might have to be identified separately otherwise, the model might create some phantom ripples at the ends of the BW profile. For control purposes it is very important that an accurate mapping of the slice screw positions to the measurements data boxes of the weight profile is established [96]. If the number of data boxes is an odd integer multiple of the number of actuators, and the actuators are
perfectly aligned to some measurements, the interaction matrix is of the form:

\[
\begin{bmatrix}
 p_{-5} & p_{-1} & p_3 & p_7 \\
 p_{-6} & p_{-2} & p_2 & p_6 \\
 p_{-7} & p_{-3} & p_1 & p_5 \\
 p_{-4} & p_0 & p_4 \\
 p_{-5} & p_{-1} & p_3 & p_7 \\
 p_{-6} & p_{-2} & p_2 & p_6 \\
 p_{-7} & p_{-3} & p_1 & p_5 \\
 p_{-4} & p_0 & p_4 \\
 p_{-5} & p_{-1} & p_3 & p_7 \\
 p_{-6} & p_{-2} & p_2 & p_6
\end{bmatrix}
\]  

(3.51)

where it is assumed that the number of measurements is four times the number of actuators. Note that in this matrix there are four kinds of rows:

- \([p_{-4} \ p_0 \ p_4]\)
- \([p_{-5} \ p_{-1} \ p_3 \ p_7]\)
- \([p_{-6} \ p_{-2} \ p_2 \ p_6]\)
- \([p_{-7} \ p_{-3} \ p_1 \ p_5]\)

If the positive and negative elements are the same, and \(p_0\) for the nonsquare matrix is the same as \(p_0\) for the square matrix, and \(p_4\) for the nonsquare matrix is the same as \(p_1\) for the square matrix, then the outputs of both systems are going to be equal for every fourth point of the nonsquare matrix and every point of the square matrix. The general shape of the output profile should therefore be the same. The nonsquare matrix gives a higher resolution of the profile than the square matrix.

Figure 3.28 shows the geometrical relationship between the actuators and the measurements. It is rare (number of measurements are odd multiple of the number of actuators) that an actuator will have a measurement box located at the same CD location. In some cases (number of measurements
Figure 3.28: Geometry when the number of actuators is not the same as the number of measurements. It is assumed that the first measurement is half a data box width from the edge.

In general, when the number of measurements is not an integer multiple of the number of actuators, each column will have different entries and each would have to be identified separately or be found by some weighted mean from an average response.

3.2.3 Interaction matrix parameterized with Gram polynomials

It was shown in chapter 2 how a profile can be approximated using Gram polynomials. The actuator settings can be written in terms of Gram polynomials as well. The number of actuators does
not have to be the same as the number of measurements so another set of orthogonal polynomials will be used. For \( N_a \) actuators and up to degree \( m_a \) the actuator settings at time \( t \) can be written as:

\[
\text{act}(t) = x^{T}_{0a}u_c(t) \tag{3.53}
\]

If the above equation and Equation 3.44 are used in the least-squares equation from last chapter (Equation 2.36), the least-squares equation can be written as

\[
\hat{y}_c(t) = (x_0x_0^T)^{-1}x_0G_g(s)x_0^T u_c(t) \tag{3.54}
\]

where the matrix \( x_0 \) is defined as \([P_{1a} P_{2a} \cdots P_{m_a}]^T\) — the Gram polynomials with number of data points equal to the number of actuators. For a linear \( g_a(s) \) an \( m \times m_a \) matrix can be defined as:

\[
G_x = (x_0x_0^T)^{-1}x_0G^T x_0 a \tag{3.55}
\]

Using Equation 2.37 the \( G_x \) matrix, for \( m_a \) actuators and \( m \) measurements, can be written as

\[
G_x = \begin{bmatrix}
\frac{P^T P_{1a}}{P^T P_1} & \frac{P^T P_{2a}}{P^T P_2} & \cdots & \frac{P^T P_{m_a}}{P^T P_m} \\
\frac{P^T P_{1a}}{P^T P_2} & \frac{P^T P_{2a}}{P^T P_2} & \cdots & \frac{P^T P_{m_a}}{P^T P_m} \\
\vdots & \vdots & \ddots & \vdots \\
\frac{P^T P_{1a}}{P^T P_m} & \frac{P^T P_{2a}}{P^T P_m} & \cdots & \frac{P^T P_{m_a}}{P^T P_m}
\end{bmatrix} \tag{3.56}
\]

Having established a form for the \( G \) matrix from the previous section, the form or the shape of the \( G_x \) matrix can be investigated. The form of the \( G_x \) matrix for a square interaction matrix, \( G \), with dimension \( 105 \times 105 \) and with the diagonal elements given by \( p_{0,1,2} = [1.0, 0.4, 0.2] \), i.e. \( k = 2 \), is plotted in Figure 3.29 and Figure 3.30. From Figure 3.29 it can be seen that \( G_x \) is almost a diagonal matrix. Figure 3.30 is a plot of every second diagonal of \( G_x \) except the main diagonal which is positive and much larger in magnitude than the off diagonals. The odd diagonals which are not shown in the figure are all zero for this particular choice of \( G \). What that signifies is that even order polynomials in the input do not contribute anything to the odd order polynomials in the output, and the odd order polynomials in the inputs do not contribute anything to the even order polynomials in the output. That property does of course depend on the structure of the \( G \) matrix. In general it can be expected of a real paper machine to contain some nonlinearities, resulting in odd order polynomials in the input affecting even order polynomials in the output, and vice versa.

From equation 3.43 (the matrix does not necessarily have to be a square matrix, if it is not a square matrix only the matrices \( G_i \)'s will change), it can be seen that the \( G \) matrix can be split.
Figure 3.29: Meshed surface of the $G_x$ matrix for an interaction matrix of size 105x105 and $p_0 = 1$, $p_1 = 0.4$ and $p_2 = 0.2$.

Figure 3.30: Every second diagonals of $G_x$ from Figure 3.29. The zeroth (main) diagonal is not shown.
into several matrices, as

\[ G = p_0 I + p_1 \begin{bmatrix} 0 & 1 & 0 \\ 1 & \ddots & 1 \\ 0 & 1 & 0 \end{bmatrix} + \cdots + p_k \begin{bmatrix} 0 & \cdots & 0 & 1 \\ \vdots & \ddots & \vdots & \vdots \\ 1 & \ddots & \ddots & 1 \\ 0 & 1 & \cdots & 0 \end{bmatrix} \] (3.57)

or more generally

\[ G = p_0 G_0 + p_1 G_1 + \cdots + p_k G_k \] (3.58)

where \( G_k \) describes the effects moving the slice actuators has on their \( k \)th neighbour. The \( G_x \) matrix can then be written as

\[ G_x = (x_0 x_0^T)^{-1} x_0 G x_0^T \]

\[ = (x_0 x_0^T)^{-1} x_0 (p_0 G_0 + p_1 G_1 + \cdots + p_k G_k) x_0^T \] (3.59)

\[ = p_0 G_{x_0} + p_1 G_{x_1} + \cdots + p_k G_{x_k} \]

where the matrices \( G_{x_k} \) are defined as

\[ G_{x_k} = (x_0 x_0^T)^{-1} x_0 G_k x_0^T \] (3.60)

The parameters of the slice lip approximation can then be written in terms of the parameters of the Gram polynomials for the slice as:

\[ y_c = G_x u_c \]

\[ = (p_0 G_{x_0} + p_1 G_{x_1} + \cdots + p_k G_{x_k}) u_c \]

\[ = \begin{bmatrix} G_{x_0} u_c & G_{x_1} u_c & \cdots & G_{x_k} u_c \end{bmatrix} \begin{bmatrix} p_0 \\ p_1 \\ \vdots \\ p_k \end{bmatrix} \] (3.61)

The parameters \( p_k \) in the above equation can be estimated in the following fashion

\[ \begin{bmatrix} p_0 \\ p_1 \\ \vdots \\ p_k \end{bmatrix} \approx (A^T A)^{-1} A^T y_c \] (3.62)

where

\[ A = \begin{bmatrix} G_{x_0} u_c & G_{x_1} u_c & \cdots & G_{x_k} u_c \end{bmatrix} \] (3.63)

This method of finding the \( G_x \) matrix could be used to find initial estimates of the \( G_x \) matrix, although its use is limited to the case when the mapping from the actuators to the measurements is well known.
Also, the structure of $G$ is assumed to be well known and all its diagonal elements identical. Having said that, some results using the equations above to estimate $G_x$ will be presented in later chapters.

The condition number for the $G_x$ matrix can also become quite large. Avoiding $G_x$ matrices with large condition numbers can become one of the designing criteria when deciding what order of polynomial to use. The condition number of the $G_x$ matrix is generally not a problem as long as the polynomial order is much smaller than the number of data points, or there is small cross coupling between adjacent measurements. Reciprocal condition numbers for a fixed dimension of $G_x$ matrix but varying $G$ matrix are shown in Figure 3.31. The $G$ matrix is created as before with the elements on the main diagonals as $p_{0,1,2} = [1, 2r, -r]$. If 10 polynomials are used, as in this example, there is no danger of a bad condition number if the $G$ matrix is at least $26 \times 26$, no matter how severe the cross coupling is between adjacent measurements. If there is very little cross coupling ($r \leq 0.15$) there is also no danger of a badly conditioned interaction matrix.

![Reciprocal condition numbers of $G_x$ (dimension 10x10)](image)

**Figure 3.31:** Reciprocal condition number of $G_x$ for a fixed number of Gram polynomials but varying the sizes of the interaction matrix, $G$, and its off diagonal elements which are $p_1 = 2r$ and $p_2 = -r$.

Similar results can be obtained by fixing the size of the interaction matrix but varying its elements and the size of the $G_x$ matrix as shown in Figure 3.32. The $G_x$ matrix will not become badly conditioned as long as its size is less than about one fifth of the size of the $G$ matrix, or if cross coupling is small. The cause of the bad condition number for the $G_x$ matrix can be seen in Figure
3.33, which shows a plot of the singular values for the $G_x$ matrix for three different dimensions of $G_x$. It can be seen that the lowest singular value approaches zero for a bigger $G_x$ matrix, and as a result the condition number grows.

Reciprocal condition numbers of $G_x$ (G is 105x105)

![Reciprocal condition numbers of Gx](image)

Figure 3.32: Reciprocal condition numbers of $G_x$ for $G$ of size $105 \times 105$ with $p_0 = 1$, $p_1 = 2r$, $p_2 = -r$ and varying the number of Gram polynomials.

It will be shown in chapter 5 how the $G_x$ matrix can be used for control. It involves finding an $m_a \times m$ matrix, $K_0$, such that $G_x K_0 = I$. The rank of $G_x$ is the minimum of $m$ and $m_a$, therefore, the rank of $G_x G_x^T$ is also the minimum of $m$ and $m_a$. The inverse of $G_x$, which can be written as

$$K_0 = G_x^T (G_x G_x^T)^{-1}$$

will only exist if $m_a \geq m$ [94], i.e. the inverse only exists if the $m \times m$ matrix $G_x G_x^T$ has a full rank. The inverse simplifies to $K_0 = G_x^{-1}$ if $G_x$ is a square matrix ($m = m_a$). Note that the dimension of $G_x$ is $m \times m_a$ and will be much smaller than the dimension of $G$ ($N \times N_a$) so calculating $K_0$ requires less computations than the plant inverse.

### 3.3 Slice lip model

The actuators are assumed to only apply force but no bending moment. Hence, the moment can be completely specified by the deflection of the slice lip at the points where the actuators are
Chapter 3: Model for cross-directional basis weight response

![Plot of singular values of G_x for an interaction matrix of dimension 105 x 105 with p_0 = 1, p_1 = 2r, p_2 = -r and with r = 0.3.]

**Figure 3.33:** Plot of singular values of $G_x$ for an interaction matrix of dimension $105 \times 105$ with $p_0 = 1$, $p_1 = 2r$, $p_2 = -r$ and with $r = 0.3$.

located. The intermediate values can be obtained by a straight line fit between the nearest neighbors because the moment is defined as a force at a particular point times the distance from the force. This is characteristic of beams with point loading.

The slice lip is assumed to have free ends, which implies zero moments and zero shear force at the ends. Furthermore, the slice lip is assumed to have a constant flexural rigidity, meaning that the bending modulus, $EI$, is constant for the whole slice. Assuming small deflection theory [27], i.e. no horizontal displacement at the ends, the bending moments can be written in terms of the second derivative of the slice lip deflection as

$$M = EI \frac{d^2 u(x)}{dx^2}$$  \hspace{1cm} (3.65)

where $EI$ can be treated as a constant. Only about 1% error is introduced if the deflection is up to one twentieth of the length of the beam. For a paper machine the deflection of the slice lip is less than $20 \text{ mm}$ for a slice lip of $10 \text{ m}$ so that the horizontal displacement can be neglected.
Chapter 3: Model for cross-directional basis weight response

The moments for a beam with forces acting on the ends \( p(0) \) and \( p(L) \) and at a distance \( a \) from one end \( p(a) \) is (see Figure 3.34)

\[
EI \frac{d^2u}{dx^2} = M(x) = p(0)x + p(a)(x - a)
\]

\[
= \frac{-p(a)b}{L}x + p(a)(x - a)
\]

(3.66)

Since the bending modulus \( EI \) is constant along the beam, integration of Equation 3.66 gives

\[
u(x) = -\frac{p(a)b}{6EI}x^3 + \frac{p(a)}{6EI}(x - a)^3 + c_1x + c_2
\]

(3.67)

Now assuming the forces acting on the beam are springs with a certain spring constant, the force created by an actuator located at \( ia \) is

\[
p(ia) = k_s(z(ia) - u(ia))
\]

(3.68)

where \( u(ia) \) is the deflection of the slice at the point \( ia \) on the beam, and \( z(ia) \) is the actuator setpoint at the same point, i.e. the positive difference \( z(ia) - u(ia) \) is the expansion of a spring with a spring constant \( k_s \). The above equation can also be written as:

\[
u(ia) = z(ia) - \frac{p(ia)}{k_s}
\]

(3.69)

For the beam in Figure 3.34 the force acting at \( x = 0 \) is \(-p(a)b/L\) so from Equation 3.67 and the above equation

\[
u(0) = c_2 = z(0) + \frac{p(a)b}{Lk_s}
\]

(3.70)

The \( c_2 \) coefficient from Equation 3.67 can therefore be found if the force at point \( a \), the length of the slice lip, the spring coefficient of the actuator, and the actuator setting at point 0, are all known. Evaluating the other boundary condition, the \( c_1 \) coefficient can be found as:

\[
u(L) = -\frac{p(a)b}{6EI}L^2 + \frac{p(a)}{6EI}b^3 + c_1L + z(0) + \frac{p(a)b}{Lk_s} = z(L) + \frac{p(a)a}{Lk_s}
\]

(3.71)
leading to

\[ c_1 = \frac{z(L) - z(0)}{L} + p(a)\left\{ \frac{b}{6EI} \left( L^2 - b^2 \right) + \frac{1}{Lk_s} \left( \frac{a - b}{L} \right) \right\} \quad (3.72) \]

Equation 3.67 can now be written as

\[ u(x) = \left[ \frac{1}{6EI} \left\{ b(L^2 - x^2 - b^2)x + L(x - a)^3 \right\} + \frac{1}{Lk_s} \left( b + \frac{a - b}{L}x \right) \right] p(a) \\
+ z(0) + \frac{z(L) - z(0)}{L}x \quad (3.73) \]

If the above equation is evaluated at \( x = a \) and \( u(a) \) is written as \( u(a) = \frac{z(a) - z(0) - z(L)}{L}a \), the force at point \( a \) can be written as:

\[ p(a) = \frac{z(a) - z(0) - \frac{z(L) - z(0)}{L}a}{\frac{1}{k_s} + \frac{1}{3EI}(a^2b) + \frac{1}{Lk_s} (b + \frac{a - b}{L}a)} \quad (3.74) \]

Therefore the slice lip deflection can be written as:

\[ u(x) = \left[ \frac{1}{6EI} \left\{ b(L^2 - x^2 - b^2)x + L(x - a)^3 \right\} + \frac{1}{Lk_s} \left( b + \frac{a - b}{L}x \right) \right] \\
\frac{1}{k_s} + \frac{1}{3EI}(a^2b) + \frac{1}{Lk_s} \left( b + \frac{a - b}{L}a \right) \\
\left( \frac{z(a) - z(0) - \frac{z(L) - z(0)}{L}a}{L} \right) + z(0) + \frac{z(L) - z(0)}{L}x \quad (3.75) \]

The slice lip deflection can thus be written in terms of the actuators locations (\( z(0), z(a) \) and \( z(L) \)).

As an example assume \( a = 1 \), \( L = 18 \), \( k_s = 1/18 \), and \( EI \) is varied from around 0.001 to 0.85.

The resulting deflections are shown in Figure 3.35. Figure 3.36 shows the detailed deflection and the three actuators working on the beam for the stiffest lip from Figure 3.35. The dotted lines show the actuators at rest when they all have setpoint equal to zero, and the beam has zero deflection. The actuator at \( x = 1 \) is then moved one unit and the solid lines show the resulting deflection as well as the actuators compression (at \( x = 0 \) and \( x = 18 \)), and the actuators expansion (at \( x = 1 \)).

If it is assumed for the slice in Figure 3.37 that superposition holds, the deflection at the points where the forces are applied can be calculated as (note that if all forces except \( p(1a) \) are equal to zero, the equation below is equal to Equation 3.73):

\[
\begin{bmatrix}
  u(1a) \\
  u(2a) \\
  \vdots \\
  u((N_a - 2)a)
\end{bmatrix} = Q \begin{bmatrix}
  p(1a) \\
  p(2a) \\
  \vdots \\
  p((N_a - 2)a)
\end{bmatrix} + z(0) + \frac{z(L) - z(0)}{L} \begin{bmatrix}
  \frac{1a}{2a} \\
  \vdots \\
  \frac{(N_a - 2)a}{2a}
\end{bmatrix} \quad (3.76)
\]

or in vector form

\[ u = Qp + z(0) + \frac{z(L) - z(0)}{L}x \quad (3.77) \]
Figure 3.35: Deflection of the beam in Figure 3.34 for different stiffnesses, when an actuator at $x = 1$ is moved one unit.

Figure 3.36: Deflection of a beam when an actuator at $x = 1$ is moved to one from the initial position of zero (shown for the actuator as a dotted line with the endpoints denoted with small circles).
where matrix $Q \in \mathbb{R}^{(N_a-2)\times(N_a-2)}$ is symmetric with elements given by

$$Q(i,j) = \frac{1}{k_s} \left\{ \frac{k_s a^3 i(N_a - 1 - j)(2j(N_a - 1) - j^2 - i^2)}{6EI (N_a - 1)} \right\}$$

for $i \leq j$, and $u = [u(1a) \ u(2a) \ \cdots \ u((N_a-2)a)]^T$, $p = [p(1a) \ p(2a) \ \cdots \ p((N_a-2)a)]^T$, and $x = [1a \ 2a \ \cdots \ (N_a-2)a]^T$. By defining a vector of actuator setpoints $z = [z(1a) \ z(2a) \ \cdots \ z((N_a-2)a)]^T$ we can write the forces created by the springs as:

$$p = k_s(z - u)$$

(3.79)

Then the deflection of the slice can be written as

$$u = (I + k_s Q)^{-1} \left( k_s Q z + z(0) + \frac{z(L) - z(0)}{L} x \right)$$

(3.80)

and the forces as

$$p = k_s(I + k_s Q)^{-1} \left( z - z(0) - \frac{z(L) - z(0)}{L} x \right)$$

(3.81)

with the values at the endpoints given by

$$p(0) = - \sum_{i=1}^{N_a-2} p(ia) \frac{N_a - 1 - i}{N_a - 1} \quad \text{and} \quad p(L) = - \sum_{i=1}^{N_a-2} p(ia) \frac{i}{N_a - 1}$$

$$u(0) = z(0) - \frac{p(0)}{k_s} \quad \text{and} \quad u(L) = z(L) - \frac{p(L)}{k_s}$$

(3.82)

(3.83)

If it is desired to find the slice lip deflection between the actuators, Equation 3.81 can be used to calculate the forces exerted on the slice by the actuators. Equation 3.77 can then be used together with the matrix $Q$ defined for noninteger values of $i$ in Equation 3.78 such that it becomes an $N_{\text{fine}} \times (N_a - 2)$ matrix where $N_{\text{fine}}$ is the desired number of points for which the slice lip deflection
Chapter 3: Model for cross-directional basis weight response

is required. The Equations above are useful if the actuator setpoints are known, and it is desired to find the slice lip deflection. On the other hand the slice lip deflection might be known, and it is required to find the actuator setpoints needed for the particular slice lip deflection. Equation 3.73 can be written in terms of the slice lip deflections at the endpoints as:

\[ u(x) = \frac{1}{6EI} \left[ b(L^2 - x^2 - b^2)x + L(x - a)^3 \right] p(a) + u(0) + \frac{u(L) - u(0)}{L} \] (3.84)

If superposition applies as before, the above equation can be written in a vector form as:

\[ \mathbf{u} = \mathbf{Q}_u \mathbf{p} + u(0) + \frac{u(L) - u(0)}{L} \mathbf{x} \] (3.85)

With the symmetric matrix, \( \mathbf{Q}_u \in \mathbb{R}^{(N_a-2) \times (N_a-2)} \) defined as

\[ \mathbf{Q}_u(i, j) = \frac{a^3}{6EI} \frac{i(N_a - 1 - j)(2j(N_a - 1) - j^2 - i^2)}{N_a - 1} \quad i \leq j \] (3.86)

using \( p = k_s(z - u) \), this can be written as

\[ \mathbf{z} = \frac{1}{k_s} \mathbf{Q}_u^{-1} \left\{ (1 + k_s \mathbf{Q}_u) \mathbf{u} - u(0) - \frac{u(L) - u(0)}{L} \mathbf{x} \right\} \] (3.87)

with the forces given by

\[ \mathbf{p} = \mathbf{Q}_u^{-1} \left( \mathbf{u} - u(0) - \frac{u(L) - u(0)}{L} \mathbf{x} \right) \]

\[ = \mathbf{Q}_u^{-1} \begin{bmatrix} \frac{-N_a-2}{N_a-1} & \frac{-1}{N_a-1} & \cdots & \frac{-1}{N_a-1} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{-1}{N_a-1} & \frac{-1}{N_a-1} & \cdots & \frac{-N_a-2}{N_a-1} \end{bmatrix} \begin{bmatrix} u(0) \\ \mathbf{u} \\ u(L) \end{bmatrix} \]

\[ = \mathbf{Q}_u^{-1} \mathbf{A}^T \mathbf{u} \] (3.88)

The forces for the whole slice, including the endpoints can be written as:

\[ \mathbf{P} = \begin{bmatrix} p(0) \\ \mathbf{p} \\ p(L) \end{bmatrix} = \begin{bmatrix} \frac{-N_a-2}{N_a-1} & \cdots & \frac{-1}{N_a-1} \\ \vdots & \ddots & \vdots \\ \frac{-1}{N_a-1} & \cdots & \frac{-N_a-2}{N_a-1} \end{bmatrix} \mathbf{p} = \mathbf{A}^T \mathbf{p} \] (3.89)

The bending moments are then written as:

\[ \begin{bmatrix} M(0) \\ M \\ M(L) \end{bmatrix} = \mathbf{M} = \sum_{j=0}^{n+1} p(ja)(x - ja) = a \begin{bmatrix} 0 & \cdots & 0 \\ 1 & \vdots & \vdots \\ \vdots & \ddots & \vdots \\ N_a - 1 & \cdots & 1 \end{bmatrix} \begin{bmatrix} p(0) \\ p(1a) \\ \vdots \\ p((N_a - 2)a) \end{bmatrix} \]

\[ = \mathbf{S} \mathbf{A}^T \mathbf{p} = \mathbf{S} \mathbf{A}^T \mathbf{Q}_u^{-1} \mathbf{A}^T \mathbf{u} \] (3.90)

Accordingly the bending moments can be calculated directly from the parameters of the Gram polynomials of the slice, which allows limiting the bending moment by modifying the parameters. Note that all the matrices in Equation 3.90 can be precomputed, as they only depend on the slice lip configuration.

58
3.3.1 Bending limit

In general, the Gram polynomial approximation of a specific profile can be found using least-squares estimates, according to Equation 2.36. The least-squares solution can result in a shape profile that is not possible to achieve with the slice lip since the stroke is outside the actuator’s range or the bending limit of the slice will be exceeded. To satisfy those restrictions, the square of the approximation error ($\varepsilon(t)$ from Equation 2.35) can be minimized subject to restrictions on the maximum and the minimum value of the slice lip position as well as the minimum and the maximum bending moments. The minimization of $\varepsilon^2(t)$ (i.e. least-squares) can be transformed into a quadratic programming problem according to

$$\min_{u_c} \| u - x_{0a}^T u_c \| = \min_{u_c} (u - x_{0a}^T u_c)^T (u - x_{0a}^T u_c)$$

$$= \min_{u_c} \left( u^T u - u_c^T x_{0a} u - (x_{0a} u)^T u_c + u_c^T x_{0a} x_{0a}^T u_c \right)$$

(3.91)

The value of $u^T u$ does not have any effect on the minimization with respect to $u_c$, therefore, it can be omitted. Also the minimization of $u_c^T x_{0a} u$ is the same as the minimization of $(x_{0a} u)^T u_c$ which can be seen from the fact that $(u_c^T x_{0a} u)^T = (x_{0a} u)^T u_c$. The above equation can then be written as

$$\min_{u_c} \| u - x_{0a}^T u_c \| = \min_{u_c} \left( u_c^T x_{0a} x_{0a}^T u_c + 2(-x_{0a} u)^T u_c \right)$$

$$= \min_{u_c} \left( \frac{1}{2} u_c^T x_{0a} x_{0a}^T u_c + (-x_{0a} u)^T u_c \right)$$

(3.92)

subject to

$$Ru_c \leq B$$

(3.93)

which can be solved using Matlab’s optimization toolbox [55]. The restrictions can be written as

$$SA^T Q_0^{-1} A x_{0a}^T u_c \leq \max M$$

$$-SA^T Q_0^{-1} A x_{0a}^T u_c \leq -\min M$$

(3.94)

$$x_{0a}^T u_c \leq \max u$$

$$-x_{0a}^T u_c \leq -\min u$$

3.4 Summary

In this chapter the paper machine has been described and a model that relates slice lip movement to basis weight response has been developed. A model has also been given that can be used to
describe the bending of the slice lip. The main feature of the model is an interaction matrix that
describes how the movement of a single actuator will effect a large section of the measured profile.
The effect is transmitted through bending of the slice lip which sets up waves on the forming wire
and spreads out on the wire. The interaction matrix can, for many paper machines, become singular
causing problems when controlling the basis weight. Some of the problems in creating the interaction
matrix are discussed and it is found that it is difficult to create a matrix that will not create ripples,
particular to the model, in the model output. Gram polynomials are then used to parameterized the
interaction matrix, resulting in a much smaller matrix which can be identified directly and will be
discussed in the next chapter. The form of the parameterized interaction matrix is discussed and
it is found that for a general interaction matrix the parameterized matrix will have nonzero entries
mainly on the diagonal. A model of the slice lip as a flexible beam with multiple forces is given
and the slice lip deflection, which is written as Gram polynomials, is related to the bending moment.
The parameters of the Gram polynomials can therefore be manipulated to limit the bending moments
for the whole slice.
Chapter 4

Identification

Basis weight, which is measured at the finishing section of the paper machine, is controlled by the slice lip actuators located on the headbox. For control purposes it is necessary to know the transportation delay from the slice lip to the measurements. The delay can be estimated by carrying out bump tests on the slice lip, or by knowing the distance the paper travels through the machine, and the speed of the machine, it can be calculated. A knowledge of the interaction matrix, $G$, which describes how each actuator affects the basis weight, is also required. The interaction matrix of a paper machine depends on the speed of the machine, the target basis weight, the consistency of the pulp in the headbox and other variables. Until such time as a model is available that will calculate the interaction matrix given all the miscellaneous variables, the interaction matrix has to be estimated. If Gram polynomials are used, the parameterized interaction matrix, $G_x$ (see previous chapter) could as well be estimated. Currently no paper machine data is available from which the $G_x$ matrix could be identified. Therefore, a paper machine was simulated using an interaction matrix identified from industrial data. The identified matrix was then used as a model test-bed to identify the $G_x$ matrix. There are two ways to identify the $G_x$ matrix from those simulations. One way is to use the estimated $G$ matrix and then calculate $G_x$ using Equation 3.55, i.e. $G_x = (x_0x_0^T)^{-1}x_0Gx_0^T$. Another way is to simulate a paper machine using a particular interaction matrix and then identify directly the $G_x$ matrix, from the parameters of the orthogonal approximation of the measured profile, $y_c$, and the parameters of the orthogonal polynomials of the slice lip, $u_c$. In this chapter the latter method, which identifies the $G_x$ matrix directly, will be used.

4.1 Identifying the interaction matrix

There are many ways to estimate the interaction matrix [22, 64]. Basically it involves perturbing several actuators and using least-squares estimation or correlation studies to find the steady-state gain from the actuators to the measurements over an area of the measured profile corresponding to the perturbed actuators (see Figure 4.38). For a single actuator, several estimations might have to be run, one for each measurement, and then from those estimations the response shape could be obtained. The experiments are then repeated for several actuators across the machine. The response shapes
from those several actuators are generally averaged to get a single response shape to represent the response to any actuator across the machine. As was mentioned in the previous chapter, there might be some problems with mapping the measurements to the actuators.

Experiments were done on a 8 m wide linerboard paper machine, with target weight at 61.1 g/m² [64, 65, 77]. The 105 slice actuators are located 76 mm apart and there are 445 measurements across the machine with data box size around 18 mm. The data were collected for 60 scans of the profile, and moisture and basis weight profiles were recorded, as well as the slice lip setpoints and basis weight profiles mapped to the 105 actuators. It was observed for this particular paper machine on this particular day that an actuator could affect an area of the paper sheet up to nine actuators wide. Therefore, in order to minimize the effect of cross-coupling on the identification, every ninth actuator is perturbed with a PRBS signal, starting away from the edge, at actuator number eight. Hence 11 actuators are perturbed with PRBS signals, at the CD positions given in the first column of Table 4.1.

Figure 4.39 shows a black and white image of the actuators and the measured basis weight, with the mean value of every CD point subtracted in order to better see the response. From the black and white image, the area of the measured profile, corresponding to the 11 actuators can be found. The model describing the effect of each of the inputs on its corresponding area can then be estimated. The centers of response as seen from the input-output data in Figure 4.39 are assumed to be the measurement points corresponding to the center of response. Those measurement data boxes are shown in Table 4.1 and will only be used to limit the area where the steady state gain from
Table 4.1 Location of the PRBS inputs and the corresponding output measurements deduced from Figure 4.39.

<table>
<thead>
<tr>
<th>Actuator number</th>
<th>Measurement number</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>27</td>
</tr>
<tr>
<td>17</td>
<td>64</td>
</tr>
<tr>
<td>26</td>
<td>105</td>
</tr>
<tr>
<td>35</td>
<td>144</td>
</tr>
<tr>
<td>44</td>
<td>182</td>
</tr>
<tr>
<td>53</td>
<td>220</td>
</tr>
<tr>
<td>62</td>
<td>259</td>
</tr>
<tr>
<td>71</td>
<td>298</td>
</tr>
<tr>
<td>80</td>
<td>339</td>
</tr>
<tr>
<td>89</td>
<td>377</td>
</tr>
<tr>
<td>98</td>
<td>416</td>
</tr>
</tbody>
</table>

the actuators to the measurements is found. The exact location of the center of response will be calculated later from the estimated responses, by calculating the cross-correlations between the 11 response shapes.

Figure 4.40 shows input and output data for one CD location. It can be seen that the dynamics are very fast, indeed the system reaches steady state within one scan, after a pure delay of two scans. The delay can also be established by looking at the cross correlation between the 11 PRBS signals and their corresponding outputs. Figure 4.41 shows the cross correlation between the 11 PRBS signals and their presumed center of response. The dynamics of the system are negligible therefore, the maximum of the correlation function indicates the delay of the system. It is evident that the maximum correlation is reached for a delay equal to two, for all the 11 sequences. The same delay is obtained when finding the maximum of the absolute value of the correlation between the 11 inputs and their closest measurements. Hence the knowledge of the delay will be assumed to be known in the estimation that follows.

From the knowledge of the process delay and dynamics, the process can be approximately modeled by

$$y(t, x) = p_0 u(t - 2, x)$$  \hspace{1cm} (4.95)
where \( y(t, x) \) and \( u(t, x) \) are the basis weight and actuator set point respectively, at time \( t \) and location \( x \) across the sheet. Each input will affect several outputs, which can be written as

\[
[y(t, x - k) \ y(t, x - k + 1) \ \cdots \ y(t, x + k)] = [p_{-k} \ p_{-k+1} \ \cdots \ p_{k}]u(t - 2, x) \quad (4.96)
\]

Since the basis weight and actuator setpoints have a nonzero DC level, it is necessary to look at the incremental inputs and outputs, to estimate all the \( p_j \)'s:

\[
\Delta_t[y(t, x - k) \ y(t, x - k + 1) \ \cdots \ y(t, x + k)] = [p_{-k} \ p_{-k+1} \ \cdots \ p_{k}]\Delta_t u(t - 2, x) \quad (4.97)
\]
All the $2k + 1$ estimations for each input, can be done separately with

$$
\Delta_t y(t, x - k) = y(t, x - k) - y(t - 1, x - k) = p_{-k}(u(t - 2, x) - u(t - 3, x)) \quad (4.98)
$$

which can be identified using batch least squares, recursive least-squares or recursive instrumental variables [87]. Recursive instrumental variables are used for these data to get unbiased estimates. Only the gain of the system is estimated and the system is assumed to be time-invariant during the
trial, therefore a forgetting factor of one is used. Figure 4.42 shows the 11 estimates using the slice lip actuators and measurement data boxes in Table 4.1. Those are 11 independent estimates, shown here to see if all the 11 inputs are sufficiently rich to give good estimates. It can be seen that the estimates converge within the 60 scans, but they do not all converge to the same value, because the outputs in Table 4.1 do not necessarily correspond exactly to the inputs. The final estimates are shown in Figure 4.43 with the plus signs.

To obtain a cross directional model, the responses for moving an actuator is estimated at 20 measurement locations on both sides of the corresponding measurement location from Table 4.1, using recursive instrumental variable as before. Figure 4.43 shows the 11 response shapes from the estimation. In order to average the 11 responses it is necessary to find the number of data boxes
Chapter 4: Identification

<table>
<thead>
<tr>
<th>PRBS signal number i</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>38</td>
<td>78</td>
<td>117</td>
<td>156</td>
<td>194</td>
<td>233</td>
<td>272</td>
<td>312</td>
<td>351</td>
<td>390</td>
</tr>
<tr>
<td>2</td>
<td>-38</td>
<td>0</td>
<td>39</td>
<td>79</td>
<td>118</td>
<td>156</td>
<td>195</td>
<td>234</td>
<td>274</td>
<td>313</td>
<td>352</td>
</tr>
<tr>
<td>3</td>
<td>-78</td>
<td>-39</td>
<td>0</td>
<td>39</td>
<td>78</td>
<td>117</td>
<td>156</td>
<td>195</td>
<td>234</td>
<td>274</td>
<td>312</td>
</tr>
<tr>
<td>4</td>
<td>-117</td>
<td>-79</td>
<td>-39</td>
<td>0</td>
<td>39</td>
<td>77</td>
<td>116</td>
<td>155</td>
<td>195</td>
<td>234</td>
<td>273</td>
</tr>
<tr>
<td>5</td>
<td>-156</td>
<td>-118</td>
<td>-78</td>
<td>-39</td>
<td>0</td>
<td>38</td>
<td>78</td>
<td>117</td>
<td>156</td>
<td>195</td>
<td>234</td>
</tr>
<tr>
<td>6</td>
<td>-194</td>
<td>-156</td>
<td>-117</td>
<td>-77</td>
<td>-38</td>
<td>0</td>
<td>39</td>
<td>78</td>
<td>118</td>
<td>157</td>
<td>196</td>
</tr>
<tr>
<td>7</td>
<td>-233</td>
<td>-195</td>
<td>-156</td>
<td>-116</td>
<td>-78</td>
<td>-39</td>
<td>0</td>
<td>39</td>
<td>79</td>
<td>118</td>
<td>157</td>
</tr>
<tr>
<td>8</td>
<td>-272</td>
<td>-234</td>
<td>-195</td>
<td>-155</td>
<td>-117</td>
<td>-78</td>
<td>-39</td>
<td>0</td>
<td>40</td>
<td>79</td>
<td>118</td>
</tr>
<tr>
<td>9</td>
<td>-312</td>
<td>-274</td>
<td>-234</td>
<td>-195</td>
<td>-156</td>
<td>-118</td>
<td>-79</td>
<td>-40</td>
<td>0</td>
<td>39</td>
<td>78</td>
</tr>
<tr>
<td>10</td>
<td>-351</td>
<td>-313</td>
<td>-274</td>
<td>-234</td>
<td>-195</td>
<td>-157</td>
<td>-118</td>
<td>-79</td>
<td>-39</td>
<td>0</td>
<td>39</td>
</tr>
</tbody>
</table>

Table 4.2 Number of data boxes between the maximum values of the cross correlations between the 11 responses in Figure 4.43.

**Figure 4.44:** The 11 estimated response shapes from Figure 4.43 shifted according to the first line in Table 4.2 such that they overlay.
Chapter 4: Identification

The cross-correlation between all the responses is calculated and the number of data boxes between the maximum of the cross-correlation is shown in Table 4.2. It can be seen that the number of data boxes between the two outermost responses, the PRBS signals number one and 11, is always 390 data boxes, no matter which PRBS signal is used as a reference. The distance between the second response and the third response is between 39 and 40 data boxes, depending on which response shape is used to calculate the difference. The first line in the table, i.e. the correlation between the first response shape to all the other responses, gives the difference between the peak of response that is very close to the mean value, therefore it is used to align the responses on top of each other to find the average response. A plot of the results are shown in Figure 4.44. The response shapes are averaged, giving a typical response shape for a paper machine as can be seen in Figure 4.45. The figure shows the mean and the standard deviation for the 11 responses from the previous figure, as well as the actuator spacing. It can be seen that beyond 16 measurement zones from the center of response, there is almost no response to the actuator bump. Therefore it will be assumed that the width of response is 33 measurement zones. To make the response symmetric the left and right hand side in Figure 4.45 are averaged, hence 17 parameters will be used in the interaction matrix. The block diagonal interaction matrix, $G$, is constructed by putting the averaged response in every column of the matrix, centered in the rows corresponding to those columns. As mentioned in
Chapter 3, it is difficult to decide in which row to place the center of the response, especially if the number of measurements is not an integer multiple of the number of actuators or if the sheet wanders.

4.2 Identifying the parameterized matrix

As previously mentioned, once the $G$ matrix has been estimated the $G_x$ matrix can be calculated using its definition (see Equation 3.55). Another way, which does not depend on the knowledge of the $G$ matrix, is to estimate the matrix directly from the knowledge of the parameters of the orthogonal approximation of the measured profile, $y_c$, and the parameters of the orthogonal polynomial of the slice lip, $u_c$. To simplify the estimation, the structure of the $G$ matrix could be used. But doing so would defy one of the reasons for using orthogonal polynomials, i.e. not having to establish the mapping from an actuator to the measurements. For that reason every element of $G_x$ has to be estimated. The problem is therefore MIMO (multiple input multiple output) estimation. A MIMO system can be written as [86 p.80]

\[
I y_c(t) + A_1 y_c(t - 1) + \cdots + A_{n_a} y_c(t - n_a) = B_1 u_c(t - d) + \cdots + B_{n_b} u_c(t - d - n_b + 1) \tag{4.99}
\]

where $y_c(t)$ is a vector of length $m$, $A_i$ is a matrix of size $m \times m$, $u_c(t)$ is a vector of length $m_a$ and $B_i$ is a matrix of size $m \times m_a$. Keeping the analogy with SISO systems, a parameter matrix $\theta$ can be defined as

\[
\theta = [A_1 \quad A_2 \quad \cdots \quad A_{n_a} \quad B_1 \quad \cdots \quad B_{n_b}]^T \tag{4.100}
\]

with the size as $[n_a m + n_b m_a] \times m$. The regression vector $\varphi$ is a vector of vectors and is written as:

\[
\varphi(t) = \begin{bmatrix}
-y_c(t - 1) \\
y_c(t - 2) \\
\vdots \\
-y_c(t - n_a) \\
u_c(t - d) \\
\vdots \\
u_c(t - d - n_b + 1)
\end{bmatrix} \tag{4.101}
\]

and will have length of $[n_a m + n_b m_a]$. The output vector $y_c$ can then be written as

\[
y_c(t) = \theta^T \varphi(t) \tag{4.102}
\]

It has been shown that a paper machine response is governed by a delay and a gain. The model for the basis weight response is not found to have any zeros, so $n_b = 1$. The time constant for the
Chapter 4: Identification

system is very small and can be ignored in most cases, so \( n_a \leq 1 \). Furthermore, the matrix \( A_1 \) can be assumed to be diagonal, i.e. the outputs do not affect each others. Every row of the \( G_x \) matrix can then be estimated separately, so the estimation has been reduced to \( m \) MISO (multiple input single output) estimation problems, each estimating \( m_a + n_a \) parameters. The \( m \) systems can be written as:

\[
y_c(t,i) = [a_i; B_1(i,:)] \begin{bmatrix} -y_c(t-1,i) \\ u_c(t-2) \end{bmatrix}
\]

(4.103)

where \( a_i \) is the pole for the \( i \)th input, \( B_1(i,:) \) is the \( i \)th row of the \( B_1 \) matrix and \( y_c(t,i) \) is the \( i \)th element of \( y_c(t) \). If there is no dynamics, for each MISO system, \( i \), where \( i \in [1, \ldots, m] \), the system is written as:

\[
y_c(t,i) = [G_x(i,1) \ G_x(i,2) \ \cdots \ G_x(i,m_a)] \begin{bmatrix} u_c(t-d,1) \\ u_c(t-d,2) \\ \vdots \\ u_c(t-d,m_a) \end{bmatrix}
\]

\[
= \theta(i)^T \varphi(t).
\]

(4.104)

The \( G_x \) matrix can be estimated using standard least squares or recursive least squares [86, 87, 113]. If the system has some dynamics (for simplicity the dynamics is assumed to be the same for all the polynomial orders), for example

\[
y_c(t) = \frac{1 + a}{1 + a q^{-1}} G_x u_c(t - d)
\]

(4.105)

the \( G_x \) matrix can be estimated from the estimates of \( B_1 \) and \( a \) such that

\[
G_x = \frac{B_1}{1 + a}
\]

(4.106)

Figure 4.46: Basis weight response to a flat slice lip (at 10\( \mu m \)) with 105 actuators for two different models of a paper machine.
For the simulation studies the plant is assumed to be described by the interaction matrix from the previously identified $G$ matrix except here it is assumed to be $420 \times 105$, i.e. the number of measurements is assumed to be an integer multiple of actuators. The reason for doing so can best be seen in Figure 4.46 where a flat input gives (apart from the edges) a flat output if the number of measurements is an integer multiple of the number of actuators. On the other hand if the number of measurements is assumed not to be an integer multiple of the number of actuators, the model output is very jagged as is evident from the solid line in Figure 4.46. This jagged profile is particular to the model chosen, where all responses are assumed to be the same and the center of every actuator is assumed to be aligned with a particular measurement data box, which is impossible for a noninteger ratio of actuators to measurements. Therefore there will be multitudes of rows and their sum will take on varying values, if all the response profiles to each actuator are assumed to be the same. It is therefore very difficult to fine tune a model to accurately describe a paper machine. It can be somewhat simplified by use of a number of measurements that is an integer multiple of the number of actuators.

The slice lip profile used for the simulation is controlled by 10th order Gram polynomials and 10th order polynomials are used to approximate the measured profile. Independent PRBS signals of magnitude $\pm 0.1$ are added to the parameters of the polynomials of the slice lip, the paper machine is simulated, the basis weight profile is measured and the Gram polynomial parameters of the simulated output profile are calculated. Figure 4.47 shows the parameters of the polynomials of the slice lip, the input $u_c$, and the parameters of the Gram polynomial approximation of the simulated output, $y_c$, used for the estimation of $G_x$. The PRBS signals added to the parameters of the slice lip results in the slice lip deflection of at most $\pm 0.6 \cdot 10 \mu m = \pm 6 \mu m$ (see the top graph in Figure 4.48) from its initial settings. Therefore the peak to peak deflection of the slice lip required by the excitation signal is $12 \mu m$ which is only about a quarter of the peak to peak value of the slice lip before applying the PRBS signals. The slice lip bending, that results from moving the slice lip for the estimation trials is shown at the bottom of Figure 4.48 with length of the slice lip, $L = 8 m$, and the bending modulus, $EI = 4500 Nm^2$. The dotted line in the graph shows the bending limit that will be imposed upon the slice in the next simulation. The measured simulated basis weight profile will then have a deflection of at most $\pm 0.24 g/m^2$ (Figure 4.48) from the initial measured profile, which amounts to about 50% of the peak to peak value of the initial basis weight profile before applying the PRBS signals. This input-output data is used to estimate the $G_x$ matrix using 10 independent recursive least squares (RLS) on Equation 4.104 with $i$, the row number, between one and 10 inclusively. Each of
Figure 4.47: Input, $u_c$, and output, $y_c$, used for the estimation of $G_z$, when there are no limits on the slice lip bending.
the 10 independent estimators, estimates 10 parameters, one for each input. Figure 4.49 shows the estimated parameters from the 10 estimators, one for each row. It can be seen that the parameters converge in about 40 scans. The estimates at scan 59 are used to make the $G_x$ matrix. It can be seen in Figure 4.50 that the estimated matrix is the same as the true matrix. At least there is no visible difference between them.

### 4.2.1 Limited bending of the slice lip

As said earlier, the estimation of the parameterized matrix can be performed as long as the regressors, $\varphi(t)$ in Equation 4.104 are linearly independent. When the bending is limited, the regressors will become correlated, because if bending limits are exceeded, all the parameters of the Gram polynomials will have to be adjusted and will therefore be correlated. For that reason the estimates might not converge to the true values. Also a limit on the slice bending will result in the higher order parameters being limited the most, because higher order polynomials have higher bending moments. Therefore the excitation for the higher order polynomial will be severely limited, resulting in difficulty in estimating the corresponding row of the $G_x$ matrix.
Figure 4.49: Estimation of each row of $G_x$ when no limit is put on the slice lip bending. The true parameters for this simulation, are shown with dotted lines.
Chapter 4: Identification

Figure 4.50: Estimate of $G_x$ after 59 scans when no limits are put on the slice lip bending. The estimated matrix is denoted with * and dotted lines between its elements. The true matrix used in the simulation is denoted with o and solid lines.

Figure 4.51 shows the input and output used for the estimation, when the bending moments are limited at ±0.23 Nm. It can be seen that the input parameters are correlated and they also have smaller values than when there was no limit on the bending moments. Because of smaller input parameters the output parameters will also be smaller as can be seen in the same figure. Figure 4.52 shows the slice lip profiles, measured basis weight profile from a simulated paper machine and the bending moments of the slice lip. The slice lip is now smoother than before (compare to Figure 4.48), especially at the edges. It shows clearly in the plot of the bending moments, where they are successfully limited at ±0.23 Nm. Figure 4.53 shows the estimated parameters of the $G_x$ matrix. It can be seen that the convergence is slower than before, and for the higher orders, the parameters of the corresponding row do not converge within the 59 scans. Figure 4.54 better illustrates the estimates of $G_x$ after the last scan. Now the difference between the true matrix and the estimated one is clearly visible, mainly for the higher order terms.
Figure 4.51: Input, \( u_c \), and output, \( y_c \), used to estimate \( G_x \), when there are limits on the slice lip bending.
4.2.2 Estimation of $G_x$ including dynamics

The paper machine might have some dynamics other than the delay. For example there might be that only the filtered profile, i.e. the profile out of the CD estimator, is available. Then the model given in Equation 4.105 can be used and one additional parameter, $a$ will be estimated for each row, and the pole averaged to give the pole of the system.

For this simulation the filter pole is set at 0.2 or $a = -0.2$. The same input as previously will be used and no limit will be used for the bending moments. The inputs and outputs are shown in Figure 4.55. The dynamics are clearly visible in the output. The estimates of $G_x$ do not seem to be affected by estimating one more parameter for each row as can be seen by comparing Figure 4.56, the estimates of $G_x$ when the filter pole is estimated as well, to Figure 4.49, the estimates of $G_x$ without the filter. The estimates of the filter pole converge quickly to the correct value of $-0.2$ as can be seen in Figure 4.57.

4.3 Identifying the inverse of the parameterized matrix

In the controller design (see Chapter 5), use is made of a matrix $K_0$, which is calculated from
Figure 4.53: Estimation of each row of $G_x$ when limit is put on the slice lip bending. The true parameters for this simulation, are shown with dotted lines.
Chapter 4: Identification

Figure 4.54: Estimate of $G_x$ after 59 scans when limits are put on the slice lip bending. The estimated matrix is denoted with * and dotted lines between its elements. The true matrix used in the simulation is denoted with o and solid lines.

$G_x K_0 = I$, where $I$ is a $m \times m$ matrix. When $G_x$ is a square matrix, the inverse of $G_x$ is equal to $K_0$. Equation 4.105 can then be written as

$$u_c(t - d) = \frac{K_0}{1 + a} y_c(t) + \frac{aK_0}{1 + a} y_c(t - 1) \quad (4.107)$$

Instead of having to calculate the inverse of the estimate of $G_x$, which might be close to singularity, the matrix $K_0$ could be identified directly. The identification of $K_0$ can only be done in open loop, because in closed loop the controller would be identified as well as the $K_0$ matrix. From Equation 4.104, for $i \in [1, \ldots, m_a]$ we can write:

$$u_c(t - d, i) = [B_1(i,:) \quad B_2(i,:)] \begin{bmatrix} y_c(t) \\ y_c(t - 1) \end{bmatrix} = \theta(i)^T \varphi(t). \quad (4.108)$$

From the above equation the dynamics and the inverse of $G_x$ can be calculated as

$$K_0 = B_1 + B_2$$

$$a = B_2 / B_1 \quad (4.109)$$
Figure 4.55: Input and output used to estimate the $G_z$ matrix when the dynamics is estimated as well.
Figure 4.56: Estimation of the parameterized interaction matrix \( G_x \) when the dynamics is estimated as well. The dotted lines show the true values.
where the division is an element by element division. Another way to estimate $K_0$ is to assume knowledge of $a$. The input $u_c$ can then be sent through the filter, with pole at $a$, and the filtered input used to identify the inverse matrix. The filtered input is

$$u_c'(t) = \frac{1 + a}{1 + aq^{-1}} u_c(t - d)$$

(4.110)

therefore

$$y_c(t) = G_z u_c'(t)$$

(4.111)

or

$$u_c'(t) = K_0 y_c(t)$$

(4.112)

Therefore the $K_0$ matrix can be identified using output and filtered input, provided the filter is known, or Equations 4.108 and 4.109 used.

Note that the elements of the vector $y_c$ are correlated as is evident from Equation 4.104, where a single input, $u_c(t, i)$, will in general affect all the outputs, $y_c(t)$. Therefore it can be expected that the estimates of $K_0$ will be slower to converge than for the estimate of $G_z$. Figure 4.58 shows the estimates of the 10 rows of $K_0$ using the same input and output data as before (see Figure 4.47). Comparing those results with Figure 4.49 it can be seen that the convergence is about 10 scans slower for the estimation of $K_0$. Nevertheless the estimates do converge to the true values and there is no visible difference after 59 scans between the estimated matrix and the true matrix used in the simulation as can be seen in Figure 4.59.

Figure 4.57: Estimation of $a$, for the 10 rows in the previous figure.
Figure 4.58: Estimation of each row of $K_0$, the inverse of $G_x$. The dotted lines show the true values.
4.4 Summary

In this chapter the interaction matrix, $G$ and the parameterized interaction matrix, $G_x$ have been identified.

It is shown that even though a response shape can be identified that is used in creating the interaction matrix, it is not trivial to construct the interaction from the response shape.

The identified $G$ matrix, with number of measurements as integer multiple of the number of actuators, is used to simulate a paper machine to use for identification of the parameterized matrix. Few examples are given on the identification of $G_x$, with or without bending limits. An example is also given on how to identify the inverse of the interaction matrix, which might be helpful in the control design.
Chapter 5
Basis weight Cross Directional Control

The main problem with controlling basis weight on paper machines is the cross coupling in the cross direction. It will be dealt with by the use of Gram polynomials introduced in chapter 2. A significant dead-time is present in many pulp and paper processes. The cross directional control of basis weight has been shown in chapter 3 to have transportation delay. The Dahlin controller, which is a dead-time compensator, is used because of its simple design and ease of use.

5.1 Dahlin controller and integrator antiwindup

The design of a Dahlin controller states that the closed loop response of the feedback system should be first-order plus dead time, with unit gain and a closed loop dead time equal to the open loop one [38]. The closed loop time constant is the only tuning parameter.

Let a discrete plant, \( H_{OL}(q^{-1}) \), be described by

\[
H_{OL}(q^{-1}) = \frac{B(q^{-1})}{A(q^{-1})} q^{-d} = \frac{b_0 (1 + b_1 q^{-1} + \cdots + b_n q^{-n_b})}{1 + a_1 q^{-1} + \cdots + a_n q^{-n_a}} q^{-d}
\]

With the use of a feedback, it is possible to design a controller, \( C = \frac{B_c(q^{-1})}{A_c(q^{-1})} \) (see Figure 5.60), such that the closed loop system is a first-order process plus a delay with steady state gain equal to one. It is written as:

\[
H_{CL}(q^{-1}) = \frac{1 - p}{1 - pq^{-1}} q^{-d}
\]

where \( p \), the tuning parameter, is the closed loop pole. The closed loop system in Figure 5.60 is then

\[
H_{CL}(q^{-1}) = \frac{H_{OL} C}{1 + H_{OL} C} = \frac{BB_c}{AA_c + BB_c q^{-d}} q^{-d} = \frac{1 - p}{1 - pq^{-1}} q^{-d}
\]

It is then straightforward to show that the controller becomes:

\[
C(q^{-1}) = \frac{B_c(q^{-1})}{A_c(q^{-1})} = \frac{1 - p}{1 - pq^{-1} - (1 - p)q^{-d}} \frac{A(q^{-1})}{B(q^{-1})}
\]

This is the Dahlin controller and it inverts the process dynamics, i.e. it cancels out all the process poles and zeros. The inverse of the process has to stable which means that all the process zeros have
Figure 5.60: Dahlin controller for a plant.

to be stable. In practice it is not possible to know the poles and zeros with infinite accuracy. The unstable process poles can therefore not be cancelled with unstable zeros, hence the process has to have stable process poles as well. The process zeros have to be well damped so as not to cause ringing in the control signal. The model used for basis weight control is a stable first order model with gain and a delay so problems with poorly damped or unstable zeros and unstable poles is not an issue.

In the slice lip controller used in the industry today, control action is typically taken every three scans. The Dahlin controller can then be designed based on the steady state value of the process, with the implicit assumption that the process reaches steady state within the three scans. Before we make that assumption, let’s look at the Dahlin controller with control taken every three scans. Assume the model is written as a first order process with delay of two, as

\[ H_{OL}(q^{-1}) = \frac{b}{1 - f_p q^{-1} q^{-2}} \]  

(5.117)

The model of the plant that the controller sees (every three scans) is then

\[ H_{OL}(q_s^{-1}) = \frac{b(1 + f_p + f_p^2 q_s^{-1})}{1 - f_p^3 q_s^{-1}} q_s^{-1} u(k) \]  

(5.118)

where \( q_s^{-1} \) is used to denote sampling every three scans. The controller, assuming the closed loop system is \( \frac{1 - E}{1 - p q_s^{-1}} q_s^{-1} \), is then

\[ C(q_s^{-1}) = \frac{(1 - p)}{1 - q_s^{-1}} \frac{(1 - f_p^3 q_s^{-1})}{b(1 + f_p + f_p^2 q_s^{-1})} \]  

(5.119)

If the dynamics of the process has also died out within the control interval, i.e. \( f_p \) is close to zero, the controller is then basically a steady state controller that takes into account the delay of the system. The controller is then written as

\[ C(q_s^{-1}) = \frac{(1 - p)A(1)}{1 - q_s^{-1}} = \frac{1 - p}{1 - q_s^{-1}} \frac{1 - f_p}{b} \]  

(5.120)

and is calculated every three scans. The control signal, which is calculated at \( k = 1, 4, 7, \ldots \), is then:

\[ u(k) = \frac{(1 - p)(1 - f_p)}{b} c(k) + u(k - 3) \]  

(5.121)
Its performance can be seen in the upper graph in Figure 5.61 for $f_p = 1 - b = 0.1$ and $p = 0.05$. It can be seen that the system does not quite reach steady state within the three scans and consequently there is a small overshoot in the output. To compare its performance to the performance of a control taken every scan, the same plant is used, but now the closed loop pole is set equal to the cubic root of 0.05 or $p = 0.37$, which gives the same time constant as 0.05 for control action every third scan. The response for the control taken every scan is shown in the lower half of Figure 5.61. It can be seen that the output is almost the same as for control action taken every third scan, except there is no overshoot. Much more frequent control action is taken in this latter case and not much is gained on performance for a system with such a fast dynamics. On the other hand much faster response is possible when control action is calculated every scan and hence it will be used in what follows.

**Figure 5.61:** Dahlin controller with control action taken every 3 scans with the closed loop pole, $p = 0.05$ and the open loop pole is, $f_p = 0.1$. Dahlin controller with control action taken every scan with the closed loop pole, $p = 0.37$ and the same open loop pole. The * denotes the output at every third scan.

An example of a Dahlin controller for a plant, $H_{OL} = \frac{1-f_p}{1-f_p q^{-1}}q^{-2}$, is according to Equation 5.116

$$C = \frac{\frac{1-p}{1-f_p} (1 - f_p q^{-1})}{1 - pq^{-1} - (1 - p)q^{-2}}$$

(5.122)

The plant has a steady state gain of one so the control signal will have to be at the same level as the output, for the output of the plant to follow the reference signal. Figure 5.62 shows the reference
Chapter 5: Basis weight Cross Directional Control

signal as well as the output and the control signal for a plant with a pole at \( f_p = 0.3 \) and with the closed loop pole chosen as \( p = 0.05 \). It can be seen that the control signal required to achieve the closed loop response is always greater than or equal to one. Now if the control signal is bounded at \( \pm 0.5 \) it is observed that the controller performance deteriorates (see middle plot in Figure 5.62), especially after long periods of constant input it takes a long time for the controller to respond. The reason is the integral action of the controller, it integrates an ever present error, which results from limited control action. When a new step is applied to the system the controller has to integrate (wind back), out of the saturation. It can be seen in the middle plot in Figure 5.62 that because of the integral action it takes considerable time for the control signal to respond to changes in the input. For example the step change applied at sample 13 is not seen in the output until at sample 19 instead of sample 15. If there would be a change in the reference signal (\( y_{des} \)) of short duration during the integrator wind back, it is possible that its effect will not be seen in the output. It is possible to guard against this windup [6]. The controller, \( C \), which is written in terms of its nominator and denominator as

\[
A_c u = B_c e
\]

The bounded control signal can be observed, with the aid of an observer \( A_o \). Adding the observed signal to both side of the previous equation results in

\[
A_o u + A_c u = B_c e + A_o u
\]

or

\[
A_o u = B_c e + (A_o - A_c)u
\]

A controller with antiwindup compensation (see Figure 5.63) is then given by [6]

\[
A_o v = B_c e + (A_o - A_c)u
\]

\[
u = \text{sat}\{v\}
\]

where \( \text{sat}\{\cdot\} \) is the saturation limit on the control signal. When there is no saturation this controller is equivalent to Equation 5.123, when it saturates it can be interpreted as an observer with dynamics given by \( A_o \). Figure 5.62 shows clearly that the performance of the controller with integrator antiwindup in place is superior to the case when no integrator antiwindup is in place. With antiwindup in place the controller responds immediately after the time delay of the process, and the output follows the input. The output will not be able to reach the reference signal because of the limited control action.
Chapter 5: Basis weight Cross Directional Control

Desired input-output response

Without integrator antiwindup

With integrator antiwindup

Figure 5.62: The signals in a closed loop control with a Dahlin controller; with and without limits on the control action, and with and without integrator antiwindup. The plant pole is $f_p = 0.3$ and the desired closed loop pole is $p = 0.05$.

Figure 5.63: Controller with antiwindup.

5.2 Control of basis weight

The basic control strategy for basis weight control using Gram polynomials is shown in block diagram in Figure 5.64. With the use of the pre-compensator, $K_0$, the controller can be designed for a SISO feedback system with controller $C$ and plant dynamics which is the same as the dynamics from $u_c$ to $y_c$. There is a delay in the model of the plant, therefore a Dahlin controller is used. For a
plant written as \( \frac{B(q^{-1})}{A(q^{-1})} q^{-d} = \frac{b}{1-f_p q^{-1}} q^{-d} \) the Dahlin controller is written as Equation 5.116. Limits will have to be placed on the control signal in order to protect the slice lip from excessive bending and the stroke of the slice lip is also limited. When the slice lip is in saturation the controller has to know, else the signal out of the controller might windup to a bigger and bigger value. To protect against the windup [6] the controller can be designed using Equation 5.126. For \( B(q^{-1}) = b \) and \( A(q^{-1}) = 1 - f_p q^{-1} \) the controller can then be written as

\[
A_o v_c(t) = K_0 \frac{1-\frac{p}{b}}{B} [e(t) - f_p e(t-1)] + \left[ A_o - \left( 1 - pq^{-1} - (1-p)q^{-d} \right) \right] u_c(t) \tag{5.127}
\]

For \( d = 1 \) and a dead beat observer, \( A_o = 1 \) the controller can be written as

\[
v_c(t) = K_0 \frac{1-\frac{p}{b}}{B} [e(t) - f_p e(t-1)] + u_c(t-1) \tag{5.128}
\]

and for a delay equal to two, the controller is

\[
v_c(t) = K_0 \frac{1-\frac{p}{b}}{B} [e(t) - f_p e(t-1)] + pu_c(t-1) + (1-p)u_c(t-2) \tag{5.129}
\]

Figure 5.65 shows the block diagram of the controller with antiwindup in place.
Estimation results from the previous chapter show that $G_x$ is close to a constant times the identity matrix. Here we use this information to initialize the estimated matrix. First the system is run in open loop to get a better estimate of the $G_x$ matrix. At scan 40 the system is put into closed loop with, $K_0$, the inverse of $G_x$ as the pre-compensator. The limit on the bending moment of the slice was arbitrarily set to be $\pm 0.23 \text{ Nm}$. The dynamics are simply a time delay of one scan and the closed loop pole, $p$, is set at 0.5. Figure 5.66 shows the inputs and outputs of the simulation both of the open loop using PRBS inputs and the closed loop with some step changes. It is observed that there are small overshoots for the parameters of 8th-10th order. Those overshoots arise when the controller needs to limit the slice lip bending. For this particular slice lip profile the above mentioned parameters need to be modified in order to limit the bending. Figure 5.67 shows the recursive estimate of the rows of the $G_x$ matrix, for the same period as the last figure. As mentioned above the estimated matrix is initially taken to be the identify matrix. Some of the parameters converge within 40 scans, when the system is in open loop. Other parameters do converge when the step changes are applied to the system at scans 60 and 90. Figure 5.68 shows the slice lip profile at scan 110 and the calculated bending moment. The figure also show the output and the desired output at the same scan. It can be seen in the figure that there is a difference between the two. The reason for the difference is that the bending moment of the slice lip is limited at $\pm 0.23 \text{ Nm}$. If those limits are removed the difference between the output profile and the desired profile disappears, at the cost of bending moments increasing substantially as can be seen in Figure 5.69.

It was mentioned in Chapter 3 that a conventional control might become unstable in the presence of uncertainty for a plant with a badly conditioned interaction matrix. The condition of the interaction matrix is of no concern if Gram polynomials are used, only the condition number of the parameterized interaction matrix. It was shown in Chapter 3 that the condition number of $G_x$ can be used as one of the selection criteria for the number of orthogonal polynomials used in the approximation of the profiles. As long as the condition number of the parameterized interaction matrix is low, the error term $G_x \Delta G_x^{-1}$ in the loop transfer matrix, $c(I + G_x \Delta G_x^{-1})$ will be small.

5.2.1 Removing polynomials from a profile

Another simulation similar to the one above was performed, with the dry weight profile assumed to be one of the profiles from the industrial data set (see the solid line in Figure 5.70). The Dahlin controller with closed loop pole at 0.5, is then to remove all the low order polynomials from the profile (the dashed line in Figure 5.70) without exceeding the bending limits of the slice. The initial
Figure 5.66: Input, $u_c$, and output, $y_c$. The system is initially in open loop but the loop is closed at scan 40.
Figure 5.67: Estimation of each row of $G_x$. The true parameters for this simulation, are shown with dotted lines. The system is initially in open loop but the loop is closed at scan 40.
Chapter 5: Basis weight Cross Directional Control

Figure 5.68: Slice lip profile, bending moments and output at scan 100.

Figure 5.69: Slice lip profile, bending moments and output at scan 100, when no limits are used for the slice lip.
position of the slice lip is assumed to be the same as one of the slice lip profile from the same industrial data set (see the dot-dashed line in Figure 5.71). The system is initially kept in open-loop to estimate the parameterized interaction matrix, and the loop is closed at time sample 40.

The initial and the final setting for the slice lip deflection, as well as their Gram polynomial approximations, are shown in Figure 5.71. It can be seen that only minor adjustments are needed for the slice lip position to remove the first 10 Gram polynomials from the profile as can be seen in Figure 5.72. Here the final output profile is shown with a solid line and its 10th order Gram polynomial approximation with a dashed line. That can be compared with Figure 5.70 to see how the low order polynomials have disappeared. The 2 sigma value of the final profile is $0.80 \, \text{g/m}^2$ down from $0.94 \, \text{g/m}^2$ for the initial profile. It is worth noting that a very little change is observed in the bending moments when the slice lip is manipulated to remove the 10 first polynomials from the output. Figure 5.73 shows the bending moments of the slice lip. The figure has actually two lines, one before control, and one after control, but the changes are so small that they cannot be distinguished on the graph. The bending moment changes at most $5 \, \text{Nm}$ from the initial bending moment, or only about 2% of the maximum value of the bending moments.
Chapter 5: Basis weight Cross Directional Control

Figure 5.71: Slice lip actuators settings.

Figure 5.72: Final weight profile.
5.2.2 Control of heavy weight paper

A paper machine with 132 actuators and 132 measurements is simulated assuming its response to the center actuator moving one unit is as shown in Figure 5.74. It can be seen that the response to a single actuator covers one fourth of the width of the machine. This is a similar response as suggested in [21] and [71].

This fictional paper machine is assumed to have an initial output profile that consists of a single sine wave across the width of the machine. This sinusoidal wave is supposed to be removed from the profile, but first the machine is run in open-loop to estimate the parameterized interaction matrix. The inputs are PRBS signals as can be seen in Figure 5.75. The outputs are also shown in the same figure. The estimates of the rows of $G_x$ are shown in Figure 5.76. It can be seen that there is a small bias in most of the parameters. The reason is the small amplitude in the input signal, which needs to be small in order to limit the stroke of the actuators. If the amplitude of the input signal is increased a better convergence is achieved. The final estimates are shown in Figure 5.77.

Figure 5.78 shows the final basis weight output, the desired basis weight, and the initial basis weight profile. The slice lip deflection required to reduce the initial basis weight profile to the final basis weight profile, is shown in the same figure, as well as its bending moment.
5.3 Comparison to conventional control

Figure 5.79 shows a block diagram of a very simple conventional controller utilizing the inverse of the interaction matrix to decouple the actuators for control. More elaborate scheme exists that take into account the control action as well as minimizing the output error [126]. For simplicity only the basic control scheme will be considered here, and the number of actuators are set to be equal to 250 and the number of measurements are four times the number of actuators, or 1000. The interaction matrix, $G$, has the same $p_m$ parameters as before (see Figure 4.45).

Figure 5.80 shows a control of a profile consisting of 5 sinusoidal waves when a conventional control is used. The five sinusoidal waves have wavelengths that are 100, 30, 10, 4, and 1 cycles per sheet width. It can be seen that the five sinusoidal wave are completely eliminated from the profile except the high frequency one which has still got a small magnitude. Notice that the slice lip deflection is very jagged and consequently the bending moments are very high.

When the process is controlled by utilizing Gram polynomials up to 40th order, it is observed that the slice lip is much smoother as can be seen in Figure 5.81. With the use of Gram polynomials in the control design, the lower frequencies have been eliminated but the higher frequencies, which would require high bending moments on the slice lip, are not altered. The lack of high frequencies...
Figure 5.75: Input and output for a heavy grade paper machine.
Figure 5.76: Estimates of each row of $G_x$ for a heavy grade paper machine.
Figure 5.77: Final estimates of $G_x$ for a heavy grade paper machine.

Figure 5.78: Final basis weight profile, the desired profile, and the initial profile. The slice lip deflection. The bending moments.
deflections of the slice lip results in the inability of the controller to eliminate the high frequencies in the output. What is gained, is that the maximum value of the bending moment is only 156Nm, compared to 5750Nm for the conventional control.

The condition number of the interaction matrix can cause problems for the control of the paper machine for certain type of disturbances, as was discussed in Chapter 3. Two examples are shown here, one for a badly condition matrix and the other for a better conditioned matrix. In the first a $105 \times 105$ interaction matrix with $p_0 = 1$, $p_1 = 0.82366$, and $p_2 = -0.41183$ is used. Those nominal values are assumed to be known and its reciprocal condition number is 0.013. The diagonal of the interaction matrix has uncertainty of the form of uniform random numbers between $-5 \cdot 10^{-9}$ and
Figure 5.81: Control with Gram polynomials.

Figure 5.82: Disturbance rejections using conventional control on a well conditioned interaction matrix.
The Dahlin controller is then to remove a disturbance, or an initial profile, in the output, which is the same as the minimum singular vector of the interaction matrix. The results are shown in Figure 5.82. It is observed that the disturbance is eliminated with the controller.

The second example uses an interaction matrix of the same size, but with the elements on first and the second diagonals slightly different. The values are: \( p_1 = 0.784935028 \), and \( p_2 = -0.392467514 \). The reciprocal condition number is now \( 1.43 \cdot 10^{-10} \), or much closer to zero (much worse) than in the first example. The same uncertainties are used and the control results are shown in Figure 5.83. It is seen that the controller goes in the wrong direction initially but is able to remove most of the disturbances from the profile. It is worth noting in order to remove the disturbances the slice lip has to make deflection that are 100 million times bigger than in the first example, or in other words not realizable. For both examples, controller using Gram polynomials will not reduce the high frequency initial profiles. The slice lip will therefore not be altered and consequently the bending moments are zero.

5.4 Chaos

From the diagonal shape of the parameterized interaction matrix, \( G_x \), which is apparent in this and the previous chapter, it is tempting to assume the matrix is a diagonal matrix with zero
entries outside its main diagonal. Then, only the elements on the diagonal of the $G_x$ matrix need to be identified. Furthermore it is assumed that the interaction matrix, $G$, is described by only two parameters; a parameter on the main diagonal, and another parameter on the diagonal above and below the main diagonal. The parameterized interaction matrix, $G_x$, can then be described by two parameters and estimated with $k = 1$ using Equation 3.62. Having obtained the $G_x$ matrix the pre-compensator matrix, $K_0$, is then used to calculate current control signal. To simplify the calculation of the pre-compensator matrix, it is further assumed that the $G_{x1}$ matrix (see Equation 3.59) is also a diagonal matrix so the matrix $K_0$ is calculated by

$$K_0(i,i) = \frac{1}{p_0 + p_1 G_{x1}(i,i)} \quad (5.130)$$

For simulation purposes and interaction matrix, $G$, is created for 105 actuators and 105 measurements with $p_0 = 1$ and $p_1 = 0.4$. The form of the $G_x$ matrix is plotted in Figure 5.84 for 10th order polynomials used to approximate the measured basis weight profile and also for the control of the slice lip. The controller is chosen as a Dahlin controller, with the model of the process as $B(q^{-1}) = 1$, i.e. all the gain will be attributed to the interaction matrix, $G$. The delay of the simulated process will be assumed to be known and equal to one. The Dahlin controller is then

$$C(q^{-1}) = \frac{1 - p}{1 - pq^{-1} - (1 - p)q^{-d}} B(q^{-1})$$

$$= \frac{1 - p}{1 - q^{-1}} \quad (5.131)$$

and $p$, the closed loop pole, is used as the design parameter. The block diagram shown in Figure 5.64 is used. The paper machine is modeled by an interaction matrix that is block diagonal with uniformly distributed random numbers on the 3 main diagonals. It is maintained constant throughout. The desired profile is a sine wave and is approximated with 10 Gram polynomials of order one to 10 inclusively. The Gram polynomial approximates the sine wave perfectly as can be seen in the upper half in Figure 5.86, where the desired profile is shown as well as the difference, magnified 10000 times, between the profile and its approximation. The lower left graph in the figure shows the parameters of the orthogonal approximation of the desired profile. When this system is simulated with the aforementioned control and estimation method, the output profile is found to be the same as the desired profile and the parameters converge to their true values.

Noise is then added to the entries on the diagonals in the interaction matrix. The noise is uniformly distributed random numbers between -0.2 and 0.2. The range of values for $p_{ij}$ are between
Chapter 5: Basis weight Cross Directional Control

Figure 5.84: Meshed surface of the $G_x$ matrix for $G$ of size $105 \times 105$ with $p_0 = 1$ and $p_1 = 0.4$.

Figure 5.85: Estimates of $p_0$ and $p_1$. 

106
Figure 5.86: The desired profile and 10000 times the difference between the profile and its approximation with 10th order Gram polynomials. The smaller graphs show the parameters of the approximation for two sets of the Gram polynomials.

0.8 and 1.2 and the range for $p_1$ is between 0.2 and 0.6. The $G_x$ matrix for this particular interaction matrix is shown in Figure 5.87. It can be compared to Figure 5.84 and it is seen that their difference is very small. The process is simulated with this particular structure of the interaction matrix. The estimated $p_0$ and $p_1$ are shown in Figure 5.85 for 100 interaction matrices with different noises. It can be seen that the estimated parameters are around the true values even though there is a small bias for each experiment. The output profile is very close to the desired profile.

The same tests are done using different form of the Gram polynomials, i.e. the polynomials described by Equations 2.12 – 2.14. The parameters needed to approximate the sinusoidal profile are shown in the lower right graph in Figure 5.86. The $G_x$ matrix for $G$ with dimension $105 \times 105$, $p_1 = 1.0$ and $p_2 = 0.4$, is plotted in Figure 5.88. It can be noted that the first order polynomial has quite a large effect on the ninth order polynomial, because the ninth order polynomial is much smaller in amplitude than the first order. The same goes for the effect the second order polynomial has on the 10th order polynomial. For an interaction matrix with no noise the output profile is found to be the same as the desired profile, just like the case when the other form of the Gram polynomials was
used. When the uniformly distributed noise is added to the interaction matrix the system is observed to go into chaos with on-line identification of the plant matrix. Figure 5.90 shows a bifurcation diagram (the output for CD point 10 as a function of the previous output for that point) for certain values of the closed loop pole, $p$. In the figure, pseudo phase plane plots for different values of $p$ are also plotted. The fractal dimension can be estimated by calculating the correlation dimension for the seven phase plane plots in Figure 5.90, using an algorithm in [97]. It is found that the dimension is zero for periodic orbit, for the limit cycle it is equal to one and for $p = 0.580$ and $p = 0.550$ the dimension is 1.5, which suggests the system is chaotic. By zooming in on Figure 5.90 it can be observed that the output goes from being at a fixed value, when $p = 0.64$, to a limit cycle, when $p = 0.60200$. Then at $p = 0.60199$ the system has period three cycle, and when $p$ is decreased further, the system has period 6, 12, 24 etc. cycle oscillation and then around $p = 0.584501$, chaos occurs. The values of $p$ where the bifurcation occurs is shown in Table 5.3. The table also lists the ratio between the difference of the points of bifurcation. The ratio is very close to Feigenbaum's number (approximately 4.669). The small difference is mainly because of errors in finding the exact point of bifurcation. For example the point where the system goes from period 96 to period 192 appears around $p = 0.58451000$, but according to previous two values of point of bifurcation, and
the Feigenbaum's number, this should appear at $p = 0.58450992$ which is within the error bounds for determining the points of bifurcation.

![Figure 5.88: Meshed surface of the $G_x$ matrix using the Gram polynomials from Figure 2.3.](image)

A 1% difference in initial condition on the desired value of the output profile does affect the response in such a way that it becomes uncorrelated to the nominal response after about 25 samples (see Figure 5.91), which further implies chaotic motion.

It should be emphasized here that the chaotic behaviour only show up in the system when the badly scaled polynomials are used.

5.5 Summary

In this chapter the Dahlin controller has been described, with and without integrator antiwindup. The Dahlin controller is then used together with Gram polynomial approximation of the measured profile to control the basis weight on a paper machine by means of adding Gram polynomials to the slice lip of the headbox. The control method is compared to a conventional method, which involves the inverse of the interaction matrix. The method using Gram polynomials is always found to require
much less bending on the slice lip, as it only controls the low frequencies in the output profile. It is also not sensitive to the condition number of the interaction matrix.

It is observed that when some approximations are made and a form of the Gram polynomials are used that has bad scaling, the control of the process can become chaotic. The chaotic behaviour has only been observed for the polynomials that are badly scaled. No instances of chaos are observed when Gram polynomials are used, which all have a comparable amplitude.
Chapter 5: Basis weight Cross Directional Control

Figure 5.90: A bifurcation diagram and pseudo phase plane plots for different values of $p$. 

\[ p = 0.64000 \]
\[ p = 0.60200 \]
\[ p = 0.55000 \]
\[ p = 0.59000 \]
\[ p = 0.60199 \]
\[ p = 0.58000 \]
\[ p = 0.58500 \]
Table 5.3 The points of bifurcation and the ratio of the differences between points of bifurcation.

<table>
<thead>
<tr>
<th>period</th>
<th>$p$</th>
<th>$\frac{p_k-p_{k-1}}{p_{k+1}-p_k}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>0.6013</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>0.5882</td>
<td>4.613</td>
</tr>
<tr>
<td>24</td>
<td>0.585360</td>
<td>4.223</td>
</tr>
<tr>
<td>48</td>
<td>0.5846875</td>
<td>4.598</td>
</tr>
<tr>
<td>96</td>
<td>0.58454125</td>
<td>4.68</td>
</tr>
<tr>
<td>192</td>
<td>0.58451000</td>
<td></td>
</tr>
</tbody>
</table>

Figure 5.91: Response of the system with $p = 0.55$, for 1% difference in initial condition on the desired value of the CD value at location 10.
Chapter 6

Summary and Future work

6.1 Summary

In this thesis it has been shown how the basis weight on a paper machine can be controlled by manipulating the slice lip on the headbox, taking into account limits on the bending moments of the slice lip. A model that relates slice lip movement to basis weight response has been developed. The main feature of the model is an interaction matrix that describes how a movement of a single actuator will effect a large section of the measured profile through bending of the slice lip which sets up spreading waves on the forming wire, that further spread out on the wire. Some of the problems in creating the interaction matrix are discussed and it is found that it is difficult to create a matrix that will not create ripples in the model output. Another more serious problem is that the interaction matrix can, for many paper machines, become singular which might create problems when controlling the basis weight.

Orthogonal polynomials with equal weight given to every data point, called Gram polynomials, are then introduced as a means of approximating the profile and then the parameters of the polynomials are fed back to the controller. Several methods are given to create the Gram polynomials, both sequentially and recursive algorithms. A unified sequential formula is given which all the other methods can be written in terms of. It is observed that a bad selection of the scaling of the polynomials can lead to problems in control. For a particular selection of scaling factors, it was shown that the Gram polynomials approach the Legendre polynomials for number of data points much larger than the order of the polynomial. The Gram polynomials are therefore sometimes referred to as discrete Legendre orthogonal polynomials. Some methods are discussed for finding the optimum number of polynomials to approximate a profile. It is found that a modified AIC gives a good balance between complexity and fit. When choosing a polynomial order, it is also important to consider the condition number of the parameterized interaction matrix, if a robust performance is desired.

The parameters of the Gram polynomial approximation of the slice lip can be related to the bending limit of a slice lip by modelling the slice lip as a flexible beam with multiple forces. The parameters can therefore be manipulated to limit the bending moments for the whole slice.
Compared to conventional control methods, the method using Gram polynomials is always found to require much less bending on the slice lip, as it only controls the low frequencies in the output profile. The method is also not sensitive to the condition number of the interaction matrix, which cannot be chosen at will. It does however depend on the condition number of the parameterized interaction matrix, which can be manipulated by adding or removing polynomial order.

It is observed that when some approximations are made and a form of the Gram polynomials are used that has bad scaling, the control of the process will become chaotic. No instances of chaos are observed when Gram polynomials are used, which all have a comparable amplitude.

6.2 Future work

As has been said in this thesis the mapping, between the actuators and the measurements, which is the key element of the current control systems, is effectively eliminated with the control system proposed in this thesis. The problem with shrinkage of the paper sheet has not been addressed in this work. If there is for example different shrinkage at the edges than the middle of the cross direction a different set of polynomials might be required.

For a paper machine with multiple actuators controlling the same CD profile orthogonal polynomials might be used in classifying the measured profile such that one set of actuators will control certain set of polynomials and another set of actuators control another set of polynomials.
References


