Ruprecht-Karls-Universität Heidelberg
Fakultät für Physik und Astronomie
Max - Planck - Institut für Kernphysik

HD-IHEP98-10
HD-ASIC-43-1098

Entwurf und Test eines optisch konfigurierbaren Widerstandsnetzwerkes

Jörg Langeheine

ASIC-Labor http://wwwasic.ihep.uni-heidelberg.de
Schröderstraße 90 D-69120 Heidelberg

Fakultät für Physik und Astronomie
Ruprecht-Karls-Universität Heidelberg

Diplomarbeit

im Studiengang Physik

ausgeführt von

Jörg Langeheine

Oktober 1998

Tel: 06221/544359, Fax: 06221/544345

E-Mail: langehei@asic.uni-heidelberg.de

Institut für Hochenergiephysik, ASIC-Labor

Schröderstraße 90, D-69120 Heidelberg

Inhalt

In dieser Arbeit wurde ein neuartiges Konzept zur optischen Konfiguration der Gewichte in einem analogen Widerstandsnetzwerk konzipiert, simuliert, gebaut und getestet. Das in einem $0.6\,\mu$m CMOS Prozeß realisierte System besteht aus einem eindimensionalen Netzwerk mit 64 Knoten und zwei Positionsdetektoren zur präzisen Ausrichtung des Gesamtsystems. Die optische Konfiguration erfolgt über die kontrollierte Ablenkung eines Laserstrahls mittels eines Galvanometer-Scanners. Das Netzwerk führt eine lokale Mittelung aus, deren Reichweite über die Gewichtsfaktoren optisch eingestellt werden kann. Bei der Evaluierung der Eigenschaften wurde das System mit einfachen Eingangsmustern elektrisch stimuliert. Die Konfiguration der Gewichte erfolgte entweder elektrisch oder optisch. Das Antwortverhalten des Netzwerkes (response) entspricht weitestgehend den Erwartungen aus Simulationen. Die allen Analogschaltungen inhärenten Offsetvariationen wurden quantifiziert.

Abstract

In this thesis a novel concept for the optical configuration of synaptic connections in an analog resistive network is developed, simulated, built and experimentally verified. The system has been realized using a $0.6\,\mu$m CMOS process and consists of a one dimensional network with 64 nodes and two position detectors for precise alignment of the complete system. The optical configuration is done with a laser beam controlled by a galvanometric scanner. The network carries out a local averaging. The range can be adjusted by optically controlled weight factors. The behaviour of the system has been investigated using simple electrical input patterns. The weights were configured either electrically or optically. The response of the network corresponds to a large extent to the expectations from simulations. Offset variations inherent to all analog circuits have been determined.

Inhaltsverzeichnis

Einführung

Heutige Computer unterscheiden sich wesentlich von biologischen Rechensystemen, wie beispielsweise dem menschlichen Gehirn, sowohl was ihre Leistungsfähigkeit bzgl. unterschiedlicher Problemstellungen angeht, als auch in ihrer Struktur. Aufgaben die in irgendeiner Weise mit dem Begriff Intelligenz in Verbindung gebracht werden, wie z.B. die Erkennung von Mustern, werden vom menschlichen Gehirn viel effektiver bewältigt, als von Computern, wohingegen einfache Berechnungen oder die Speicherung großer Datenmengen von Computern genauer und schneller erledigt werden können als vom Menschen. Um die außerordentlichen Leistungen des menschlichen Gehirns auf die maschinelle Verarbeitung von Daten zu übertragen, wurde in den letzten zehn Jahren verstärkt versucht, biologische Berechnungsprinzipien in Form *neuronaler Netze* nachzuahmen.

Bei der Realisierung neuronaler Netze in heutiger CMOS-Technologie ist die Komplexität der zu implementierenden Netzwerke durch den Platzbedarf der synaptischen Verbindungen zwischen den Neuronen beschränkt, durch deren Stärke der Algorithmus, den das Netzwerk ausführt, bestimmt wird. Daher scheint es attraktiv, den Platzbedarf der synaptischen Verbindungen dadurch zu reduzieren, daß man ihre Stärken optisch durch ein räumlich und zeitlich moduliertes Lichtintensitätsmuster einstellt.

Um die Machbarkeit dieses Ansatzes zu überprüfen, ist ein eindimensionales Netzwerk aus als Transistorschaltungen implementierten Widerständen in CMOS-Technologie realisiert worden, dessen Verhalten sich durch Variation der elektrisch oder optisch einstellbaren Widerstandswerte verändern läßt. Das Netzwerk führt für ein aus 64 Eingangsspannungen bestehendes räumliches Eingangsmuster eine lokale Mittelung durch. Die Reichweite der Mittelung wird durch die Wahl der einzelnen Widerstände bestimmt. Das Ziel ist nun das Verhalten des Widerstandsnetzwerkes optisch zu konfigurieren.

Die vorliegende Arbeit umfaßt fünf Kapitel, die die Beschreibung des Projektes wie folgt einteilen: Im ersten Kapitel wird das Projekt zuerst motiviert und das Verhalten eindimensionaler Widerstandsnetzwerke diskutiert. Im Schlußteil dieses Kapitels wird das Konzept der optischen Programmierung beschrieben. Im zweiten Kapitel werden an einem Testchip durchgeführte Voruntersuchungen zu der Realisierbarkeit des vorgestellten Konzeptes geschildert. Kapitel drei beschreibt den entwickelten Chip zusammen mit seinem auf Grund von Simulationen erwarteten Verhalten. Die an diesem Chip durchgeführten elektrischen Messungen und der verwendete Meßaufbau werden in Kapitel vier dargelegt. Die Arbeit endet mit der Darstellung der an dem entworfenen Chip durchgeführten optischen Messungen. Dazu wird zuerst der optische Aufbau erläutert, und anschließend werden die Ergebnisse interpretiert.

Kapitel 1

Idee und Motivation des Projektes

1.1 Neuronale Netze

Der Versuch, Maschinen, oder genauer Computer, mit menschlichen bzw. 'intelligenten' Fähigkeiten auszustatten, hat unter anderem zur Theorie neuronaler Netze geführt. Im Unterschied zu heutigen Computern, die sequentiell mit sehr hoher Geschwindigkeit Byte für Byte verarbeiten, besteht das menschliche Gehirn aus sehr vielen (etwa 100 Milliarden) 'Rechenzellen' (Neuronen) einfacher Funktionalität, die wiederum mit vielen anderen dieser Neuronen verbunden sind. Mit künstlichen neuronalen Netzen wird versucht, den Biocomputer Gehirn nachzubauen, einerseits um mehr über die Funktionsweise des Gehirns zu erfahren, andererseits um Computern, oder allgemeiner Maschinen, menschliche Fähigkeiten wie z.B. das Erkennen von Mustern beizubringen. Auf Grund der sequentiellen Natur heutiger Rechnerarchitekturen scheint es unvermeidlich, neuronale Netze als Hardware zu implementieren, wenn man ihre parallele Struktur voll ausnutzen möchte.

1.1.1 Biologische Motivation

Berechnende biologische Systeme, wie z.B. das menschliche Gehirn, bestehen aus miteinander verbundenen Neuronen, die untereinander sehr stark vernetzt sind. Abb. 1.1 ([Ritter 90]) zeigt den schematischen Aufbau eines typischen Neurons. Seine Hauptbestandteile sind Dendritenbaum, Zellkörper, Synapse und Axon. Beim Axon handelt es sich um eine Nervenfaser, die den Ausgang des Neurons bildet. Das Axon verbindet das Neuron mit anderen Neuronen durch Synapsen, die sich entweder direkt am Zellkörper oder an den Dendriten des Neurons befinden.

Die Übertragung der elektrischen Signale findet durch positiv und negativ geladene Ionen statt. Die Zellwand des Neurons ist dabei nur für positive Ionen (z.B. K^+ oder Ca^{2+}) durchlässig, so daß sich durch Diffusion ein negatives Ruhepotential von etwa -70mV einstellt. Dabei wirken Zellinneres und -äußeres wie die Elektroden eines Kondensators mit der Zellwand als Dielektrikum und die Ionenkanäle wie Widerstände für den (Ionen-) Strom. Nach einer Änderung des Zellpotentials gegenüber dem Ruhepotential wird sich dieses jenem wieder exponentiell nähern. Die entsprechende Zeitkonstante (RC) beträgt bei Nervenzellen etwa 2.4 ms ([Rojas 93]). Wird nun das Potential im Zellkörper über eine bestimmte Schwelle (etwa -40 mV) erregt, sendet das Neuron einen elektrischen Spannungsimpuls aus. Je stärker ein Neuron durch andere Neuronen erregt wird um so öfter 'feuert' es solche Impulse ab. Die Aktivität eines Neurons ist also eine analoge Größe, deren Wert durch eine Frequenz ausgedrückt

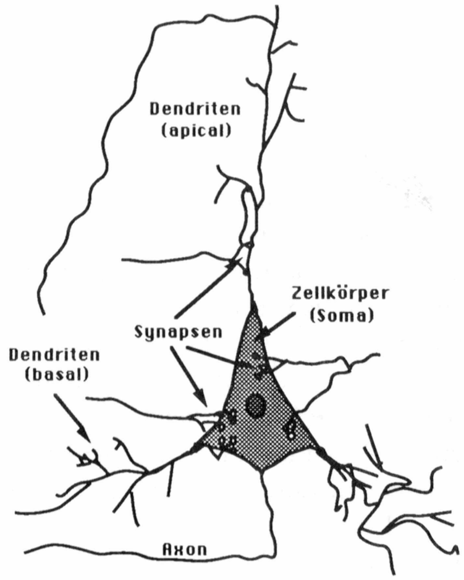

Abbildung 1.1: Schematischer Aufbau eines Neurons (Aus [Ritter 90]). Dargestellt ist hier eine Pyramidenzelle, die etwa 60 % der Neuronen der Hirnrinde des Menschen ausmachen, die restlichen 40 % sind sogenannte Sternzellen.

wird.

Auf der Eingangsseite bewirken die Aktionspotentiale bei den Synapsen die Ausschüttung von sogenannten Neurotransmittern. Diese öffnen bei der Zielzelle je nach Art bestimmte Ionenkanäle. Dadurch können dann entweder positive oder negative Ionen in die Zielzelle eindringen und ihr Potential erhöhen oder erniedrigen. Im ersten Fall spricht man von erregenden oder exzitatorischen Synapsen, im letzteren von hemmenden oder inhibitorischen. Die Ladungsmenge, die durch einen elektrischen Impuls an der Synapse in die Zielzelle ausgeschüttet wird, hängt von der Ausprägung der synaptischen Verbindung ab, z.B. realisiert durch die Anzahl vorhandener Ionenkanäle, die geöffnet werden können.

Im Zellkörper des Neurons kommt es also zu einer Summierung der durch die Synapsenstärken gewichteten Signale anderer Neuronen. Überschreitet diese Summe eine gewisse Schwelle, wird das Neuron selbst aktiv, feuert, und kann so zur Erregung oder Hemmung anderer Neuronen beitragen.

Bei dem in Abb. 1.1 gezeigten Neuron handelt es sich um eine sogenannte Pyramidenzelle, die 60% der Neuronen in der Hirnrinde des Menschen ausmachen und langreichweitige Axonen mit meist erregend wirkenden Synapsen besitzen. Ihnen wird die eigentliche Informationsverarbeitung zugeschrieben. Die restlichen 40% werden von sogenannten Sternzellen gebildet, die gewöhnlich nur mit ihrer unmittelbaren Umgebung sternförmig verbunden und von inhibitorischer Wirkung sind. Sie dienen der Stabilisierung des nicht nur massiv parallelen sondern auch stark zurückgekoppelten Systems.

Es sollte darauf hingewiesen werden, daß es sich hier um eine vereinfachte, exemplarische Darstellung handelt. So unterscheidet man beispielsweise viel mehr Neuronenarten als nur die Pyramiden- und Sternzellen und kennt auch Synapsen, die rein elektrisch funktionieren.

Untersuchungen der Hirnrinde des Menschen haben ergeben, daß die Neuronen dort in höheren funktionalen Einheiten organisiert sind, die wegen ihrer Form Mikrosäulen genannt werden. Diese sind nebeneinander zu sogenannten Rindenfeldern zusammengefügt, die bestimmten Aufgaben- bzw. Körpergebieten zugeordnet werden können.

1.1.2 Mathematische Abstraktion

Inspiriert durch das biologische Vorbild kann man versuchen die *wesentlichen* Merkmale der Neuronen in einer funktionale Beschreibung zusammenzufassen. Abb. 1.2 zeigt eine solche Abstraktion des Neurons, das hier mit dem Index i bezeichnet ist. Das Neuron erhält die

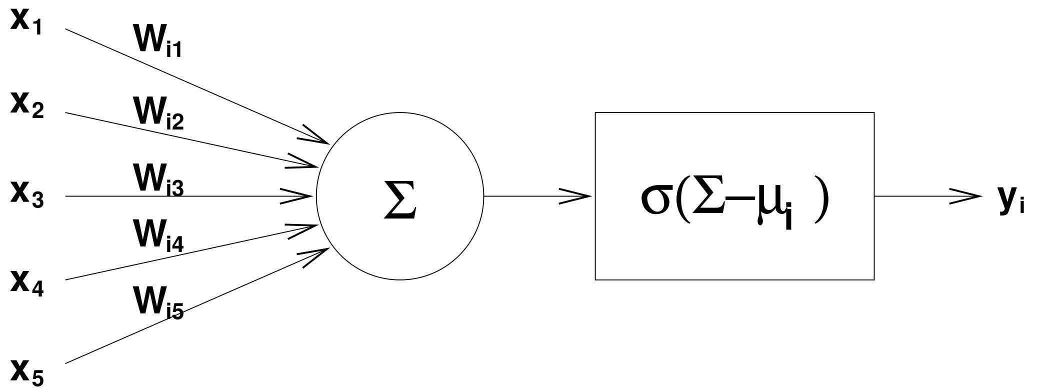

Abbildung 1.2: Abstrakte Darstellung eines Neurons.

Eingangssignale x_1 bis x_5 und gibt den Wert y_i aus. Die Eingangssignale werden durch Multiplikation mit den Synapsenstärken W_{ij} gewichtet. Das μ_i stellt den Schwellwert, der für ein Feuern des Neurons nötig ist, dar. Die Stärke des Feuerns wird aus der Differenz zwischen den summierten und gewichteten Eingängen und dem Schwellwert durch die Funktion σ berechnet. Mathematisch formuliert ergibt sich:

$$y_i = \sigma\left(\sum_j w_{ij} - \mu_i\right) \tag{1.1}$$

Aus dem biologischen Vorbild ergeben sich folgende Forderungen an die Funktion $\sigma(x)$.

$$\begin{aligned}
\sigma(x) &\geq 0 \qquad \forall x \in \mathcal{R} && \text{(1.2a)}\\
\lim_{x\to\infty} \sigma(x) &= const && \text{(1.2b)}\\
\lim_{x\to-\infty} \sigma(x) &= 0 && \text{(1.2c)}\\
\sigma'(x) &\geq 0 \qquad \forall x \in (R) && \text{(1.2d)}
\end{aligned}$$

Die erste Forderung ist klar, da das Neuron nicht mit negativer Frequenz feuern kann. Ebenso scheinen eine asymptotisch erreichte Maximalaktivität des Neurons und ein Zurückgehen der Aktivität auf 0 bei inhibitorischer Stimulation wie in (1.2b) und (1.2c) erfaßt sinnvoll. Schließlich drückt (1.2d) die Forderung nach einem eindeutigen Zusammenhang zwischen Reizung und Aktivität des Neurons aus. Verlangt man in (1.2c) daß der Grenzwert gegen 1 als Konstante strebt, so spricht man von der Klasse der sigmoiden Funktionen. Beispiele sind die Heaviside'sche Sprung- und die Fermifunktion, die beide in Abb. 1.3 dargestellt sind.

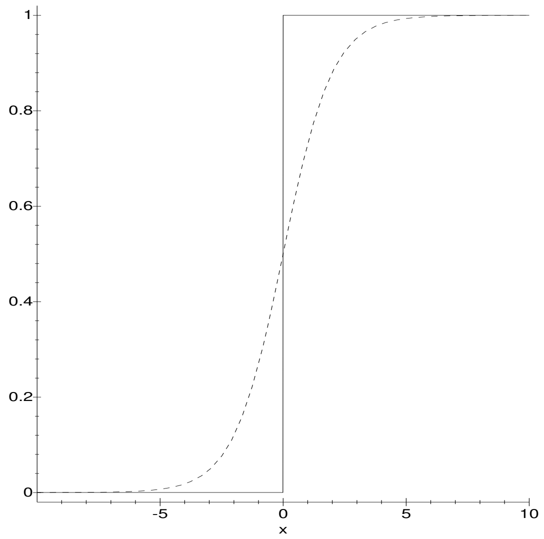

Abbildung 1.3: Graph der Heaviside'schen Sprungfunktion

$$\theta(x) = \begin{cases} 0 & x \leq 0 \\ 1 & x > 0 \end{cases}$$

(durchgezogen) und der Fermifunktion

$$f(x) = \frac{1}{1 + exp(-x)}$$

(gestrichelt).

In der Theorie neuronaler Netzwerke wird davon ausgegangen, daß es nicht auf die genaue Form der sigmoiden Funktion ankommt, sondern nur auf deren Nichtlinearität. (siehe z.B. [Hertz 91]).

Für eine mechanische Berechnung (z.B. einer Funktion durch einen Computer) werden von der Berechenbarkeitstheorie drei Mechanismen als wesentlich erachtet. Die Verarbeitung von Daten, ihre Übertragung, und ihre Speicherung. In heutigen Computern ist die Verarbeitung unter anderem in der CPU realisiert, Speicherung findet im Hauptspeicher, auf der Festplatte oder CD-ROMs statt. Wie sind nun Daten, oder Informationen zur Datenverarbeitung, also das, was man im Zusammenhang mit Computern das Programm nennen würde, in einem neuronalen Netz gespeichert? Hier ist die strikte Trennung zwischen Information und 'Programm' aufgehoben: Während die zu verarbeitenden Daten durch den Aktivitätszustand der Neuronen dargestellt werden, ist der Algorithmus durch die Verteilung und Stärke der synaptischen Verbindungen W_{ij} und die Schwellwerte μ_i codiert. Wenn man jetzt noch die nicht vorhandenen Verbindungen durch Synapsenstärken $W_{ij} = 0$ beschreibt und die Schwellwerte μ_i als für ein Neuron i globalen Offset der Synapsenstärken auffaßt, wird die Funktion eines neuronalen Netzes ausschließlich durch die Synapsenstärken bestimmt.

Biologische Systeme werden aber nicht extern programmiert, sondern zeichnen sich durch ihre Lernfähigkeit aus, die es ihnen ermöglicht, sich selbständig den Anforderungen der Umgebung anzupassen. Im Bild der neuronalen Netze kann man das so verstehen, daß die richtigen Synapsenstärken, oft auch Gewichte genannt, im Hinblick auf die jeweilige Aufgabe erst 'gelernt' werden müssen.

1.1.3 Hardware Implementierung

Wie in der Einleitung schon angedeutet, möchte man neuronale Netze nicht nur mit Hilfe von Computern simulieren, sondern, um ihre parallele Struktur wirklich ausnutzen zu können, auch in Hardware nachbilden. Aus mehreren Gründen bietet sich dafür ein Entwurf in VLSI[1] Technologie an: Geringer Platzbedarf und vergleichsweise niedriger Stromverbrauch ermöglichen einen variablen Einsatz und einen hohen Grad an Komplexität (gemessen an dem Maß-

[1] VLSI steht für **V**ery **L**arge **S**cale **I**ntegration

stab heutiger künstlicher neuronaler Netze) bei moderatem technologischen Aufwand. Ein einfaches analoges elektronisches Modell eines Neurons zeigt Abb. 1.4. Auf der linken Seite

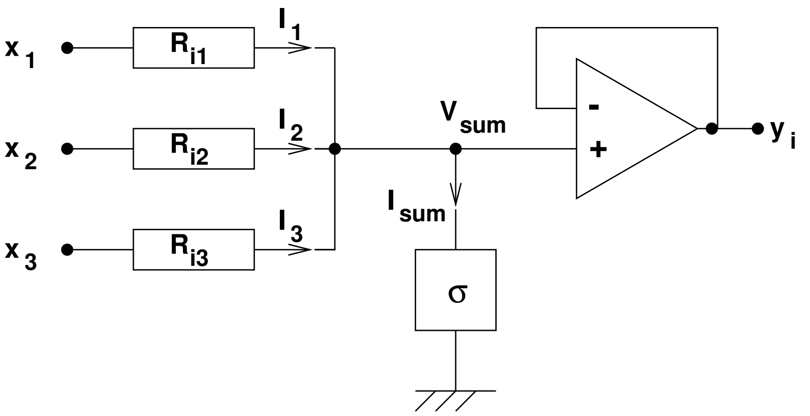

Abbildung 1.4: Elektronische Realisierung eines Neurons mit Widerständen als Synapsen.

sind die Eingangsspannungen der anderen Neuronen mit $x_1 \ldots x_3$ gekennzeichnet. Diese erzeugen den Strom I_j durch die Widerstände R_{ij}, der dann zu

$$I_{sum} = \sum_{j=1}^{3} \frac{x_j - V_{sum}}{R_{ij}} \tag{1.3}$$

summiert wird. Der mit σ bezeichnete und hier nicht näher spezifizierte Transimpedanzverstärker wandelt diesen entsprechend einer sigmoiden Kennlinie in eine Spannung um, deren Impedanz dann noch von einem Spannungsfolger so verringert wird, daß die nächste Stufe von Neuronen damit angesteuert werden kann.

Das obige Konzept ermöglicht zwar die lineare Gewichtung der Einganssignale und deren Summation, läßt aber die Frage nach der Variation der Synapsenstärken offen. Inhibitorische Synapsen lassen sich so gar nicht nachbilden. Man könnte diesen Mangel aber durch einen entsprechenden Offset des die sigmoide Funktion darstellenden Teils ausgleichen. Für eine VLSI Version eines neuronalen Netzes wäre es also wünschenswert eine Schaltung zu haben, die eine zumindest teilweise lineare Strom-Spannungs-Kennlinie besitzt, deren Steilheit elektronisch vorgegeben und beliebig lange abgespeichert werden kann.

Wenn jedes Neuron mit jedem anderen Neuron verbunden wird, wächst die Anzahl der Synapsen quadratisch mit der der Neuronen. Typischerweise wird man dem Netzwerk eine Struktur vorgeben, die bestimmte Verbindungen a priori ausschließt. Trotzdem wird für kleine Netzwerke mit einigen bis einigen zig Neuronen gelten:

$$\#(Synapsen) = \alpha \cdot \#(Neuronen)^2 \qquad \alpha \in]0,1] \tag{1.4}$$

Folglich wird der Platzbedarf der Synapsen für eine Implementierung der limitierende Faktor sein. In der Natur ist dieses Problem weniger gravierend, da dort drei Dimensionen zur Verfügung stehen, wogegen es bei der Benutzung heutiger VLSI-Technologie, die auf die Erzeugung zweidimensionaler Strukturen beschränkt ist, ein schwerwiegendes Problem darstellt.

Eine mögliche Lösung dieses Problems wäre, die Information der Gewichtsfaktoren optisch auf den Chip aufzubringen, z.B. durch Abscannen des Chips mit einem intensitätsmodulierten Laserstrahl. Das Problem der Speicherung und Berechnung der nötigen Gewichte wäre somit auf einen Computer ausgelagert. Gegenüber einer reinen Softwarelösung wäre der (Zeit-)

Aufwand beim Trainieren des neuronalen Netzes zwar eher noch höher, bei der Anwendung auf ein konkretes System könnte man aber alle Vorteile der parallelen Implementierung des Netzes nutzen.

In der voliegenden Diplomarbeit soll untersucht werden, inwieweit eine solche 'optische' Programmierung mit einem CMOS[2] Chip möglich ist. Als Beispiel wurde ein Widerstandsnetzwerk mit optisch einstellbaren Widerständen entworfen und getestet.

1.2 Widerstandsnetzwerke

Zukünftige intelligente Systeme bedürfen eines Instruments, um ihre Umwelt erkennen zu können. In diesem Zusammenhang existiert reges Interesse, eine solche Erkennung durch sogenannte *Vision-Chips* zu leisten, durch Mikrochips also, die Bildaufnahme und -verarbeitung auf einem einzigen Stück Silizium vereinigen (Einen Überblick über *Vision-Chips* findet man bei [Moini 97]).

Der erste Schritt zu einer räumlichen Bildverarbeitung besteht meist in einem Kantenerkennungsalgorithmus. Typischerweise wird das Bild dazu zuerst geglättet, z.B. durch einen Gaußfilter, und danach ein Differentialoperator, z.B. ein Laplacefilter, darauf angewandt. Der erste Schritt dient dazu, Bildfehler, die bei der Bildaufnahme entstanden sind, wieder auszumitteln. Der zweite sorgt für die Kantenerkennung, würde aber ohne den ersten auch in Wirklichkeit nicht vorhandene Kanten hervorbringen. Um eine Glättung im obigen Sinne zu erreichen benutzt man in analogen *Vision-Chips* oft Widerstandsnetzwerke in der von C. Mead ([Mead 89]) vorgeschlagenen und unten diskutierten Form. Obwohl sein Faltungskern nicht die für optimal erachtete Gaußform besitzt (Eine Implementation, die das leistet, ist in [Kobayashi 91] beschrieben), wird es seiner Einfachheit wegen komplizierteren Netzwerken oft vorgezogen.

1.2.1 Eindimensionales kontinuierliches Widerstandsnetzwerk

Eine Glättung oder Ausschmierung eines räumlichen Musters entspricht einer Mittelung über mehrere Punkte. Allerdings möchte man nicht einfach den Mittelwert aus einer bestimmten Anzahl von Punkten bilden, sondern den Einfluß des erregenden Signals mit dem Abstand abnehmen lassen. Gesucht ist also eine Struktur, die ein einzelnes Eingangssignal mit wachsendem Abstand immer stärker unterdrückt. Das ist genau die Eigenschaft eines schlecht isolierten Leiters. Je weiter man sich von der Quelle entfernt, desto schwächer wird das gemessene Spannungssignal. Abb. 1.5 zeigt das Ersatzschaltbild für einen solchen Draht, der hier der Einfachheit halber gerade, beliebig dünn und unendlich lang angenommen wird.

Wie sieht nun der Spannungsverlauf entlang des Drahtes aus, wenn am Knoten 0 eine Spannung V_0 anliegt? Da das Widerstandsnetz als unendlich angenommen wird, sieht es an jeder Stelle x gleich aus. Für $x \geq 0$ gilt also nach dem Ohm'schen Gesetz:

$$V(x+dx) - V(x) = -I(x+dx)\, r\, dx \tag{1.5a}$$

$$I(x+dx) - I(x) = -V(x)\, g\, dx \tag{1.5b}$$

Dabei gibt $r\,dx$ den Widerstand zwischen der Position x und $x+dx$ und $g\,dx$ die Leitfähigkeit für das Leitungsstückchen der Länge dx zu Erde an. Umstellen der Gleichungen (1.5a) und

[2]CMOS steht für **C**omplementary **M**etal-**O**xide-**S**emiconductor, also eine Technologie, in der auf einem Chip sowohl n- als auch p-Kanal MOS Transistoren erzeugt werden können.

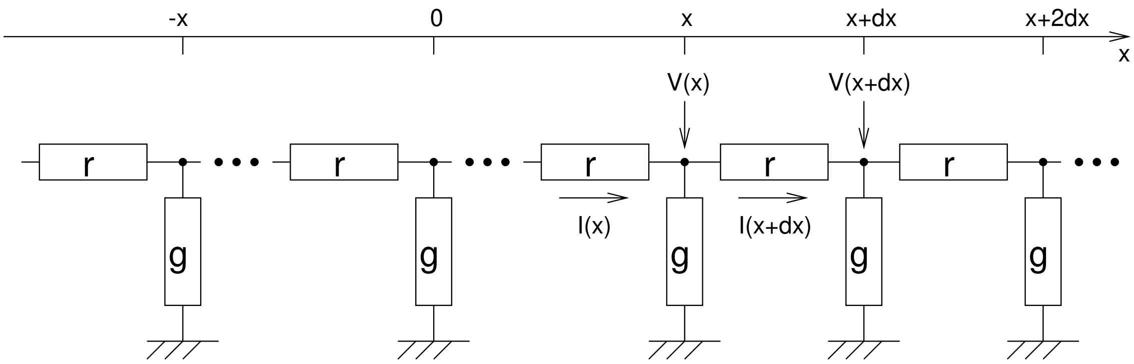

Abbildung 1.5: Ersatzschaltbild für eine gerade unendlich ausgedehnte Leitung mit Widerstand r$[\frac{\Omega}{m}]$ und Leitfähigkeit g$[\frac{1}{\Omega \cdot m}]$ zu Erde.

(1.5b) führt zu:

$$\frac{dV(x)}{dx} = -I(x)\,r \tag{1.6a}$$

$$\frac{dI(x)}{dx} = -V(x)\,g \tag{1.6b}$$

Differentiation von (1.6a) und Einsetzen von (1.6b) liefert:

$$\frac{d^2V}{d^2x} = rg\,V \tag{1.7}$$

Diese Gleichung wird durch

$$V(x) = \begin{cases} 0 \\ V_0\, e^{\pm\sqrt{rg}\,x} \end{cases} \tag{1.8}$$

gelöst. Unter Berücksichtigung der Anfangsbedingung $V(x) = V_0$ scheidet die triviale Lösung $V(x) = 0$ aus. Da bis jetzt nur der Fall für $x \geq 0$ betrachtet wurde, kommt die divergierende Lösung mit dem positiven Vorzeichen im Exponenten nicht in Frage. Für $x < 0$ ändert sich aber nur das Vorzeichen des dx in obiger Betrachtung. Die gesuchte Lösung lautet daher

$$V(x) = V_0 e^{-\frac{|x|}{l}} \tag{1.9}$$

mit

$$l = \frac{1}{\sqrt{rg}}. \tag{1.10}$$

l wird charakteristische Länge oder auch Diffusionslänge genannt und gibt an, in welchem Abstand zur Quelle ein Spannungssignal auf $\frac{1}{e}$ abgefallen ist. Daß l tatsächlich die Dimension einer Länge hat, folgt aus der Definition von r und g als Widerstand bzw. Leitfähigkeit pro Längeneinheit.

1.2.2 Eindimensionales diskretes Widerstandsnetzwerk

Wie sieht nun die Antwort eines *diskreten* Widerstandsnetzwerkes (siehe Abb. 1.6) aus? Für den diskreten Fall wird aus der Differentialgleichung (1.5a) eine Differenzengleichung, so

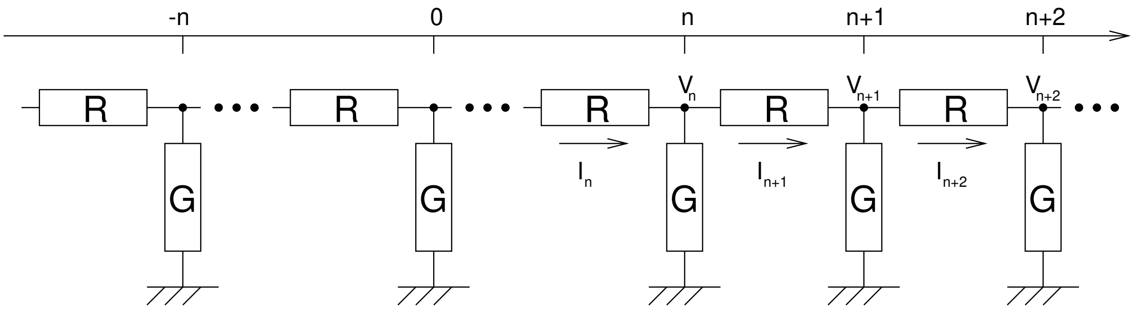

Abbildung 1.6: Das gleiche Schaltbild wie in Abb. 1.5, aber diesmal als diskretes Widerstandsnetzwerk.

daß für den Strom durch den Widerstand R gilt:

$$V_{n+1} - V_n = -R\, I_{n+1} \qquad \text{bzw.} \tag{1.11a.i}$$

$$V_{n+2} - V_{n+1} = -R\, I_{n+2} \tag{1.11a.ii}$$

Aus (1.5b) wird dann:

$$I_{n+2} - I_{n+1} = -GV_{n+1} \tag{1.11b}$$

Multipliziert man Gl. (1.11b) mit R und ersetzt I_n und I_{n+1} mit Hilfe der Gleichungen (1.11a.i) und (1.11a.ii), so ergibt sich

$$V_n - V_{n+1} + V_{n+2} - V_{n+1} = RGV_{n+1}\ , \tag{1.12}$$

was sich noch zu

$$V_{n+2} - (2 + RG)V_{n+1} + V_n = 0 \tag{1.13}$$

vereinfachen läßt. Gl. (1.13) ist eine Differenzengleichung zweiter Ordnung. Da sich das diskrete Netzwerk zumindest für große L ähnlich verhalten sollte wie das kontinuierliche, erwartet man auch hier wieder, daß das Signal mit wachsendem Abstand von der Quelle exponentiell abfällt. Daraus ergibt sich folgender Ansatz:

$$V_n = V_0 \gamma^{|n|} \tag{1.14}$$

Setzt man diesen in (1.13) ein, so ergibt sich eine quadratische Gleichung für γ:

$$\gamma^2 - (2 + RG)\gamma + 1 = 0 \tag{1.15}$$

Diese wird durch

$$\gamma = 1 + \frac{RG}{2} \pm \sqrt{RG + \frac{R^2G^2}{4}} \tag{1.16}$$

gelöst. Da V_n für große n verschwinden muß, muß $\gamma < 1$ gelten, so daß nur die Lösung mit dem negativen Vorzeichen vor der Wurzel in Frage kommt. Ersetzt man jetzt wieder $\sqrt{RG}$ durch $\frac{1}{L}$ (jetzt dimensionslos), so ergibt sich schließlich:

$$\gamma = 1 + \frac{1}{2L^2} - \frac{1}{L}\sqrt{1 + \frac{1}{4L^2}} \tag{1.17}$$

Allerdings ist zu beachten, daß das l im kontinuierlichen Fall wirklich eine Länge darstellt, während sich L hier auf die natürliche Einheit des diskreten Netzwerkes bezieht, nämlich auf eine RG-Einheit. Wie unterscheidet sich nun die diskrete von der kontinuierlichen Lösung? Dazu kann man Gl. (1.9) umschreiben zu:

$$\frac{V(x)}{V_0} = e^{-\frac{|x|}{l}} = \left(e^{-\frac{1}{l}}\right)^{|x|} \tag{1.18}$$

Um (1.14) mit (1.18) vergleichen zu können, muß man l durch L ersetzen, und sowohl γ, als auch $e^{-\frac{1}{L}}$ in eine Postenzreihe entwickeln:

$$\gamma = 1 - \frac{1}{L} + \frac{1}{2L^2} - \frac{1}{8L^3} + O\left(\frac{1}{L^4}\right) \tag{1.19}$$

$$e^{-\frac{1}{L}} = 1 - \frac{1}{L} + \frac{1}{2L^2} - \frac{1}{6L^3} + O\left(\frac{1}{L^4}\right) \tag{1.20}$$

Die Differenz von (1.20) und (1.19) ergibt:

$$\gamma - e^{\frac{1}{L}} = \frac{1}{24L^3} + O\left(\frac{1}{L^4}\right) \tag{1.21}$$

Die Lösung für das diskrete Netzwerk scheint also sehr schnell gegen diejenige für das Kontinuierliche zu konvergieren, wie auch aus Abb. 1.7 ersichtlich ist. Daher ist es möglich, auch die Reaktion des diskreten Netzwerks als exponentiellen Abfall mit der charakteristischen Länge L zu beschreiben, ohne einen zu großen Fehler zu machen, solange L nicht zu klein (z.B. $L > 1$) ist. Von einer charakteristischen Länge L zu sprechen wird dadurch erst sinnvoll. Bei der Herleitung der Diffusionslänge L wurde eine Spannung direkt am Knoten 0 angelegt.

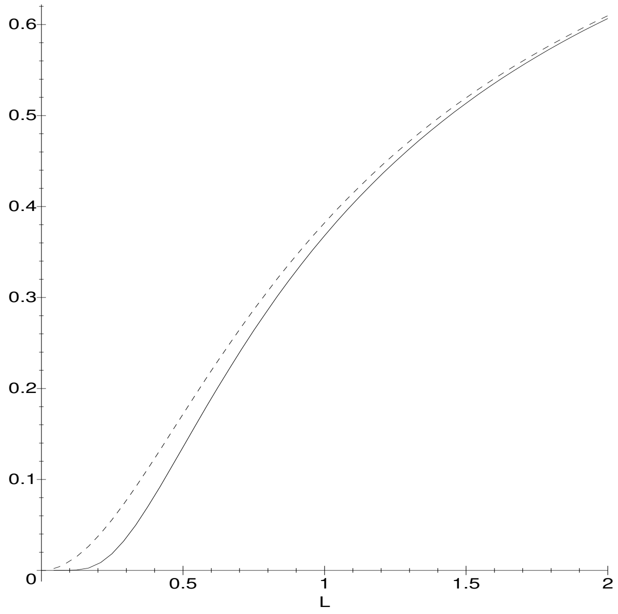

Abbildung 1.7: Zum Vergleich sind hier $e^{-\frac{1}{L}}$ (durchgezogen) und γ (gestrichelt) gegen L aufgetragen. Selbst für kleine L ist der Unterschied erstaunlich gering.

Um die gewünschte lokale Mittelung zu erreichen muß man die Spannungssignale an den unteren Enden der G's einkoppeln, wie in Abb. 1.8 gezeigt. Eine interessante Größe ist der Spannungsabfall über dem Widerstand mit der Leitfähigkeit G zum Knoten 0, oder anders

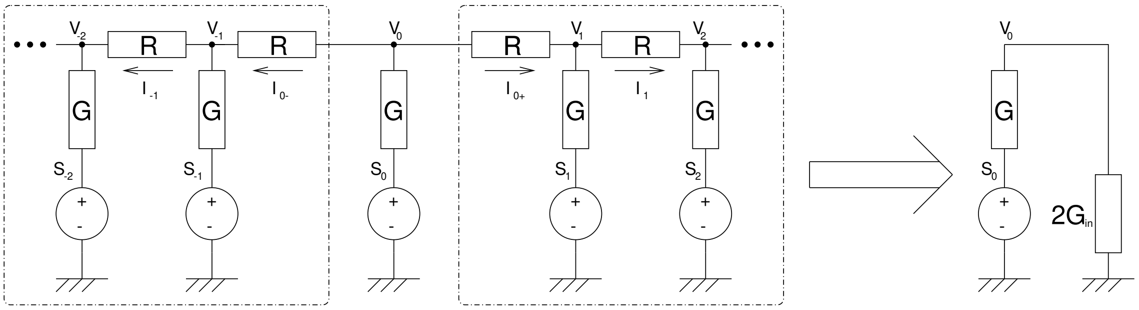

Abbildung 1.8: Das Netzwerk aus Abb. 1.6 ist hier noch mit Signalspannungsquellen S_i versehen. Außerdem ist noch ein Ersatzschaltbild für die Berechnung der Spannung V_0 in Abhängigkeit von S_0 und L angegeben.

formuliert die Leitfähigkeit des gesamten Netzwerkes von einer der Spannungsquellen S_i aus gesehen. Die nötige Vorgehensweise deutet Abb. 1.8 an. Zuerst muß die Leitfähigkeit G_{in} vom Knoten 0 zu Erde durch eine der in Abb. 1.8 markierten Netzwerkhälften bestimmt werden. Gl. (1.11a.i) beschreibt aber genau Spannung und Ströme durch eine solche Netzwerkhälfte. Dividiert man nun diese Gleichung durch V_n

$$\frac{I_n}{V_n}R = 1 - \frac{V_{n+1}}{V_n} = 1 - \gamma \tag{1.22}$$

und setzt dann für $n = 0$ ein, so ergibt sich:

$$G_{in} = \frac{I_0}{V_0} = \frac{1-\gamma}{R} \tag{1.23}$$

Einsetzen von γ liefert:

$$G_{in} = \frac{G}{2}\left(-1 + \sqrt{\frac{4}{RG} + 1}\right) \tag{1.24}$$

Für den Spannungsteiler auf der rechten Seite von Abb. 1.8 gilt:

$$\frac{V_0}{S_0} = \frac{G}{G + 2G_{in}} \tag{1.25}$$

Mit 1.24 ergibt sich also:

$$\frac{V_0}{S_0} = \frac{1}{\sqrt{\frac{4}{RG} + 1}} = \frac{1}{\sqrt{4L^2 + 1}} \tag{1.26}$$

Die pathologischen Fälle $L = 0$ und $L = \infty$ ergeben $V = S$, d.h. gar keine Mittelung bzw. $V = 0$, also eine vollständige Mittelung (unendlich viele Eingänge sind auf dem Potential $0V$).

Generell werden sich die Signale S_i mehrerer Quellen linear überlagern, da das Netzwerk nur aus linearen Bauteilen besteht. Trotzdem soll hier noch kurz die Antwort des Netzwerkes auf ein kantenförmiges Eingangssignal (z.B. $S_i = V_{in}$ für $i \leq 0$ und $S_i = 0$ für $i > 0$) angegeben werden. Um die Form der Ausgangskurve zu bestimmen wird das Netzwerk wie in Abb. 1.9 aufgeteilt. Auf Grund der Eingangsspannungen muß durch den Widerstand R in der Mitte ein Strom fließen. Auf beiden Seiten des Widerstandes befindet sich jetzt ein

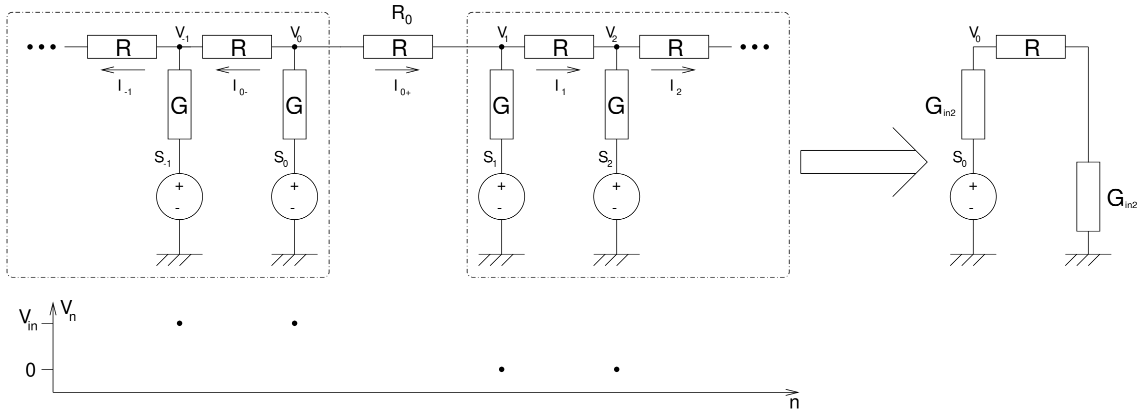

Abbildung 1.9: Darstellung der Eingangssignale für eine Kante von V_{in} auf 0 zwischen den Knoten 0 und 1. Außerdem ist das Ersatzschaltbild zur Berechnung des Spannungsabfalls über dem R zwischen den Knoten 0 und 1 eingezeichnet.

in jeweils nur eine Richtung ausgedehntes Widerstandsnetzwerk, in das ein Strom eingespeist bzw. aus dem ein Strom herausgezogen wird. Wie groß die Spannungen an den beiden Enden des Widerstandes und der Strom durch diesen sind, hängt von den drei im Ersatzschaltbild eingetragenen Widerständen ab. Die Spannung entlang der Knoten wird sich auch hier wieder exponentiell mit zunehmendem Abstand von R_0 dem jeweiligen Wert der Eingangsspannung annähern.

Der Gesamtwiderstand einer der beiden Netzwerkhälften aus Abb. 1.9 ergibt sich mit G_{in} zu:

$$R_{in2} = \frac{1}{G_{in}} - R \tag{1.27}$$

Der Spannungsabfall V_R über dem R in der Mitte im Verhältnis zu der gesamten Spannungsänderung zwischen den Knoten 0 und 1 folgt aus dem Ersatzschaltbild von Abb. 1.9:

$$\frac{V_R}{V_{in}} = \frac{R}{2R_{in2} + R} = \frac{R}{2\left(\frac{1}{G_{in}} - R\right) + R} = \frac{R}{\frac{4}{G(\sqrt{4L^2+1}-1)} - R} \tag{1.28}$$

Dabei wurde im letzten Schritt G_{in} aus Gl. (1.24) eingesetzt. Dies läßt sich noch vereinfachen zu:

$$\frac{V_R}{V_{in}} = \frac{\sqrt{4L^2+1} - 1}{4L^2 - (\sqrt{4L^2+1} - 1)} \tag{1.29}$$

Abb. 1.10 zeigt $\frac{V_R}{V_{in}}$ als Funktion von L. Für ein großes L erwartet man also einen langsamen exponentiellen Verlauf der Knotenspannungen mit wachsendem Abstand von der Eingangskante und einen kleinen Spannungssprung von Knoten 0 zu 1. Je größer das L desto stärker die Glättung der Kante.

1.2.3 Simulation des eindimensionalen linearen Widerstandsnetzwerkes

Die obigen Betrachtungen gingen alle von einem unendlich ausgedehnten Netzwerk aus. Natürlich ist es nicht möglich, ein solches Netzwerk zu implementieren. Um das Verhalten

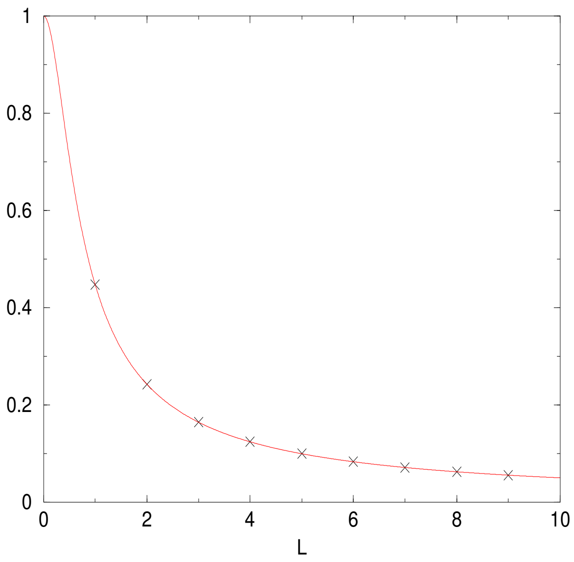

Abbildung 1.10: Graph der Funktion $\frac{V_R}{V_{in}}$ (Gl. (1.29) für L zwischen 0 und 10 für das unendlich ausgedehnte diskrete Netzwerk (rote Kurve) und der Simulation entnommene Datenpunkte des gleichen Verhältnisses (schwarze Kreuze).

eines endlichen Netzwerkes zu untersuchen, wurde ein Netzwerk mit 64 Knoten mit dem Simulator *SpectreS*, integriert in das Softwarepaket *Cadence*, simuliert.

In der den Abbildungen 1.11 bis 1.14 zu Grunde liegenden Simulationen wurde nur am Eingang zu Knoten 31 eine endliche Spannung von 0.5 V angelegt. Für R wurde ein Wert von 300 kΩ angenommen, G wurde so variiert, daß $L = \frac{1}{\sqrt{RG}}$ die Werte $1, 2, \ldots, 10$ annahm. Abb. 1.11 zeigt die resultierenden 10 Kurven. Das scharfe Eingangssignal wird mit wachsendem L immer stärker verschmiert. Abb. 1.12 zeigt die Ergebnisse in logarithmischer Darstellung. In der Umgebung von Knoten 31 ergeben sich für alle Werte von L lineare Verläufe, das Signal zerfällt also exponentiell. An den Rändern ergeben sich für die größeren L-Werte allerdings Abweichungen von diesem linearen Verhalten: Hier beeinflußt die endliche Ausdehnung des Netzwerkes also sein Verhalten. Der Effekt ist allerdings absolut gesehen sehr klein, da das Signal selbst für $L = 10$ auf etwa 2 mV an den beiden Randknoten abgefallen ist.

Um zu untersuchen, ob L ein gutes Maß für die charakteristische Länge des Netzwerkes ist, wurde für jede der Kurven aus Abb. 1.12 eine lineare Regression durchgeführt. Dazu wurden nur diejenigen Datenpunkte benutzt, die in erster Näherung auf einer Geraden lagen. Setzt man in Analogie zu 1.18 für

$$\frac{V_n}{V_0} = \left(e^{-\frac{1}{L_{sim}}}\right)^{|n|} \tag{1.30}$$

an, so kann man aus den Geradensteigungen L_{sim} bestimmen. In Abb. 1.13 ist L_{sim} gegen L aufgetragen. Der theoretisch zu erwartende Verlauf für ein unendlich ausgedehntes diskretes Netzwerk ergibt sich mit 1.14 zu:

$$m_\infty = -\frac{1}{\frac{1}{n}\ln\frac{V_n}{V_0}} = -\frac{1}{\frac{1}{L}\ln\gamma(L)} \tag{1.31}$$

Die Kurve, die sich ergibt, wenn man γ aus 1.17 in 1.31 einsetzt, ist auch in Abb. 1.13 eingezeichnet, genauso wie die Winkelhalbierende $y = L$, die den theoretischen Wert für ein unendlich ausgedehntes *kontinuierliches* Netzwerk angibt. Die Abweichung der theoretischen Kurve für das diskrete Netzwerk weicht nur minimal von der für das kontinuierliche ab und

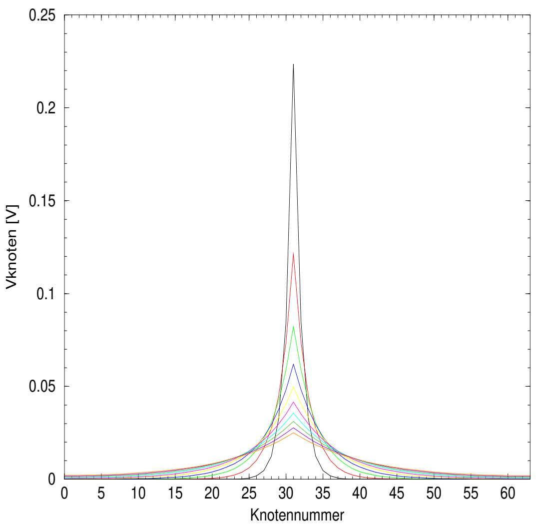

Abbildung 1.11: Simulation eines Netzwerkes mit 64 Knoten: Die Eingangsspannung wurde auf 0.5 V für Knoten 31 und 0 V für alle anderen Knoten eingestellt. Der Widerstand R betrug 300 kΩ. Die Werte für $\frac{1}{G}$ wurden so variiert, daß $L = \frac{1}{\sqrt{RG}}$ die Werte $1, 2, \ldots, 10$ annahm.

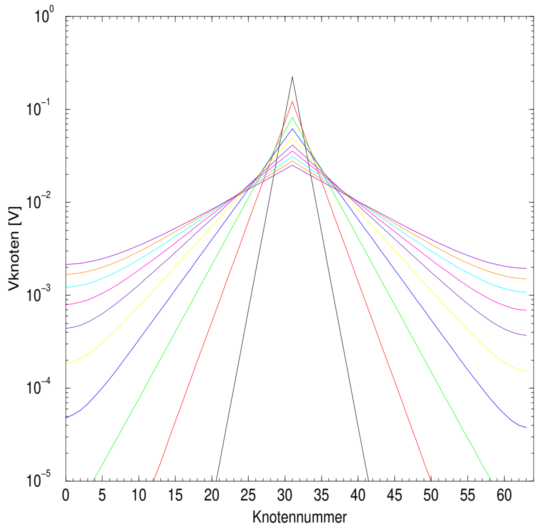

Abbildung 1.12: Ergebnisse derselben Simulation wie in Abb. 1.11 in logarithmischer Darstellung.

stimmt mit der aus der Simulation erhaltenen Kurve für kleine Werte von L sehr gut überein. Für höhere Werte von L weicht das L_{sim} allerdings deutlich nach oben ab. Das ist wahrscheinlich darauf zurückzuführen, daß vor der linearen Regression nicht alle durch die endliche Größe des Netzwerkes vom idealen Verhalten abweichenden Datenpunkte entfernt wurden.

Der Spannungsabfall über dem Eingangswiderstand im Verhältnis zum Eingangssignal ist durch Gl. (1.26) gegeben und ein Maß für den Eingangswiderstand des Netzwerkes. In Abb. 1.14 ist dieses Verhältnis für die Simulation des endlichen und den theoretischen Verlauf des unendlichen diskreten Netzwerkes dargestellt. Beide stimmen sehr gut überein.

Abb. 1.15 zeigt die Antwort des endlichen Netzwerkes auf die Einzelstimulation verschiedener Knoten für $L = 1$. Deutliche Unterschiede zum Verhalten bei der Stimulation einer der mittleren Knoten ergeben sich nur für die ersten 3 Kurven. Deutlicher treten die Randeffekte dagegen für $L = 10$ zu Tage, wie man Abb. 1.16 entnehmen kann.

Wie in der Diskussion des unendlichen diskreten Netzwerkes wurde als weiteres Eingabemuster ein kantenförmiges untersucht. Dabei wurde an den ersten 32 Eingängen jeweils 0.5 V, an den übrigen 32 0 V angelegt. L wurde wieder von 1 bis 10 in ganzzahligen Schritten variiert. Aus Abb. 1.17 läßt sich entnehmen, daß die Kante durch das Netzwerk mit wachsendem L immer mehr aufgeweicht wird. In Abb. 1.18 sind die Spannungen an den Knoten 32 bis 63 logarithmisch aufgetragen. Wie in Kapitel 1.2.2 für das unendliche diskrete Netzwerk ergibt sich auch hier, abgesehen von den Randeffekten, ein exponentielles Abnehmen des Signals entlang des Netzwerkes.

Aus den Simulationsdaten wurde der Spannungabfall über dem Widerstand zwischen Knoten 31 und 32 im Verhältnis zum Gesamthub der Kante extrahiert und in Abb. 1.10 mit dem

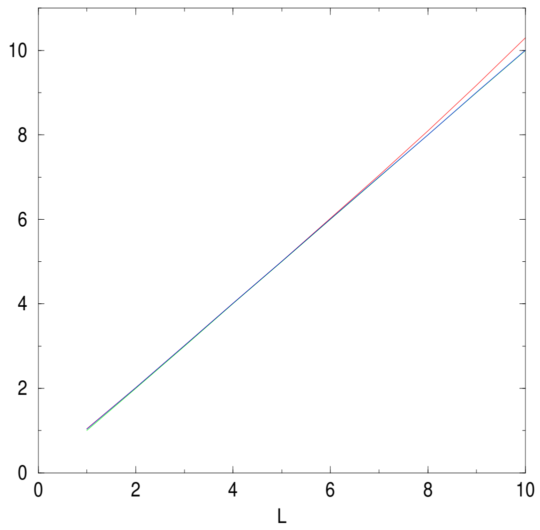

Abbildung 1.13: Vergleich der Diffusionslängen: Die schwarze Kurve zeigt den theoretisch zu erwartenden Verlauf, die Rote den aus der Simulation abgeleiteten (Erklärungen siehe Text). Zum Vergleich ist noch die Winkelhalbierende $L = \frac{1}{\sqrt{RG}}$ eingezeichnet (blau).

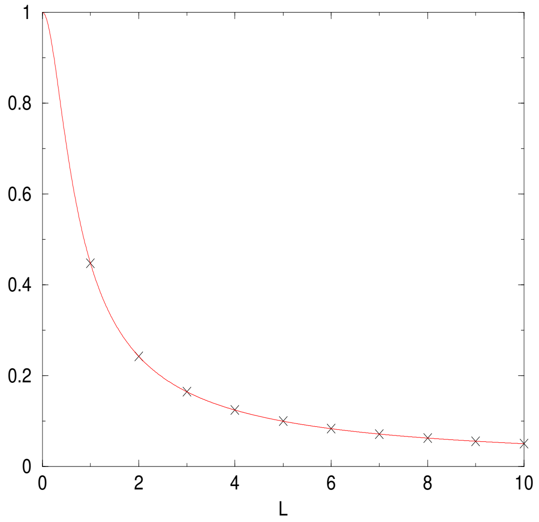

Abbildung 1.14: Vergleich von Gl. (1.26) mit den Simulationsergebnissen (schwarze Kreuze).

theoretischen Verlauf für das unendliche diskrete Netzwerk verglichen. Die Übereinstimmung ist auch hier wieder sehr gut.

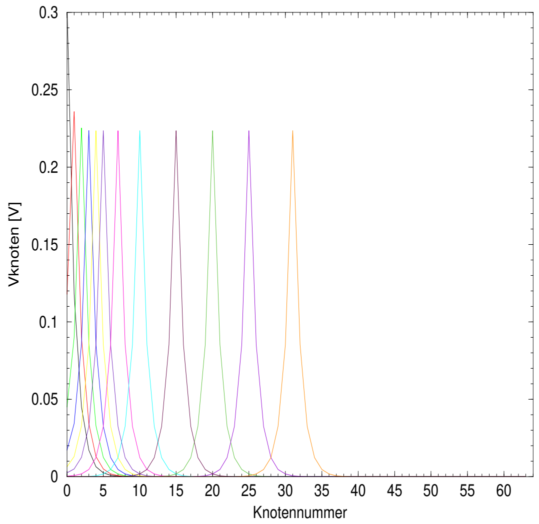

Abbildung 1.15: Antwort des Netzwerks auf Stimulation jeweils eines Eingangs. Stimuliert wurden die Eingänge. (1,2,...,5,7,10,15,20,25,31). Die Eingansspannungen betrugen wieder 0 und 0.5 V, $R = \frac{1}{G} = 300\,\mathrm{k}\Omega$ (also $L = 1$).

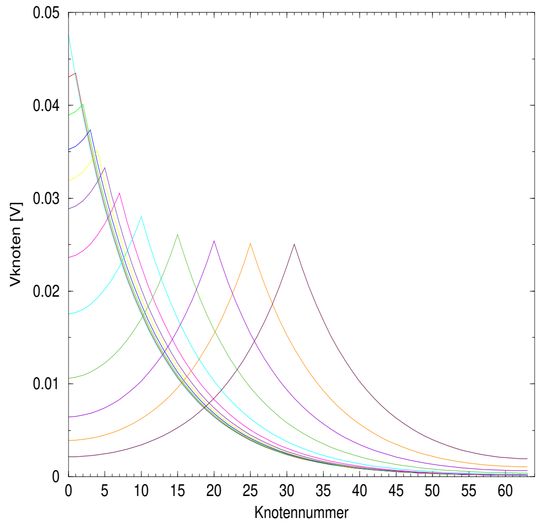

Abbildung 1.16: Die gleiche Simulation wie für Abb. 1.15, aber mit $\frac{1}{G} = 30\,\mathrm{M}\Omega$ (also $L = 10$).

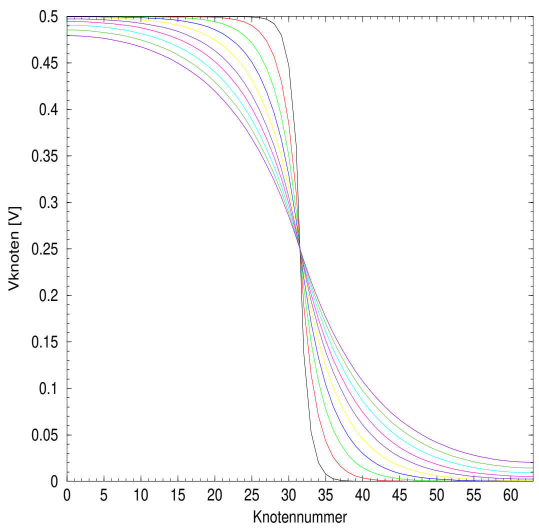

Abbildung 1.17: Simulation des Netzwerkes für ein kantenförmiges Eingangsmuster: Die Eingangsspannungen der ersten 32 Knoten betrug 0.5 V, die der letzten 32 0 V. R und G hatten die gleichen Werte wie in der in Abb. 1.11 beschriebenen Simulation.

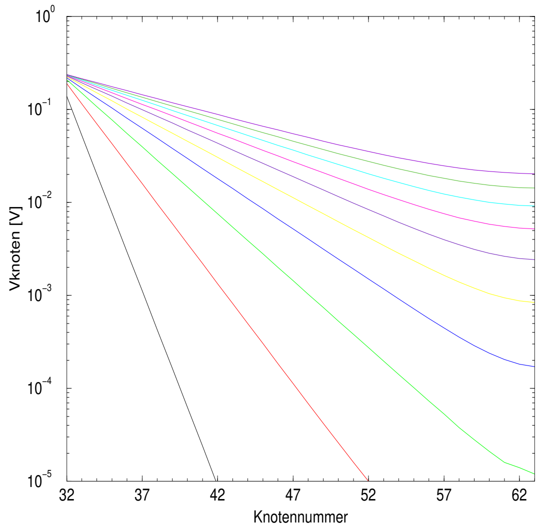

Abbildung 1.18: Logarithmische Darstellung der Simulationsergebnisse für das kantenförmige Eingangsmuster. Nur die Knotenspannungen V_{32} bis V_{63} sind dargestellt.

1.3 Optische Konfiguration

Wie schon in Kapitel 1.1.3 erwähnt, soll untersucht werden, inwieweit sich Information optisch auf einen Mikrochip übertragen läßt. Dazu ist einerseits das Lichtsignal in eine Spannung umzuwandeln und, zumindest für eine begrenzte Zeit, zu speichern und andererseits ein räumlich und zeitlich ausgedehntes Muster auf dem Chip zu erzeugen.

1.3.1 Elektrisches Konzept

Photodioden wandeln das einfallende Licht gemäß

$$I_{ph} = \frac{\lambda e}{hc} JAQ \propto J \tag{1.32}$$

in einen Strom um ([Loose 96]). Hier bezeichnet I_{ph} den erzeugten Photostrom, J die Lichtintensität, A die Fläche der Photodiode, Q die Quanteneffizienz, die angibt mit welcher Wahrscheinlichkeit ein Photon ein zum Strom beitragendes Ladungspaar erzeugt, λ die Wellenlänge des auftreffenden Lichtes und e, h, c die drei Naturkonstanten Elementarladung, Planck'sches Wirkungsquantum und Lichtgeschwindigkeit.

Integriert man den erzeugten Photostrom über die Zeit auf, indem man, wie in Abb. 1.19 gezeigt, damit einen Kondensator entlädt, so folgt, falls der Kondensator zum Zeitpunkt $t = 0$ mit Hilfe des Reset-Schalters entladen wurde, für die Spannung V_C am unteren Anschluß der Diode:

$$V_C = Vdd - \frac{Q}{C} = Vdd - \frac{1}{C}\int_0^t I_{ph} dt' \tag{1.33}$$

Damit ist ein Lichtsignal mit Intensität J und Dauer t in eine Spannung umgewandelt worden, die außer von J und t noch von Q und A abhängt. Außerdem wird diese Spannung zumindest für eine begrenzte Zeit nahezu konstant bleiben. Wie schnell sich V_C ändert, hängt davon ab, wie groß der Dunkelstrom der Photodiode und der Leckstrom des durch einen Transistor realisierten Schalters sind. Eine genauere Untersuchung des zeitlichen Verhaltens von V_C wird in Kapitel 2.1 durchgeführt.

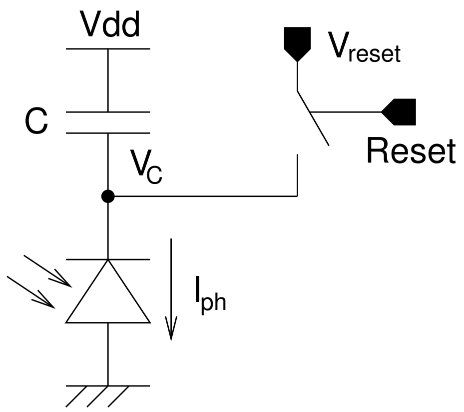

Abbildung 1.19: Schaltung, um einen Lichtpuls in eine Spannung zu konvertieren und diese zu speichern

1.3.2 Optisches Konzept

Um eine gute räumliche Auflösung zu erreichen muß eine sehr gut fokussierbare und ausreichend starke Lichtquelle benutzt werden. Weiterhin wäre eine Modulierbarkeit der

Lichtintensität wünschenswert. Es bietet sich der Einsatz eines modulierbaren Diodenlasers an. Zur dynamischen Positionierung des Laserstrahls auf dem Chip werden sogenannte Galvanometer-Scanner eingesetzt. Diese erlauben eine sehr genaue Auslenkung eines drehbar gelagerten Spiegels abhängig z.B. von einer analogen Eingangsspannung. Die maximal möglichen Ablenkungen betragen typischerweise $\pm 20°$. Die Positioniergenauigkeit wird von den Herstellern in der Regel mit einigen μrad angegeben. Für eine Auslenkung des Spiegels um einige Grad, benötigt ein solcher Scanner typischerweise einige ms.

Prinzipiell gibt es nun zwei Möglichkeiten der Fokussierung: Vor dem Ablenkspiegel oder dahinter. Da die Größe der Strahltaille des fokussierten Laserstrahls linear von der Brennweite der fokussierenden Linse abhängt, eine Linse zwischen Scanner und Chip aber andererseits den Abstand zwischen Ablenkspiegel und Chip vergrößert und damit die Positioniergenauigkeit einschränkt, ist hier ein Kompromiß nötig. Abb. 1.20 zeigt einen optischen Aufbau mit einer Fokussierung hinter dem Scanner. Da der Fokus einer sphärischen Linse für das wie eingezeichnet einfallende Licht auf einem Kreis (3-dimensional auf ein Kugelschale) liegt, benutzt man für andere Laserpositionierungs-Anwendungen meistens sogenannte Fθ-Objektive, die das Licht in einer Ebene hinter dem Chip fokussieren.

Damit die Schaltung aus Abb. 1.19 richtig funktioniert, muß die den Chip treffende Lichtmenge genau stimmen, muß die Laserintensität also auf das richtige Maß abgeschwächt werden können. In Abb. 1.20 sind dazu neben Laser, Scanner und Chip noch 4 Neutralfilter eingezeichnet. Mit ihnen kann die Intensität des Lasers um eine Dekade pro Filter abgeschwächt werden.

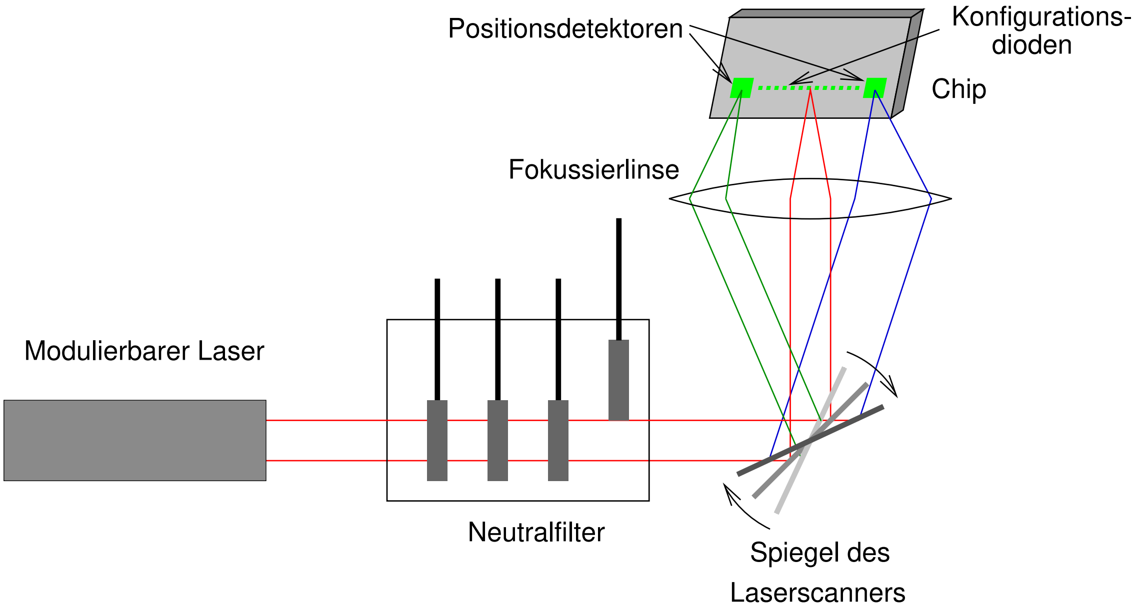

Abbildung 1.20: Schematischer Aufbau für die optische Konfiguration eines Mikrochips

Eine Messung kann z.B. folgendermaßen ablaufen: In einem ausreichend abgedunkelten Raum werden nacheinander die entsprechenden Photodioden angefahren. Die aufgebrachte Lichtmenge richtet sich nach der Verweildauer des Strahls auf der jeweiligen Diode und seiner Lichtintensität während dieser Zeit. Nach dieser optischen Programmierung kann das Netzwerk Eingabemuster verarbeiten.

Damit man mit dem Laserspot gezielt bestimmte Positionen auf dem Chip anfahren kann, muß eine Beziehung zwischen dem Eingangssignal des Scanners, der Position des Chips und dem Koordinatensystem auf dem Chip hergestellt werden. Man benötigt also Posititionsdetektoren, die Information über die aktuelle Position des Laserspots zurückliefern, so daß der Laserspot mit Hilfe einer Rückkopplungsschleife auf die richtige Startpostition justiert werden kann. Desweiteren sollte es möglich sein, die Spotgröße und -gestalt auf dem Chip zu bestimmen, sowie ein Maß für die einfallende Laserintensität zu erhalten. Ein Postitionsdetektor, der diese Funktionen erfüllt, wird in Kapitel 3.7 entwickelt.

Kapitel 2

Messungen an *VisionTest96*

In 1.3.1 wurde eine Schaltung vorgestellt, mit der eine Lichtmenge in eine Spannung umgewandelt und diese Spannung gespeichert werden kann. Auf dem Chip *VisionTest96*[1] befinden sich zwei Implementierungen einer solchen Schaltung, anhand derer das zeitliche Verhalten der Spannung V_C (siehe Abb. 1.19) studiert wurde. Außerdem konnte aus den Messungen eine Abschätzung der Diffusionslänge freier Ladungsträger im Substrat für den benutzten Prozeß vorgenommen werden.

2.1 Schaltung

2.1.1 Optoelektronische Grundlagen

In dem benutzten Herstellungsprozeß der Firma AMS[2] stehen drei verschiedene pn-Übergänge und damit Arten von (Photo-) Dioden zur Verfügung (siehe Abb. 2.1). Diese unterscheiden

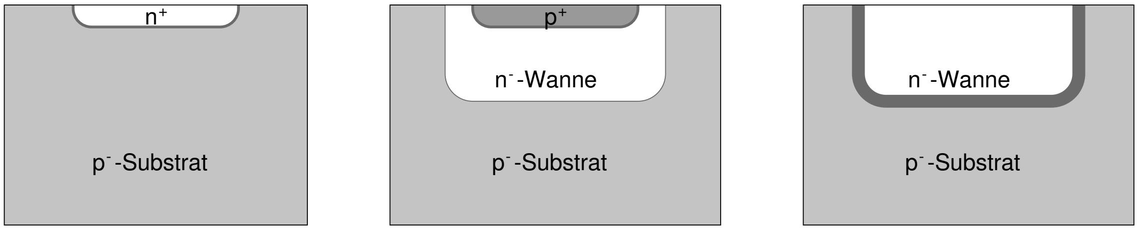

Abbildung 2.1: Die drei verschiedenen pn-Übergänge im benutzten CMOS-Prozeß. Man beachte, daß das Substrat in Wirklichkeit viel tiefer reicht als hier eingezeichnet. Außerdem sind die unterschiedlich weit ausgedehnten Verarmungszonen zwischen den Übergängen angedeutet.

sich sowohl in der spektralen Empfindlichkeit, als auch in ihren elektrischen Eigenschaften, was seine Begründung einerseits in der Dotierungsstärke der beteiligten p- und n-Schichten, als auch in ihrer Dicke findet. Eine etwas ausführlichere Diskussion findet sich in der Diplomarbeit von M. Loose ([Loose 96]). Tabelle 2.1 faßt einige der optischen und elektrischen

[1]Der Chip *VisionTest96* besteht aus optischen Teststrukturen, die 1996 von D. Droste, M. Keller und M. Theodori entworfen wurden. Der Chip wurde in einem 0.8 μm double poly double metal Prozeß der Firma Austria Mikro Systeme hergestellt.

[2]AMS steht für **A**ustia **M**ikro **S**ysteme International AG

Diode	Q	J_{F3} $[\frac{fA}{\mu m^2}]$	J_{U3} $[\frac{fA}{\mu m}]$	J_{F5} $[\frac{fA}{\mu m^2}]$	J_{U5} $[\frac{fA}{\mu m}]$
n^+-Substrat	0.49	0.01	0.08	0.01	0.12
p^+-Wanne	0.14	0.04	0.3	0.04	0.5
Wanne-Substrat	0.53	0.01	-	0.01	-

Tabelle 2.1: Quanteneffizienz Q verschiedener CMOS-Dioden bei 675 nm (aus [Loose 96]) und ihre **S**ättigungssperrstromdichten bei 3 und 5 V Sperrspannung aufgeschlüsselt nach dem mit der **F**läche, und dem mit dem **U**mfang der Diode skalierendem Teilbeitrag (aus [AMS 0.8]).

Eigenschaften der verschiedenen pn-Übergänge im zu Grunde liegenden CMOS-Prozeß zusammen. Die Quanteneffizienz Q wurde mit einem Diodenlaser der Wellenlänge 675 nm an einem in einem 1.2 μm-Prozeß der Firma AMS produzierten Chip gemessen([Loose 96]).

Die in der Tabelle angegebenen Sperrstromdichten J_{F3} und J_{F5} geben den Sperrstrom einer Photodiode des jeweiligen Typs an, der bei 3 bzw. 5 V Sperrspannung senkrecht zur Chipoberfläche pro $(\mu m)^2$ Diodenfläche ins Substrat fließt. Die Parameter J_{U3} und J_{U5} geben den entsprechenden Strom an, der parallel zur Chipoberfläche durch die Seitenwand der Diffusionsoberfläche fließt. Der Sperrstrom einer Diode mit Umfang U und Fläche A bei 5 V Sperrspannung ergibt sich also mit den Konstanten aus Tabelle 2.1 zu :

$$I = J_{U5}U + J_{F5}F \tag{2.1}$$

Licht, das ungehindert auf die Chipoberfläche fällt, erzeugt im Silizium Elektron-Loch-Paare. Liegt nun in der dotierten Schicht ein elektrisches Feld an, so werden sich die Ladungsträger entlang (Löcher) oder entgegen (Elektronen) der elektrischen Feldlinien bewegen: Für den Fall der Photodiode geht man davon aus, daß das elektrische Feld fast ausschließlich in der Verarmungszone des pn-Übergangs anliegt. Dort sollte also die Ladungstrennung erfolgen.

Im Substrat erzeugte Ladungsträger bewegen sich zufällig durch das (annähernd) feldfreie Silizium, werden aber bis zur Zeit τ nach ihrer Entstehung mit der Wahrscheinlichkeit $1 - e^{-1}$ durch Rekombination wieder vernichtet sein. Die Ausbreitung der freien Ladungsträger ohne elektrisches Feld kann man durch einen Diffusionsprozeß beschreiben. Die Diffusionslänge L_D hängt von der Dotierung des Siliziums ab, und beträgt bei den üblichen Substratdotierungen bis über 100 μm.

Treffen die durch Licht erzeugten Ladungsträger auf andere aktive Bauelemente, so können sie Ströme und Spannungen dort verändern und folglich das Verhalten der betroffenen Schaltung bis hin zur Fehlfunktion beeinträchtigen. Eine wirkungsvolle Methode zur Reduktion dieses Effektes besteht darin, alle Flächen, die nicht optisch sensitiv sein sollen, mit mindestens einer Metalllage zu bedecken. Der Dämpfungsfaktor ist definiert als

$$\frac{\text{Photostrom ohne Metallabdeckung}}{\text{Photostrom mit Metallabdeckung}} \tag{2.2}$$

und wurde für den 1.2 μm-Prozeß der Firma *AMS* bei 675 nm([Loose 96]) und für den 0.8 μm-Prozeß des gleichen Herstellers für verschiedene Wellenlängen ([Theodori 97]) gemessen. Die Ergebnisse für 675 nm sind in Tabelle 2.2 wiedergegeben.

2.1.2 Prinzipielle Idee

Die Abbildungen 2.2 und 2.3 zeigen die beiden implementierten Teststrukturen. Die Schaltung *Optstore_C* besteht aus einer Photodiode, die den Photostrom in die Kapazität

Prozeß	Dämpfungsfaktor		
	Metall-1	Metall-2	Metall-1 + Metall-2
1.2 μm-Prozeß	1783	764	4976
0.8 μm-Prozeß	2350	6739	18065

Tabelle 2.2: Dämpfungsfaktoren für 2 verschiedene Herstellungsprozeße mit 2 Metalllagen bei 675 nm. Die Unterschiede der gemessenen Dämpfungsfaktoren, sowie die Tatsache, daß die Dämpfungsfaktoren sich für zwei Metallagen nicht multiplikativ verhalten, ist auf Interferenzeffekte zurückzuführen. Die Dicke der Metallschichten ist nämlich der Wellenlänge des benutzten Lichts vergleichbar.

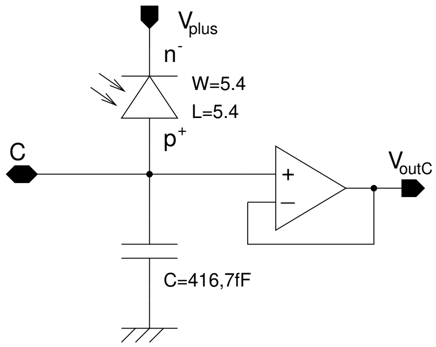

Abbildung 2.2: Teststruktur *Optstore_C*. Für die Diode sind die beteiligten Diffusionen angegeben. W und L geben Breite und Länge des pn-Übergangs in μm an.

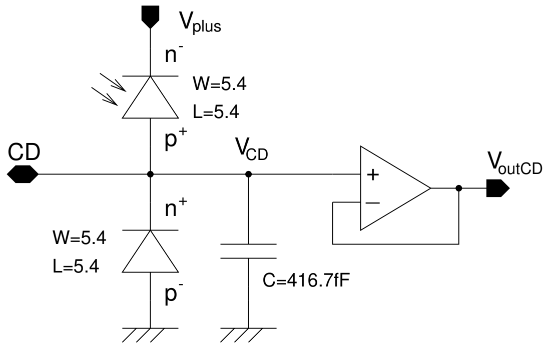

Abbildung 2.3: Teststruktur *Optstore_CD*. Angaben der Diffusionstypen und Maße der einzelnen Photodioden wie in Abb. 2.3. Die untere Diode ist mit Metall abgeschirmt.

hineindrückt und diese damit auflädt, und einem als Spannungsfolger geschalteten Operationsverstärker. Dieser dient nur dazu, die Spannung über der Diode auslesen zu können ohne die Ladung auf der Kapazität zu beeinflussen. Die Schaltung *Optstore_CD* aus Abb. 2.3 besitzt neben der p^+-Wanne-Diode noch eine gleich große, aber mit Metall abgedeckte, n^+-Substrat-Diode. Diese ist dazu gedacht, den Dunkelstrom den die p^+-Wanne-Diode liefert abzuführen, so daß im Speicherzustand, also ohne Beleuchtung, die Spannung über dem Kondensator möglichst konstant bleibt. Die mit C bzw. CD und V_{plus} bezeichneten Pins sind mit Pads verbunden, so daß der Strom durch die Photodioden auch direkt gemessen werden kann.

2.1.3 Schaltung inklusive Schutzstrukturen

Die Pads des Chips *VisionTest96* sind gegen elektrostatische Entladung (ESD[3]) geschützt, d.h. damit durch Ladungsübertragung keine zu hohen Spannungen zwischen den Transistor-Gates auf dem Chip und dem Substrat entstehen, sind Dioden mit jedem Pad verbunden, die es bei zu kleiner Spannung mit Erde, bei zu großer mit der Versorgungsspannungsleitung Vdd kurzschließen. Die Abbildungen 2.4 und 2.5 zeigen die Schaltungen aus Abb. 2.2 und Abb. 2.3 mit diesen Schutzdioden.

Für die Ermittlung des zeitlichen Verlaufes der Spannung über dem Kondensator ohne

[3]ESD steht für **E**lectro **S**tatical **D**ischarge

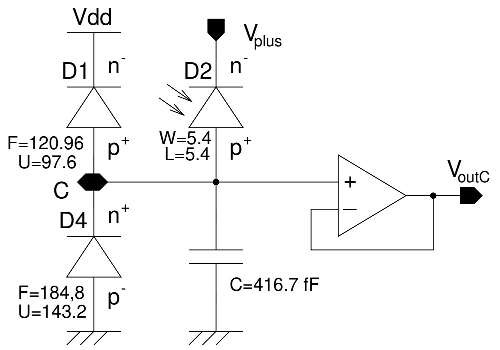

Abbildung 2.4: Teststruktur *Optstore_C* mit Schutzdioden. Alle Längenmaße sind in μm angegeben. Für die Schutzdioden sind Umfang und Länge angegeben.

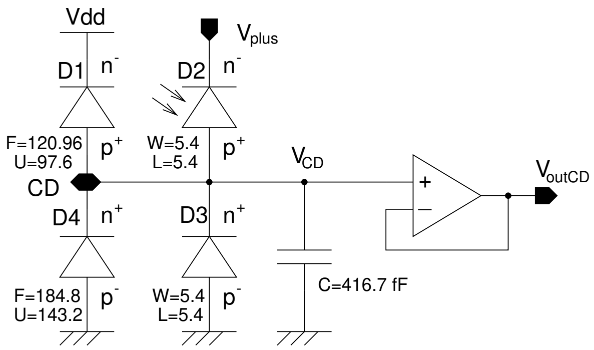

Abbildung 2.5: Teststruktur *Optstore_CD* mit Schutzdioden. Alle Längenangaben in μm. Die Schutzdioden sind die gleichen wie in Abb. 2.4.

Beleuchtung (Speicherzustand) kann man alle Dioden durch eine einfacheres Modell ersetzen, nämlich eine Stromquelle und einen dazu parallelen Widerstand. Als weitere Vereinfachung kann man annehmen, daß durch das Gate des Eingangstransistors des Operationsverstärkers kein Strom fließt.

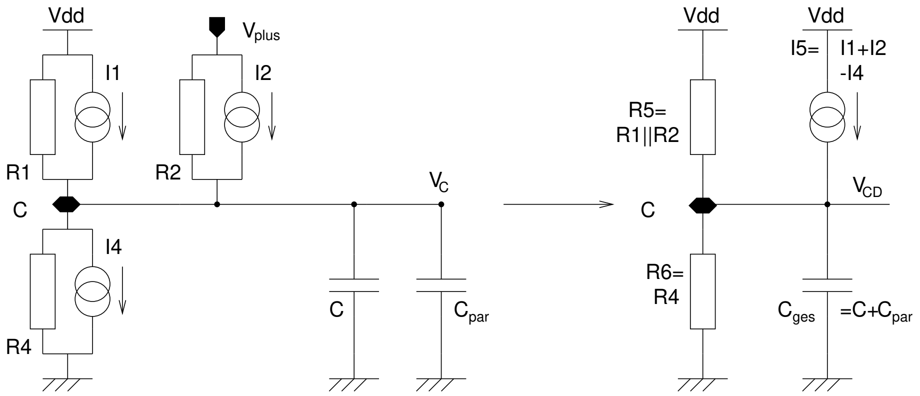

Abbildung 2.6: Ersatzschaltbild der Schaltung *Optstore_C*. Auf der rechten Seite ist die Vereinfachung des Ersatzschaltbildes gezeigt, wobei $V_{plus} = Vdd$ angenommen wurde.

Die Abbildungen 2.6 und 2.7 zeigen die entsprechenden Ersatzschaltbilder für die Schaltungen *Optstore_C* und *Optstore_CD*. Die Widerstände lassen sich dabei noch zu

$$R_5 = R_1 \parallel R_2 = \frac{R_1 R_2}{R_1 + R_2} \tag{2.3}$$

$$R_6 = \begin{cases} R_4 & \text{für } \textit{Optstore\_C} \quad \text{bzw.} \\ R_3 \parallel R_4 = \frac{R_3 R_4}{R_3 + R_4} & \text{für } \textit{Optstore\_CD} \end{cases} \tag{2.4}$$

zusammenfassen. Zusätzlich zu C ist noch die parasitäre Kapazität C_{par} eingezeichnet. Diese faßt alle parasitären Kapazitäten parallel zu C, nämlich die Eingangskapazität des Gates des Eingangstransistors und die Kapazität des Pads und der Leitungen zusammen. C_{par} beträgt

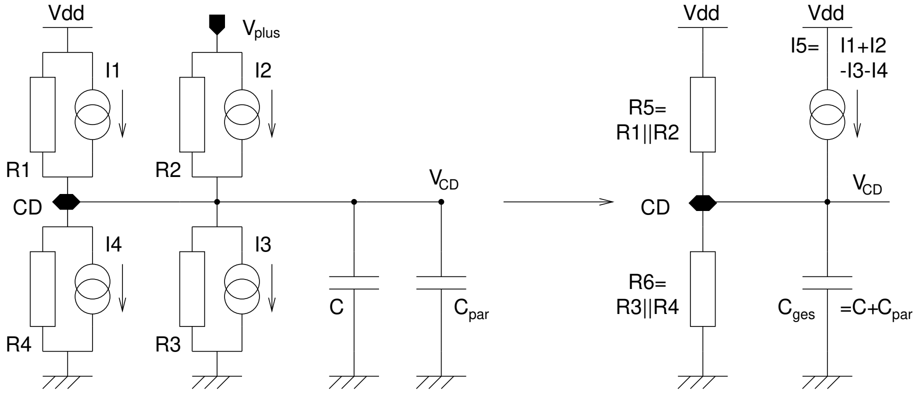

Abbildung 2.7: Ersatzschaltbild der Schaltung *Optstore_CD*. Die Abbildung unterscheidet sich von Abb. 2.6 nur durch die zusätzliche Diode $D4$.

etwa 480 pF. Für die insgesamt zu berücksichtigende Kapazität ergibt sich also:

$$C_{ges} = C + C_{par} \simeq 900\,\mathrm{pF} \tag{2.5}$$

Bei der Vereinfachung des Ersatzschaltbildes wurde entsprechend der Situation während der Messung davon ausgegangen, daß $V_{plus} = Vdd$ gilt. Die Werte für die Dunkelströme I_1 bis I_4 und die Widerstände R_1 bis R_4 können aus den in Tabelle 2.1 aufgeführten Prozeßparametern und den in den Abbildungen 2.4 und 2.5 angegebenen Abmessungen der einzelnen Dioden ermittelt werden.

Das Gleichstrom-Verhalten der vereinfachten Ersatzschaltbilder aus den Abbildungen 2.6 und 2.7 kann durch zwei Parameter beschrieben werden: Die Gleichgewichtsspannung V_{Gl} und die Zeitkonstante τ, mit der die Schaltung diese Spannung nach einer Störung erreicht. Für den Gleichgewichtsfall kann man die Kapazität C ignorieren. Die Anwendung der Kirchhoff'schen Knotenregel und des Ohm'schen Gesetzes liefert:

$$V_{Gl} = \frac{R_6}{R_5 + R_6} Vdd + \frac{R_5 R_6}{R_5 + R_6} I_5 \tag{2.6}$$

Um die Zeitkonstante τ zu bestimmen, kann andererseits die Stromquelle I_5 ignoriert werden, die nur für einen konstanten Beitrag zur Spannung V_C bzw. V_CD sorgt. Faßt man die parallelen Widerstände R_5 und R_6 zu einem Widerstand zusammen, so ergibt sich:

$$\tau = R_5 \parallel R_6\, C = \frac{R_5 R_6}{R_5 + R_6} C \tag{2.7}$$

Ist die Spannung am Knoten C bzw. CD also zum Zeitpunkt $t = 0$ durch V_0 gegeben, dann gilt für $V_{C\setminus CD}$:

$$V_{C/CD} = V_{Gl} + (V_0 - V_{Gl}) e^{-\frac{t}{\tau}} \tag{2.8}$$

Die Zahlenwerte, die sich aus den angegebenen Parametern für τ und V_{Gl} für die beiden Schaltungen ergeben, sind in Tabelle 2.3 in Kapitel 2.3.2 aufgeführt.

Abb 2.8 zeigt das Layout der beiden Schaltungen *Optstore_C* und *Optstore_CD*. Die beiden Operationsverstärker sind durch eine gestrichelte Box eingefaßt, die interessierenden Dioden und Kondensatoren durch Pfeile markiert. Die Kantenlänge eines Pads beträgt 105 μm, der gezeigte Ausschnitt aus dem Layout etwa $500 \times 780\,\mu\mathrm{m}^2$.

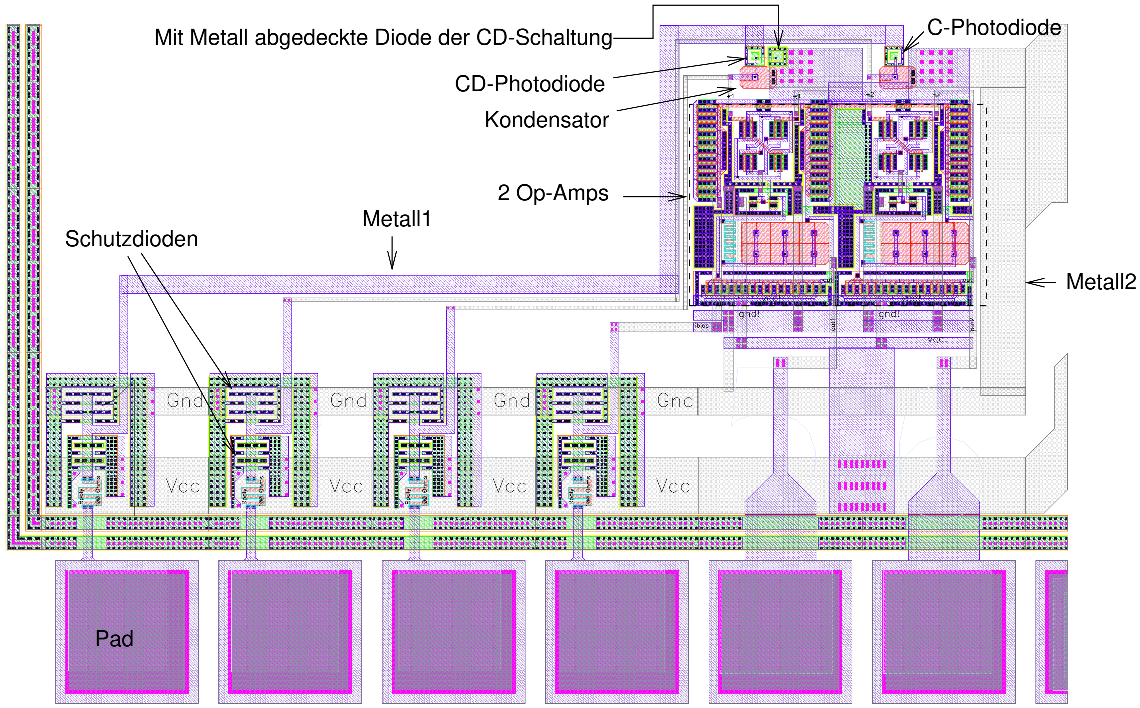

Abbildung 2.8: Layout der Schaltung *Optstore_C* und *Optstore_CD*.

2.2 Optischer Aufbau

Der optische Aufbau zur Messung der Schaltungen *Optstore_C* und *Optstore_CD* ist in Abb. 2.9 dargestellt. Da auf dem Chip nur die n^+-Substrat Diode mit Metall abgeschirmt ist, muß der Chip lokal beleuchtet werden. Die ersten Versuche den Chip mit einer LED zu beleuchten haben nämlich gezeigt, daß bei Beleuchtung des gesamten Chips die Photoströme der Schutzdioden eine Aufladung des Kondensators verhindern. Zur Beleuchtung wurde daher ein Diodenlaser mit der Wellenlänge 675 nm benutzt. Die Laserintensität konnte mit Hilfe von Neutralfiltern mit einem Transmissionskoeffizienten von etwa 10% pro Filter variiert werden. Hinter den Filtern mußte das Licht ein Raumfilter passieren. Dieses dient dazu, höhere Beugungsordnungen des Strahls, verursacht z.B. durch Linsenfehler oder Reflexionen zwischen den Neutralfiltern, aus dem Strahl wieder herauszufiltern. Kurz vor dem Chip wurde der Strahl mit einem Mikroskopobjektiv mit 16 mm Brennweite auf den Chip fokussiert. Damit die entspechenden Photodioden mit dem Laserstrahl genau getroffen werden konnten, wurde der Chip samt einer kleinen Meßplatine auf einem XY-Fahrtisch angebracht, der eine (fast) mikrometergenaue Positionierung des Chips erlaubt.

2.3 Ergebnis

2.3.1 Diffusionslänge und Laserspotgröße

Um die Größe des Laserspots auf dem Chip festzustellen wurde der Chip mit Hilfe des Fahrtisches mit konstanter Geschwindigkeit von $1\,\mu m/s$ dergestalt in vertikaler Richtung bewegt, daß der Laserstrahl die CD-Photodiode von links nach rechts überstrich. Dabei wurde

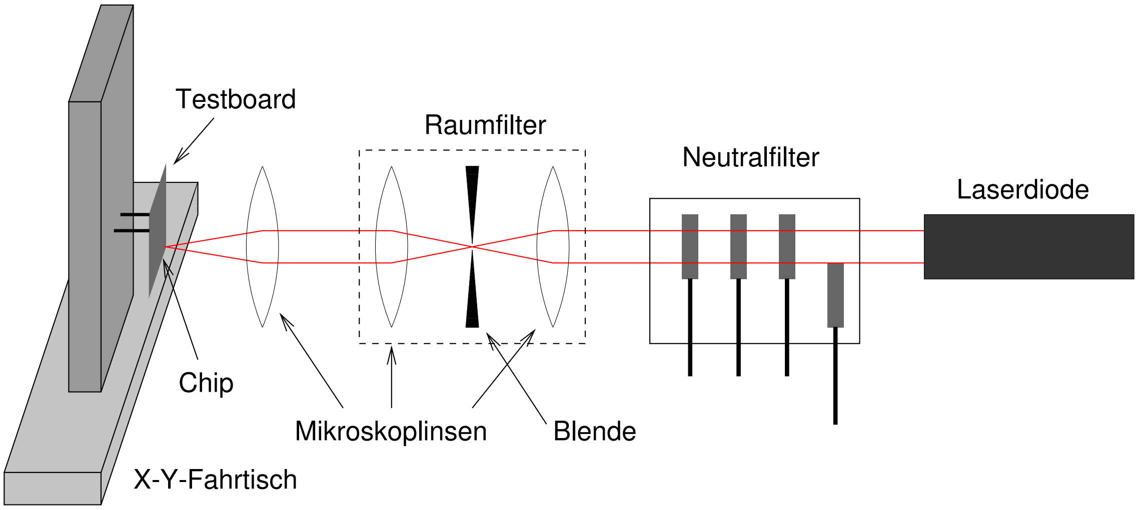

Abbildung 2.9: Optischer Aufbau für die Mesungen an *VisionTest96*.

sowohl der Strom gemessen, der in den Anschluß V_{plus} hineingeflossen, als auch derjenige, der aus Anschluß CD wieder hinausgeflossen ist. Das Ergebnis ist in Abb. 2.10 dargestellt.

Der Strom I_{CD} wird von der p^+-Wanne-Diode der CD-Schaltung erzeugt. Zu diesem können nur Ladungsträger beitragen, die in der N-Wanne bzw. in der p^+-Diffusion erzeugt werden, da die im Substrat erzeugten Löcher die N-Wanne durchqueren müßten, um zu der auf Masse liegenden p^+-Diffusion zu kommen. Wie in Abb. 2.11 zu sehen, ist I_{CD} für etwa $6\,\mathrm{s} \simeq 6\,\mu\mathrm{m}$ konstant bei $-600\,\mathrm{nA}$. Das heißt, daß der Spot sich in dieser Zeit ganz in der N-Wanne befunden haben muß. Dementsprechend läßt sich die Spotgröße aus der Differenz der Seitenlänge der N-Wanne und der Weglänge, während der I_{CD} konstant war, abschätzen. Bei einer Seitenlänge von $14.4\,\mu\mathrm{m}$ ergibt sich ein Spotdurchmesser von etwa $8\,\mu\mathrm{m}$.

Zum Strom I_{plus} können sowohl die in der N-Wanne als auch die im Substrat erzeugten Ladungsträger beitragen, gesammelt an der C- und den beiden CD-Dioden. Da letztere im Substrat umherdiffundieren können, und erst nach relativ langer Zeit wieder rekombinieren, kommt es zu dem langreichweitigen Einfluß des Lichtes auf I_{plus}. Im eindimensionalen Fall, wenn also alle durch das Licht erzeugten Ladungsträger nur in die Richtung der Photodiode diffundieren könnten, ergäbe sich (vgl. [Sze 85]) ein exponentieller Abfall:

$$I \propto e^{-\frac{x}{L_D}} \tag{2.9}$$

mit der Diffusionslänge

$$L_D = \sqrt{D\tau}\ . \tag{2.10}$$

Dabei bezeichnet D die Diffusionskonstante und τ die mittlere Lebensdauer eines Ladungsträgers.

Hier ist das Problem allerdings komplizierter, da die Ladungsträger, hier sind es Elektronen, in einen ganzen Halbraum diffundieren können. Außerdem wird ein Teil der Elektronen auch an anderen Anoden, z.B. N-Wannen der beiden Operationsverstärker, sowie an einer der anderen Photodioden, abfließen.

Die Einschnitte in der Kurve für I_{plus} sind auf die Abschirmung des Lichtes durch Metalleitungen zurückzuführen, wie auch an zwei Beispielen in Abb. 2.10 angedeutet ist. Man beachte, daß der Hauptbeitrag zu I_{plus} für den Lichteinfall links von den 3 Dioden von der CD-Diode geliefert wird, während er für Lichteinfall rechts der 3.Diode von der C-Diode stammt.

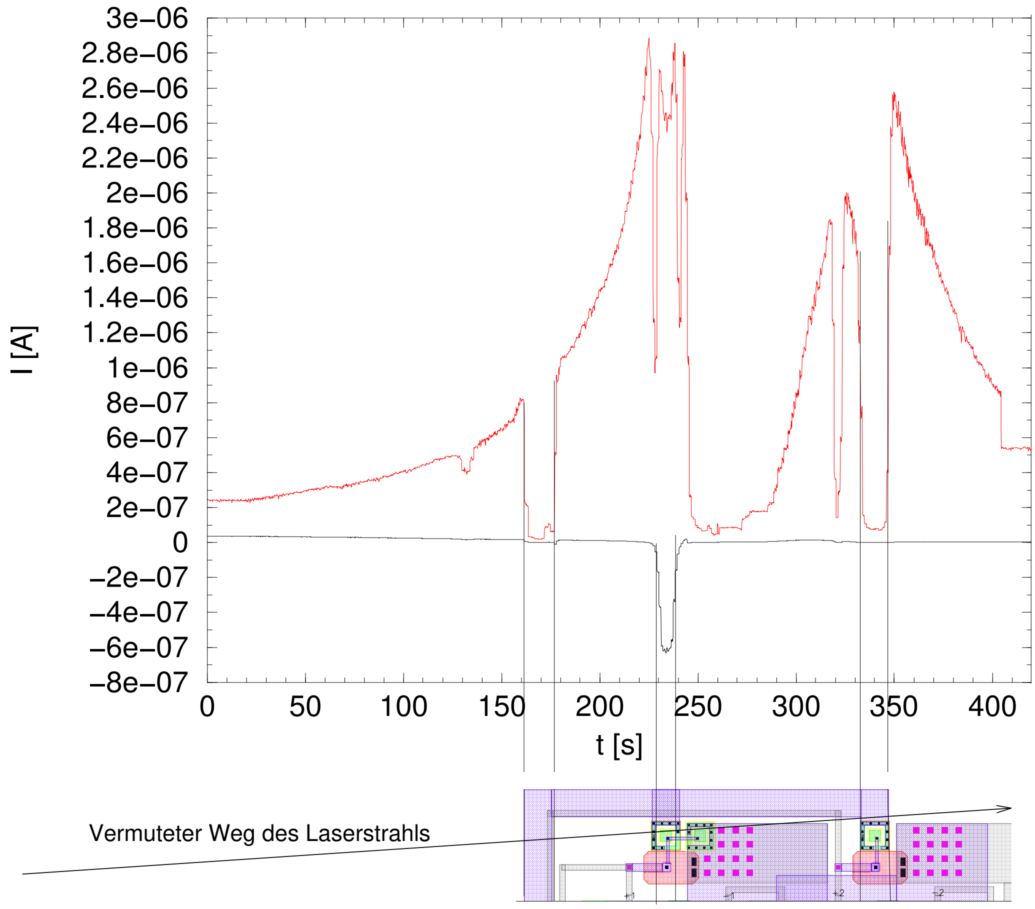

Abbildung 2.10: Photostrom I_{plus} in den Anschluß V_{plus} hinein (obere Kurve) und Strom I_{CD} aus Anschluß CD heraus (untere Kurve) in Abhängigkeit vom Abstand des Laserspots. Im unteren Teil ist der mit dem Laserstrahl überstrichene Teil des Chips gezeigt. Eine Sekunde entspricht etwa einem μm.

Daß I_{plus} zwischen der abgeschirmten CD-Diode und der C-Diode überhaupt ansteigt, ist nur so zu erklären, daß der Laserspot den Chip nicht ganz parallel zu der durch die 3 Dioden vorgegebenen Richtung, sondern wie in Abb. 2.10 gezeigt, überquert hat. Insofern sollte der linke Teil der Stromkurve für die Abhängigkeit der Ladungsträgerkonzentration vom Abstand ihrer Entstehung repräsentativer sein.

In Abb. 2.12 ist der Logarithmus Naturalis von I_{plus} aufgetragen. Wie nach obiger Diskussion erwartet ergibt sich für den linken Teil der Kurve keine Gerade, der Photostrom nimmt also nicht exponentiell mit dem Abstand ab. Für den rechten Teil ergibt sich ein mehr oder weniger gerader Verlauf. Um einen Schätzwert für die Diffusionslänge zu erhalten wurde für beide Entfernungsrichtungen eine lineare Regression durchgeführt. Für die CD-Diode ergibt sich eine Diffusionslänge L_D von $80.3\,\mu$m und für die C-Diode eine von $49.6\,\mu$m. Daraus kann man zumindest folgern, daß man Lichteinfall in einer Umgebung einiger $100\,\mu$m nach Möglichkeit vermeiden sollte.

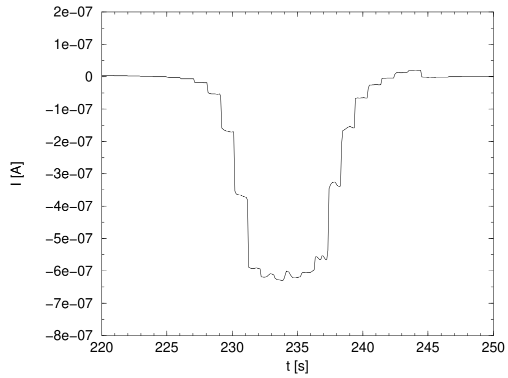

Abbildung 2.11: Ausschnitt aus Abb 2.10. Dargestellt ist nur der entscheidende Teil der Kurve von I_{CD}.

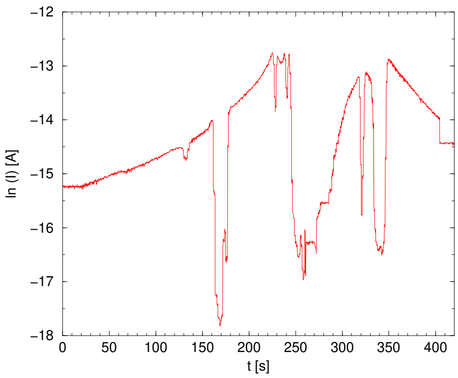

Abbildung 2.12: Logarithmus Naturalis des in Abb. 2.10 dargestellten Stromes I_{plus}

2.3.2 Zeitkonstanten

Um sowohl die Lade- als auch die Entladekurve der Kondensatoren der Schaltungen *OptstoreC* und *OptstoreCD* zu messen, wurden für jede Schaltung jeweils zwei Versuche durchgeführt. Beim ersten wurde die Photodiode der Schaltung mit dem Laserstrahl beleuchtet und so der Kondensator auf $V_{plus} = 5\,\mathrm{V}$ aufgeladen. Danach wurde das Laserlicht vom Chip abgeschirmt und die Schaltung sich selbst überlassen. Die Resthelligkeit wurde mit einem Photometer gemessen und war jeweils geringer als $0.1\,\frac{\mathrm{nW}}{\mathrm{m}^2}$ (Grenze der Nachweisempfindlichkeit des Photometers). Beim zweiten Mal wurde V_{plus} erst auf $0\,\mathrm{V}$ gesetzt, um den Kondensator zu entladen, und dann auf $5\,\mathrm{V}$ erhöht. Bei gleicher Resthelligkeit wurde wieder der Verlauf von $V_{C/CD}$ aufgezeichnet.

Schaltung	Entladekurve		Ladekurve		Theorie	
	$V_{Gl}[V]$	$\tau[s]$	$V_{Gl}[V]$	$\tau[s]$	$V_{Gl}[V]$	$\tau[s]$
C	2.98	716.2	2.86	756.3	4.07	61
CD	2.83	30.9	2.94	383.2	3.96	59

Tabelle 2.3: Gleichgewichtsspannung V_{Gl} und Zeitkonstante τ für *Optstore_C* und *Optstore_CD*. Es sind sowohl die Meßergebnisse aus Kapitel 2.3.2 als auch die mit dem linearen Diodenmodell aus Kapitel 2.1.3 errechneten Werte angegeben.

Das Ergebnis ist in den Abbildungen 2.13 bis 2.16 dargestellt. Um die Parameter τ und V_{Gl} zu extrahieren, wurde die Funktion

$$V(t) = \begin{cases} V_{Gl} + V_0 e^{-\frac{t}{\tau}} & \text{für das Entladen} \\ V_{Gl} - V_0 e^{-\frac{t}{\tau}} & \text{für das Laden} \end{cases} \tag{2.11}$$

an die Daten angepaßt. Dazu mußten die ersten Datenpunkte, die die konstante Anfangsspannung wiedergeben, aus den jeweiligen Datensätzen entfernt werden. Die sich ergebende Funktion ist ebenfalls in den Abbildungen dargestellt, die Parameter τ und V_{Gl} sind in Tabelle 2.3 der theoretischen Vorhersage gegenübergestellt. Der Parameter V_0 hatte sich in dem

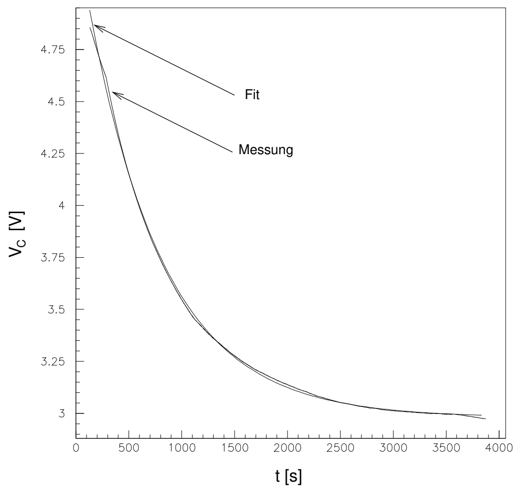

Abbildung 2.13: Entladekurve für den Kondensator aus *OptstoreC*.

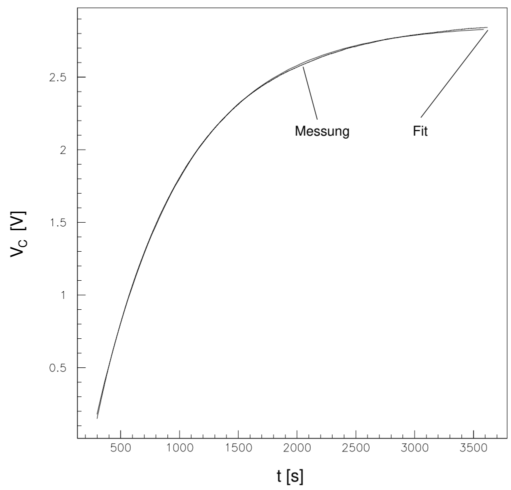

Abbildung 2.14: Ladekurve für den Kondensator aus *OptstoreC*.

linearisierten Modell zu $Vdd - V_{Gl}$ ergeben, mußte bei der Durchführung der Parameteranpassung aber variabel gehalten werden, um den zeitlichen Offset zwischen dem Beginn der Messung und dem Beginn des Entladevorgangs auszugleichen.

Die Meßergebnisse der Schaltung *OptstoreC* stimmen qualitativ sehr gut mit dem gewählten Modell überein: Wie an den Abbildungen 2.13 und 2.14 zu sehen ist, liegen die Datenpunkte fast perfekt auf der der Anpassung zu Grunde liegenden Exponentialfunktion. Außerdem stimmen die aus der Lade- und aus der Entladekurve extrahierten Parameter bis auf wenige Prozent miteinander überein.

Dagegen werden die Lade- und Entladekurven, die sich für die CD-Schaltung ergeben, nur noch sehr schlecht durch die angepaßten Exponentialfunktionen beschrieben. Daher haben die beiden τ-Parameter nur einen sehr beschränkten Aussagewert. Trotzdem ist der Unterschied von einer Größenordnung zwischen Ladezeit und Entladezeit frappant. Die sich einstellenden Gleichgewichtsspannungen V_{Gl} stimmen dagegen sowohl untereinander als auch mit denjenigen, die sich für *OpstoreC* ergeben, gut überein. Eigentlich erwartet man, daß die Gleichgewichtsspannungen der CD-Schaltung unter denen der C-Schaltung liegen, da der Widerstand zu Erde durch die zusätzliche n^+-Substrat-Diode kleiner werden sollte. Um das zu beobachten, ist die Streuung der angepaßten Parameter allerdings zu groß. Desweiteren ist der Effekt so klein, daß er leicht durch die Schwankungen der Prozeßparameter auf dem Chip überlagert werden kann.

Die diskutierten Lade- und Entladekurven wurden alle an dem selben Chip gemessen. Die Schaltung *OptstoreC* wurde noch auf einem anderen Chip ausgemessen. Dort ergab sich qualitativ der gleiche Verlauf. Die Zeitkonstanten betrugen dort allerdings 17 s für den Lade- und 75 s für den Entladevorgang.

Der krasse Unterschied zwischen dem Verhalten von *OpstoreC* und *OpstoreCD* kann nicht erklärt werden, da die zusätzliche abgeschirmte Photodiode der CD-Schaltung wesentlich kleiner als die in beiden Schaltungen vorhandene Schutzdiode des gleichen Typs ist, also

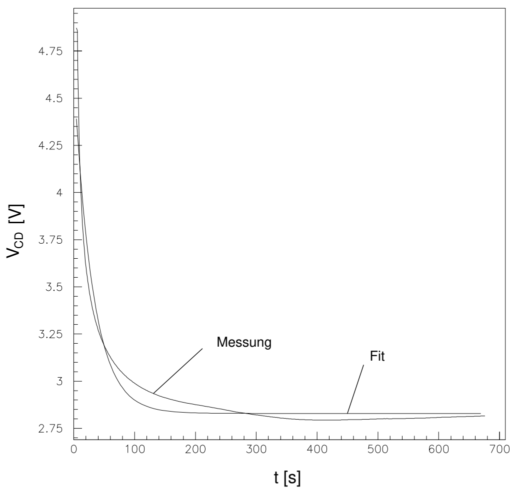

Abbildung 2.15: Entladekurve für den Kondensator aus *OptstoreCD*.

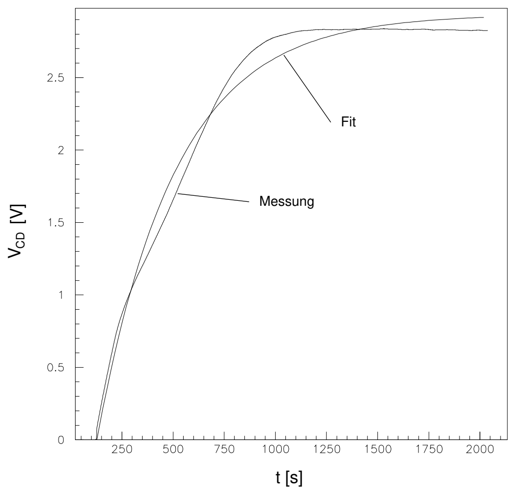

Abbildung 2.16: Ladekurve für den Kondensator aus *OptstoreCD*.

weder einen großen quantitativen noch einen qualitativen Unterschied darstellt.

Die aus den Prozeßparametern ermittelten Parameter für die linearen Modelle der beiden Schaltungen stimmen nicht sehr gut mit den aus der Messung gewonnenen Werten überein. Die für *OpstoreC* gemessene Zeitkonstante ist sogar um einen Faktor 12 größer als der vorhergesagte Wert. Die schlechte Übereinstimmung hängt sicherlich zu einem großen Teil damit zusammen, daß die in Tabelle 2.2 angegebenen Prozeßparameter das Verhalten der jeweiligen Photodiode nur sehr grob beschreiben. So ist es z.B. physikalisch nicht unmittelbar einsichtig, warum sich der Sperrstrom durch die Seitenwand der Diffusion mit der Spannung ändert, derjenige durch die Grundfläche aber nicht.

Zusammenfassend ist festzustellen, daß der Kondensator die Spannung für wenigstens einige hundert Millisekunden, vielleicht sogar einige Sekunden bis auf ein Prozent genau speichern kann.

Kapitel 3

Design und Layout von *Oasys_RN*

Um das Konzept eines optisch konfigurierbaren Widerstandsnetzwerkes zu testen, wurde der Chip *Oasys_RN*[1] mit den folgenden Bestandteilen entworfen: Ein Widerstandsnetzwerk der in Kapitel 1.2 beschriebenen Form mit 64 Knoten, dessen Widerstände durch CMOS-Schaltungen realisiert sind und ein Positionsdetektor.

Die Widerstände sind durch CMOS-Schaltungen realisiert, und lassen sich elektrisch bzw. optisch einstellen. Für die optische Konfiguration wird ein Lichtsignal mit der in Abb. 1.19 gezeigten Schaltung in eine Spannung und diese schließlich in einen Bias-Strom umgewandelt.

Die Eingangssignale können entweder elektrisch oder durch Beleuchtung des Chips mit einem geeigneten Strichmuster erzeugt werden. Im elektrischen Fall werden die analogen Eingangsspannungen sequentiell durch ein Schieberegister gesteuert auf Kondesatoren geschrieben, die als kurzzeitige Analogspeicher dienen. Im optischen Fall wird das einfallende Licht durch Photodioden in Ströme und diese dann durch Transistoren in Spannungen, die proportional zum Logarithmus der Photoströme sind, umgewandelt.

Um die Knotenspannungen auslesen zu können, werden diese durch einen Impedanzwandler verstärkt. Mit einem Multiplexer kann eine beliebige Knotenspannung auf den Ausgangsverstärker durchgeschaltet werden, dessen Ausgangsspannung extern ausgelesen wird.

Als zweiter Baustein des Chips wurde ein Positionsdetektor entwickelt, der, wie in Kapitel 1.3.2 angedeutet, sowohl eine Ortung, als auch eine Vermessung des Laserspots ermöglichen soll.

Alle Simulationen wurden mit dem Simulator *SpectreS* durchgeführt. Falls kein CMOS-Modell angegeben ist, wurde das *CMOS47*-Modell[2] benutzt. Für die Herstellung des Chips wurde der $0.6\,\mu$m double poly double metal Prozeß der Firma Austria Mikro Systeme verwendet.

---

[1]Die Bezeichnung Oasys steht **O**ptically **a**djustable **sys**tem. Der submittierte Chip enthält neben dem Widerstandsnetzwerk noch einen von Markus Loose entworfenen Kamerachip mit dem Namen Oasys_Vichi, den Temperatursensor Oasys_t von Michael Keller und einige Teststrukturen , die von Martin Feuerstack-Raible entworfen wurden und unter dem Namen Oasys_T zusammengefaßt sind. Die Endung _RN steht für **R**esistive **N**etwork und dient zur Unterscheidung von den anderen auf dem Chip implementierten Systemen.

[2]Das CMOS47 Modell ist ein semiempirisches Modell für das Verhalten von CMOS-Transistoren und wird auch als Bsim3 Modell bezeichnet. Nähere Informationen finden sich unter [Cad].

3.1 Design des eindimensionalen Widerstandsnetzwerkes

Im Rahmen diskreter analoger elektrischer Schaltungen gehören Widerstände zu den billigen und einfach zu benutzenden Bauteilen. Die Implementierung von Widerständen in einem CMOS-Prozeß gestaltet sich, zumindest für hochohmige Widerstandswerte, schwieriger. Der Designer kann im Wesentlichen folgende Möglichkeiten dazu benutzen:

- Polysiliziumleitungen, deren Widerstand durch das Verhältnis von Länge und Breite der Leitung bestimmt ist.
- N-Wannen mit der gleichen Abhängigkeit des Widerstands von der Geometrie.
- Geschaltete Kondensatoren: Bei jedem Schaltvorgang wird ein durch die Kapazität der Kondensatoren vorgegebene Ladungsmenge übertragen. Die Schaltfrequenz bestimmt dann den Widerstand.
- Transistorschaltungen, die den linearen Bereiche der Ausgangskennlinie des Transistors benutzen.

Auf Grund des geringen Schichtwiderstandes kostet die Verwendung einer der ersten beiden Möglichkeiten bei der Realisierung großer Widerstände mit $R > 10\,\mathrm{k}\Omega$ viel Platz. Außerdem kann man den Wert eines solchen Widerstandes nur digital, durch Dazu- oder Wegschalten kleiner Widerstandstückchen, variieren, was wiederum zu hohem Platzbedarf und zu einem eingeschränktem Dynamikbereich führt. Gegen den Einsatz von N-Wannen spricht außerdem der mit Umfang und Fläche anwachsende Leckstrom in das Substrat, der dem Strom durch den Widerstand verloren geht. Bei der Benutzung geschalteter Kondensatoren erweist es sich als schwierig, die durch die Schaltvorgänge ausgelöste Ladungsinjektion sowohl in die Kondensatoren selbst als auch in benachbarte Leitungen zu unterdrücken. Desweiteren ist auch mit dieser Lösung eine Regelbarkeit einzelner Widerstände nicht ohne Weiteres zu erreichen. Bei der Verwendung von Transistoren ergibt sich dagegen meist ein sehr beschränkter Eingangsspannungsbereich in dem der 'Widerstand' annähernd linear ist. Desweiteren stellt der auf lokale Schwankungen der Prozeßparameter zurückzuführende Fixed-Pattern-Noise[3] ein schwerwiegendes Problem dar.

Für den Entwurf des eindimensionalen Widerstandsnetzwerkes wurde wegen der am leichtesten zu realisierenden Einstellbarkeit des Widerstandswertes eine Transistorschaltung verwendet. Das Konzept dieser Realisierung des Netzwerkes stammt von C. Mead ([Mead 89]) und ist für den Betrieb im Subthreshold-Bereich entwickelt worden, sollte aber prinzipiell auch für höhere Gate-Source-Spannungen funktionieren. Da der Betrieb im Subthreshold-Bereich den für die Implentierung neuronaler Systeme großen Vorteil eines geringen Leistungsverbrauchs mit sich bringt, wurde das Netzwerk auf den Betrieb in diesem abgestimmt.

3.1.1 Transkonduktanz-Verstärker

In Abb. 1.8 treten zwei Arten von Widerständen auf, deren eine mit R und deren andere mit G gekennzeichnet ist. Für die mit G gekennzeichneten ist hier eine Schaltung gesucht,

[3]Fixed-Pattern-Noise bezeichnet die Schwankungen der elektrischen Eigenschaften von Bauteilen des gleichen Typs, die sich an unterschiedlichen Orten auf dem Chip befinden. Die elektrischen Unterschiede sind von Chip zu Chip verschieden, ändern sich für einen Chip mit der Zeit aber nicht mehr.

die einen zur Differenz zwischen Eingangsspannung S_i und Knotenspannung V_i proportionalen Strom liefert. Da die Eingangsspannung wie oben beschrieben auf einem Kondensator gespeichert werden soll, darf die Schaltung eingangsseitig keinen Strom verbrauchen. Diese Aufgabe erfüllt der in Abb. 3.1 dargestellte Transkonduktanz-Verstärker.

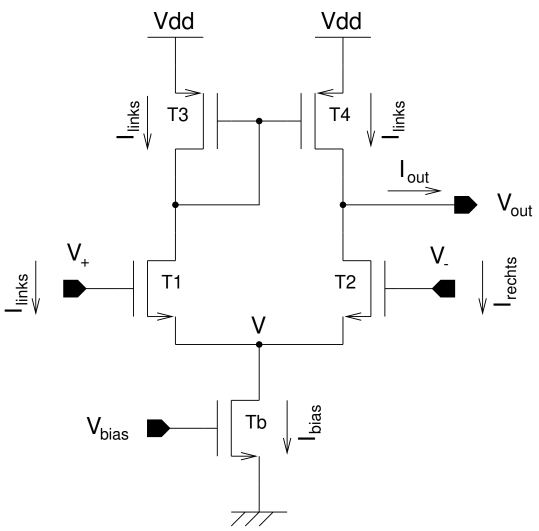

Abbildung 3.1: Schaltplan des Transkonduktanz-Verstärkers. Die Transistoren T1 bis T4 haben alle 2 Gates, deren Breite und Länge jeweils $W = L = 2\,\mu$m betragen. Der Biastransistor Tb hat nur ein Gate derselben Größe.

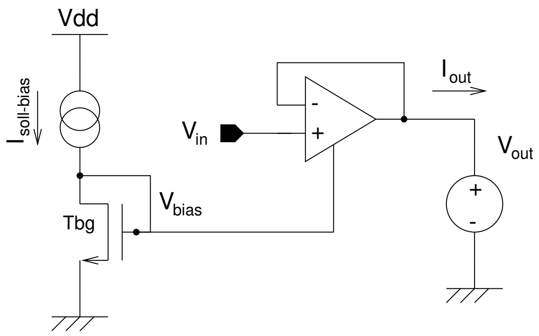

Abbildung 3.2: Beschaltung des Transkonduktanz-Verstärkers für die Simulation. Die Biasspannung V_{bias} wird durch den Transistor Tbg generiert. Breite und Länge von Tbg sind genauso wie für Tb gewählt: $W = L = 2\,\mu$m.

Die Transistoren $T1$ und $T2$ bilden ein differentielles Paar, die Transistoren $T3$ und $T4$ einen Stromspiegel. Wird der Ausgang durch eine Spannungsquelle auf V_{out} gehalten, so ergibt sich, falls der Stromspiegel den Strom I_{links} tatsächlich 1 : 1 in den rechten Zweig spiegelt, für den Ausgangsstrom:

$$I_{out} = I_{rechts} - I_{links} \tag{3.1}$$

Andererseits muß wegen der Kirchhoff'schen Knotenregel aber auch gelten:

$$I_{bias} = I_{rechts} + I_{links} \tag{3.2}$$

Faßt man Gl. (3.1) und (3.2) zusammen, so folgt für den Ausgangsstrom:

$$I_{out} = I_{bias} \frac{I_{rechts} - I_{links}}{I_{rechts} + I_{links}} \tag{3.3}$$

Der Ausgangsstrom ist also das *Produkt* aus dem Biasstrom und dem Verhältnis aus Differenz und Summe der Ströme durch den linken und rechten Zweig. Diese werden im Allgemeinen von den Gate-Source-Spannungen der Transistoren $T1$ und $T2$ in monoton steigender Weise abhängen, so daß sich *qualitativ* ein Verlauf wie für eine der in Abb. 3.3 dargestellten Kurven ergeben sollte.

Im Subthreshold-Bereich kann der Drainstrom I_D eines Transistors durch folgende Gleichung ([Geiger 90]) beschrieben werden:

$$I_D = \frac{W}{L} I_{D0} \quad \underbrace{e^{-V_{BS}(\frac{1}{nV_t} - \frac{1}{V_t})}}_{\text{Bulk-Effekt}} \cdot \underbrace{(1 - e^{-\frac{V_{DS}}{V_t}})}_{V_{DS}\text{-Term}} \cdot \underbrace{e^{\frac{V_{GS}-V_T}{nV_t}}}_{V_{GS}\text{-Term}} =: \quad I_0 e^{\frac{V_{GS}}{nV_t}} \tag{3.4}$$

Hier bezeichnet V_{GS} die Gate-Source-Spannung, V_{DS} diejenige zwischen Drain und Source, V_T die Threshold-Spannung ($V_T \simeq 0.8\,\mathrm{V}$ in dem benutzten Prozeß), $V_t = \frac{kT}{e}$ das thermische Spannungsäquivalent und n einen prozeßabhängigen Parameter, dessen Wert zwischen 1 und 2 liegt. I_{DO} hängt nur von den Prozeßparametern ab, und kann näherungsweise durch

$$I_{D0} = 2K' \left(\frac{nV_t}{e}\right)^2 \tag{3.5}$$

beschrieben werden. Dabei steht K' für den Verstärkungsfaktor und e für die Euler'sche Zahl.

Wenn V_{DS} groß ist gegen V_t, also beispielsweise $V_{DS} \geq 5V_t \approx 130\,\mathrm{mV}$ bei Raumtemperatur, kann man den V_{DS}-Term vernachlässigen: Der Transistor befindet sich im Sättigungsbereich. Dafür sollte man dann noch dem Kanallängenmodulationseffekt[4] Rechnung tragen. In erster Näherung multipliziert man den Ausdruck für I_D aus Gl. (3.4) dazu noch mit $(1 + \lambda V_{DS})$. Der Parameter λ wird als Kanallängenmodulationsfaktor bezeichnet, und hängt sowohl von den Prozeßparametern als auch von der Länge des Transistors ab. Je länger der Transistor desto kleiner das λ. Typische Werte für λ liegen zwischen 0.1 und $0.01\,\mathrm{V}^{-1}$.

Wendet man Gl. (3.4) auf die Eingangstransistoren $T1$ und $T2$ des Transkonduktanz-Verstärkers an, so ergibt sich:

$$I_{links} = I_0 e^{\frac{V_+ - V}{nV_t}} \tag{3.6}$$

$$I_{rechts} = I_0 e^{\frac{V_- - V}{nV_t}} \tag{3.7}$$

Einsetzen dieser Ausdrücke für I_{links} und I_{rechts} in Gl. (3.3) führt zu:

$$I_{out} = I_{bias} \tanh \frac{V_+ - V_-}{2nV_t} \tag{3.8}$$

Durch Differentiation nach $(V_+ - V_-)$ erhält man die differentielle Leitfähigkeit des Verstärkers:

$$G = \frac{d}{d(V_+ - V_-)} I_D = \frac{I_{bias}}{2nV_t} \cdot \frac{1}{\cosh^2(\frac{V_+ - V_-}{2nV_t})} \tag{3.9}$$

Die Leitfähigkeit des Transkonduktanz-Verstärkers ist also auch proportional zum Biasstrom und kann durch diesen kontrolliert werden. Für kleine Spannungsdifferenzen $(V_+ - V_-) < 2nV_t$ ist der Cosinus hyperbolicus ungefähr gleich eins, die Leitfähigkeit also nahezu konstant.

Bei der Verwendung des Transkonduktanz-Verstärkers als Widerstand G in dem Widerstandsnetzwerk aus Abb. 1.8 soll ein zur Spannungsdifferenz $V_i - S_i$ proportionaler Strom in

[4] Wenn man die Drain- bzw. Sourcepotentiale eines MOS-Transistors gegenüber dem Substratpotential erhöht, vergrößern sich die Raumladungszonen der pn-Übergänge an Drain und Source, so daß die Länge des leitenden Kanals verkürzt wird. Da der Strom durch den Tranistor proportional zu $\frac{1}{L}$ ist, ändert dieser sich abhängig von der anliegenden V_{DS}-Spannung. Diesen Effekt bezeichnet man als Kanallängenmodulationseffekt.

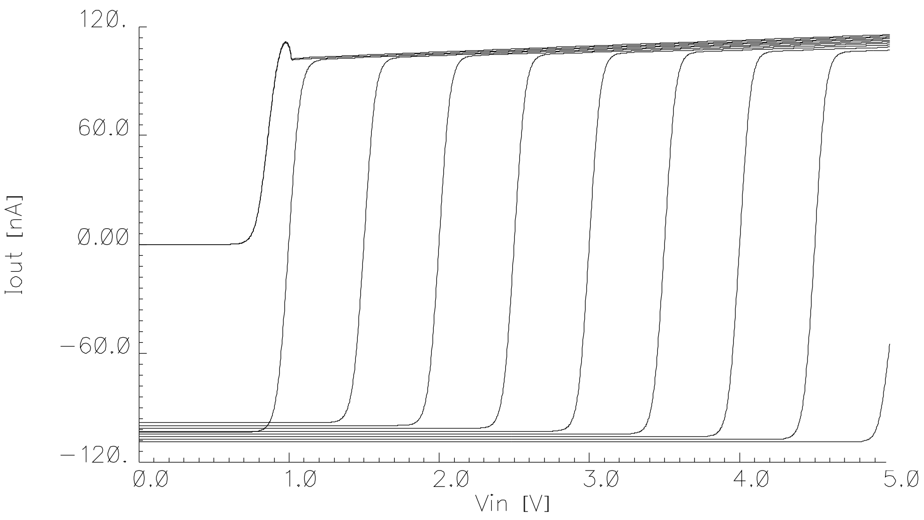

Abbildung 3.3: Simulation des Transkonduktanz-Verstärkers aus Abb. 3.1. Die gewählte Beschaltung ist in Abb. 3.2 gezeigt. V_{in} wurde für $V_{out} = 0, 0.5, \ldots, 5\,\mathrm{V}$ jeweils von 0 bis 5 V variiert. Der Soll-Wert für den Biasstrom betrug $I_{soll-bias} = 100\,\mathrm{nA}$. Vollständige tanh-artige Verläufe ergaben sich für $1 < V_{out} < 4.5\,\mathrm{V}$.

den Knoten i eingespeist werden. Folglich muß der Ausgang mit dem V_--Eingang kurzgeschlossen werden, der Transkonduktanz-Verstärker also als Spannungsfolger verschaltet werden. Für die Simulation wurde der Transkonduktanz-Verstärker deshalb wie in Abb. 3.2 beschaltet. Bei der Generation der Biasspannung durch den Stromspiegel ist zu beachten, daß der hineingespiegelte Strom durch den Biastransistor Tb auf Grund des Kanallängenmodulationseffektes in der Regel nicht ganz dem Soll-Biasstrom $I_{soll-bias}$ entspricht.

In der Simulation wurde die Ausgangspannung V_{out} auf 2 V gehalten und die Eingangsspannung V_{in} von 1.7 bis 2.3 V variiert. Abb. 3.4a zeigt die Kurven, die sich für verschiedene Biasströme $I_{soll-bias} = 25, 50, \ldots, 250\,\mathrm{nA}$ ergeben.

In Abb. 3.4b sind die sich aus diesen Simulationen ergebenden differentiellen Leitfähigkeiten gegen die Eingangsspannungen V_{in} aufgetragen. Prinzipiell bestätigt die Simulation also das vorhergesagte Tangens hyperbolicus-Verhalten des Ausgangsstromes und die $\frac{1}{\cosh^2}$-Abhängigkeit der Leitfähigkeit. Um die Abhängigkeit der differentiellen Leitfähigkeit von I_{bias} zu untersuchen wurde $G|_{V_{in}=2\,\mathrm{V}}$ für die 11 verschiedene Biasströme gegen $I_{soll-bias}$ aufgetragen. Die Kurve besitzt zwei nahezu gerade Teilbereiche unterschiedlicher Steigung, die bei etwa 50 nA ineinander übergehen. Das ist darauf zurückzuführen, daß die Transistoren sich unterhalb dieses Übergangs im Subthreshold-Bereich befinden, oberhalb aber in der Übergangsregion zwischen schwacher und starker Inversion, in der ihr Verhalten sich nicht mehr durch Gl. (3.4) beschreiben läßt.

Um den sinnvollen Ein- und Ausgangsspannungsbereich zu bestimmen, wurde der Transkonduktanz-Verstärker mit der Beschaltung aus Abb. 3.2 für verschiedene Ausgangsspannungen $V_{out} = 0, 0.5, \ldots, 5\,\mathrm{V}$ simuliert. V_{in} wurde von 0 bis 5 V variiert, und $I_{soll-bias}$ betrug 100 nA. Das Ergebnis ist in Abb. 3.3 gezeigt.

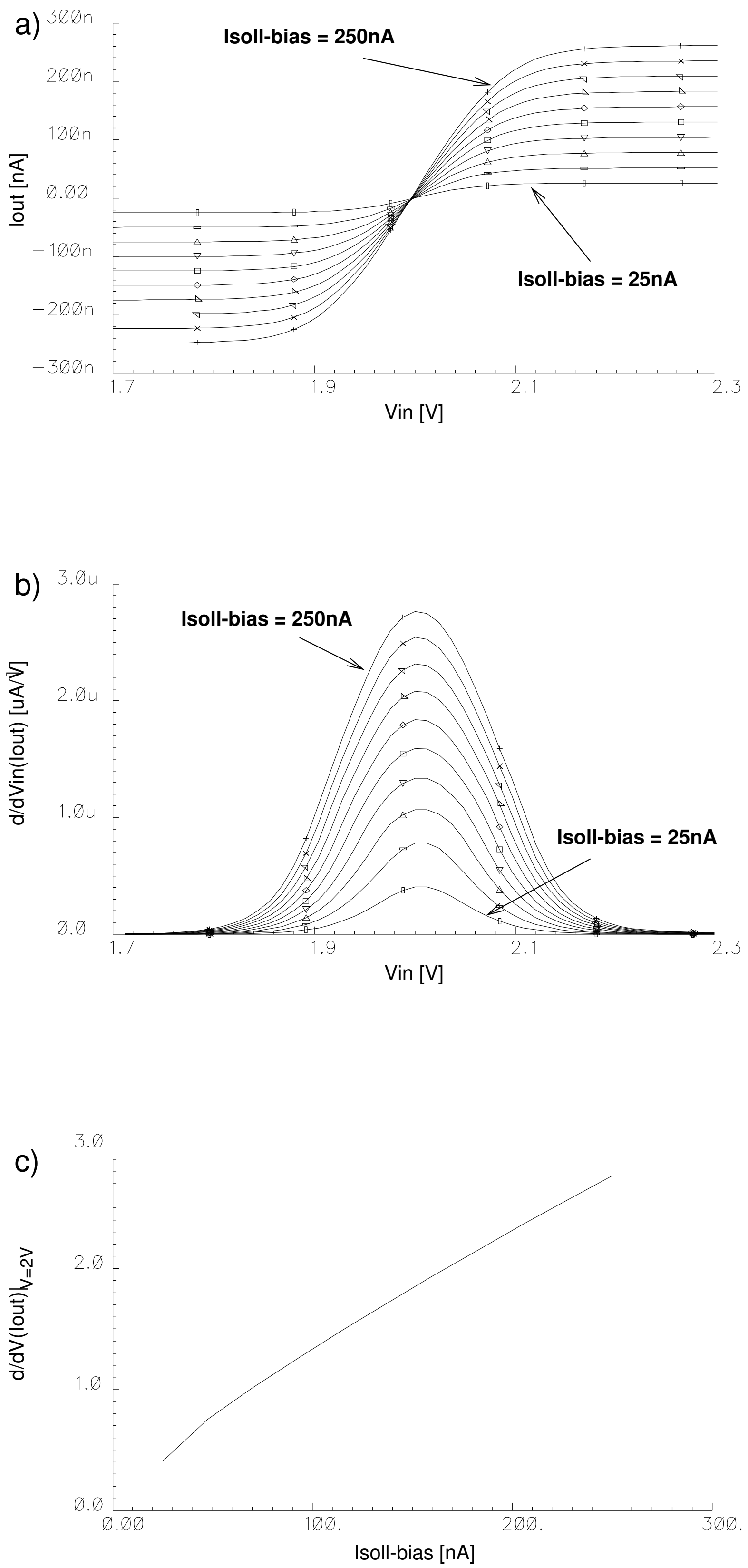

Abbildung 3.4: Simulation des Transkonduktanz-Verstärkers für verschiedene Ströme von $I_{soll-bias}$. Die Eingangsspannung V_{in} wurde für eine konstante Ausgangsspannung $V_{out} = 2\,\mathrm{V}$ zwischen 1.7 und 2.3 V variiert. Simuliert wurde für $I_{bias-soll} = 25, 50, \ldots, 250\,\mathrm{nA}$. a) Kurvenschar für den Ausgangsstrom I_{out}. b) Leitfähigkeit des Transkonduktanz-Verstärkers $G = \frac{d}{dV_{in}} I_{out}$. c) Leitfähigkeit des Transkonduktanz-Verstärkers an der Stelle $V_{out} - V_{in} = 0$.

Der Transkonduktanz-Verstärker funktioniert offensichtlich nicht, wenn die beiden Eingangsspannungen V_{in} und V_{out} kleiner als etwa 0.8 V sind. Das ist zu erwarten, denn damit ein im Sinne der Abbildung nennenswerter Ausgangsstrom fließen kann, muß mindestens einer der Eingangstransistoren leiten. Dazu muß seine Gate-Source-Spannung aber im Bereich der Threshold-Spannung von etwa 0.8 V liegen.

In der Herleitung des Ausdrucks für I_{out} aus Gl. (3.8) hatten sich sowohl der Einfluß von V_{BS} als auch der der Kanallängenmodulation der Eingangstransistoren herausgekürzt. Für die restlichen drei Transistoren des Transkonduktanz-Verstärkers ist $V_{BS} = 0$. Allerdings macht sich sowohl für den Stromspiegel als auch für den Biastransistor die Kanallängenmodulation bemerkbar: Der positive flache Teil der Kennlinien für den Ausgangsstrom aus Abb. 3.3 ist t mit wachsendem V_{out} insgesamt zu negativeren Werten hin verschoben. Das liegt daran, daß das V_{DS} von Transistor $T4$ mit wachsendem V_{out} kleiner wird und so weniger Strom in den linken Zweig der Schaltung herübergespiegelt wird.

Die positive Steigung der Kennlinien im Sättigungsbereich für $V_{in} > V_{out}$ und die mit wachsendem V_{out} negativer werdenden Ausgangsströme für den anderen Sättigungsbereich finden ihre Erklärung in dem mit V (aus Abb. 3.1 !) steigendem Biasstrom durch Tb. In den beiden Sättigungsbereichen wird V nämlich nur durch den jeweils leitenden Eingangstransistor bestimmt.

3.1.2 Hres-Widerstand

In Kapitel 3.1.1 wurde eine Möglichkeit, die G-Widerstände aus Abb. 1.8 durch CMOS-Transistoren zu realisieren, vorgestellt. Für die R-Widerstände soll eine von C. Mead entwickelte Schaltung ([Mead 89]), die in der Literatur mit Hres[5] bezeichnet wird, verwendet werden.

Die Kennlinie eines Widerstandes im I-V-Diagramm ist definitionsgemäß linear. Bei der Realisierung eines Widerstandes mit Hilfe von CMOS-Transistoren ist das für einen größeren Spannungsbereich nur schwerlich zu erreichen. Reduziert man die Anforderung an die Widerstandskennlinie ein wenig, sollte diese monoton steigend bezüglich der angelegten Spannungsdifferenz, punktsymmetrisch bezüglich des Ursprungs der I-V-Ebene und unabhängig vom gemeinsamen Offset der beiden Eingangsspannungen sein. Die einfachste Idee ist nun, einen Transistor im linearen Bereich als Widerstand zu benutzen, dessen Gatespannung in Abhängigkeit von den Spannungen an Source und Drain auf geeignete Weise gesteuert wird. Da für die Steuerspannung die Spannungen an beiden Enden des Widerstandes auf die gleiche Weise berücksichtigt werden müssen, um die Symmetrie-Forderung zu erfüllen, ist es sinnvoller zwei Transistoren zu verwenden. Abb. 3.5 zeigt eine solche Anordnung. Um die Kennlinie des in Abb. 3.5 gegebenen CMOS-Widerstandes zu berechnen, geht man wieder von Gl. (3.4) aus. Faßt man $\frac{W}{L} I_{D0} e^{\frac{V_T}{nV_t}}$ zu I_c zusammen, drückt alle Spannungen in V_t aus und vermischt den V_{GS}-Term mit dem für den Bulk-Effekt unter Berücksichtigung von $V_B = 0$, so vereinfacht sich diese zu:

$$I_D = I_c(1 - e^{-V_{DS}}) e^{(\frac{V_G}{n} - V_S)} \tag{3.10}$$

[5]Hres ist eine Abkürzung für **H**orizontal **res**istor. Die Widerstandschaltung wurde für die Nachahmung der *Horizontalen Zellen*, die als eine Art Widerstandsnetzwerk für eine erste Bildverarbeitung direkt hinter der Netzhaut verantwortlich sind, entwickelt.

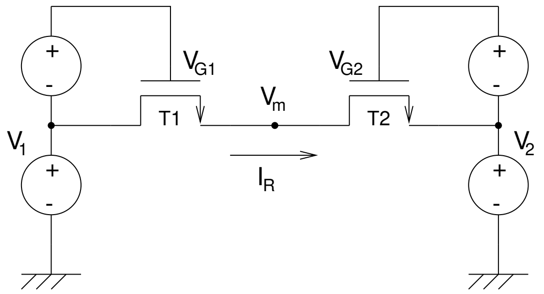

Abbildung 3.5: CMOS-Widerstand bestehend aus zwei Transistoren. Die Wahl und Erzeugung der Biasspannungen $V_{G1} - V_1$ und $V_{G2} - V_2$ wird im Text erklärt. Auf dem Chip *Oasys_RN* sind die Transistoren mit $W = L = 2\,\mu$m und jeweils 2 Gates implementiert.

Die Gatespannungen der beiden Transistoren seien so eingestellt, daß gilt:

$$V_q := \frac{V_{G1}}{n} - V_1 = \frac{V_{G2}}{n} - V_2 \tag{3.11}$$

Der Transistor, dessen Source-Anschluß mit der niedrigeren Eingangsspannung verbunden ist, ist jeweils derjenige, der den Stromfluß begrenzt. Für diesen ist nämlich $\frac{V_G}{n} - V_S = V_q$, während für den Transistor, dessen Source durch V_m gegeben ist, $\frac{V_G}{n} - V_S = V_q + V_D - V_S \geq V_q$ gilt.

Da der Strom I_R durch $T1$ und $T2$ gleich sein muß, ergibt sich für $V_1 > V_2$ mit Gl. (3.10):

$$I_R = I_c(1 - e^{-(V_1 - V_m)})e^{(V_q + V_1 - V_m)} = I_c(1 - e^{-(V_m - V_2)})e^{V_q} \tag{3.12}$$

Aus dieser Gleichung kann man die Spannung V_m zwischen den Transistoren bestimmen:

$$e^{V_m} = \frac{e^{V_1} + e^{V_2}}{2} \tag{3.13}$$

Setzt man V_m in den rechten Teil von Gl. (3.12) ein, so ergibt sich für den Strom durch den CMOS-Transistor:

$$I_R = I_C e^{V_q} \frac{e^{V_1} - e^{V_2}}{e^{V_1} - e^{V_2}} = I_C e^{V_q} \tanh\left(\frac{V_1 - V_2}{2}\right) \tag{3.14}$$

Für den letzten Schritt wurden Zähler und Nenner mit $\frac{1}{2}e^{-\frac{V_1+V_2}{2}}$ multipliziert.

Der Wert des Widerstandes hängt also von $I_C e^{V_q}$ ab. Die Idee zur Erzeugung von V_q besteht darin, einen Transistor gleicher Bauart wie $T1$ bzw. $T2$ aus Abb. 3.5 als Diode zu verschalten, den Strom durch diesen zu kontrollieren und dafür zu sorgen, daß er die gleiche Sourcespannung besitzt. Abb. 3.6 zeigt eine Schaltung, die genau das leistet. Wenn man von dem Transistor Td absieht, handelt es sich wieder um einen Transkonduktanz-Verstärker, der als Spannungsfolger geschaltet ist. Solange wie die Eingangsspannung V_{knoten} in einem sinnvollen Abstand von Vdd und Masse liegt, wird also die Source-Spannung von Td mit V_{knoten} übereinstimmen. Da keine Last am Ausgang des Transkonduktanz-Verstärkers angeschlossen ist, wird sich der Biasstrom I_{bias} zu gleichen Teilen auf die beiden Zweige aufteilen. Für den Strom durch Td gilt dann mit Gl. (3.10):

$$\frac{I_{bias}}{2} = I_c e^{(\frac{V_{G_d}}{n} - V_{knoten})} = I_c e^{V_q} \tag{3.15}$$

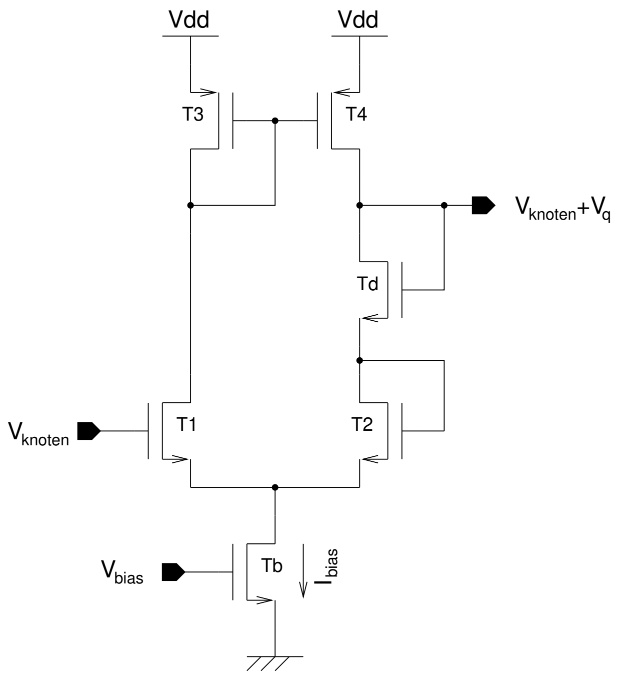

Abbildung 3.6: Bias-Schaltung für den CMOS-Widerstand aus Abb. 3.5. Die Knotenspannung an einem der beiden Enden des Widerstandes wird am Eingang V_{knoten} gemessen und am Gate des als Diode verschalteten Transistors Td wird die Bias-Spannung $V_{knoten} + V_q$ ausgegeben, die den Wert des Widerstandes bestimmt. Die auf dem Chip *Oasys_RN* realisierte Schaltung verwendet für alle Transistoren $W = L = 2\,\mu$m. Außer Tb besitzen alle Transistoren 2 Gates.

Da $V_{DS} = V_{GS} \gg V_t$ ist, wurde der V_{DS}-Term durch 1 ersetzt. Mit Gl. (3.15) läßt sich der Strom durch den CMOS-Widerstand I_R aus Gl. (3.14) schreiben als:

$$I_R = \frac{I_{bias}}{2} \tanh\left(\frac{V_1 - V_2}{2V_t}\right) \tag{3.16}$$

Da die Spannungen V_1 und V_2 in den Gl. (3.16) voraussgehenden Gleichungen in Einheiten von V_t ausgedrückt wurden, taucht im Nenner der Arguments des Tangens hyperbolicus wieder das thermische Spannungsäquivalent V_t auf.

Der Strom durch den CMOS-Widerstand ist also nur von der Größe der angelegten Spannungsdifferenz im Verhältnis zu V_t und von dem Bias-Strom durch den Bias-Schaltkreis abhängig. Gl. (3.16) hat große Ähnlichkeit mit Gl. (3.8), die den Ausgangsstrom des Transkonduktanz-Verstärkers beschreibt. Die einzigen Unterschiede sind der Faktor $\frac{1}{2}$ vor I_{bias} und das Fehlen von n im Nenner des Arguments des Tangens hyperbolicus. Beide Schaltungen werden sich also nur für Eingangsspannungsdifferenzen kleiner als einige V_t annähernd linear verhalten.

Für den differentiellen Widerstand der Hres-Schaltung folgt mit Gl. (3.16):

$$R = \frac{1}{\frac{d}{d(V_1 - V_2)} I_R} = \frac{4V_t}{I_{bias}} \cosh^2\left(\frac{V_1 - V_2}{2V_t}\right) \tag{3.17}$$

Analog zum Transkonduktanz-Verstärker ist die differentielle Leitfähigkeit ($G = \frac{1}{R}$) wieder proportional einem Bias-Strom.

Die Hres Schaltung wurde in der in Abb. 3.7 gezeigten Beschaltung simuliert. Jeweils zwei Hres-Hälften sind zu einem Symbol zusammengefaßt, da diese die Einheit darstellen, aus der durch Aneinanderreihung eine Widerstandskette erzeugt werden kann. Wie bei der Simulation des Transkonduktanz-Verstärkers wird der Bias-Strom durch Spiegelung eines Soll-Bias-Stromes erzeugt.

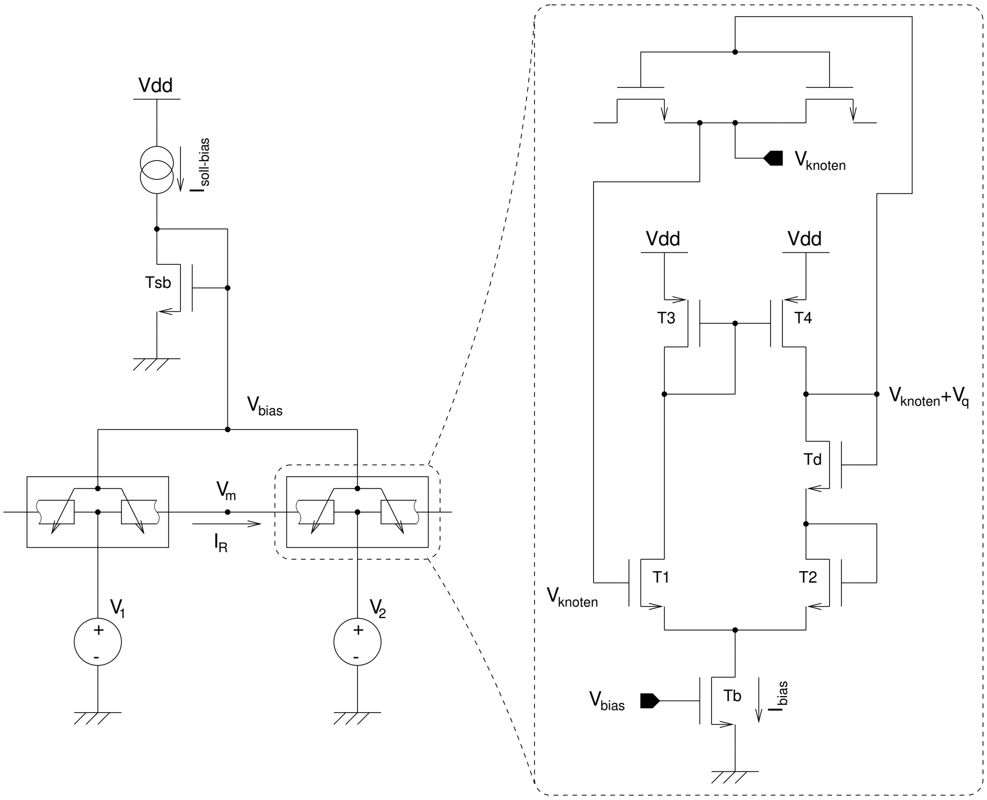

Abbildung 3.7: Beschaltung zweier Hres-Hälften für die Simulation. **Links:** Beide Widerstandshälften werden mit der gleichen Bias-Spannung versorgt, die mit dem Strom $I_{soll-bias}$ erzeugt wird. Der Transistor Tsb ist der gleiche wie Tb, besitzt also zwei Gates mit den Abmessungen $W = L = 2\,\mu$m. **Rechts:** Die Schaltung zeigt was sich hinter dem Symbol für die Hres-Hälfte verbirgt.

Um die Abhängigkeit des Widerstandes von I_{bias} zu testen, wurde die Spannung V_2 auf 2 V fixiert und die Spannung V_1 für $I_{soll-bias} = 50, 100, \ldots, 500\,$nA jeweils zwischen 1.7 und 2.3 V variiert. Ausnahmsweise wurde für diese Simulation das *CMOS15*-Modell[6] verwendet. Das Ergebnis zeigt Abb. 3.8a. Man beachte, daß der Verlauf von I_R hier nur für $V_1 = 1.9\ldots, 2.1\,$V gezeigt ist, der Bereich, in dem sich die Hres-Schaltung linear verhält, also nur einige 10 mV umfaßt. Abb. 3.9 zeigt die gleiche Kurvenschar für eine Simulation mit dem *CMOS47*-Modell. Der lineare Bereich ist hier deutlich größer (Man beachte die andere Skala von V_1). Allerdings besitzen die I_R-Kurven in dieser Simulation unrealistische Extrema an den Enden des linearen Bereichs. Der Vergleich der Simulationen mit den beiden CMOS-Modellen zeigt, daß das Verhalten der CMOS-Transistoren im Subthreshold- und im Übergangsbereich von schwacher und starker Inversion mit den zur Verfügung stehenden Modellen nicht immer ganz wirklichkeitsgetreu simuliert werden kann.

Der differentielle Widerstand, der sich aus der Simulation mit dem *CMOS15*-Modell ergibt, ist in Abb. 3.8b gegen V_1 aufgetragen. Nach Gl. (3.17) sollte sich eine $\cosh^2$-Abhängigkeit von $V_1 - V_2$ ergeben. Insofern stimmt die Simulation qualitativ mit dem analytischen Ausdruck überein. Die Nichtdifferenzierbarkeit bei $V_1 = V_2$ ist allerdings nicht erklärlich, und

[6] Das CMOS15 Modell beschreibt das Verhalten eines CMOS-Transistors und wurde von der Firma AMS entwickelt. Nähere Informationen finden sich unter [Cad].

deutet wieder auf Mängel des Modells hin.

Abb. 3.8c zeigt in Analogie zu Abb. 3.4c den Maximalwert der differentiellen Leitfähigkeit $G_{|V_1=V_2}$. Nach Gl. (3.17) sollte $G = \frac{1}{R}$ linear von I_{bias} und damit auch fast linear von $I_{soll-bias}$ abhängen. Daß das nicht der Fall ist, kann wieder so ausgelegt werden, daß sich die entsprechenden Transistoren für die größeren der benutzten Bias-Ströme nicht mehr im Subthreshold-Bereich befinden. Diese These wird auch von der Beobachtung unterstützt, daß der lineare Bereich der I_R-Kurven aus Abb. 3.8a umso größer wird, je größer der Soll-Bias-Strom gewählt war, bzw. daß in Abb. 3.8b die Kurven für den differentiellen Widerstand mit wachsendem $I_{soll-bias}$ breiter werden.

Zur Abschätzung des Eingangsspannungsbereichs, in dem die Hres-Schaltung wie gewünscht funktioniert, wurde in einer weiteren Simulation mit dem *CMOS15*-Modell für $V_2 = 0, 0.5, \ldots, 5\,\mathrm{V}$ die Spannung V_1 zwischen 0 und 5 V verändert. Das Ergebnis ist in Abb. 3.10 dargestellt. Der Abbildung kann man entnehmen, daß die Hres-Schaltung nur für $V_1, V_2 \in [0.9, 3.2]$ sicher funktioniert. Die untere Grenze für die Eingangsspannungen ergibt sich wie beim Transkonduktanz-Verstärker daraus, daß die beiden Eingangstransistoren $T1$ und $T2$ der Bias-Schaltung ein V_{GS} von der Größenordnung der Thresholdspannung benötigen, um ordnungsgemäß arbeiten zu können. Da für die Transistoren des Stromspiegels und Td die gleiche Forderung zu erfüllen ist, ergibt sich als obere Grenze für den Eingangsspannungsbereich etwa $Vdd - 2V_T$. Die endliche Steigung der I_R-Kurven in den Sättigungsbereichen unter- und oberhalb von $V_1 = V_2$ ist auf den Kanallängenmodulationseffekt sowohl des Bias-Transistors Tb, als auch des Transistors der Widerstandsverbindung, der mit der niedrigeren der beiden Eingangsspannungen verbunden ist, zurückzuführen.

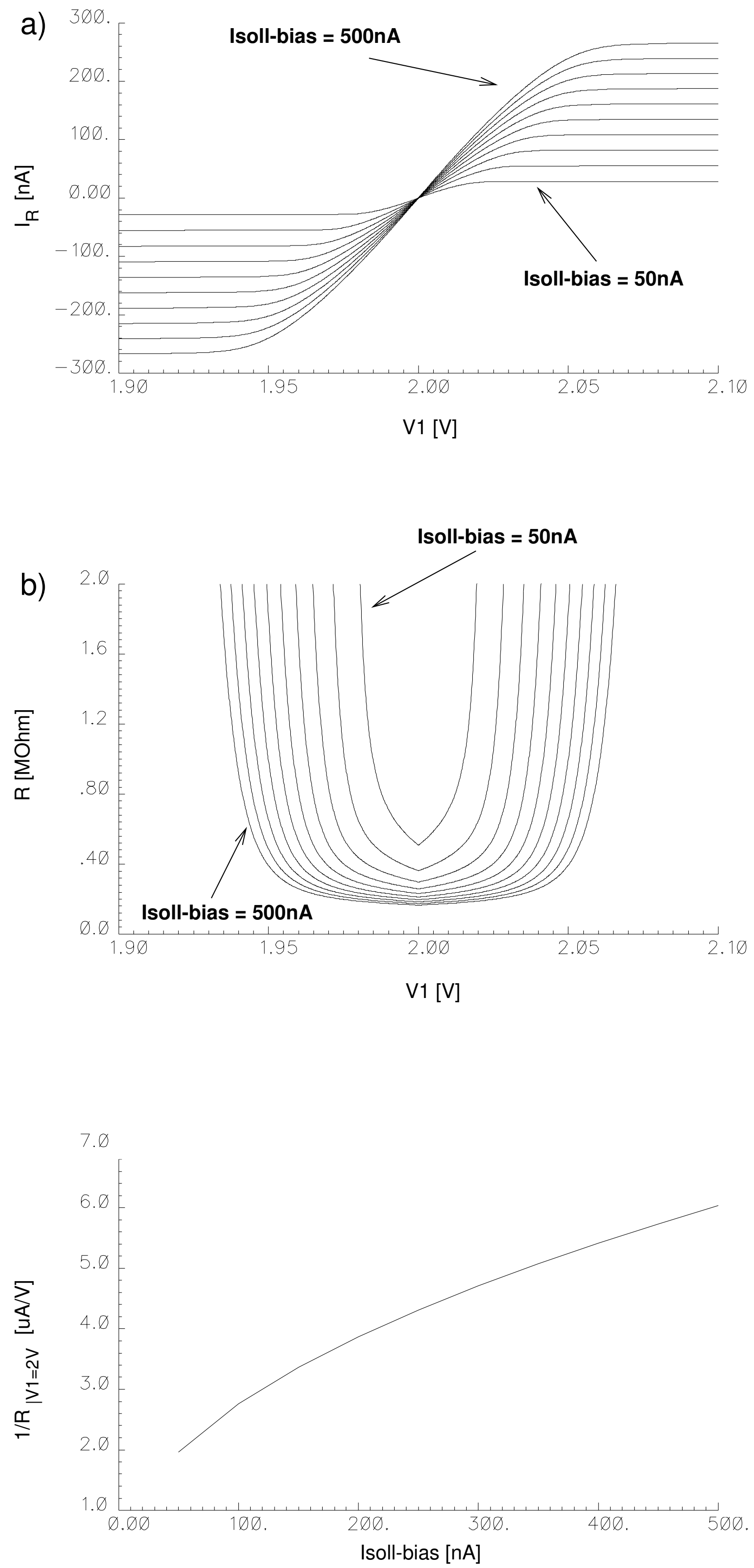

Abbildung 3.8: Simulation des Hres-Widerstandes für verschiedene Ströme von $I_{soll-bias}$. Die Eingangsspannung V_1 wurde für eine konstante Ausgangsspannung $V_{out} = 2\,\mathrm{V}$ zwischen 1.7 und 2.3 V variiert. Simuliert wurde für $I_{bias-soll} = 50, 100, \ldots, 500\,\mathrm{nA}$. a) Kurvenschar für den Strom durch den Widerstand I_R. b) Widerstand des Transkonduktanz-Verstärkers $R = \frac{1}{\frac{d}{dV} I_R}$. c) Widerstand des Transkonduktanz-Verstärkers an der Stelle $V_1 = V_2$.

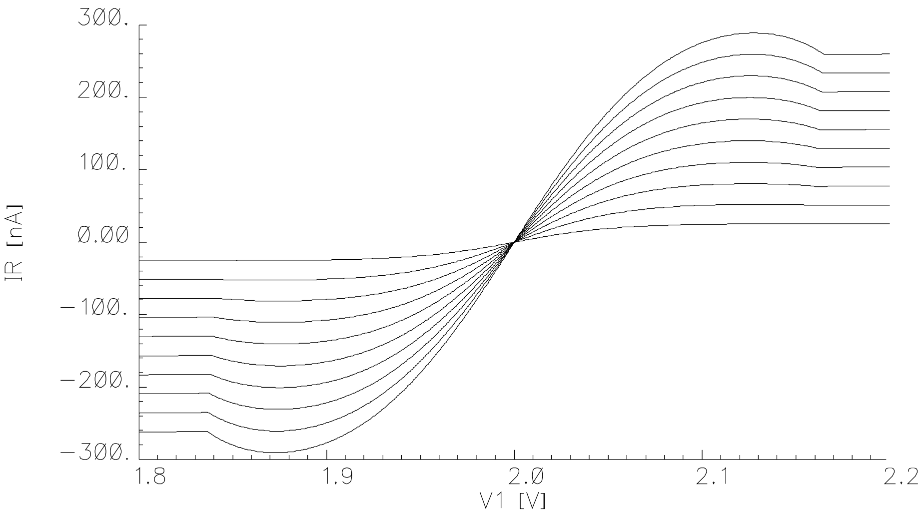

Abbildung 3.9: Die gleiche Simulation wie für Abb. 3.8 wurde mit dem *CMOS47*-Modell durchgeführt. Im Vergleich zu den Kurven in Abb. 3.8a ist der lineare Bereich der Kurven größer. Die Extrema an den Enden des linearen Bereichs scheinen unrealistisch, da die Schaltung dann für zwei endliche Eingangsspannungsbereiche einen negativen Widerstand besitzen müßte.

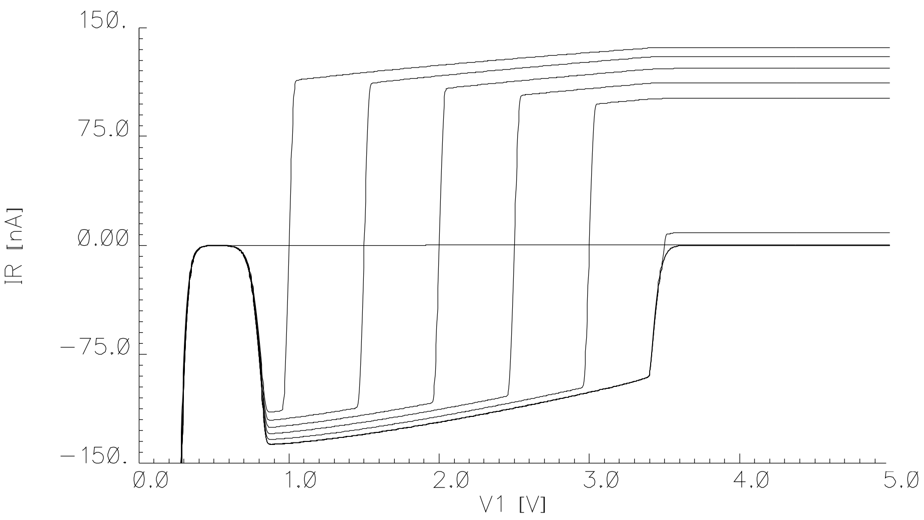

Abbildung 3.10: Simulation der Hres-Schaltung aus Abb. 3.7. V_1 wurde für $V_{out} = 0, 0.5, \ldots, 5\,\mathrm{V}$ jeweils von 0 bis 5 V variiert. Der Soll-Wert für den Biasstrom betrug $I_{soll-bias} = 200\,\mathrm{nA}$.

3.2 Optische Konfiguration der CMOS-Widerstände

Nachdem in Kapitel 1.3.1 eine Möglichkeit gefunden wurde, Lichtsignale in Spannungen zu verwandeln, und im Kapitel 3.1 Implementierungen für die mit R und G bezeichneten Widerstände diskutiert wurden, deren Widerstandswerte sich durch einen Bias-Strom einstellen lassen, bleibt noch die Umsetzung der gespeicherten Spannung in einen Strom zu erörtern.

Abb. 3.11 zeigt die realisierte Lösung des Problems. Die auf dem Kondensator C gespei-

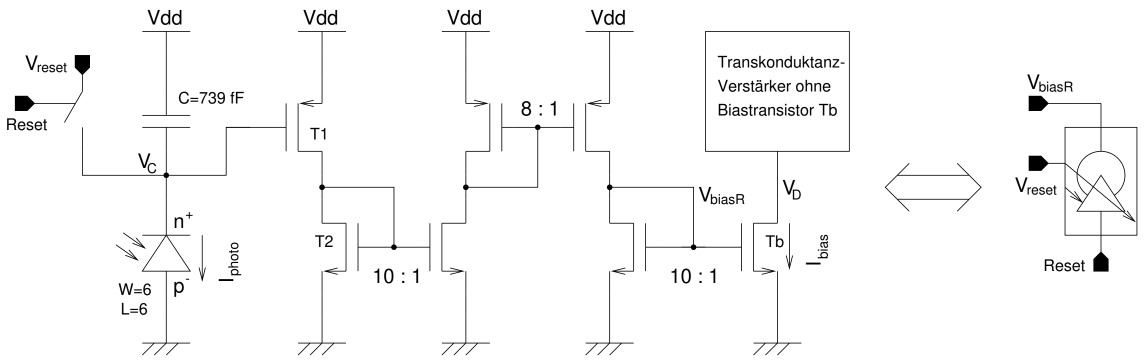

Abbildung 3.11: Schaltung zur Transformation eines Lichtpulses in einen Biasstrom für den Transkonduktanz-Verstärker. Auf der rechten Seite ist das im Folgenden für die Schaltung verwendete Symbol gezeigt. Länge und Breite der Photodiode, im Folgenden Steuerdiode genannt, sind durch $W = L = 6\,\mu$m gegeben. Der Reset-Schalter ist als *transmission gate* realisiert, dessen Transistoren beide die minimale Gate-Größe von $W = 0.8\,\mu$m und $L = 0.6\,\mu$m besitzen. Mit Hilfe dieses Schalters ist es möglich, den Kondensator auf eine beliebige Spannung vorzuladen, bzw. den zu erzeugenden Bias-Strom elektrisch zu kontrollieren. Die drei Stromspiegel reduzieren den Strom durch $T1$ um den Faktor 800. Der resultierende Strom dient als Bias-Strom für den Transkonduktanz-Verstärker. Alle verwendeten Transistoren haben die Gate-Abmessungen $W = L = 2\,\mu$m. Die Zahlen ober- und unterhalb der Stromspiegel geben die Anzahl der auf der jeweiligen Seite verwendeten Transistoren obigen Typs und damit das Soll-Verhältnis der gespiegelten Ströme an.

cherte Spannung V_C wird von dem Transistor $T1$ in einen Strom gewandelt. Da, wie weiter oben schon erwähnt, die Bestandteile des Widerstandsnetzwerkes im Subthreshold-Bereich, bzw. kurz darüber arbeiten sollen, muß der von $T1$ erzeugte Strom noch um den Faktor 800 reduziert werden, was durch die entsprechenden Stromspiegel erreicht wird. Der letzte Stromspiegel enthält den Biastransistor Tb des Transkonduktanz-Verstärkers aus Abb. 3.1. Den Bias-Strom, der sich in Abhängigkeit von V_C ergibt, zeigt Abb. 3.12.

Für 6 verschiedene Werte von V_D, der Drain-Spannung des Bias-Transistors Tb, wurde die Spannung V_C von 0 bis V durchlaufen. Das Auffächern der 6 Kurven zeigt wieder den Kanallängenmodulationseffekt von Tb, der aber in den Simulationen des Transkonduktanz-Verstärkers schon berücksichtigt war. Die durch die Kanallängenmodulation hervorgerufene Streuung zeigt, daß das angegebene Verhältnis von 1:800 der Ströme durch $T1$ und Tb nur ein Richtwert sein kann.

Die Stromkurven aus Abb. 3.12 können in drei verschiedene Bereiche eingeteilt werden, die mit den Arbeitsbereichen des Transistors $T1$ zusammenfallen. Der Drain-Strom eines MOS-Transistors im linearen Bereich kann in erster Näherung durch folgende Gleichung beschrieben

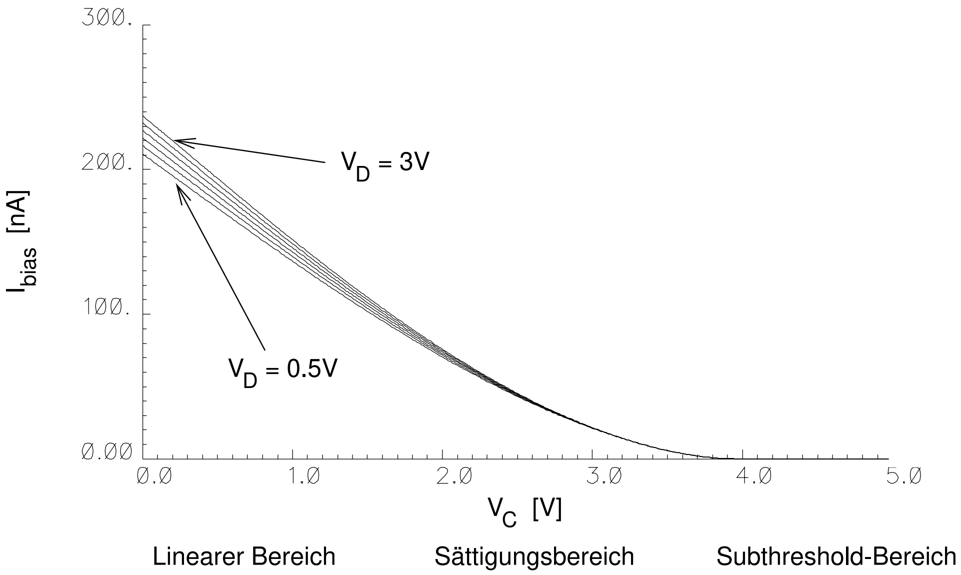

Abbildung 3.12: Simulation der Schaltung aus Abb. 3.11. Für $V_D = 0.5, 1, \ldots, 3\,\mathrm{V}$ wurde die Spannung V_C von 0 bis 5 V variiert. Unterhalb des Graphen sind dem Wertebereich von V_C qualitativ die Arbeitsbereiche des Transistors $T1$ zugeordnet.

werden ([Geiger 90]):

$$I_D = \frac{K'W}{L}\left(V_{GS} - V_T - \frac{V_{DS}}{2}\right) V_{DS}(1 + \lambda V_{DS}) \quad \text{falls} \quad \begin{array}{c} V_{GS} > V_T \\ 0 < V_{DS} < V_{GS} - V_T \end{array} \tag{3.18}$$

Für den Sättigungsbereich gilt, ebenso in erster Näherung:

$$I_D = \frac{K'W}{2L}(V_{GS} - V_T)^2(1 + \lambda V_{DS}) \quad \text{falls} \quad \begin{array}{c} V_{GS} > V_T \\ V_{DS} > V_{GS} - V_T \end{array} \tag{3.19}$$

Der Drainstrom I_D im Subthresholdbereich, also für $V_{GS} < V_T$ ist in Gl. (3.4) angegeben.

Für große Gate-Spannungen V_C von $T1$ ist der erzeugte Strom klein, der Spannungsabfall über $T2$ also in der Größenordnung einer Threshold-Spannung. $T1$ befindet sich dann im Sättigungsbereich. Für Spannungen kleiner als die Thresholdspannung V_T wird der Strom durch $T1$ daher exponentiell mit der Gatespannung V_C anwachsen. Für $V_C > V_T$ ist dagegen eine quadratische Abhängigkeit zu erwarten. In der Simulation ergibt sich für $V_C = 0\,\mathrm{V}$ ein Spannungsabfall über $T2$ von etwa 1.5 V, d.h. für Gatespannungen $V_C \geq 2.7\,\mathrm{V}$ befindet sich $T1$ im linearen Bereich. Der Strom durch $T1$ und damit auch Tb steigt daher für V_C oberhalb von 2.7 V nur noch linear mit V_C an.

Der für V_C sinnvolle Spannungsbereich liegt etwa zwischen 0 und 4 V. Das heißt, daß bei einer Resetspannung $V_{reset} = 5\,\mathrm{V}$ mit der aufgebrachten Lichtmenge erst ein Offset von 20% des zur Verfügung stehenden Dynamikbereichs überwunden werden muß, bevor die Leitfähigkeit G beeinflußt werden kann. Da der zur Verfügung stehende Laser nur zwischen 20 und 100% moduliert werden kann, wird die Photodiode andererseits bei jedem Scanvorgang mit einer Mindestmenge Licht bestrahlt werden, so daß der 'tote' Bereich der Bias-Schaltung sogar notwendig ist, um den vollen Dynamikbereich des Bias-Stroms I_{bias} optimal ausnutzen zu können.

Um die für eine Spannungsänderung ΔV_C der Spannung V_C benötigte Lichtintensität zu bestimmen muß man Gl. (1.32) folgendermaßen umstellen:

$$J = \frac{hc}{\lambda e} \cdot \frac{I_{ph}}{AQ} = 1.84 \cdot \frac{I_{ph}}{AQ} \tag{3.20}$$

Da für den Kondensator

$$C = \frac{q}{|Vdd - V_C|} = \frac{I_{ph}t}{|Vdd - V_C|} \tag{3.21}$$

gilt[7], kann man Gl. (3.20) umschreiben zu:

$$\frac{J}{\Delta V_C} = 1.84 \cdot \frac{C}{\Delta t AQ} \tag{3.22}$$

Auf Grund der Spezifikationen des zur Verfügung stehenden Laserscanners ist für einen Scan aller 64 Photodioden, die im Folgenden als *Kontrolldioden*[8] bezeichnet werden sollen, eine Zeit zwischen 10 ms und $1s$ einzuplanen. Für die Belichtungszeit einer Photodiode ergibt sich daraus $\Delta t \in [0.15\,\mathrm{ms}, 15\,\mathrm{ms}]$. Setzt man $A = WL$ mit den in Abb. 3.11 angegebenen Werten für W und L und $Q = 0.49$ aus Tabelle 2.1 in 3.22 ein, so ergibt sich

$$\frac{J}{\Delta V_C} \in [4.9\,\frac{\mathrm{W}}{\mathrm{m}^2}, 490\,\frac{\mathrm{W}}{\mathrm{m}^2}] \tag{3.23}$$

Um die Spannung V_C um 5 V zu ändern, muß die Lichtintensität für die minimal angenommene Belichtungszeit also etwa 2500 $\frac{\mathrm{W}}{\mathrm{m}^2}$ betragen. Zum Vergleich: Die Intensität senkrecht einfallenden Sonnenlichtes außerhalb der Erdatmosphäre entspricht etwa 1300 $\frac{\mathrm{W}}{\mathrm{m}^2}$. An klaren Sommertagen an Orten der geographischen Breite von Heidelberg reduziert sich diese durch Absorption und Reflexion in der Atmosphäre auf einige 100 $\frac{\mathrm{W}}{\mathrm{m}^2}$. Mit einem fokussierten Diodenlaser sind allerdings problemlos um zwei oder drei Dekaden höhere Intensitäten erreichbar.

Damit die Schaltung aus Abb. 3.11 möglichst unempfindlich gegen einfallendes Restlicht ist, sind möglichst kleine Photodioden wünschenswert. Andererseits sind die zu erwartenden geometrischen Variationen der Diodenfläche im Herstellungsprozeß relativ zur Gesamtfläche umso größer, desto kleiner die Diodenfläche gewählt wird. Insofern stellt die gewählte Dimensionierung einen Kompromiß zwischen der Genauigkeit der Photoströme und der Unempfindlichkeit gegen einfallendes Restlicht dar.

Das Layout eines Widerstandsknotens ist in Abb. 3.13 dargestellt. Ganz rechts befinden sich die zwei Widerstandshälften der Hres-Schaltung, links daneben der Transkonduktanz-Verstärker und der die Bias-Spannung erzeugende Teil der Hres-Schaltung. Der Transistor Td ist in der Mitte der 4 Transistoren der Widerstandsverbindung angeordnet, da sein Verhalten möglichst gut mit dem der ihn umgebenden vier Transistoren übereinstimmen soll. Links von diesem Knotenelement ist der 800 : 1-Stromspiegel zusammen mit dem Transistor $T1$ aus

[7] Um eine Verwechselung mit der Quanteneffizienz zu vermeiden, wird die Ladung hier mit q bezeichnet.

[8] Mit Hilfe dieser Photodioden können die Leitfähigkeiten der Transkonduktanz-Verstärker eingestellt werden, die das Verhalten des Widerstandsnetzwerkes bestimmen. Die Funktion des Netzwerkes kann also durch diese Dioden kontrolliert werden.

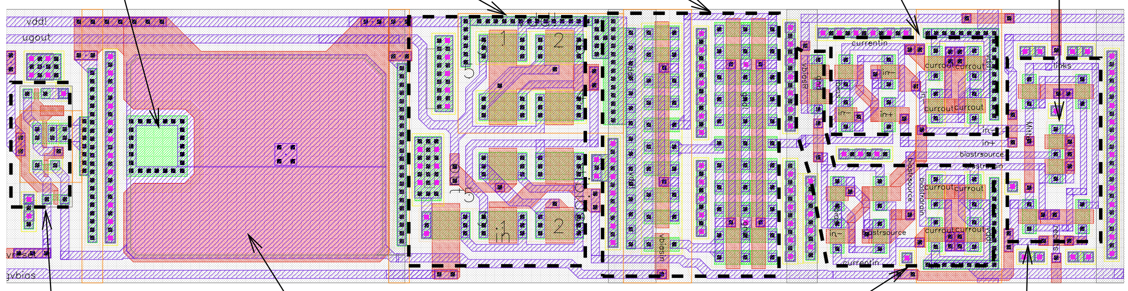

Abbildung 3.13: Ausschnitt aus dem Layout einer einzelnen Zelle des Widerstandsnetzwerkes. Die Höhe des dargestellten Ausschnitts entspricht auf dem Chip 39.8 μm.

Abb. 3.11 plaziert. Der Kondensator und die Photodiode, die zur optischen Programmierung dieser Spannung eingesetzt werden, sind im linken Teil abgebildet. Zwischen dem Kondensator und dem Stromspiegel befindet sich der in Kapitel 3.5 beschriebene Spannungsfolger.

3.3 Simulation des Netzwerkes

Die in Kapitel 3.1 vorgestellten Implementierungen der Widerstände für R und G aus Abb. 1.8 waren nur über einen sehr kleinen Eingangsspannungsbereich von einigen zig mV annähernd linear. Inwiefern sich die vorgestellten Realisierungen nun für ein Widerstandsnetzwerk eignen soll an Hand der folgenden Simulationen untersucht werden. Abb. 3.14 zeigt den Aufbau des Widerstandsnetzwerkes und seine Verschaltung mit Eingangs- und Bias-Spannungen. Die Bias-Spannungen für die Transkonduktanz-Verstärker werden mit Hilfe der in Kapitel 3.2 diskutierten Schaltung erzeugt. Diese wird im 'elektrischen' Modus betrieben, d.h. der Resetschalter ist geschlossen und die Spannung V_C auf dem Kondensator der Bias-Schaltung kann über Spannungsquellen kontrolliert werden.

Die Abhängigkeit der charakteristischen Länge des Netzwerks aus Abb. 3.14 ergibt sich mit (1.10) im Limes kleiner Signale zu:

$$L = \frac{1}{\sqrt{R|_{V_1=V_2}\, G|_{V_+=V_-}}} = \sqrt{\frac{nI_{biasR}}{2I_{biasG}}} \tag{3.24}$$

Im zweiten Schritt wurden hier die Ausdrücke für R und G aus den Gleichungen 3.17 und 3.9 eingesetzt. I_{biasR} bezeichnet dabei den Bias-Strom für die Hres-Schaltung, I_{biasG} denjenigen für den Transkonduktanz-Verstärker. Solange die Widerstände R und G im linearen Bereich arbeiten, ist L also proportional zur Wurzel aus dem Verhältnis ihrer Bias-Ströme und durch diese kontrollierbar.

Zur Untersuchung der Reaktion des Netzwerks auf 'kleine' und 'große' Signale wurden die einfachsten räumlichen Eingangsmuster, nämlich die Stimulation eines einzelnen Eingangs

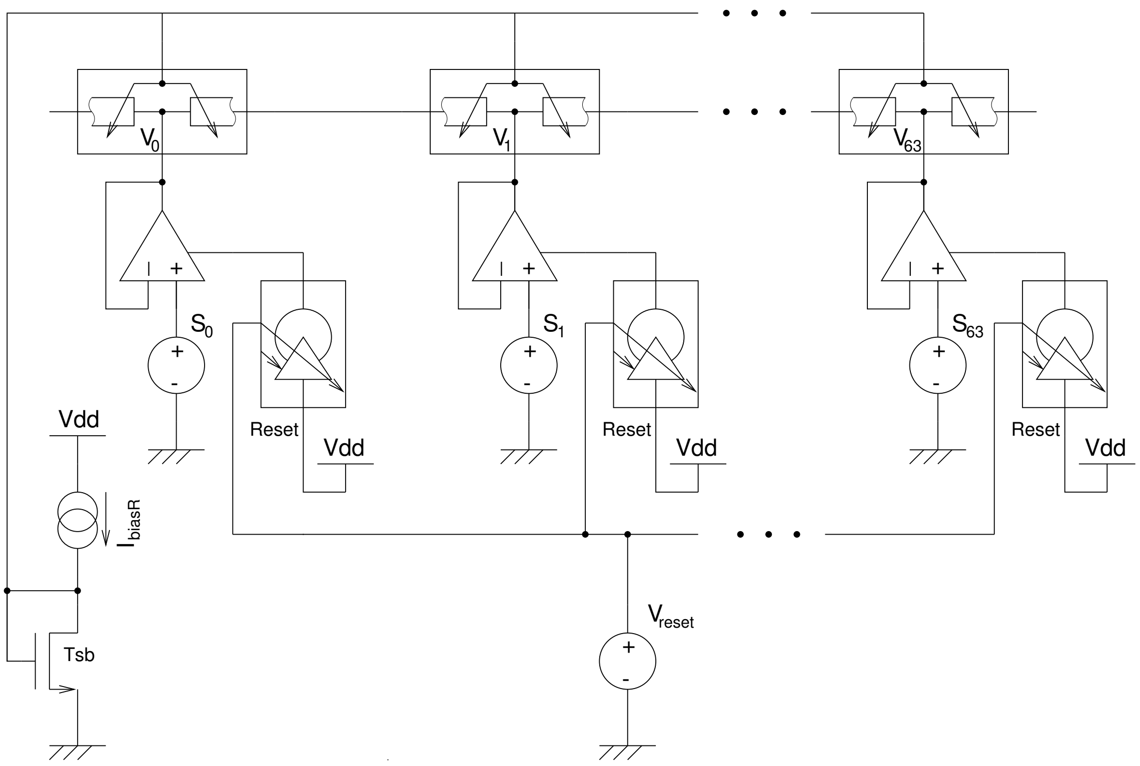

Abbildung 3.14: Schaltplan eines Widerstandsnetzwerkes mit 64 Knoten. Die Transkonduktanz-Verstärker haben die Rolle der G's, die Hres-Schaltungen die der R's aus Abb. 1.8 übernommen. Die Hres-Schaltungen erhalten ihre Bias-Spannung alle durch den Spannungsabfall über Tsb. Die Spannung V_{reset} ist durch $Reset = Vdd$ auf alle Kondensatoren der Bias-Schaltungen für die Transkonduktanz-Verstärker durchgeschaltet, d.h. alle Bias-Ströme der Transkonduktanz-Verstärker werden global durch die Spannung V_{reset} kontrolliert.

gegenüber dem gleichen Potential aller anderen Eingänge und ein kantenförmiges Muster, bei dem die Eingänge der linken Hälfte des Netzwerkes auf einem niedrigeren Potential liegen, als die der rechten Hälfte, verwendet. 'Klein' soll dabei bedeuten, daß die den R's und G's aus Abb. 1.8 entsprechenden Schaltungen alle noch annähernd im linearen Bereich arbeiten, die Spannungen über den Widerständen also weniger als etwa 50 mV betragen. Um das sicherzustellen, wurden für die Simulationen die beiden Eingangspotentiale 2 und 2.1 V benutzt. Mit 'großen' Eingangssignalen sind dann dementsprechend solche gemeint, für die zumindest ein Teil der Widerstände im Sättigungsbereich arbeitet.

Abb. 3.15 zeigt das Ergebnis für die Simulation einer 'kleinen' Einzelstimulation von Knoten 31 in linearer und logarithmischer Darstellung. Die 6 gezeigten Kurven stammen aus Simulationen mit den Reset-Spannungen $V_{reset} = 0,\ 1,\ 2,\ 3,\ 3.5,\ 3.7\,\mathrm{V}$. Der obere Teil der Abbildung zeigt das Pendant zu Abb. 1.11. Qualitativ haben die gezeigten Kurven einen ähnlichen Verlauf: Je kleiner der Wert von G, also je niedriger der Biasstrom I_{biasG} und damit höher die Bias-Spannung V_{reset} ist, desto stärker wird das scharfe Eingangssignal verschmiert.

Die gezeigten Kurven weisen an den Rändern alle einen Offset von einigen mV zu den angelegten 2 V auf, der zu größeren Werten von V_{reset} hin abnimmt. Dieser ist auf den Offset der als Spannungsfolger beschalteten Transkonduktanz-Verstärker zurückzuführen. Auf

Grund des endlichen Ausgangswiderstandes von Transistor $T4$ in Abb. 3.1 unterscheidet sich der Strom durch $T4$ abhängig von der Spannung V_{out} geringfügig von I_{links}. Dadurch entsteht bei gleichen Eingangsspannungen $V_- = V_+$ ein endlicher Ausgangsstrom. Wenn der Ausgang nun mit dem negativen Eingang kurzgeschlossen wird, und kein Strom I_{out} fließen darf, wird die Gleichheit der Ströme im rechten Zweig durch eine Offset-Spannung zwischen V_- und V_+ erzwungen. Da der Strom durch die Eingangstransistoren im Subthreshold-Bereich exponentiell von der anliegenden V_{GS}-Spannung abhängt, im Bereich starker Inversion dagegen nur quadratisch, ist der Offset für kleinere Bias-Ströme im Subthreshold-Bereich kleiner. Verglichen mit den Offsets die sich durch *transistor-mismatch*[9] der realen Transistoren auf dem Chip ergeben, ist der so entstehende Offset allerdings zu vernachlässigen.

Um die logarithmische Darstellung, die im unteren Teil der Abbildung gezeigt ist, zu erhalten, wurde von den Simulationsdaten sowohl die 2 V Eingangsspannung als auch der jeweilige Offset, gegeben durch die Spannung an den äußeren Knoten 0 und 63, abgezogen. Die sich ergebenden Kurven weisen angesichts des hochgradig nichtlinearen Verhaltens der benutzten Widerstands-Schaltungen einen erstaunlich geraden Verlauf auf. Die charakteristische Länge der gezeigten Kurven variiert zwischen $L \approx 1$ für $V_{reset} = 0\,\mathrm{V}$ und $L \approx 7$ für $V_{reset} = 3.7\,\mathrm{V}$.

Wenn man die Antwort des Netzwerkes auf 'große' Spannungssprünge im Eingangsmuster untersuchen möchte, muß man einen im Verhältnis zu den R's und G's großen Eingangsstrom in das Netzwerk einspeisen. Eine solche Situation ist einerseits mit einem kantenförmigen Eingangsmuster zu erreichen, bei dem zumindest für große Werte von L für jede Stromrichtung viele Transkonduktanz-Verstärker zu diesem Strom beitragen, oder indem man einen einzelnen Transkonduktanz-Verstärker mit einem höheren Bias-Strom versorgt.

Für die Simulation, deren Ergebnisse Abb. 3.16 zeigt, wurde die Eingangsspannung der Bias-Schaltung für den Transkonduktanz-Verstärker am Knoten 31 gleich 0 V gesetzt. Die Eingänge aller anderen Bias-Schaltungen liegen wie in Abb. 3.14 auf dem Potential $Vreset$. Die Spannung S_0 betrug 2.5 V, die an allen anderen Eingängen 2 V. V_{reset} wurde für die einzelnen Kurven wieder auf 0, 1, 2, 3, 3.5, und 3.7 V eingestellt und $IbiasR$ betrug wieder 200 nA.

Der obere Teil der Abb. 3.16 zeigt wieder die lineare Darstellung der sich ergebenden Antworten des Netzwerkes. Qualitativ ergibt sich die gewünschte Aufweichung des Eingangsmusters, nur daß hier im Vergleich zu Abb. 3.15 die Spannung am Knoten 31 zu größeren

Werten von V_{reset} zu- und nicht abnimmt, da der Beitrag von G_{31} zur Spannung am Knoten 31 für ein größeres V_{reset} ein größeres Gewicht bekommt.

Um die logarithmische Darstellung im unteren Teil der Abbildung zu erhalten, wurden wieder die 2 V Eingangsspannung und die Offsets der einzelnen Kurven von den Knotenspannungen jedes Datensatzes abgezogen und vom Ergebnis der Logarithmus Naturalis gebildet. Die sich für $V_{reset} \leq 3\,\mathrm{V}$ ergebenden Kurven weisen alle für jede Seite zwei lineare Abschnitte unterschiedlicher Steigung auf. Eine mögliche Erklärung dafür ist, daß für größere Spannungsabfälle über den R's und G's die entsprechenden Widerstände im Sättigungsbereich arbeiten, für kleinere dagegen im linearen Bereich, so daß sich zwei verschiedene Werte für $L = \frac{1}{\sqrt{RG}}$ ergeben.

Zusammenfassend läßt sich Folgendes sagen: Insgesamt behält das Netzwerk zwar seine glättende Wirkung auch für 'große' Spannungssprünge im Eingangsmuster, verliert aber sein exponentielles Antwortverhalten.

[9] *Transistor mismatch* bezeichnet das unterschiedliche Verhalten von gleichartig entworfenen Transistoren auf Grund von Variationen der Prozeßparametern.

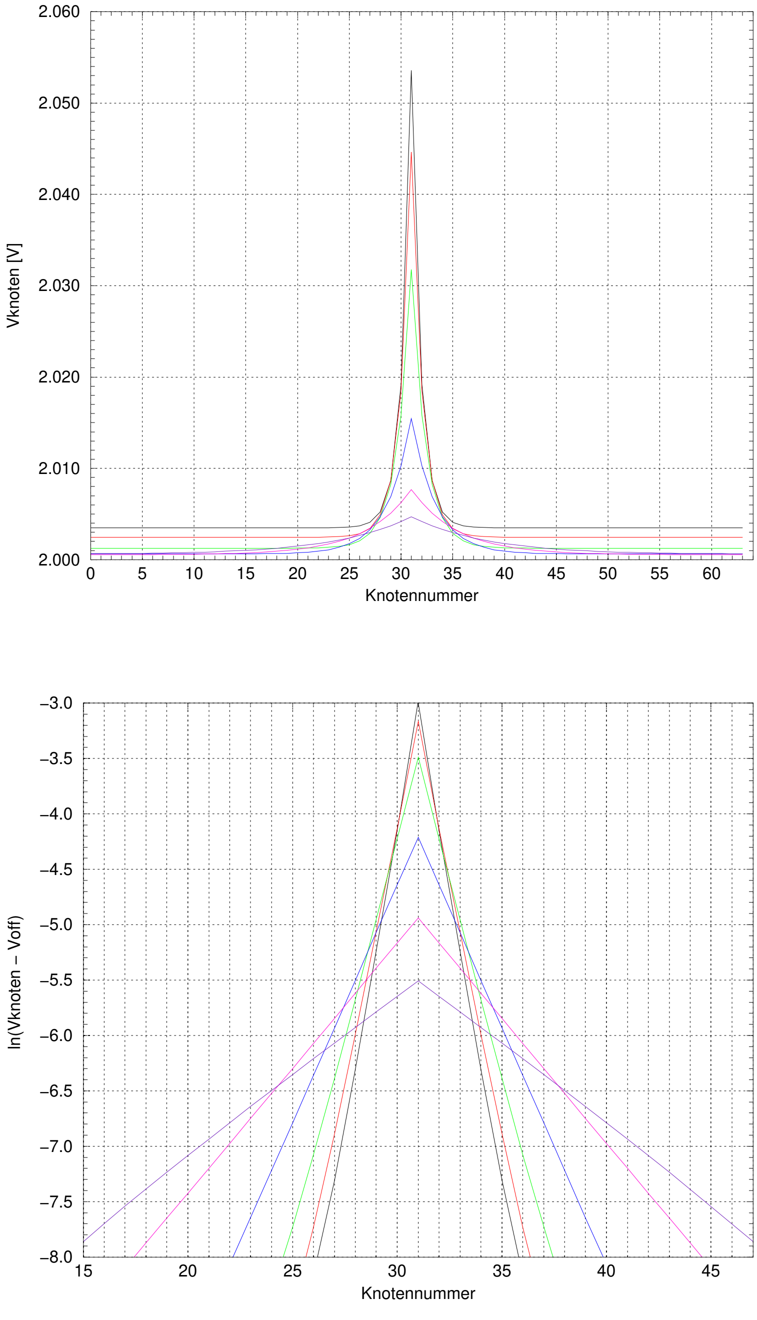

Abbildung 3.15: Statische Simulation der Schaltung aus Abb. 3.14. Der Biasstrom für die Hres-Widerstände I_{biasR} betrug 200 nA. Die Biasspannug für die Transkonduktanz-Verstärker V_{reset} betrug für die verschiedenen Kurven 0, 1, 2, 3, 3.5, 3.7 V. Alle Eingangsspannungen S_i außer S_{31} betrugen 2 V, S_{31} betrug 2.1 V. **Oben:** Räumliche Antwort des Widerstandsnetzwerkes für die 6 verschiedenen Reset-Spannungen V_{reset}. **Unten:** Logarithmische Darstellung der räumlichen Antwort des Netzwerkes. Dargestellt ist die Funktion $\ln(V_i - V_{63})$.

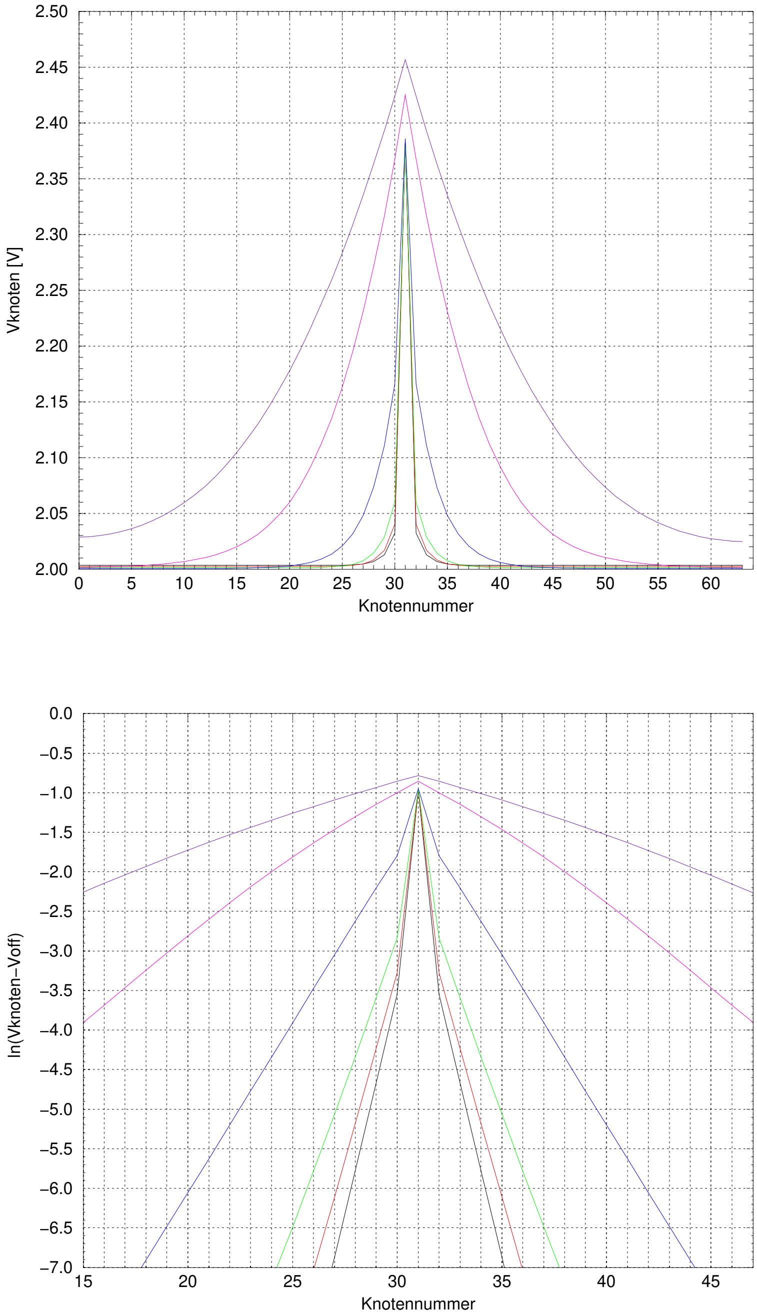

Abbildung 3.16: Die gleiche Simulation wie in Abb 3.15, nur daß S_{31} hier 2.5 V betrug, und die Biasspannung $V_{reset,31}$ am Knoten 31 immer 0 V entsprach, V_{reset} also nur für die restlichen Transkonduktanz-Verstärker wie in der in Abb. 3.15 illustrierten Simulation von 0 is 3.7 V variiert wurde. **Oben:** Räumliche Antwort des Widerstandsnetzwerkes für die 6 verschiedenen Reset-Spannungen V_{reset}. **Unten:** Logarithmische Darstellung der räumlichen Antwort des Netzwerkes. Dargestellt ist die Funktion $\ln(V_i - V_{63})$.

Schließlich wurde das Netzwerk aus Abb. 3.14 noch für kantenfömige Eingangsmuster simuliert. Dazu wurde die linke Hälfte aller Eingangsspannungen auf 2 V, die rechte Hälfte auf 2.1 V bzw. 2.5 V eingestellt. Die Ergebnisse sind in Abb. 3.17 dargestellt (links für die Kante mit 100 mV, rechts für diejenige mit 500 mV Spannungsdifferenz). Alle anderen Parameter entsprachen denjenigen der Abb. 3.15 zu Grunde liegenden Simulation.

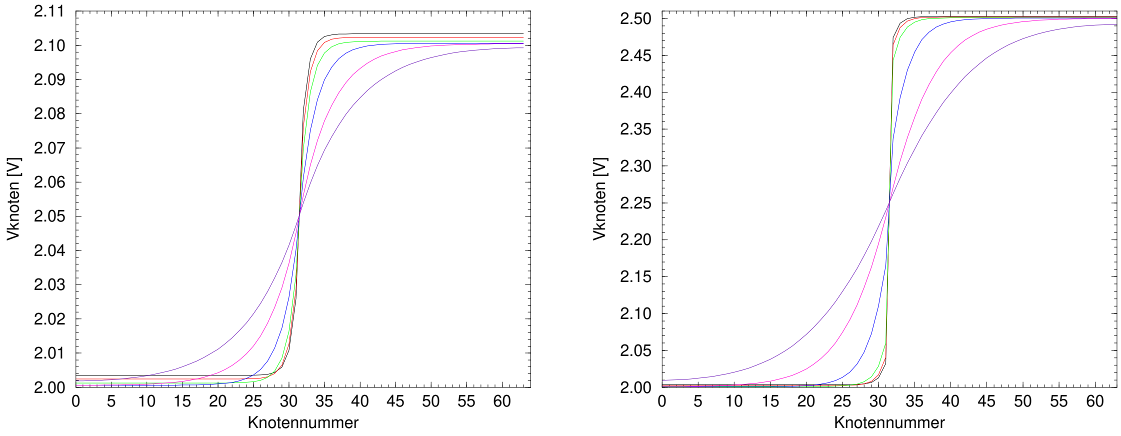

Abbildung 3.17: Vergleich der räumlichen Antworten des Netzwerkes auf eine 'kleine' und auf eine 'große' Kante zwischen den Knoten 31 und 32. Für die Simulation wurde wieder $I_{biasR} = 200\,\text{nA}$ und $V_{reset} = 0,\ 1,\ 2,\ 3,\ 3.5,\ 3.7\,\text{V}$ benutzt. Die Eingangsspannungen S_0 bis S_{31} betrugen 2 V, S_{32} bis S_{63} für die linke Abbildung 2.1, für die rechte 2.5 V.

Die in den zwei Graphen der Abb. 3.17 dargestellten Kurven unterscheiden sich qualitativ nur für $V_{reset} = 0,\ 1,\ 2\,\text{V}$. Für diese Werte von V_{reset} fällt bei dem Eingangsmuster mit dem 0.5 V- Spannungssprung der größte Teil der Spannung über dem Widerstand zwischen den Netzwerkhälften ab. Diese 3 Kurven wirken daher im rechten Teil der Abbildung steiler und einander ähnlicher, als im linken.

3.4 Optische Stimulation des Netzwerkes

Die Eingangsspannungen S_0 bis S_{63} des Widerstandsnetzwerkes können elektrisch oder optisch eingestellt werden. Bei der elektrischen Einspeisung wird die Spannung für den entsprechenden Eingang auf einem Kondensator mit $C = 656\,\text{fF}$ gespeichert. Für jeden Knoten steht ein solcher Kondensator zur Verfügung, der mit Hilfe eines *transmission gates* mit der globalen Eingangsleitung V_{in} kurzgeschlossen werden kann. Die *transmission gates* werden über ein aus 128 Flip-Flops zusammengesetztes Schieberegister angesteuert, so daß die Speicherkondensatoren nacheinander beschrieben werden können. Dadurch, daß nur jedes zweite Flip-Flop ein *transmission gate* ansteuert, soll das Timing beim Beschreiben der Kondensatoren erleichtert werden.

Zur optischen Erzeugung eines Eingangsmusters dient die in Abb. 3.18 gezeigte Schaltung, die hier *Signalpixel* genannt werden soll. Der von der Photodiode erzeugte Strom I_{photo} wird von den beiden als Dioden verschalteten PMOS-Transistoren in eine Spannung V_{in} verwandelt. Für die V_{GS}-Spannungen der beiden Transistoren gilt nach Gl. (3.4):

$$V_{GS} = nV_t(\ln(\frac{L}{W} \cdot \frac{I_D}{I_{D0}}) + V_{BS}(\frac{1}{nV_t} - \frac{1}{V_t})) + V_T = nV_t(\ln(\frac{I_D}{I_{D0}}) + \ln(\frac{L}{W})) + V_T \qquad (3.25)$$

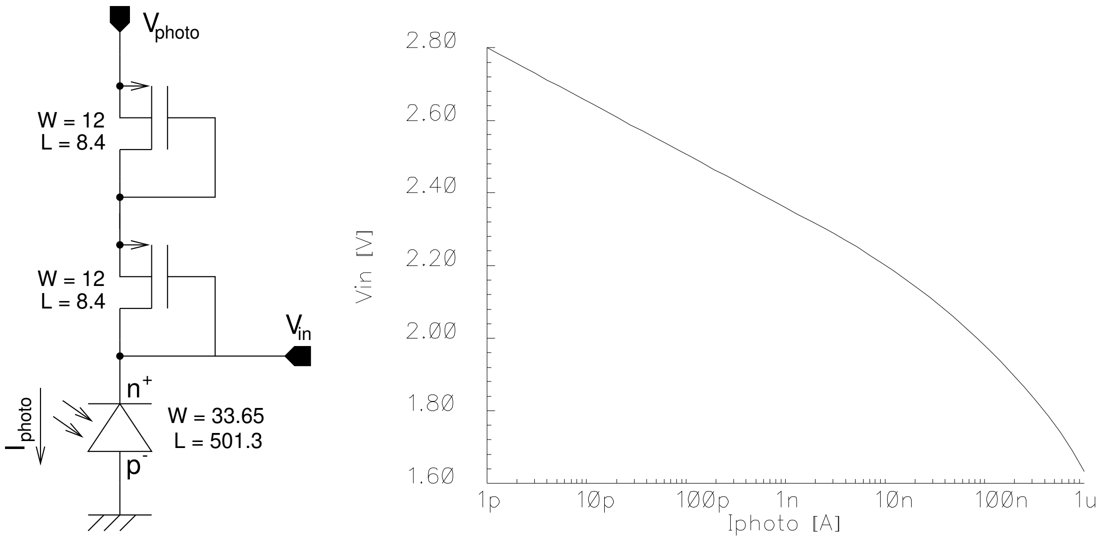

Abbildung 3.18: Schaltplan und Simulationsergebnis für den *Signalpixel*. **Links:** Zwei als Dioden verschaltete PMOS-Transistoren wandeln den in der Photodiode erzeugten Strom in die Spannung V_{in} um. Mit der Spannung V_{photo} kann die sich ergebende $V-I$-Kurve in den richtigen Spannungsbereich verschoben werden. Alle Maße sind in μm angegeben. **Rechts:** Ausgangsspannung V_{in} aufgetragen gegen den Photostrom.

Im zweiten Schritt wurde hier $V_{BS} = 0$ eingesetzt, so daß sich für den Subthresholdbereich eine logarithmische Abhängigkeit von I_{photo} ergibt:

$$V_{in} = V_{photo} - 2nV_t(\ln(\frac{I_D}{I_{D0}}) + \ln(\frac{L}{W})) - V_T \tag{3.26}$$

In der dem rechten Teil von Abb. 3.18 zu Grunde liegenden Simulation wurde die Photodiode durch eine Stromquelle ersetzt, und I_{photo} variiert. Bis zu einem Strom von 10 nA gehorcht V_{in} der logarithmischen Abhängigkeit aus Gl (3.26), für höhere Ströme befinden sich die Transistoren offensichtlich nicht mehr im Subthreshold-Bereich. Die schwache Variation von V_{in} mit dem Photostrom scheint dem Verhalten der benutzten Implementierung des Widerstandsnetzwerkes angemessen zu sein, da sich ja in Kapitel 3.3 herausgestellt hatte, daß dieses nur für Variationen der Eingangsspannung in der Größenordnung einiger 10 mV wie ein Netzwerk ohm'scher Widerstände reagiert.

Die Abmessungen der Transistoren sind unter Berücksichtigung der durch das Layout vorgegebenen Grenzen maximiert, um den zu erwartenden *mismatch* zwischen den zu jedem Knoten gehörenden Signalpixeln zu minimieren.

Für eine optische Erzeugung des Eingangsmusters soll der Chip mit einem Muster aus Strichen unterschiedlicher Lichtintensität beleuchtet werden. Dazu könnte man beispielsweise ein Dia mit dem geeigneten Strichmuster erzeugen, von hinten belichten, und das Muster dann auf den Chip abbilden. Die Photodioden des Signalpixels sind für diesen Zweck aus zwei Gründen möglichst groß zu entwerfen. Erstens wird die Aufgabe das Strichmuster auf die Signalphotodioden abzubilden für große Photodioden leichter. Zweitens sollen diese Photodioden möglichst lichtempfindlich sein, um die zur Erzeugung der Eingangsmuster benötigte Lichtintensität gering zu halten, so daß das auf die Kontrolldioden fallende Restlicht möglichst geringe Intensität besitzt.

Die Frage nach den für eine optische Stimulation sinnvollen Lichtintensitäten läßt sich mit Gl. (3.20) beantworten. Wenn man den logarithmischen Teil der Kennlinie des Signalpixels benutzen möchte, ist die obere Grenze für I_{photo} durch das Ende des Subthreshold-Bereichs, also etwa 10 nA, gegeben. Mit Gl. (3.20) entspricht das einer Lichtintensität von $J = 2.2\,\frac{\mathrm{W}}{\mathrm{m}^2}$. Auf der anderen Seite sollte die Lichtintensität so groß sein, daß der erzeugte Photostrom sich signifikant von dem Dunkelstrom I_{dunkel} der Photodiode unterscheidet. Für die in Frage stehende Diode ergibt sich mit den Parametern aus Tabelle 2.1 als Abschätzung[10] für den Dunkelstrom ein Wert von 254.3 fA. Setzt man dieses Ergebnis für I_{photo} in Gl. (3.20) ein, ergibt sich eine Lichtintensität von $61\,\frac{\mu\mathrm{W}}{\mathrm{m}^2}$. Dementsprechend sollten zur optischen Stimulation Lichtintensitäten aus dem Intervall $[1\,\frac{\mathrm{mW}}{\mathrm{m}^2}, 2.2\,\frac{\mathrm{W}}{\mathrm{m}^2}]$ verwendet werden.

3.5 Ausgangstreiber

Damit die Knotenspannungen beim Auslesen nicht beeinflußt werden, muß eine Impedanzwandlung vorgenommen werden. Das geschieht für jeden Knoten einzeln mit Hilfe eines als Spannungsfolger geschalteten Transkonduktanz-Verstärkers. Dieser unterscheidet sich von dem in Abb. 3.1 gezeigten lediglich in der Dimensionierung der Gates, für die hier $W = L = 4\,\mu$m benutzt wurde, um die Offsets zwischen den Transkonduktanz-Verstärkern der jeweiligen Knoten gering zu halten.

Die 64 Ausgänge der Transkonduktanz-Verstärker können mit Hilfe eines 64 zu 1 Multiplexers auf eine globale Ausgangsleitung durchgeschaltet werden. Die so erhaltene Ausgangsspannung wird mit einem zweistufigen Miller-Operationsverstärker als Impedanzwandler so verstärkt, daß sie mit Hilfe diskreter Elektronik auf der Meßplatine weiterverarbeitet werden kann.

Die Theorie einfacher Operationsverstärker wird in den Lehrbüchern zum Analogen VLSI-Entwurf ausführlich behandelt (siehe z.B. [Laker 94] oder [Allen 90]). Daher soll hier nur kurz erläutert werden für welchen Zweck der entworfene Operationsverstärker optimiert wurde, und welchen Spezifikationen er den Simulationen nach genügen soll. Die wichtigsten Spezifikationen eines Operationsverstärkers sind Eingangs- und Ausgangsspannungsbereich, Verstärkungs-Bandbreiten-Produkt GB, Phasenreserve PM für die gewünschte Lastkapazität C_{load}, Anstiegsgeschwindigkeit SR, Einschwingzeit t_s[11], Leerlaufverstärkung A_0, Ruhestrom I_q, sowie Eingangs- und Ausgangswiderstand und der benötigte Platz. Die oben genannten Forderungen sind weder voneinander unabhängig noch existiert eine Lösung für beliebige Punkte dieses Parameterraumes. Beim Entwurf des Operationsverstärkers versucht man dann unter Einhaltung der aufgestellten Bedingungen den Entwurf auf einen der Parameter, z.B. den Platzbedarf oder den Ruhestrom, hin zu optimieren.

Der entworfene Operationsverstärker soll sowohl als Ausgangstreiber für die Knotenspannungen dienen, als auch in den Positionsdetektoren als Ladungs-Verstärker und für eine *sample and hold*-Schaltung benutzt werden. Da die Knotenspannungen mit einem ADC mit einer maximalen Abtastrate von 500 kHz digitalisiert werden sollen, sollte der Operationsverstärker in

[10] Die in der Tabelle 2.1 angegebenen Sperrstromdichten gelten für den 0.8 μm-Prozeß der Firma AMS. Da die entsprechenden Werte für 3 V Sperrspannung in den vorliegenden Prozeßparametern für den 0.6 μm-Prozeß nicht angegeben waren, die für 5V Sperrspannung angegebenen aber mit denen für den 0.8 μm-Prozeß übereinstimmen, stellen die Sperrstromdichten für 3 V Sperrspannung im 0.8 μm-Prozeß wohl eine sinnvolle Abschätzung dar.

[11] Da eine Meßgenauigkeit von etwa 1 mV angestrebt wird, ist die Einschwingzeit hier immer auf das Erreichen eines 1 mV großen Bandes um den Sollwert bezogen.

der Lage sein, einem Spannungssprung von 1 V in höchstens 2 μs zu folgen. Dementsprechend sollte für die Anstiegsgeschwindigkeit $SR \geq 1\,\frac{\mathrm{V}}{\mu\mathrm{s}}$ und für die Einschwingzeit $t_s < 1\,\mu$s gelten. Da der Verstärker wenigstens ein Stückchen Leitung und einen externen Operationsverstärker treiben können muß, sollten diese Anforderung bei einer Lastkapazität von mindestens 50 pF erfüllt werden. Der Spannungsbereich, in dem der Operationsverstärker als Spannungsfolger beschaltet betrieben werden kann, muß zumindest den möglichen Ausgangsspannungsbereich des Widerstandsnetzwerkes, also das Intervall [1 V, 3.5 V] beinhalten. Als Optimierungskriterium sollte der Stromverbrauch des Verstärkers möglichst gering gehalten werden, um eine Störung der empfindlichen Schaltungsteile des Widerstandsnetzwerkes nicht durch einen zu großen Temperaturgradienten, verursacht durch unnötig hohen Leistungsverbrauch der Operationsverstärker, zu vermeiden.

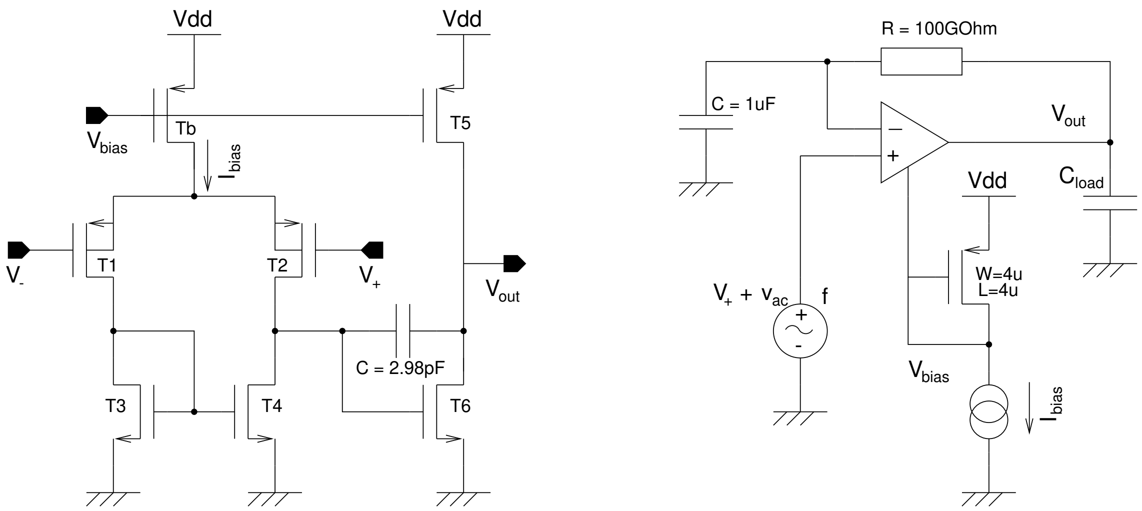

Abbildung 3.19: Miller-Operationsverstärker und Beschaltung für die AC-Simulation. **Links:** Schaltplan des zweistufigen Miller-Operationsverstärkers. $T1$ bis $T4$ sind als Cross-Coupled-Pairs mit zwei Gates mit $W = L = 16\,\mu$m entworfen. Die Gates von Tb, $T5$ und $T6$ sind jeweils $4\,\mu\text{m} \times 4\,\mu$m groß. Tb besitzt ein Gate, $T5$ 51 und $T6$ 52. **Rechts:** Beschaltung des Operationsverstärkers zur Messung seines Frequenzganges. Am positiven Eingang ist ein Sinussignal (v_{ac}) mit $100\,\mu$V Amplitude angelegt. Über einen Transistor der gleichen Bauart wie Tb wird der Strom I_{bias} in den Operationsverstärker hineingespiegelt. Der Widerstand R und der Kondensator C bilden einen Tiefpaß mit einer Zeitkonstanten $\tau = RC = 100$ ks. Der Tiefpaß sorgt dafür, daß sich sowohl V_- als auch V_{out} auf dem Eingangspotential V_+ befinden, beeinflußt das Frequenzverhalten des Operationsverstärkers bei den interessierenden Frequenzen aber nicht.

Als Lösung des oben angedeuteten Anforderungsprofils wurde der im linken Teil von Abb. 3.19 gezeigte zweistufige Miller-Operationsverstärker entwickelt. Um den Offset des Operationsverstärkers möglichst gering zu halten, wurden sowohl die Eingangstransistoren $T1$ und $T2$, als auch diejenigen des Stromspiegels $T3$ und $T4$ als *cross coupled pairs* mit großem W und L entworfen. Die Abmessungen sowie die aus Simulationen gewonnenen Kenngrößen des Operationsverstärkers sind in Tabelle 3.1 zusammengefaßt.

In der Beschaltung des Operationsverstärkers als Spannungsfolger ist die Rückkopplung des Ausgangs auf den Eingang maximal. Abhängig von der Lastkapazität C_{load} und der sogenannten Miller-Kapazität C ändert sich die Phase des rückgekoppelten Ausgangssignals. Beträgt diese Phasenverschiebung 180 °, so sind die Signale an den beiden differentiellen

C_{load} [pF]	$SR\,[\frac{\text{V}}{\mu\text{s}}]$	$t_s\,[\mu\text{s}]$	$PM\,[^\circ]$	GB [kHz]	A_0 [dB]	V_{in} [V]	$I_q\,[\mu\text{A}]$	$L \times B\,[\mu\text{m}^2]$
50	0.85	0.9	66	664	101	[0.2, 3.8]	153	183.5×90.7
100	0.89	2.1	63	"	"	"	"	"
150	0.87	2.7	55	"	"	"	"	"
200	0.74	5.4	47	"	"	"	"	"

Tabelle 3.1: Abmessungen und aus Simulationen gewonnene Kenndaten des entworfenen Miller-Operationsverstärkers. Zur Bestimmung von SR und t_s wurde die Antwort des als Spannungsfolger beschalteten Operationsverstärkers auf einen Spannungssprung von 1 V auf 3.5 V simuliert.

Eingängen des Verstärkers genau gegenphasig und die Differenz der Eingangsspannungen oszilliert mit maximaler Amplitude. Tritt dieser Fall bei einer Verstärkung größer als eins auf, fängt der Operationsverstärker an zu schwingen und ist, wenn die Lastkapazität nicht verringert werden kann, nutzlos. Als Daumenregel sollte die Phasenverschiebung zwischen Ein- und Ausgangssignal bei der Verstärkung eins deshalb 135 °, besser 120 ° nicht überschreiten ([Allen 90]. Für die als

$$PM = 180^\circ + (\text{Phase}(V_{out}) - \text{Phase}(V_{in})) \tag{3.27}$$

definierte Phasendifferenz sollte also $PM > 60^\circ$ gelten. Der entworfene Operationsverstärker wurde deshalb wie in Abb. 3.19 gezeigt simuliert.

Die sich ergebenden Verläufe von Verstärkung und Phasendifferenz in Abhängigkeit von der Frequenz sind für vier verschiedene Lastkapazitäten C_{load} in Abb. 3.20 gezeigt. Um die Phasendifferenz in Abhängigkeit der Lastkapazität zu erhalten, wurde die Frequenzabhängigkeit für 16 Werte von C_{load} simuliert. Das Ergebnis zeigt Abb. 3.21.

Nach oben genannter Daumenregel sollte der Operationsverstärker bei maximaler Rückkopplung also auf jeden Fall für Lastkapazitäten bis zu 100 pF, möglicherweise auch noch bei $C_{load} = 200$ pF funktionieren[12].

[12] Genaugenommen muß man sicherstellen, daß die Phasenreserve auch bei der Simulation mit den Prozeßparametern der *worst cases* ausreicht. Mit Hilfe der *worst case* Parameter sollen die vom Hersteller garantierten Obergrenzen der Prozeßparameterschwankungen abgedeckt werden.

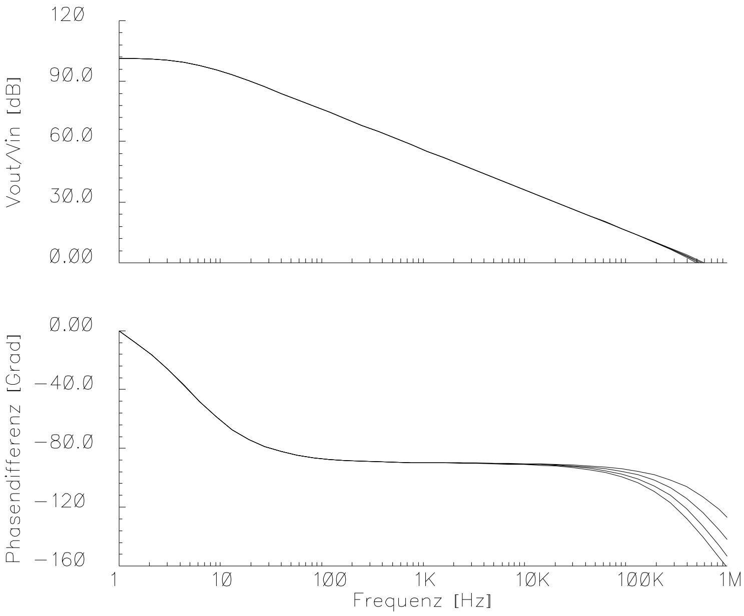

Abbildung 3.20: Verstärkungsfaktor und Phasendifferenz des Miller-Operationsverstärkers für vier verschiedene Lastkapazitäten C_{load}. Die Frequenz des Sinusgenerators aus Abb. 3.19 wurde zwischen 1 Hz und 1 MHz variiert. **Oben:**Verstärkungsfaktor $A = \frac{V_+}{V_{out}}$. **Unten:** Phasendifferenz zwischen V_+ und V_{out}.

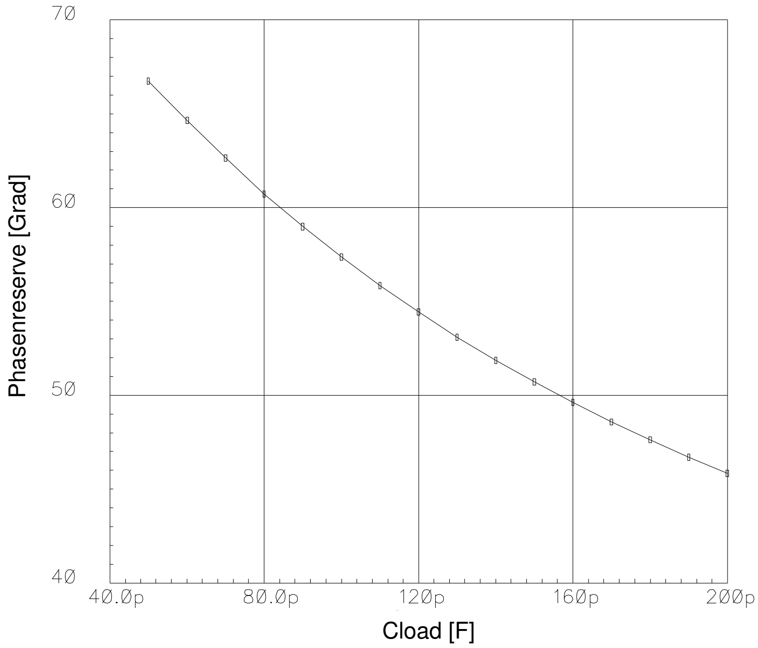

Abbildung 3.21: Phasenreserve PM des entworfenen Operationsverstärkers in Abhängigkeit von der Lastkapazität C_{load}.

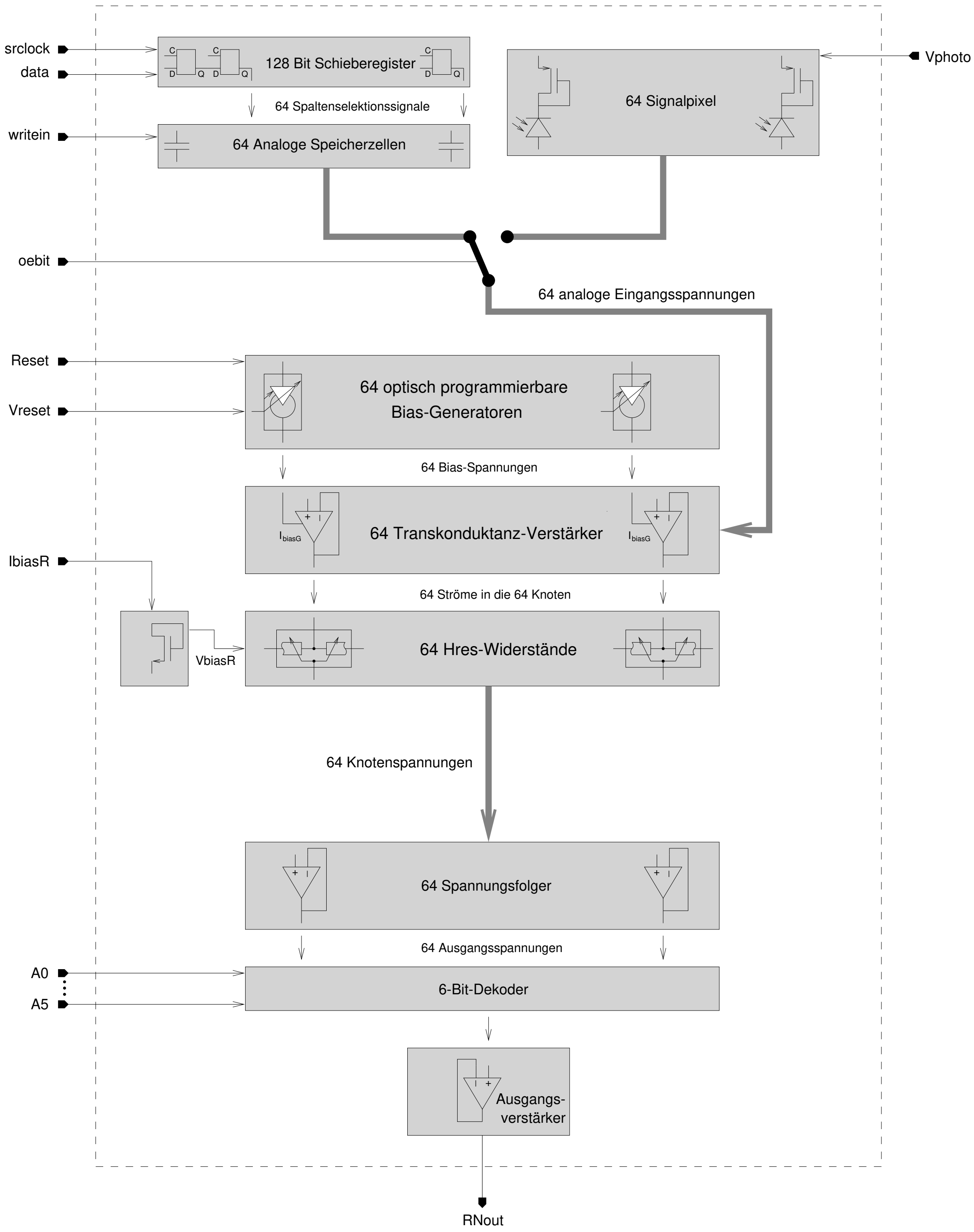

Abbildung 3.22: Blockschaltbild des implementierten Widerstandsnetzwerkes. Alle aus dem gestrichelten Rechteck herausgeführten Leitungen enden auf dem Chip auf Pads. Die Bias-Transistoren und Leitungen für die Spannungsfolger und den Ausgangsverstärker sind der Übersichtlichkeit halber nicht dargestellt. Die nach außen geführten Leitungen sind mit den im Design verwendeten Namen bezeichnet.

3.6 Blockschaltbild des gesamten Netzwerkes

Abb. 3.22 zeigt wie die in den vorangegangenen Kapiteln diskutierten Komponenten des Netzwerkes auf dem Chip zusammengefügt sind. Das Blockschaltbild gliedert sich in die drei Bestandteile *Signalerzeugung*, *optisch konfigurierbares Netzwerk* und *Ausgabe der Knotenspannungen*. Aus den beiden Möglichkeiten für die digitalen Kontrollspannungen *oebit* und *Reset* ergeben sich die vier in Tabelle 3.2 aufgeführten Betriebsmodi des Chips.

oebit [V]	Reset [V]	Erzeugung des	
		Eingangsmusters	G der Transkonduktanz-Verstärker
0	0/5	optisch	optisch, für jedes G einzeln
0	5	optisch	elektrisch, global durch V_{reset}
5	0/5	elektrisch	optisch, für jedes G einzeln
5	5	elektrisch	elektrisch, global durch V_{reset}

Tabelle 3.2: Betriebsmodi von *Oasys_RN*.

3.7 Positionsdetektor

Wie schon in Kapitel 1.3.2 angedeutet wurde, ist es für das Konzept der 'optischen Programmierung' unabdingbar, die Position des Laserspots auf dem Chip genau orten zu können. Mit Hilfe des Positionsdetektors soll es möglich sein den Laserstrahl in einer Regelschleife bis auf einen μm genau auf dem Chip zu positionieren. Die Größe der zu detektierende Lichtflecken beträgt dabei zwischen $5\,\mu$m und $50\,\mu$m. Technisch steht für diese Aufgabe der Vergleich der durch das Laserlicht verursachten Photoströme zur Verfügung. Als Maß für die Verschiebung des Lichtfleck-Schwerpunktes bietet sich die Differenz der Photoströme an. Abb. 3.23 zeigt das Prinzip der Positionsdetektion in einer und zwei mögliche Übertragungen auf zwei Dimensionen.

Die *Kleeblattlösung* oben rechts in der Abbildung liefert gleichzeitig Information über die vertikale und die horizontale Abweichung. Für größere Abweichungen kann allerdings der Fall eintreten, daß sich der Lichtfleck ganz in einem der Dreiecke befindet. In diesem Fall liegt nur für eine Richtung eine sinnvolle Information über die Abweichung der Lichtfleckposition vom Mittelpunkt des Kleeblattes vor. Allgemein muß ein geschlossener Regelkreis erst die Abweichung in einer (der *richtigen*!) Richtung vornehmen, und dann die nötige Korrektur in der anderen Dimension durchführen.

Der unten rechts in Abb. 3.23 dargestellte Ansatz benötigt zwei Taktzyklen um zwei verschiedene Differenzbildungen zwischen den Photoströmen durchzuführen, gibt aber für jede mögliche Position des Lichtflecks innerhalb der vier Quadrate die richtige Abweichung in beiden Richtungen an. Für *Oasys_RN* wurde die *Kleeblattlösung* gewählt, da diese etwas einfacher zu realisieren ist, und die Positionierung des Lichtflecks nur in einer Richtung mit Hilfe einer Regelschleife gesteuert werden muß.

Die Photodioden wurden als n^+-Substrat-Dioden realisiert. Diese Wahl ermöglicht den im Vergleich der drei möglichen Diodentypen geringsten durch die *design rules* erlaubten Abstand von $1.2\,\mu$m zwischen den benötigten Diffusionen, der bei kleinen Lichtfleckgrößen eine entscheidende Bedeutung für die maximal mögliche Auflösung besitzt. Außerdem besitzen die

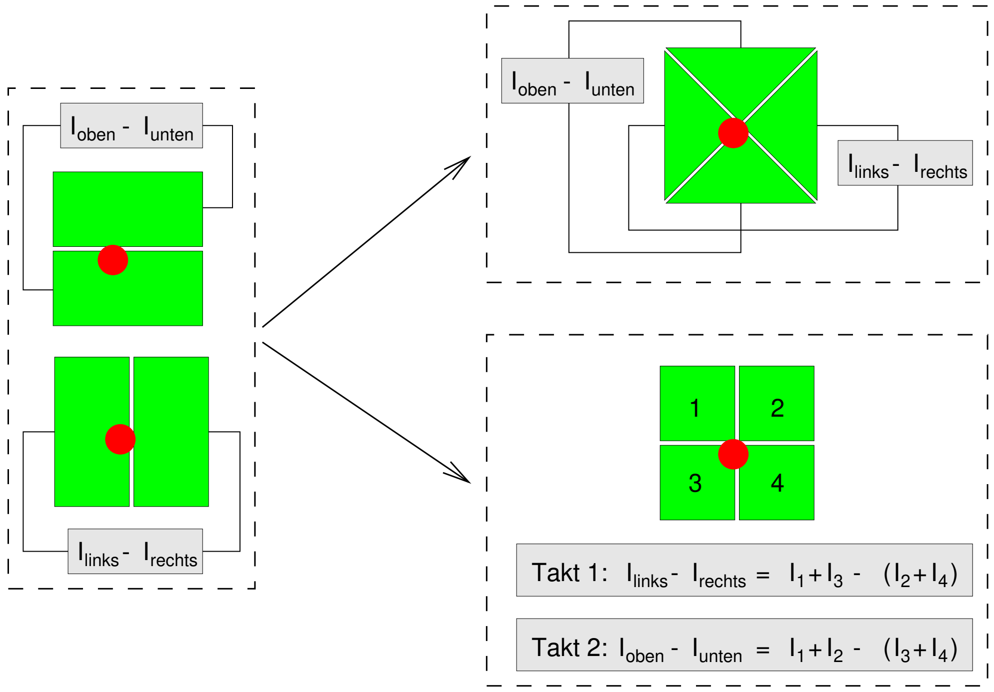

Abbildung 3.23: Schema zur Detektion eines Lichtflecks durch Subtraktion zweier Photoströme. Die Photodioden sind grün, der Lichtfleck rot dargestellt. **Links:** Lösung des Problems durch eindimensionale Differenzbildung zwischen zwei Photodioden, **Rechts:** Zwei mögliche Übertragungen in zwei Dimensionen.

n^+-Substrat-Dioden wie in Kapitel 2.1.1 diskutiert eine vergleichsweise hohe Quanteneffizienz und damit Lichtempfindlichkeit.

Um die Differenz der Ströme zu bilden und diese in eine zu dieser proportionalen Spannung umzuwandeln, wird hier das Prinzip der Ladungsverstärkung benutzt. Dazu werden ein invertierender und ein nichtinvertierender Ladungsverstärker so kombiniert, daß sich ein zur Differenz der Ströme proportionales Spannungssignal ergibt.

3.7.1 Invertierender Ladungsverstärker

Abb. 3.24 zeigt den Schaltplan eines invertierenden Ladungsverstärkers. Die Symbole $\phi1$ und $\phi2$ an den Schaltern symbolisieren die beiden Signale einer *nonoverlapping clock*, die in Abb. 3.25 gezeigt sind. Da die Clocksignale $\phi1$ und $\phi2$ nie gleichzeitig aktiv sind, sind entweder $S1$ und $S3$ geöffnet und $S2$ ist geschlossen, oder nur $S2$ ist offen. Wenn $\phi1$ aktiv ist, wird der Kondensator durch den Strom I_{photo} aufgeladen. Gleichzeitig werden die beiden Enden des Kondensators $C2$ kurzgeschlossen, so daß sich alle seine Ladungen neutralisieren können. Wird danach $S2$ geschlossen lädt der Operationsverstärker den Kondensator $C2$ so lange auf, bis sein negativer Eingang die Spannung V_{ref} erreicht hat. Dazu ist genau die Ladungmenge nötig, die dem Kondensator $C1$ während der aktiven Phase von $\phi1$ durch I_{photo} entzogen worden ist. Für die Ausgangsspannung V_{out} nach dem oben beschriebenen

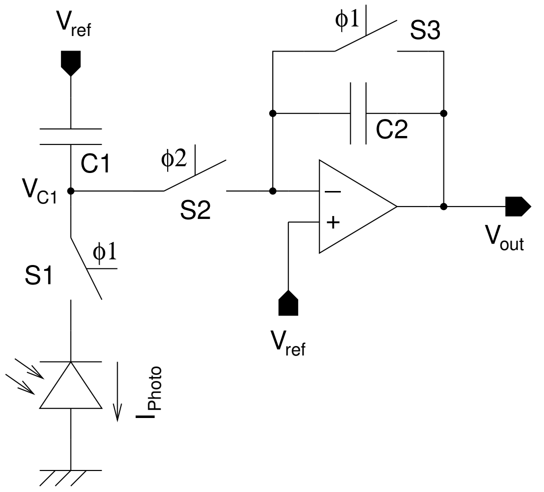

Abbildung 3.24: Schaltplan des invertierenden Ladunsverstärkers. $\phi 1$ und $\phi 2$ bezeichnen die Ausgänge der *nonoverlapping clock* aus Abb. 3.25.

Ladungsaustausch in der aktiven Phase von $\phi 2$ gilt also:

$$V_{out} = V_{ref} + (V_{ref} - V_{C1})\frac{C1}{C2} = V_{ref} + \frac{I_{photo}\Delta t}{C2} \tag{3.28}$$

Dabei bezeichnet Δt die Zeit, für die $S1$ geschlossen wird, die Zeit also, in der die Clock $\phi 1$ logisch eins ist.

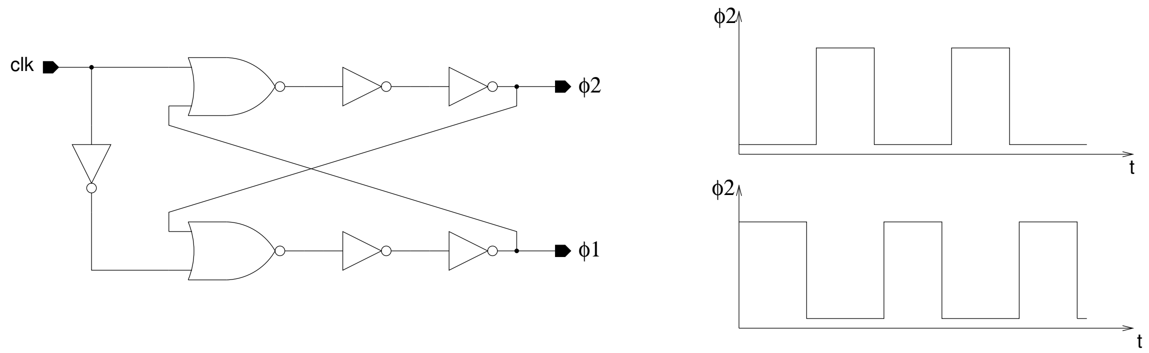

Abbildung 3.25: Links: Schaltplan der verwendeten *nonoverlapping clock.* **Rechts:** Schema der erzeugten Clock-Signale $\phi 1$ und $\phi 2$.

Die für die Erzeugung der Clock-Signale $\phi 1$ und $\phi 2$ benutzte Schaltung ist in Abb. 3.25 gezeigt. Das Prinzip beruht darauf, daß die logischen Signale eine endliche Zeit benötigen, um die NOR-Gatter und die Inverter zu durchlaufen, da die Gate-Kapazitäten der Transistoren darin mit einem endlichen Strom aufgeladen werden müssen. Beim Design der Schaltung müssen die Transistoren so dimensioniert werden, daß sich $\phi 1$ und $\phi 2$ auch bei der größten zu erwartenden Lastkapazität schneller ändern, als das Signal die beiden NOR-Gatter durchlaufen kann. Ist das nicht der Fall, so sind $\phi 1$ und $\phi 2$ für eine endliche Zeit verschieden von null. Für den Ladungsverstärker aus Abb. 3.24 bedeutete dies, daß nicht immer wenigstens einer der Schalter $S1$ und $S2$ geschlossen wäre. Die auf dem Kondensator $C1$ gesammelte Ladung würde dann neutralisiert werden, ohne daß sich in aktiven Phase von $\phi 2$ die gewünschte Ausgangsspannung einstellte.

3.7.2 Nichtinvertierender Ladungsverstärker

Der vorgestellte Ladungsverstärker integriert den Strom zu einer Ladung auf und transformiert diese dann in eine Spannung. Da schließlich die Differenz zweier Ströme zu bilden ist, liegt es nahe die aus dem einen Strom erhaltene Ladung *umzupolen* und dann zu der aus dem anderen Strom erhaltenen zu addieren. Was mit *Umpolen* gemeint ist zeigt Abb. 3.26. Während $\phi 1$ aktiv ist, wird wie beim invertierenden Ladungsverstärker der Kondensator $C1$ geladen und werden die beiden Anschlüsse von $C2$ kurzgeschlossen. In der folgenden $\phi 2$-aktiven Phase wird jetzt der negativere Anschluß von $C1$ auf das Potential V_{ref} gehoben, so daß die Spannung V_{out} sich verringern muß, um den negativen Eingang des Operationsverstärkers wieder auf die Referenzspannung zurückzuholen. Für V_{out} gilt also:

$$V_{out} = V_{ref} + (V_{C1} - V_{ref})\frac{C1}{C2} = V_{ref} - \frac{I_{photo}}{2 f_{clk} C2} \tag{3.29}$$

Hier wurde Δt durch $2\frac{1}{f_{clk}}$ ausgedrückt, wobei f_{clk} die Frequenz, mit der die *nonoverlapping clock* getaktet wird, bezeichnet.

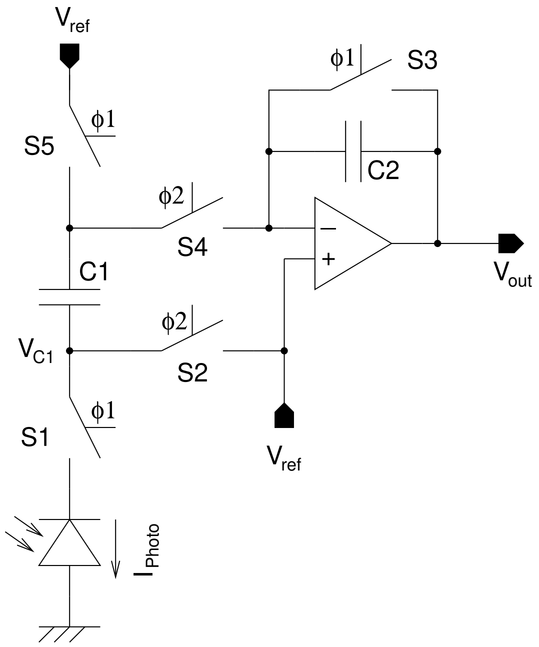

Abbildung 3.26: Schaltplan des nichtinvertierenden Ladungsverstärkers .

3.7.3 Differentieller Ladungsverstärker

Um die benötigte Differenz zwischen zwei Strömen, hier I_{links} und I_{rechts} genannt, zu bilden, werden der invertierende und der nichtinvertierende Ladungsverstärker wie in Abb. 3.27 dargestellt zusammengefügt. Die nach dem Ladungsausgleich in der $\phi 2$-aktiven Phase resultierende Spannung am Ausgang des ersten Operationsverstärkers ergibt sich durch Addition von (3.28) und (3.29) zu:

$$V_{out} = V_{ref} + \frac{(I_{rechts} - I_{links})}{2 f_{clk} C} \tag{3.30}$$

Dabei wurde C anstatt $C2$ verwendet, da in Abb. 3.27 wie in der tatsächlichen verwendeten Schaltung alle Kondensatoren die Kapazität C besitzen.

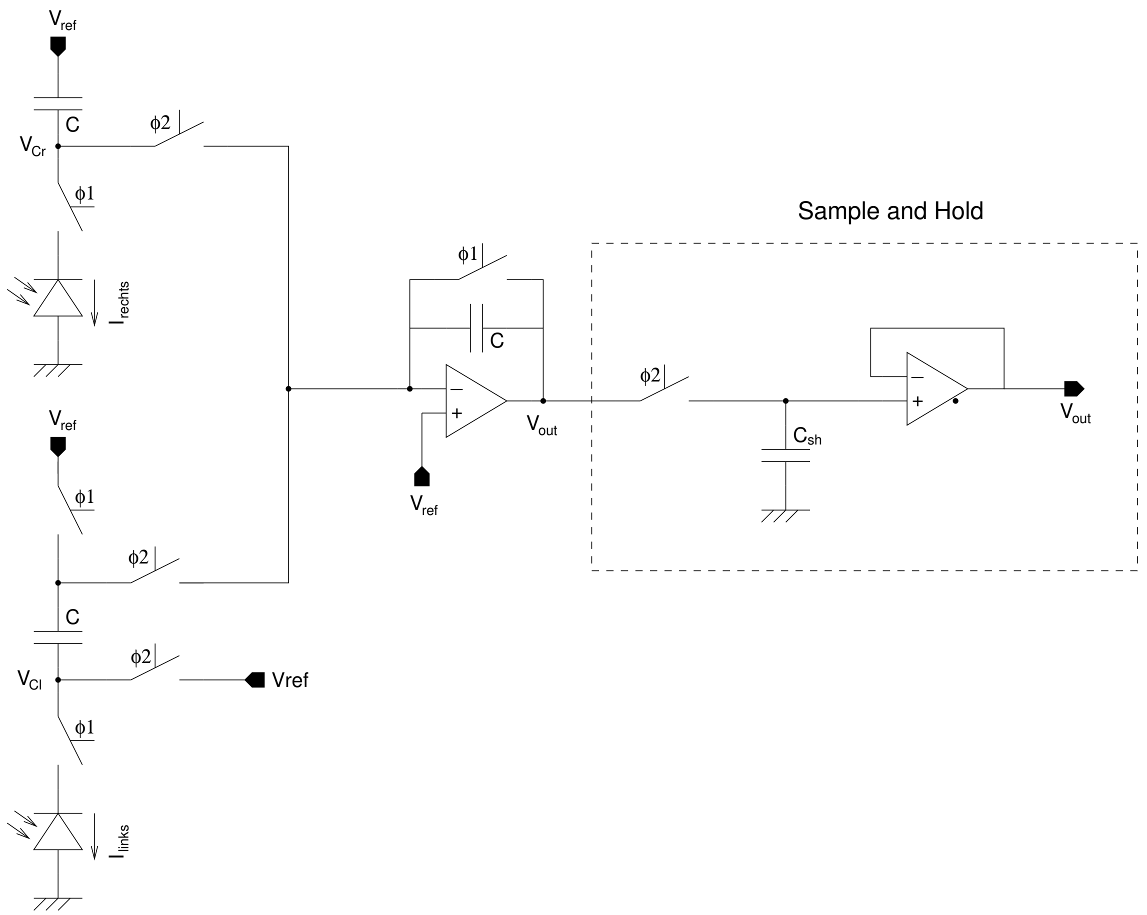

Abbildung 3.27: Schaltbild des *differentiellen Ladungsverstärkers* Die Kondensatoren der Ladungsverstärker haben, entsprechend der tatsächlichen Implementierung auf *Oasys_RN*, alle die gleiche Kapazität C.

Damit eine mögliche Auslese der analogen Ausgangsspannung nicht auf die $\phi 2$-aktive Phase beschränkt werden muß, folgt dem ersten Operationsverstärker noch eine *sample and hold* Schaltung, die das Signal der letzten $\phi 2$-aktiven Phase auf dem Kondensator C_{sh} während der $\phi 1$-aktiven Phase zwischenspeichert. Da der Operationsverstärker des Ladungsverstärkers am Beginn der $\phi 2$-aktiven Phase eine endliche Zeit benötigt, um die nötige Ausgangsspannung einzustellen, kann V_{out} in dieser Zeit nicht ausgelesen werden.

3.7.4 Implementiertes Positionsdetektionssystem

In Kapitel 1.3.2 wurde schon angedeutet, daß neben der Position des Laserspots auf dem Chip auch dessen Größe abgeschätzt werden können soll. Dazu ist es sinnvoll, die Ströme der gegenüberliegenden Photodioden auch *addieren* zu können. Um das zu erreichen, kann man analog zu Abb. 3.27 zwei invertierende Ladungsverstärker so zusammenfügen, daß sich die zu Ladungen integrierten Ströme addieren. Die Schaltung aus Abb. 3.27 wurde deswegen für den Positionsdetektor mit einigen Schaltern so erweitert, daß man mit einer digitalen Kontrollspannung (im Design *miplubit* genannt) zwischen Addition und Subtraktion der Photoströme wählen kann.

Für den Positionsdetektor wurde dann jeweils eine dieser Schaltungen für die vertikale und eine für die horizontale Richtung zusammengefügt. Auf dem Chip wird das Widerstandsnetzwerk links und rechts von einem solchen Positionsdetektor dergestalt flankiert, daß die Mittelpunkte der Steuerphotodioden und diejenigen der kleeblattfömigen Anordnung der Positionsdetektorphotodioden auf einer Geraden zu liegen kommen.

Für die benötigten Operationsverstärker wurden der in Kapitel 3.5 vorgestellte Miller-Operationsverstärker verwendet. Aus den in Tabelle 3.1 angegebenen Werten für Anstiegsgeschwindigkeit und *settling time* folgt somit eine maximal zu benutztende Clock-Frequenz von einigen 100 kHz. Außerdem ergibt sich aus dem Ein- bzw. Ausgangsspannungsbereich des Operationsverstärkers die ideale Referenzspannung V_{ref} zu 2 V.

Der Abstand der Katheten der dreieckigen Photodioden entspricht mit 1.2 μm dem minimalen von den *design rules* geforderten Abstand zweier n^+-Diffusionen. Alle in Abb. 3.27 mit C bezeichneten Kondensatoren haben eine Kapazität von 1.02 pF. Sowohl die Photodioden der Positionsdetektoren als auch die Kontrolldioden sind als n^+-Substrat-Dioden implementiert, besitzen also die gleiche Quanteneffizienz. Um die in Frage kommenden Clock-Frequenzen abschätzen zu können, muß man die für eine Spannungsänderung ΔV nötigen Integrationszeiten Δt_{ph} und Δt_{kon} für den Positionsdetektor und die Spannung V_C der Biasschaltung aus Abb. 3.11 vergleichen. Mit Gl. (3.22) ergibt sich für dieses Verhältnis:

$$\frac{\Delta t_{ph}}{\Delta t_{kon}} = \frac{A_{kon} C_{ph}}{A_{ph} C_{kon}} = 1.35 \frac{A_{kon}}{A_{ph}} \tag{3.31}$$

Unter A_{ph} ist dabei der Anteil der Fläche zweier gemeinsam auszulesener Photodioden des Positionsdetektors, der durch den Laser beleuchtet wird, zu verstehen. Die nötige Integrationszeit ist also sowohl von der Lichtfleckgröße, als auch vom Strahlprofil abhängig. Mit der Abschätzung von Δt_{kon} aus Kapitel 3.2 und einem zu erwartenden Strahldurchmesser von 10 bis 50 μm sollte die nötige Clockfrequenz etwa zwischen 100 Hz und 100 kHz betragen.

3.7.5 Layout

Abb. 3.28 zeigt einen Ausschnitt aus dem Layout von *Oasys_RN*. Auf der rechten Seite ist der Positionsdetektor abgebildet, auf der linken die ersten 10 Zellen des Widerstandsnetzwerkes. Über dem Positionsdetektor befindet sich der Ausgangsverstärker für die Knotenspannungen des Netzwerkes. Die Abmessungen des Positionsdetektors betragen 211 × 780 μm. Am unteren Ende sind die vier kleeblattförmig angeordneten Photodioden (grün) zu sehen. Darüber befinden sich die 6 Kondensatoren des Ladungsverstärkers (rot) gefolgt von den 2 dazugehörigen Operationsverstärkern. Über diesen sind die zwei *sample and hold* Kondensatoren und die *nonoverlapping clock* angeordnet. Ganz oben sind die Operationsverstärker der *sample and hold* Stufe zu sehen.

3.8 Layout von *Oasys_RN*

In Abb. 3.29 ist das gesamte Layout von *Oasys_RN* dargestellt. Auf dem Chip entsprechen die abgebildeten Strukturen etwa einer Fläche von 4.05×1.28 mm. Den größten Teil der Fläche nimmt das Widerstandsnetzwerk in der Mitte zwischen den beiden Positionsdetektoren ein. Der rechte und der untere Rand sind mit Bondpads belegt. Ganz oben ist eine separate

Ausgabe des entworfenen Operationsverstärkers plaziert, die einen genauen elektrischen Test des Verstärkers ermöglicht. Deutlich zu erkennenn ist die strichförmige Anordnung der 64 Photodioden der Signalpixel. Die Kontrollphotodioden liegen auf der Geraden durch die Mittelpunkte der Detektionsflächen der Positionsdetektoren, sind in dieser Auflösung aber nicht mehr erkennbar.

Die obersten 10 Zellen des Widerstandsnetzwerkes sind in Abb. 3.28 (gegenüber Abb. 3.29 um 90° gedreht) abgebildet. Hier sind die Kontrolldioden und die dazugehörigen Kondensatoren aus Abb. 3.11 besser zu sehen. Zwischen diesen Kondensatoren und denen für die Speicherung der elektrischen Eingangssignale sind die analogen Bausteine des Widerstandsnetzwerkes untergebracht. Abgesehen von den durch Photodioden belegten Flächen ist der ganze Chip mit Metall 2 abgedeckt, um die in Kapitel 2 erwähnten Störeinflüsse einfallenden Lichts so gering wie möglich zu halten. Aus dem gleichen Grund ist der zu der Kontrolldiode gehörende Kondensator zwischen der Kontrolldiode, die mit dem intensiven Laserlicht bestrahlt werden soll, und der analogen Elektronik plaziert.

Abb. 3.30 zeigt ein mit Hilfe eines Mikroskops aufgenommenes Photo des Chips. Der gezeigte Bildausschnitt entspricht etwa dem in Abb. 3.29 gezeigten Layout. Obwohl fast der ganze Chip mit einer Metallage überzogen ist, sind die groben Strukturen des Chips sichtbar. Am rechten und am unteren Rand des Bildes sind die Enden der etwa 25 μm dicken Bonddrähte zu erkennen, mit denen der Chip an die Kontakte des sogenannten Chip-Sockels angeschlossen wird.

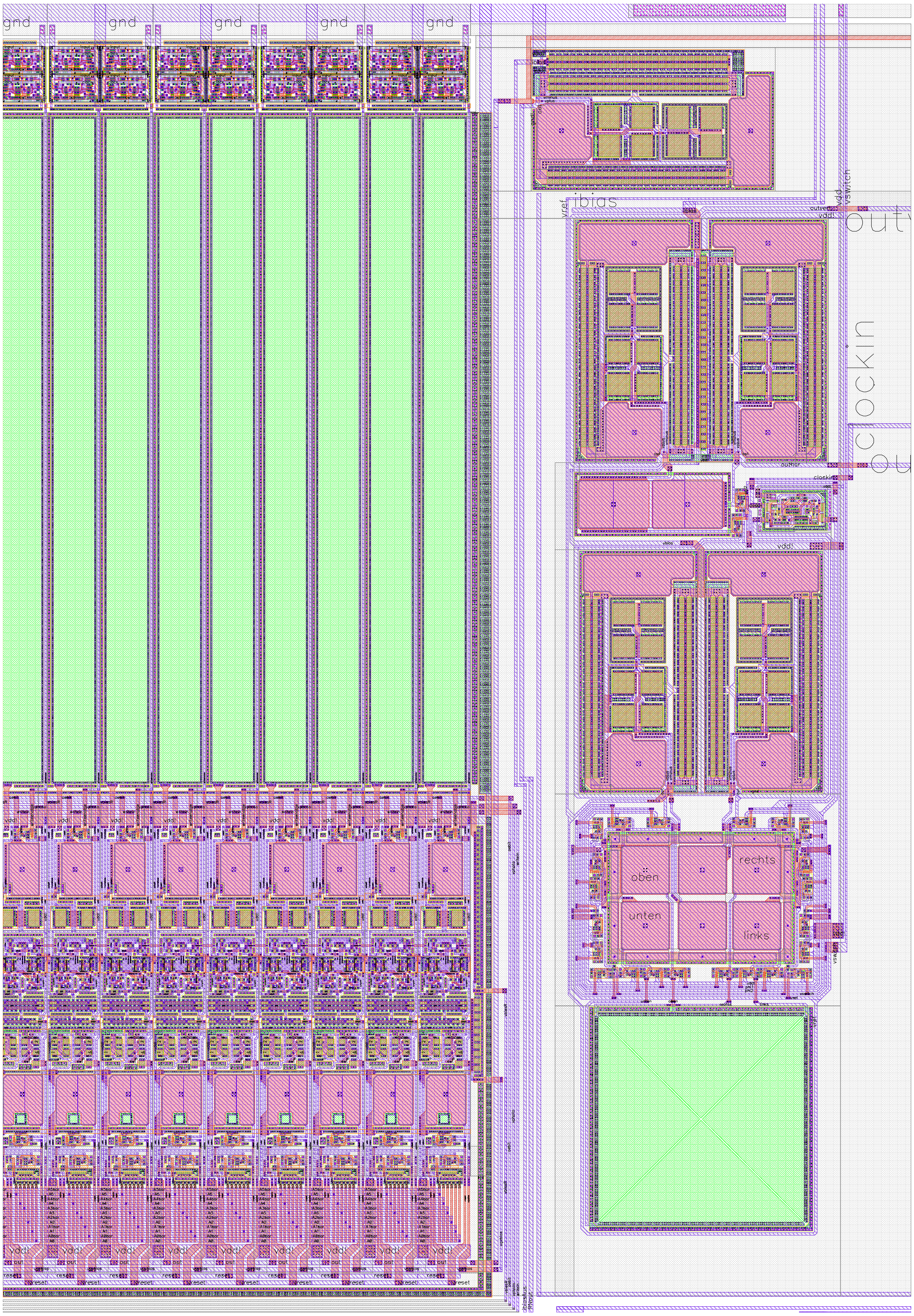

Abbildung 3.28: Layout des Positionsdetektors

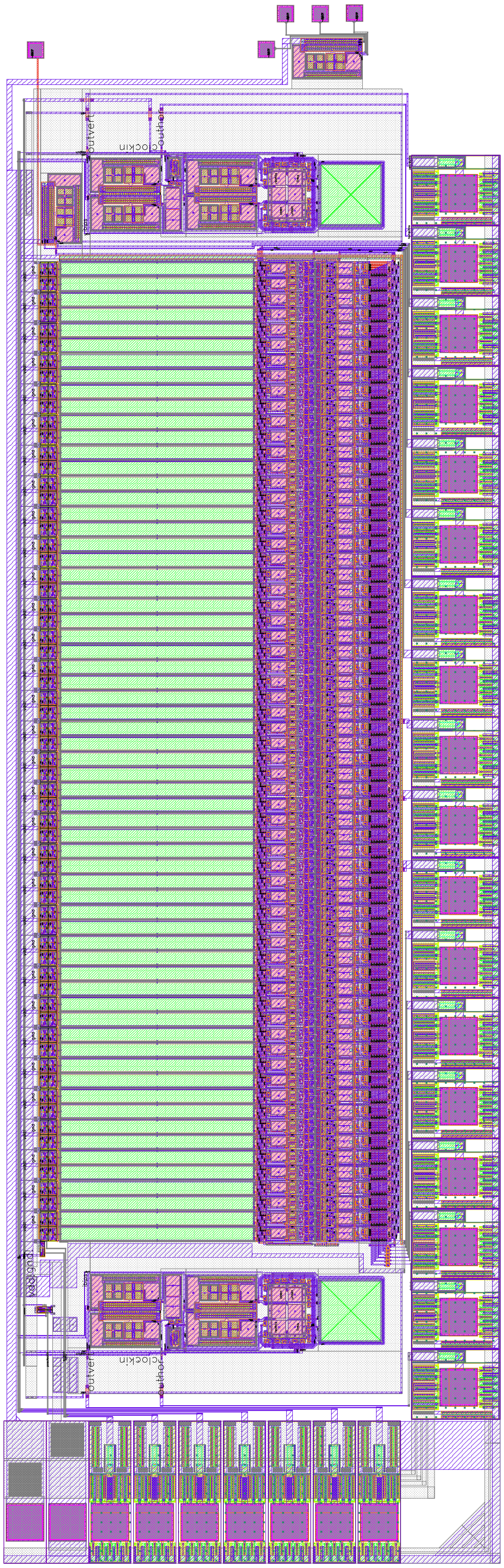

Abbildung 3.29: Layout von *Oasys_RN*.

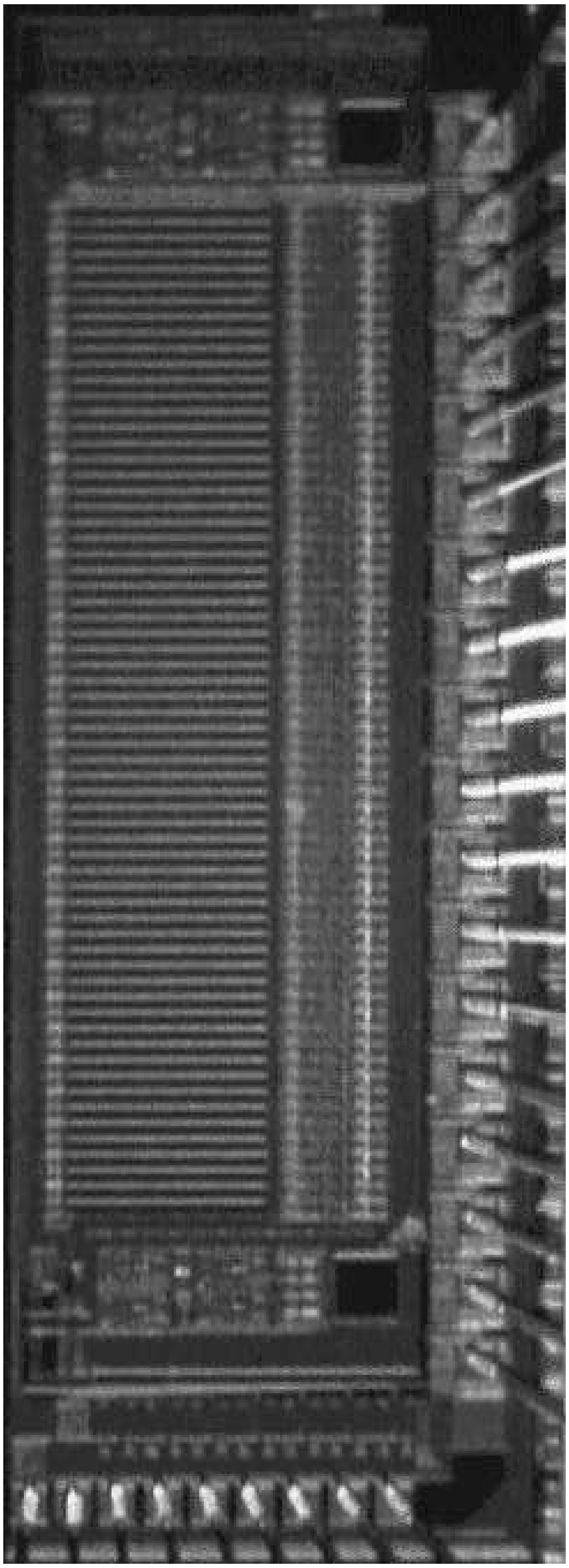

Abbildung 3.30: Mikroskopaufnahme von *Oasys_RN*.

Kapitel 4

Elektrische Messungen an *Oasys_RN*

In Kapitel 3.5 wurden die verschiedenen Betriebsmodi von *Oasys_RN* vorgestellt. Im Rahmen dieser Diplomarbeit wurden Messungen in den zwei Betriebsmodi durchgeführt, in denen das Eingangsmuster elektrisch erzeugt wird. Der elektrische Aufbau, und die Messungen mit elektrischer Einstellung der Leitwerte G der Transkonduktanz-Verstärker des Netzwerkes werden in diesem Kapitel diskutiert. Optischer Aufbau, Test der Positionsdetektoren und Messungen bei optischer Einstellung der Leitwerte der Transkonduktanz-Verstärker sind Thema des nächsten Kapitels.

4.1 Testaufbau

Der Aufbau für den Test von *Oasys_RN* muß folgende Aufgaben erfüllen:

- Elektrische Erzeugung des Eingangsmusters für das Netzwerk.
- Sequentielle Auslese der Knotenspannungen des Netzwerkes.
- Einstellung der analogen und digitalen Parameter.
- Erzeugung der notwendigen Bias-Ströme und Spannungen.
- Steuerung des Laser-Scanners.
- Auslese der Positionsdetektoren.

Zur Steuerung des Systems wurde ein mit einer Meßkarte ausgestatteter PC benutzt. Die Meßkarte kann über den PC programmiert werden und enthält 16 Analog zu Digital Konverter, 2 Digital zu Analog Wandler, 8 digitale Ein-/Ausgänge, sowie eine Reihe von Timing-IO's, die die Trigger-Signale der ADC's und DAC's ausgeben, bzw. mit Hilfe derer diese getriggert werden können. Die Auflösung der DAC's und ADC's beträgt jeweils 12 Bit. Die Verwendung des Systems PC/Meßkarte ermöglicht eine kompakte und flexible Steuerung des Systems und eine bequeme Erfassung der Daten. Die *settling time*[1] der DAC's und die Konversionszeit der ADC's beträgt maximal $3\,\mu$s, das System ist also für die Messung und Steuerung des entworfenen Chips ausreichend schnell.

[1]Damit ist hier die Zeit gemeint, die der DAC benötigt um eine beliebige Spannungsänderung innerhalb seines Ausgangsspannungsbereiches bis auf 0.5 LSB genau zu vollziehen.

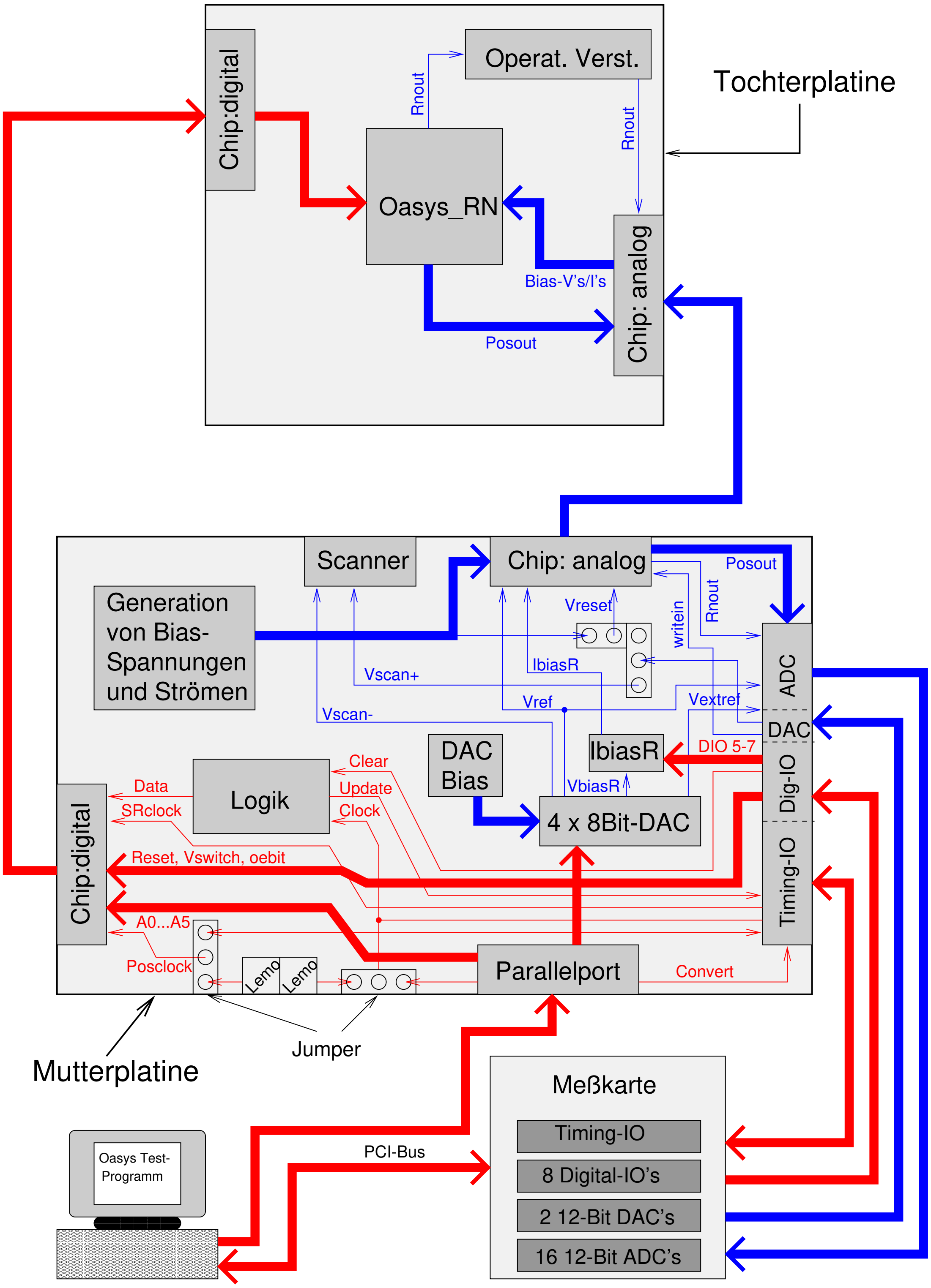

Abbildung 4.1: Elektrischer Testaufbau. Der PC steuert die Messung über die Meßkarte und den Parallelport. Die Mutterplatine erzeugt einige Bias-Spannungen und Ströme und sorgt für die Verteilung der Signale zwischen Nidaq-Karte, Computer, Laser-Scanner, externen Taktsignalen und der Tochterplatine, die den Chip beherbergt. Die dick gezeichneten Pfeile fassen jeweils mehrere Signale zusammen. Die Farbe der Pfeile kodiert die Art der Signale: Rot steht für digitale und blau für analoge Signale.

4.1.1 Elektrischer Testaufbau

Mit der Meßkarte können nicht alle erforderlichen Spannungen, Signale und Ströme erzeugt werden. Außerdem sollen die Ausgangssignale des Chips erst verstärkt werden, bevor sie von der Meßkarte digitalisiert werden. Schließlich müssen die Signale zwischen den Systemkomponenten verteilt werden. Deshalb wurden zwei Testplatinen aufgebaut, eine Tochterplatine, die den Chip und die Verstärker enthält, und eine Mutterplatine, auf der die zur Beschaltung des Chips notwendige Elektronik untergebracht ist. Die Aufteilung der Elektronik auf zwei Platinen hat folgenden Grund: Damit der Chip bei den optischen Messungen ausreichend genau positioniert werden kann, muß er zusammen mit der Platine, auf der er eingesetzt ist, an dem Fahrtisch befestigt werden. Da dieser sich bewegen kann, sollte die Platine möglichst klein und die Anzahl der Kabel möglichst gering sein, damit die entstehenden Zugkräfte die Positionierung nicht unnötig erschweren. Ferner ist die an dem Fahrtisch angebrachte Platine sehr schwer zugänglich, so daß eine Veränderung einer der Biasspannungen oder Oszilloskop-Messungen einen erheblichen Aufwand darstellten und mit großer Wahrscheinlichkeit eine Dejustierung der Optik zur Folge hätten. Der elektrische Testaufbau ist in Abb. 4.1 dargestellt. Die Anordnung der Bauteile und Anschlüsse auf den in der Abbildung schematisch dargestellten Platinen ist dem realen Aufbau nachempfunden. Um kapazitives oder induktives Übersprechen der Digitalsignale (rote Leitungen) auf die analogen Signale (blaue Leitungen) zu vermeiden, wurden analoger und digitaler Teil der Elektronik räumlich so weit wie möglich getrennt. Die PC-Meßkarte wird in den PC eingesteckt und kommuniziert mit diesem über den PCI-Bus.

Die parallele Schnittstelle des Computers wird im SPP[2]-Mode betrieben. Die acht Datenleitungen sowie einige Steuerleitungen können so direkt durch Beschreiben entsprechender Register gesteuert werden. Da die parallele Schnittstelle im SPP-Modus *open collector*-Ausgänge besitzt, benötigt man, wie in Abb. 4.2 gezeigt, für jede Ausgangsleitung einen sogenannten Pull-up-Widerstand. Sechs der insgesamt elf benutzten Ausgänge der parallelen Schnittstelle

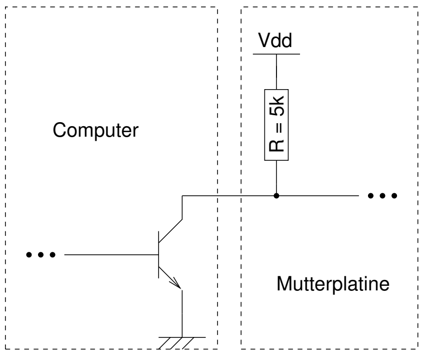

Abbildung 4.2: Ausgang des Parallel-Ports mit *pull up Widerstand* auf der Testplatine.

dienen der Kodierung der Adresse des auszulesenden Knotens über die Leitungen A_0 bis A_5. Mit der mit *Convert* bezeichneten Leitung kann über die parallele Schnittstelle die Konversion der selektierten Knotenspannung *RNout* ausgelöst werden. Auf diese Weise kann das Timing beim Auslesen des Netzwerkes durch die Software kontrolliert werden.

Die Programmierung des externen 8-Bit-DAC's erfolgt über drei weitere Ausgänge der parallelen Schnittstelle. Mit diesem können die vier Spannungen V_{biasR}, V_{ref}, V_{extref} und V_{scan-}

[2]SPP steht für **S**tandard **P**arallel **P**ort.

programmiert werden. Für jede Ausgangsspannung benötigt der DAC eine Referenzspannung. Die Referenzspannungen für V_{biasR}, V_{ref} und V_{scan-} werden durch Spannungsteiler, in Abb. 4.1 zusammengefaßt zu *DAC Bias*, erzeugt. Für V_{extref} wird V_{ref} als Referenz benutzt.

Die analoge Eingangsspannung *writein*, mit der das Eingangsmuster bei elektrischer Erzeugung sequentiell in den Chip geschrieben wird, wird mit einem der beiden 12-Bit-DAC's der Meßkarte erzeugt. Die nötige Synchronisation des auszugebenden DAC-Wertes mit der Schieberegister-Clock und dem Data-Bit für das Schieberegister wird folgendermaßen erreicht: In dem mit *Logik* bezeichneten Baustein wird einerseits das *Update*-Signal für den DAC erzeugt, in dem das Clock-Signal halbiert wird. Andererseits wird das Data-Bit für das Schieberegister erzeugt, indem nur jeder 128'te Clock-Puls weitergegeben wird. Mit Hilfe einer der beiden Zähler der Meßkarte wird aus dem Clock-Signal die zeitlich versetzte Schieberegister-Clock erzeugt. Dadurch wird sichergestellt, daß *Data* während der steigenden Flanke von *SRclock* aktiv ist, also genau ein Bit in das Schieberegister geschrieben wird. Das resultierende Timing ist in Abb. 4.3 dargestellt. Die Ausgangsspannung des DAC's ist für das

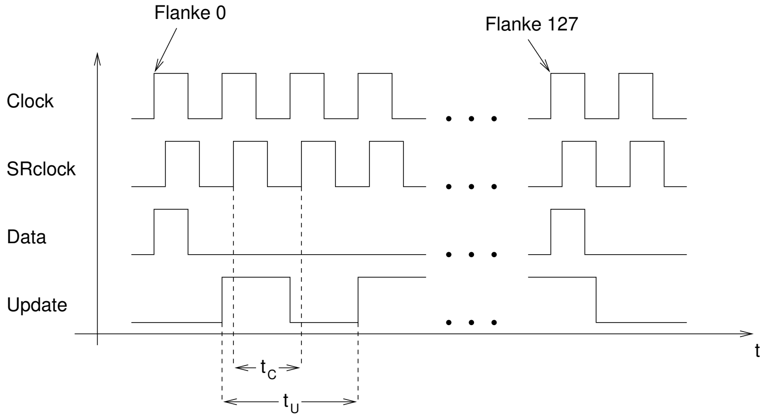

Abbildung 4.3: Timing für die Erzeugung des Eingangsspannungsmuster. t_C bezeichnet die Zeit, in der *writein* mit dem Kondensator des jeweiligen Knotens (hier 0) verbunden ist. t_U ist die Zeit zwischen einer Änderung der Ausgangsspannung des DAC's.

gewählte Timing in der ganzen Zeit konstant, in der sie an dem zu beschreibenden Kondensator anliegt, so daß das Eingangsmuster korrekt in den Chip eingelesen wird.

Durch geeignetes Plazieren eines Jumpers kann zwischen einem Ausgang der parallelen Schnittstelle und einer externen Clock als *Clock*-Signal gewählt werden. Anfänglich wurde das Eingangsmuster während der Messung kontinuierlich aufgefrischt. Die von diesen digitalen Signalen verursachten Störungen der Analogsignale waren aber deutlich größer als diejenigen, die durch den Abfall der Kondensatorspannungen in der Zeit zwischen Beschreiben der Kondensatoren und dem Auslesen der Knotenspannungen des Netzwerkes zu erwarten sind. Deswegen wurde bei allen dargestellten Messungen die Programmierung der Eingangsspannungen und die Auslese der Knotenspannungen nacheinander ausgeführt.

Die zur Erzeugung des Bias-Stromes I_{biasR} verwendete Anordnung ist in Abb. 4.4 dargestellt. Der erzeugte Strom hängt sowohl von der Spannung V_{biasR} als auch von dem mit Hilfe der Digitalausgänge *DIO5* bis *DIO7* gewählten Widerstand ab. Der Strom I_{bias} durch die Bias-Schaltung aus Abb. 3.6 sollte sich so etwa zwischen 20 nA und 200 μA variieren lassen. Der genaue Wert für I_{biasR} ergibt sich aus dem Spannungsabfall über dem gewählten Wider-

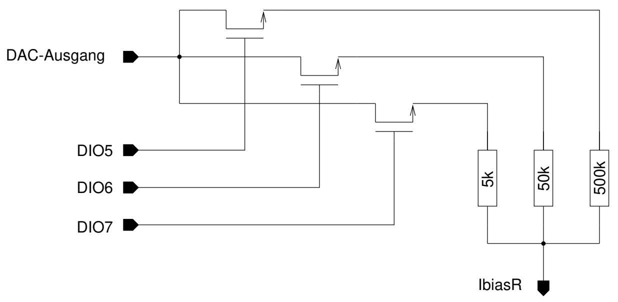

Abbildung 4.4: Erzeugung des Bias-Stromes I_{biasR}.

stand, der mit Hilfe zweier ADC's der Meßkarte differentiell gemessen wird (nicht in Abb. 4.1 dargestellt).

Die Spannung V_{reset} für die elektrische Einstellung der Leitwerte der Transkonduktanz-Verstärker des Netzwerkes kann wahlweise durch den zweiten DAC der Meßkarte (elektrische Einstellung der R's) oder durch ein Potentiometer(optische Programmierung der R-Werte) eingestellt werden. Für die optische Programmierung des Netzwerkes wird der zweite 12-Bit-DAC zur Steuerung des Galvanometer-Scanners benutzt. Da dessen Auslenkung proportional zur Spannungsdifferenz zwischen V_{scan+} und V_{scan-} ist, kann die Auslenkung des Scanners mit Hilfe des 8-Bit-DAC's grob voreingestellt werden, so daß der ganze Dynamikbereich des 12-Bit-DAC's zur Positionierung des Strahls auf dem Chip benutz werden kann. Die Spannung V_{ref} hat zwei Funktionen: Einerseits entspricht sie der Referenzspannung für die Ladungsverstärker der Positionsdetektoren, andererseits dient sie als Bezugspunkt für die Messung von *RNout*, der vier unter *Posout* zusammengefaßten Ausgangssignale der Positionsdetektoren und der externen Referenzspannung für die 12-Bit-DAC's der Meßkarte V_{extref}. Letztere wird von dem externen 8-Bit-DAC erzeugt. Zur Einstellung der absolut zu erzeugenden Ausgangsspannung der Meßkarten-DAC's wird V_{extref} gegen V_{ref} gemessen und der um 2 V erhöhte Wert dieser Spannung an die *Nidaq*-Software übergeben. V_{ref} muß daher möglichst genau auf 2 V eingestellt werden.

Alle Digitalleitungen, die die beiden Platinen verbinden, sind an beiden Enden mit Tiefpaßfiltern abgeschlossen um Reflexionen durch Impedanzsprünge an den Kabelenden zu vermeiden. Für R wurden 100 Ω und für C 20 pF gewählt. Alle Analogspannungen sind mit einem Kondensator gegen Störsignale aus der Umgebung geblockt. Die Mutterplatine besitzt zwei getrennte Versorgungsspannungsnetze: Eins für die digitalen Signale und das andere für die analogen. Auf diese Weise sollen Störungen durch Übersprechen der digitalen Signale auf die analogen über die Versorgungsleitungen vermieden werden.

Ein größer Nachteil der gewählten Meßumgebung besteht darin, daß das für den PC benutzte Betriebssystem *Windows 95* nicht echtzeitfähig arbeiten kann. Die Meßkarte umgeht dieses Problem, indem auszugebende DAC-Werte und von den ADC's konvertierte Spannungsdaten in der Karte gepuffert werden. Die digitalen Ausgänge der Meßkarte können dagegen nur durch direkten Softwareaufruf programmiert werden und erreichen dadurch höchstens eine Ausgabefrequenz von 5 kHz. Mit der parallelen Schnittstelle, mit der das Ein- und Auslesen der Daten getaktet wird, lassen sich zwar um zwei Größenordnungen höhere Frequenzen erreichen, jeglicher Echtzeitbezug geht aber spätestens beim ersten *Interrupt* verloren. Diese treten mehr oder weniger zufällig auf (vom Standpunkt des Meßprogramms) und unterbrechen den Fortgang des Programms für die Größenordnung von einer ms. Sowohl die Zeit für das Beschreiben der Analogspeicher mit dem Eingangsmuster, als auch diejenige für das

Auslesen der Knotenspannungen beträgt einige Millisekunden, variiert aber abhängig von der Anzahl und Dauer der vom Computer während dieser Vorgänge bearbeiteten *Interrupts*. Die Eingangsspannungen sollten jedoch während dieser Zeit wenigstens bis auf ein oder zwei mV konstant bleiben (vgl. Kapitel 2.3.2).

4.1.2 Testsoftware

Für die Steuerung der Messungen wurde ein C-Programm[3] geschrieben. Das Programm läuft unter *Windows 95* und benutzt ein DOS-Fenster zur Interaktion mit dem Benutzer. Im einzelnen erledigt das Programm folgende Aufgaben:

- Eingabe, Verwaltung und Speicherung der Meßparameter.
- Steuerung folgender Meßabläufe:
  - Beschreiben der Analogspeicher.
  - Auslese der Knotenspannungen.
  - Durchführung eines gesamten Scanvorganges.
- Messung und Korrektur der Offsets.
- Speicherung, Formatierung und einfache Bearbeitung der gemessenen Daten.
- Automatische Positionierung des Laserspots in x-Richtung.
- Steuerung der Laserintensität über einen externen AWG[4].

Die für die jeweiligen Messungen relevanten Programmteile werden im Folgenden zusammen mit den Messungen selbst diskutiert.

4.2 Ergebnisse der Elektrischen Messungen

Da der Aufwand für optische Messungen deutlich größer ist, als derjenige für Messungen im rein elektrischen Betriebsmodus, wurde das Widerstandsnetzwerk zuerst nur elektrisch getestet. Im ersten Teil wird das sogenannte *fixed-pattern-noise* untersucht, im zweiten die Antwort des Netzwerkes auf Stimulation eines einzigen Einganges und auf kantenförmige Eingangsmuster.

4.2.1 Offsets des Widerstandsnetzwerkes

Warum sollte der Chip sich – im Rahmen der gegebenen Meßgenauigkeit – nicht genau so verhalten wie die Simulationen es vorhersagen? Auf diese Frage gibt es mehrere Antworten: Erstens sind möglicherweise nicht alle elektrischen Eigenschaften des Layouts, wie z.B. parasitäre Kapazitäten, in der Simulation berücksichtigt. Zweitens werden viele Einflüsse, wie z.B. die Temperaturverteilung auf dem Chip, oder einfallendes Licht, in der Simulation nicht berücksichtigt. Drittens ist die Fertigungsgenauigkeit begrenzt, so daß sich für die geometrischen Abmessungen Abweichungen von dem im Layout definierten Wert ergeben. Viertens

[3] Zur Erstellung des Programms wurde eine C^{++} Entwicklungsumbebung der Firma *Borland* verwendet.
[4] AWG steht für **A**rbitrary **W**aveform **G**enerator.

unterscheiden sich die Prozeßparameter, die die Eigenschaften der verwendeten Materialien beschreiben, sowohl räumlich auf dem Chip, als auch von Wafer zu Wafer. Die beiden letzten Effekte werden unter dem Namen *fixed pattern noise* zusammengefaßt. Im Gegensatz zu den üblichen *noise*-Arten handelt es sich hier um räumliche Schwankungen. Die Bezeichnung *fixed pattern* rührt daher, daß die räumliche Verteilung der Prozeßparameterschwankungen und Geometrievariationen bei der Herstellung erzeugt werden und sich danach nicht mehr ändern. (sofern man die Strukturen auf dem Chip nicht nochmals physikalisch oder chemisch beeinflußt).

Konkret bedeuten die beschriebenen räumlichen Schwankungen, daß sich Transistoren der *gleichen* Dimensionierung und Art in ihrem elektrischen Verhalten *unterscheiden*. Auf das Verhalten des Widerstandsnetzwerkes hat das drei verschiedene Auswirkungen:

Offsets der Eingangssignale

Der emphMismatch zwischen den Transistoren $T3$ und $T4$ des Transkonduktanz-Verstärkers aus Abb. 3.1 sorgt dafür, daß die Ströme I_{links} und I_{rechts} im Allgemeinen nicht identisch sind. Auf Grund der Unterschiede von $T3$ und $T4$ sind andererseits unterschiedliche Gate-Spannungen V_+ und V_- notwendig, damit der gleiche Strom durch beide Zweige der Schaltung fließt. Diese Effekte können zu einer Offset-Spannung zusammengefaßt werden, die sich in der Beschaltung des Transkonduktanz-Verstärkers als Spannungsfolger ohne Last, die in Abb. 4.5 illustriert ist, zwischen Ein- und Ausgang einstellt. Die als Transkonduktanz-Verstärker reali-

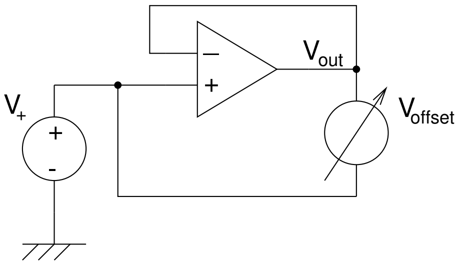

Abbildung 4.5: Offsets eines Transkonduktanz-Verstärkers.

sierten G's des Widerstandsnetzwerkes werden sich also so verhalten, als ob die Spannung am Eingang i nicht S_i sondern $S_i + O_{trans,i}$ betrüge, wo $O_{trans,i}$ den Offset des Transkonduktanz-Verstärkers i bezeichnet.

Offsets der Ausgabesignale

Die 64 Spannungsfolger, mit denen die Knotenspannungen verstärkt werden, sowie der interne und der externe Operationsverstärker sind ebenfalls mit Offsets behaftet. Die beiden letztgenannten verursachen allerdings nur einen globalen Offset für alle Knotenspannungen. Dagegen bewirken die 64 Spannungsfolger eine Verschiebung der Knotenspannungen untereinander. Bezeichnet man die Offsets des gesamten Auslesesystems mit O_{ug}[5], so ergeben sich die wirklichen Knotenspannungen aus den gemessenen Spannungen $V_{mes,i}$ zu:

$$V_i = V_{i,mes} - O_{ug} \tag{4.1}$$

[5]ug steht für **u**nity **g**ain buffer, weist also auf die 64 Spannungsfolger hin.

Variation der R's und G's

Die Variationen der Transistoren Tb und Td der Hres-Bias-Schaltung aus Abb. 3.6, sowie die der Transistoren $T1$ und $T2$ der in Abb. 3.5 dargestellten Widerstandsverbindung sorgen zusammen dafür, daß sich die resultierenden differentiellen Widerstände der 64 als Hres-Widerstände implementierten R's voneinander unterscheiden. Das gleiche gilt für die durch die Transkonduktanz-Verstärker realisierten G's des Netzwerkes, da die 64 Bias-Ströme der Transkonduktanz-Verstärker sich entsprechend der Variationen der Transistoren Tb aus Abb. 3.1 und aller Transistoren in der Schaltung aus Abb. 3.11 voneinander unterscheiden.

Messung der Offsets V_{ug} und V_{trans}

Um die Offsets der Transkonduktanz-Verstärker von denjenigen der Ausgangsspannungsfolger zu trennen, muß man das Verhalten des Netzwerkes für $L = 0$ und $L = \infty$ ausnutzen: Im zweiten Fall sind die Ausgänge der Transkonduktanz-Verstärker über die Hres-Widerstände kurzgeschlossen. Daher sollte sich an allen Knoten der Mittelwert aus der Summe aller Eingangsspannungen und Offsets der Transkonduktanz-Verstärker einstellen. Wenn an allen Knoten das gleiche Eingangssignal anliegt, können die verbleibenden Schwankungen der Ausgangsspannungen also den Offsets der Spannungsfolger zugeordnet werden.

Für $L = 0$ ist die Verbindung der verschiedenen Knoten durch die Hres-Widestände dagegen unterbrochen, so daß keine Mittelung der benachbarten Knotenspannungen auftritt. Zieht man von den unter diesen Bedingungen gewonnen Daten die vorher für die Spannungsfolger bestimmten Offsets ab, ergeben sich die Offsets der Transkonduktanz-Verstärker.

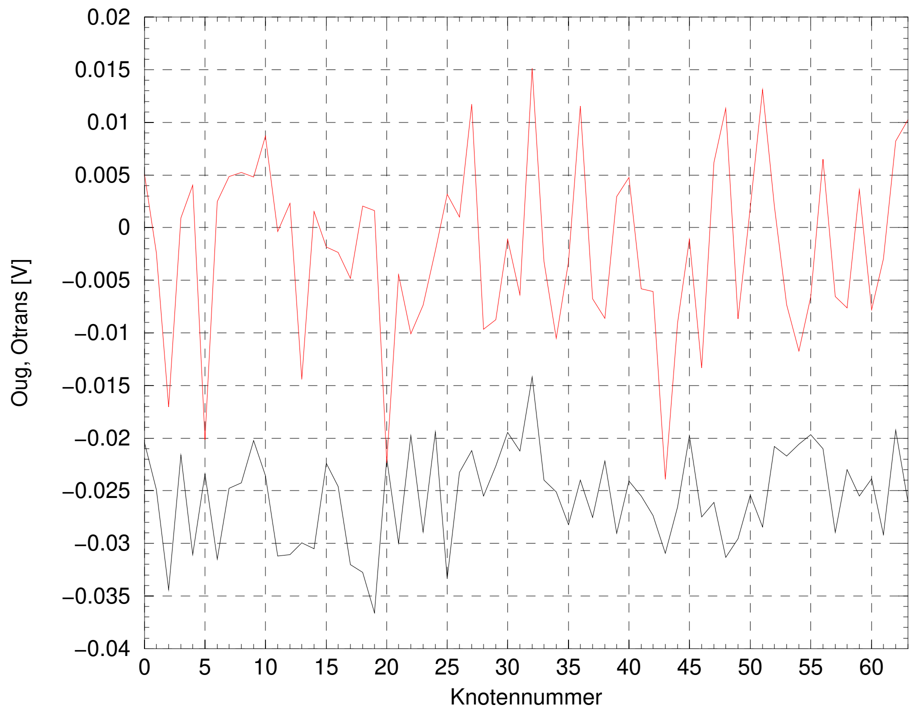

Abbildung 4.6: Offsets der Ausgangstreiber V_{ug} (rot, oben) und der Transkonduktanz-Verstärker V_{trans} (schwarz, unten). Die Auflösung des ADC's betrug 0.5 mV.

		σ [mV]	V_{pp} [mV]	V_{mittel} [mV]
Chip 1	Spannungsfolger	4,89	28.02	-29.81
	Transkonduktanz-Verstärker	7.63	34.64	-0.5
Chip 2	Spannungsfolger	4.57	10.51	-25.53
	Transkonduktanz-Verstärker	8.49	39.05	-2.03

Tabelle 4.1: Standardabweichung σ, *peak to peak*-Spannung V_{pp} und Mittelwert V_{mittel} der an zwei Chips gemessenen Offsetkurven.

Zur Bestimmung der Offsets der Spannungsfolger wurde $V_{reset} = 3.6\,\mathrm{V}$ und $I_{biasR} = 219\,\mu\mathrm{A}$ gewählt. Aus Abb. 3.12 ergibt sich der Biasstrom der Transkonduktanz-Verstärker I_{biasG} zu etwa 10 nA. Mit Gl. (3.24) erhält man damit $L \approx 200$, so daß die Offsets der Transkonduktanz-Verstärker durch das Netzwerk gut ausgeglichen werden sollten. Um $L = 0$ für die Ermittlung der Offsets der Transkonduktanz-Verstärker zu erreichen, wurde $IbiasR = 0\,\mathrm{nA}$ und $V_{reset} = 0\,\mathrm{V}$ benutzt.

Bei beiden Messungen wurden die Eingänge des Netzwerkes jeweils zehn mal mit der Eingangsspannung 2 V beschrieben. Nach jedem Einlesevorgang wurden die Knotenspannungen zehn mal ausgelesen. Aus den resultierenden 100 Datensätzen wurden dann die in Abb. 4.6 gezeigten Mittelwerte gebildet.

Die Offsets der Transkonduktanz-Verstärker sind deutlich größer als die der Spannungsfolger. Das ist insofern zu erwarten, als daß die relativen Geometrievariationen für die Spannungsfolger auf Grund ihrer vierfach größeren Gatefläche kleiner ausfallen als für die Transkonduktanz-Verstärker. Der globale Offset der Kurve für die Spannungsfolger von etwa $-30\,\mathrm{mV}$ kann seine Ursachen in der gesamten Ein-und Auslese-Elektronik haben, kann aber im Einzelnen nicht erklärt werden.

Die Messung der Offsets wurde für zwei Chips durchgeführt. Die sich ergebenden Werte für die Standardabweichung, den Mittelwert und die *peak to peak*-Spannung der Offsetkurven sind in Tabelle 4.1 zusammengefaßt. Die Werte für Chip 2 gehören zu den in Abb. 4.6 dargestellten Daten.

Um die Qualität des *fixed pattern noise* der Transkonduktanz-Verstärker zu untersuchen, wurde auf die Offsetdaten von Chip 2 eine diskrete Fouriertransformation angewendet. Das Ergebnis ist in Abb. 4.7 dargestellt. Da die zu transformierende Kurve nur 64 Datenpunkte besitzt, kann der sich ergebende Verlauf als zufällige Ausprägung eines weißen Spektrums interpretiert werden. Die Schwankungen von Geometrie und Prozeßparametern besitzen also entweder kleinere Ortsfrequenzen als die sich aus dem Abstand benachbarter Zellen des Netzwerkes ergebenden, oder hängen von so vielen unabhängigen Parametern ab, daß sich keine Regelmäßigkeit erkennen läßt.

Meßgenauigkeit

Ausgehend von der Streuung der einzelnen Knotenspannungen in den 100 gleichen Messungen zur Offsetbestimmung kann auf die Genauigkeit, mit der das Netzwerk beschrieben und wieder ausgelesen werden kann, zurückgeschlossen werden. Für die Messungen der Offsets der Spannungsfolger unterscheiden sich die einzelnen Knotenspannungen maximal um 2.5 mV. Die Messungen der Offsets der Transkonduktanz-Verstärker ergaben die in Abb. 4.8 dargestellten Differenzen zwischen maximaler und minmaler gemessener Knotenspannung.

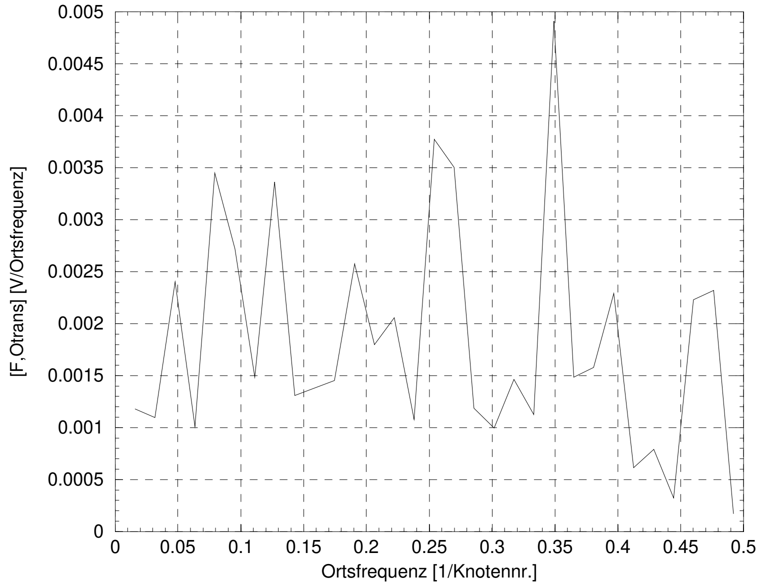

Abbildung 4.7: Fouriertransformierte der Offsetspannung V_{trans} der Transkonduktanz-Verstärker aus Abb.4.6.

Die Abweichungen sind hier wesentlich größer als für die Messung bei $L \approx \infty$. Die Ursache hierfür ist vermutlich darin zu suchen, daß die Gate-Spannungen der Transistoren $T1$ und $T2$ der Widerstandsverbindung aus Abb. 3.5 für $I_{biasR} = 0$ nicht durch die Bias-Schaltung bestimmt werden, da durch diese kein Strom fließt. Die Hres-Widerstandsverbindungen unterliegen damit nur Störeinflüssen und dem Rauschen. Daher ist es möglich, daß die R's in den betrachteten Messungen nicht immer den sehr großen Widerstand besaßen, der beabsichtigt war. Für die zur Korrektur der Eingangsspannungen verwendeten Offsets sollten die Fehler aber nicht zu groß sein, da über 100 Messungen gemittelt wurde.

Variation der Widerstände R und G

Die Bestimmung der Offsets der R- und G-Werte der einzelnen Netzwerkknoten gestaltet sich schwieriger als die der Spannungsoffsets. Die Implementierung des Netzwerkes verbietet eine direkte Messung der in Frage stehenden Größen, so daß nur das Verhältnis der Widerstandwerte untereinander bestimmt werden kann. Die dazu benutzte Vorgehensweise ist in Abb. 4.9 illustriert: Die Eingangsspannungen werden so gewählt, daß sich alle Knotenspannungen V_i für $i < n$ auf den Wert V_{inhigh} einstellen, diejenigen mit dem Index $i > n$ dagegen den Wert V_{inlow} annehmen. Das dazu nötige Eingangsspannungsmuster ist unten links in Abb. 4.9 gezeigt. Da der Strom in horizontaler Richtung in diesem Zustand auf den Widerstand R_n beschränkt bleibt, können die drei Widerstände in der gestrichelten Box vom Rest des Netzwerkes unabhängig als Spannungsteiler betrachtet werden. Aus den drei Spannungsdifferenzen über G_n, R_n und G_{n+1} läßt sich dann das Verhältnis der drei Widerstände ermitteln.

Die Messung wurde für $0 \leq n \leq 62$ mit Ausgangsspannungen $V_{inhigh} = 2.05\,\mathrm{V}$ und $V_{inlow} = 2\,\mathrm{V}$ sowohl für den gezeigten Verlauf des Ausgangsmusters als auch für das sich unter Vertauschung von V_{inlow} und V_{inhigh} ergebende Muster durchgeführt. Um bei dem

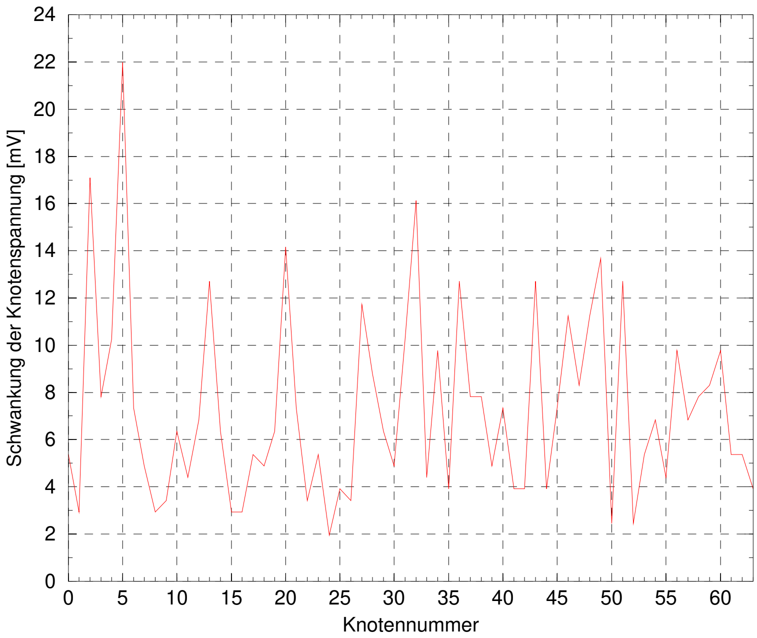

Abbildung 4.8: Differenz der maximal und minimal in den 100 Messungen zur Bestimmung der Offsets der Transkonduktanz-Verstärker auftretenden Knotenspannungen.

gewählten Spannungshub eine möglichst hohe Linearität der Widerstände zu erreichen, wurde $V_{reset} = 0$ gesetzt, was nach Abb. 3.12 einem Bias-Strom für die Transkonduktanz-Verstärker $I_{biasG} = 220\,\mathrm{nA}$ entspricht. Damit die Widerstände der R's und G's möglichst ähnliche Werte annehmen, wurde der Bias-Strom der Hres-Widerstände auf $I_{biasR} = 252\,\mathrm{nA}$ eingestellt. Zur Ermittlung der Eingangsspannungen S_n und S_{n+1} wurde ein Bisektionsverfahren verwendet. Um die restlichen Eingangsspannungen zu erhalten, wurden die oben ermittelten Offsets der Transkonduktanz-Verstärker von dem Sollwert der Eingangsspannung (also V_{inlow} oder V_{inhigh}) abgezogen. Die gemessenen Knotenspannungen sind um die Offsets der Spannungsfolger $O_{ug,i}$ korrigiert. In den folgenden Kapiteln wird diese Vorgehensweise als *kalibrierte* Messung bezeichnet. In Abb. 4.10 sind zwölf der erzeugten Ausgangsmuster dargestellt.

Aus den Spannungsdifferenzen wurden jeweils die Widerstandsverhältnisse abgeleitet und alle R's und G's in Einheiten von R_0 bzw. G_0 ausgedrückt. Die sich ergebenden Werte waren weder reproduzierbar, noch realistisch, da sich für manche R_i's Werte von bis zu $2 \cdot 10^4 R_0$ ergaben, die mit den erwarteten Prozeßparameterschwankungen nicht mehr erklärt werden können. Mögliche Erklärungen des Verhaltens sind einerseits die Ungenauigkeit, mit der die Eingangs- und Knotenspannungen kontrolliert bzw. ausgelesen werden können und andererseits die Tatsache, daß die R's und G's sich für die anliegenden Potentialdifferenzen nicht mehr linear verhalten[6]. Die aus diesen beiden Punkten resultierenden Ungenauigkeiten der Widerstandswerte multiplizieren sich bei der Zurückführung auf R_0 bzw. G_0, und können so zu immensen Fehlern in den Verhältnissen weit voneinander entfernter Widerstände führen.

Daß man aus den Messungen dennoch etwas lernen kann zeigt der rechte Teil von Abb. 4.10:

[6]Man könnte versuchen die Daten mit einem realistischeren Modell für die R's und G's zu auszuwerten. Im Rahmen dieser Diplomarbeit wurde aber den in Kapitel 5 beschriebenen optischen Messungen eine höhere Priorität eingeräumt.

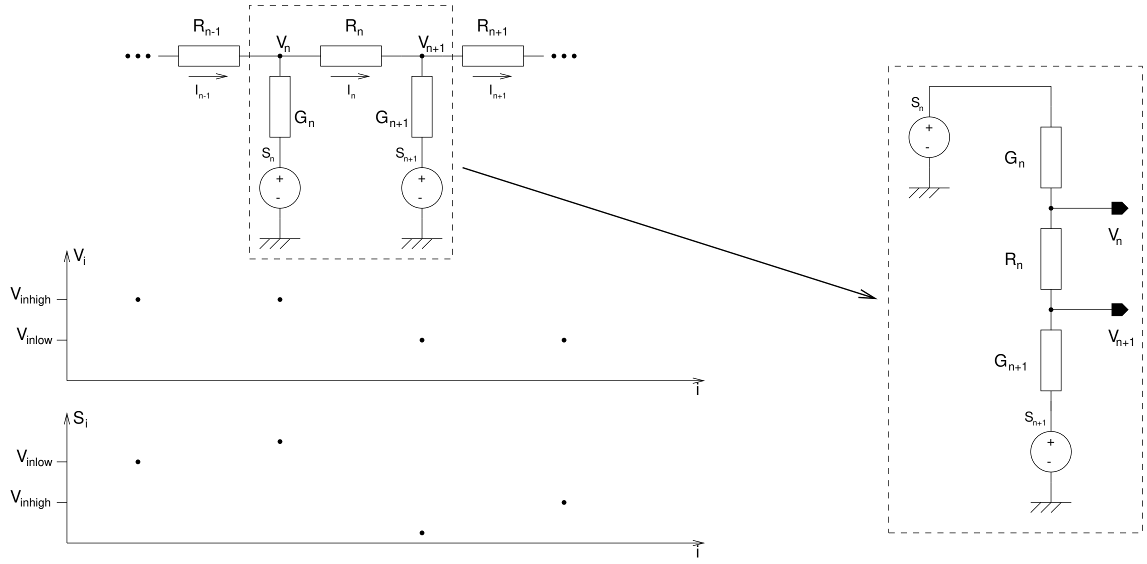

Abbildung 4.9: Beschaltung des Netzwerkes zur Bestimmung der Verhältnisse der einzelnen Widerstände zueinander.

Die Knotenspannungen im mittleren Teil des Netzwerkes sind für alle 12 gezeigten Kurven bis auf etwa 1 mV identisch. Die Wiederholgenauigkeit der Messungen mit endlichen Werten für I_{biasR} ist also nicht so schlecht wie man auf Grund von Abb. 4.8 erwarten würde, sondern in der Größenordnung von 1 mV. Andererseits schwanken die Knotenspannungen in einem etwa 3.5 mV großen Intervall um den vorgegebenen Sollwert von 50 mV. Die Korrektur des *fixed pattern noise* funktioniert also nicht ganz perfekt, sondern reduziert diesen nur auf eine *peak to peak* Spannung von etwa 3.5 mV.

4.2.2 Reaktion auf die Einzelstimulation eines Eingangs

Um die glättende Wirkung des Netzwerkes zu testen, wurde wie in den Simulationen in den beiden Kapiteln 1.2.3 und 3.3 eine *räumliche* Spannungsspitze als Eingangsmuster verwendet. Die Spannung des stimulierten Eingangs betrug 2.1 V, die der restlichen 63 Eingänge 2 V. Die Messung wurde bei einem Bias-Strom von $I_{biasR} = 252\,\mathrm{nA}$ für Resetspannungen $V_{reset} = 0$, $0.5, \cdots, 3.5\,\mathrm{V}$ für alle 64 Knoten durchgeführt. Die Offsets der Eingangs- und Ausgangsstufe wurden im Sinne der in Kapitel 4.2.1 definierten Kalibration korrigiert. Das Ergebnis für die Stimulation von Knoten 31 ist in Abb. 4.11 gezeigt. Die Meßwerte liegen zwischen 0 und 50 mV, da die Ausgangsspannung differentiell gegen $V_{ref} = 2\,\mathrm{V}$ gemessen wird.

Die Messungen stimmen abgesehen von dem nach der Kalibration verbleibenden *fixed pattern noise* qualitativ mit den in Abb. 3.15 gezeigten Ergebnissen der Simulation überein. Mit Hilfe von Abb. 4.11, in der der Logarithmus der Knotenspannungen aufgetragen ist, kann die exponentielle Abhängigkeit der Knotenspannung vom erregten Knoten verifiziert werden. In der Abbildung sind nur diejenigen Datenpunkte gezeigt, die für eine lineare Regression benutzt wurden, mit der die charakteristische Länge L für die 8 Kurven bestimmt wurde. Die resultierenden L's sind im linken Teil von Abb. 4.13 für beide Hälften des Netzwerkes gegen V_{reset} aufgetragen. Aus den Abbildungen 4.11 und 4.13 geht hervor, daß die charakteristische Länge für die rechte Seite bei allen Resetspannungen etwas größer ist als für die linke.

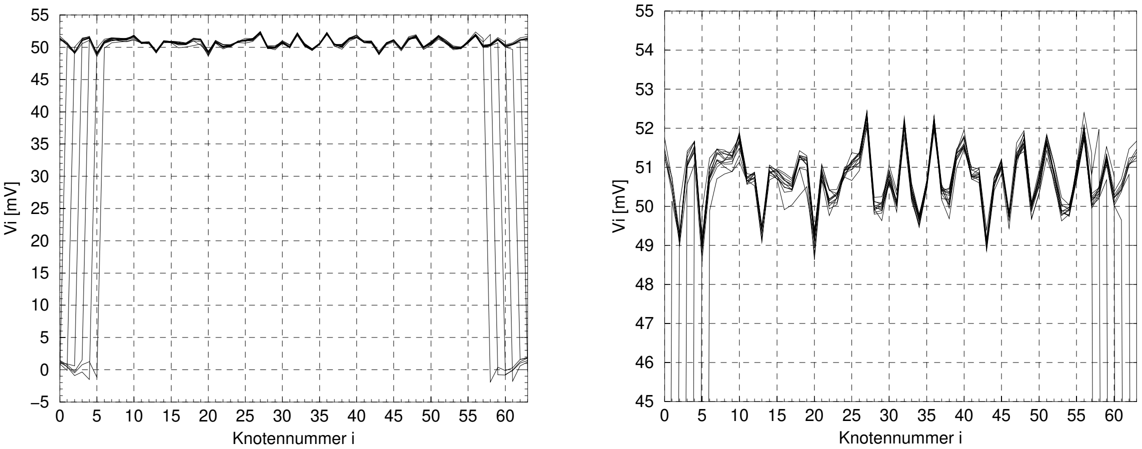

Abbildung 4.10: Verlauf einiger der mit dem Bisektionsverfahren eingestellten Kanten. **Links:** Für $0 < n < 6$ sind sechs von links nach rechts steigenden Kanten, für $56 < n < 62$ sechs von rechts nach links fallende Kanten dargestellt. **Rechts:** Die gleichen Kurven wie in dem linken Graphen sind hier in vergrößertem Maßstab zwischen 45 und 55 mV dargestellt.

Um einen Vergleich der dargestellten L-Werte mit den aus der Simulation vorhergesagten anstellen zu können, kann man Gl. (3.24) nach I_{biasG} auflösen:

$$I_{biasG} = \frac{nI_{biasR}}{2L^2} \tag{4.2}$$

Der Bias-Strom für die Transkonduktanz-Verstärker I_{biasG} wurde aus den L-Werten berechnet und ist in Abb. 4.13 gegen die Resetspannung aufgetragen. Da n aus den durchgeführten Messungen nicht bestimmt werden kann, wurde der in den Simulationsparametern von *AMS* angegebene Wert von 1.17 benutzt. Im Vergleich zu den in Abb. 3.12 gezeigten Kurven weisen die in Abb. 4.13 dargestellten einen eher linearen Verlauf über den ganzen Spannungsbereich auf. Der aus den Meßdaten berechnete Strom für $V_{reset} = 0\,\mathrm{V}$ ist etwas kleiner, als der simulierte, der aus den Meßwerten links vom Knoten errechnete sogar um den Faktor 1.7. Die beiden aus den Meßwerten ermittelten Kurven sind zwar vom Verlauf her ähnlich, unterscheiden sich aber in ihren Werten etwa um den Faktor 1.3.

Um einen Eindruck von den Variationen der R's und G's des gesamten Netzwerkes zu bekommen, wurden in Abb. 4.14 die Knotenspannungen des jeweils stimulierten Einganges für alle Messungen gegen die Nummer des erregten Einganges aufgetragen. Die acht Kurven entsprechen wieder den acht verschiedenen Werten für die Resetspannung. Das sich ergebende Muster ist für alle Kuven sehr ähnlich und wird nur für $V_{reset} \geq 3\,\mathrm{V}$, also $L \geq 3$, durch den mittelnden Einfluß des Netzwerkes in seinem Dynamikbereich deutlich beschränkt. Das ist folgendermaßen zu verstehen: Mit der gewählten Stimulation des Nezwerkes können die Variationen der verschiedenen R's und G's nicht ohne weiteres getrennt werden. Die Einflüsse sollten vielmehr exponentiell mit der charakteristischen Länge L abnehmen, so daß sich für die größeren Werte von L eine Mittelung über mehr Knoten ergibt, die zu einer Glättung der Offsetvariationen führt. Der Anstieg der Kurven zu den Seiten hin ist wieder als durch die Endlichkeit des Netzwerkes bedingter Randeffekt zu verstehen.

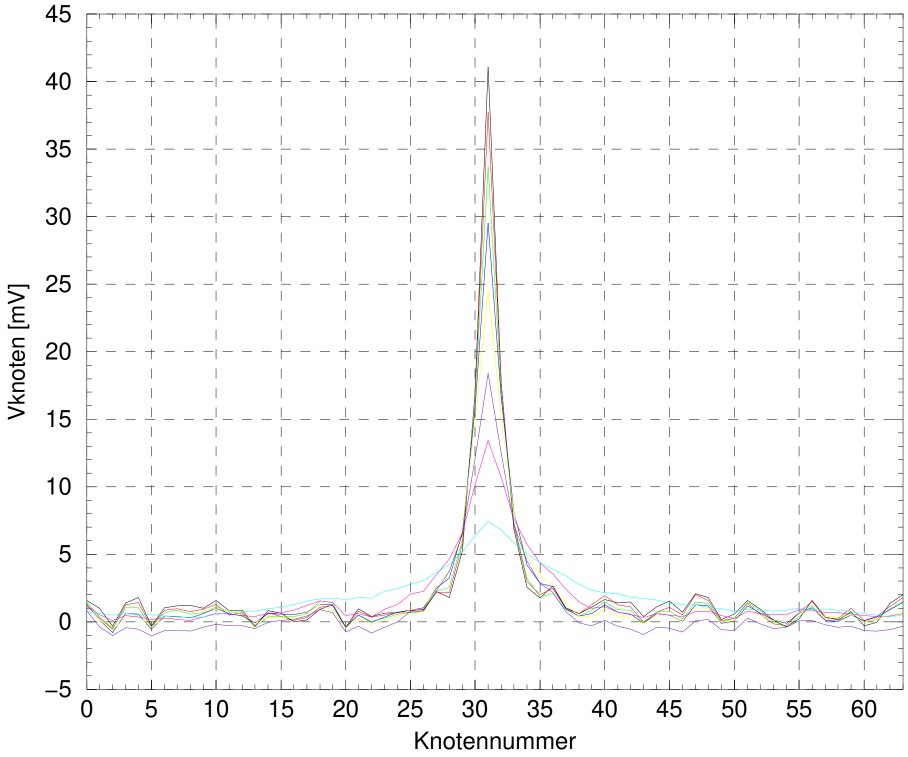

Abbildung 4.11: Antwort des Netzwerkes auf eine einzelne 100 mV große Spannungsspitze für $V_{reset} = 0, 0.5, \cdots, 3.5\,\mathrm{V}$.

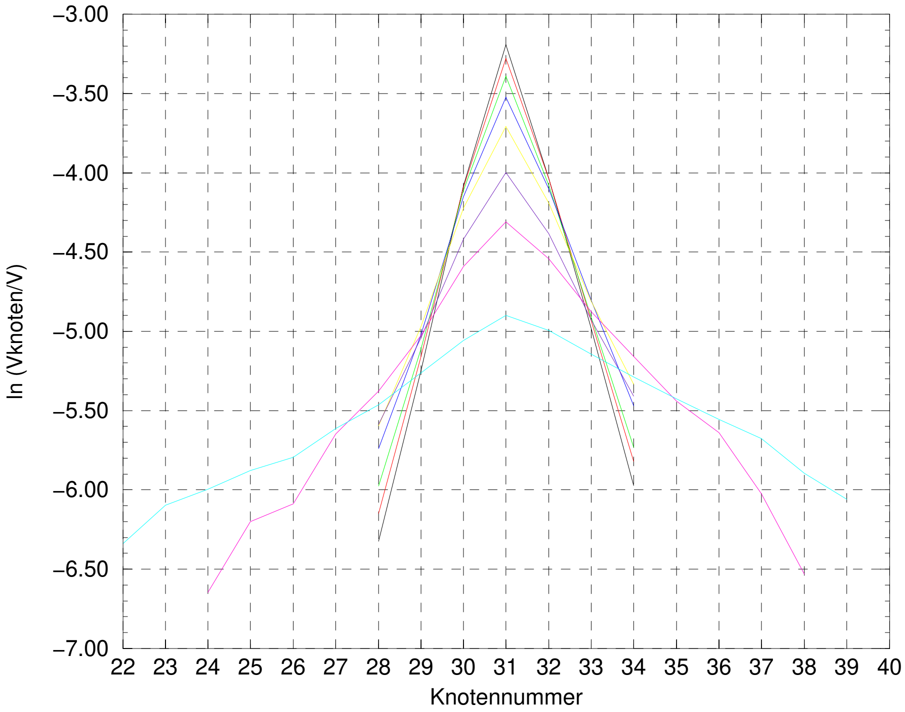

Abbildung 4.12: Logarithmische Darstellung der in Abb. 4.11 gezeigten Kurven. Gezeigt sind nur die Datenpunkte, die für die jeweilige lineare Regression verwendet wurden.

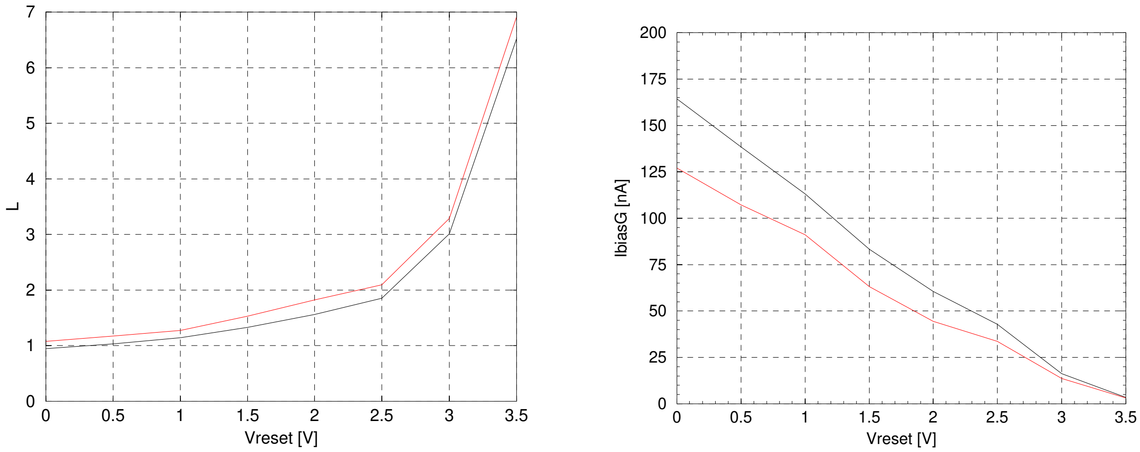

Abbildung 4.13: Links: Durch lineare Regression der in Abb. 4.12 gezeigten Kurven gewonnene charakteristische Längen L in Abhängigkeit von V_{reset}. Die obere Kurve ergibt sich aus den Meßwerten rechts des stimulierten Knotens, die untere aus denen links vom erregten Knoten. **Rechts:** Bias-Strom I_{biasG} der Transkonduktanz-Verstärker aufgetragen gegen V_{reset}.

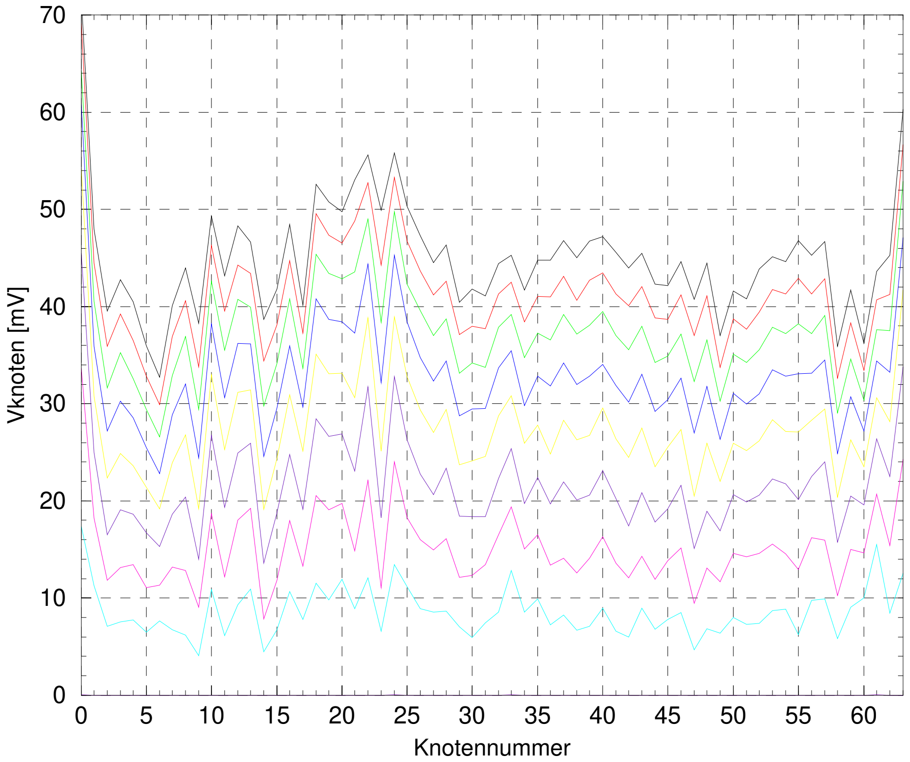

Abbildung 4.14: Spannung der Netzwerkknoten, deren Eingang jeweils mit einer 100 mV großen Spannungsspitze stimuliert wurde, für alle 64 Knoten bei $V_{reset} = 0, 0.5, \cdots, 3.5\,\mathrm{V}$.

Kapitel 5

Optische Messungen an Oasys_RN

Nachdem das Widerstandsnetzwerk in Kapitel 4 elektrisch getestet und teilweise durch seine Offsets charakterisiert werden konnte, beschäftigt sich dieses Kapitel mit ersten Tests der optischen Komponenten. Zuerst wird der optische Meßaufbau vorgestellt und diskutiert, um ein Gefühl für die erreichbare Genauigkeit und ihre Abhängigkeit von der Größe des Laserspots zu erhalten. Die tatsächliche Lichtfleckgröße wird mit Hilfe der Positionsdetektoren bestimmt. Schließlich wird als erster Test der optischen Konfigurierbarkeit des Netzwerkes versucht, einen einzelnen Eingang stärker zu gewichten als die restlichen und dann noch das der optischen Programmierung folgende zeitliche Verhalten der Ausgangssignale betrachtet.

5.1 Optischer Meßaufbau

In Abb. 1.20 wurde bereits der prinzipielle Aufbau für eine optische Programmierung des Widerstandsnetzwerkes vorgestellt. Der tatsächlich gewählte Aufbau ist in Abb. 5.1 dargestellt und unterscheidet sich von dem aus Abb. 1.20 nur durch das zwischen den Neutralfiltern und dem Scanner eingefügten Raumfilter. Dieses dient hier wie in den in Kapitel 2 beschriebenen Messungen dazu, höhere Beugungsordnungen des Laserstrahls auszufiltern. Insbesondere sollten durch Mehrfachreflexionen zwischen den Neutralfiltern verursachte Nebenmaxima aus dem Strahl entfernt werden. Die Tochterplatine mit dem Chip ist an dem x-y-Fahrtisch befestigt, der mit Hilfe eines OS-9 Rechners gesteuert[1] wird. Mit Hilfe einer Mikrometerschraube kann die Platine in der Bewegungsebene des Fahrtisches gedreht werden. Das ist nötig, um die X-Achse des Chips (die durch die Anordnung der Kontrolldioden vorgegebene Achse) und die Achse, auf der sich der abgelenkte Laser-Spot bewegt, in Einklang zu bringen.

Die Scanner-Elektronik wird wie in Kapitel 4.1.1 beschrieben mit den von der Meßkarte und dem externen DAC erzeugten Spannungen V_{scan+} und V_{scan-} angesteuert. Zur Steuerung der Laserintensität wird ein *Arbitrary Waveform Generator* benutzt, der die nötige Modulationsspannung erzeugt. Der AWG kann über die serielle Schnittstelle programmiert werden.

Die Wahl des optischen Systems, das das Laserlicht auf den Chip fokussiert, bestimmt sowohl den Durchmesser des fokussierten Lichtflecks als auch dessen Positioniergenauigkeit durch den Scanner. In dem gezeigten Aufbau wurde für die Fokussierung eine asphärische

[1] Das C-Programm für die Fahrtischsteuerung wurde von Markus Loose im Rahmen einer Diplomarbeit entwickelt (vgl. [Loose 96]).

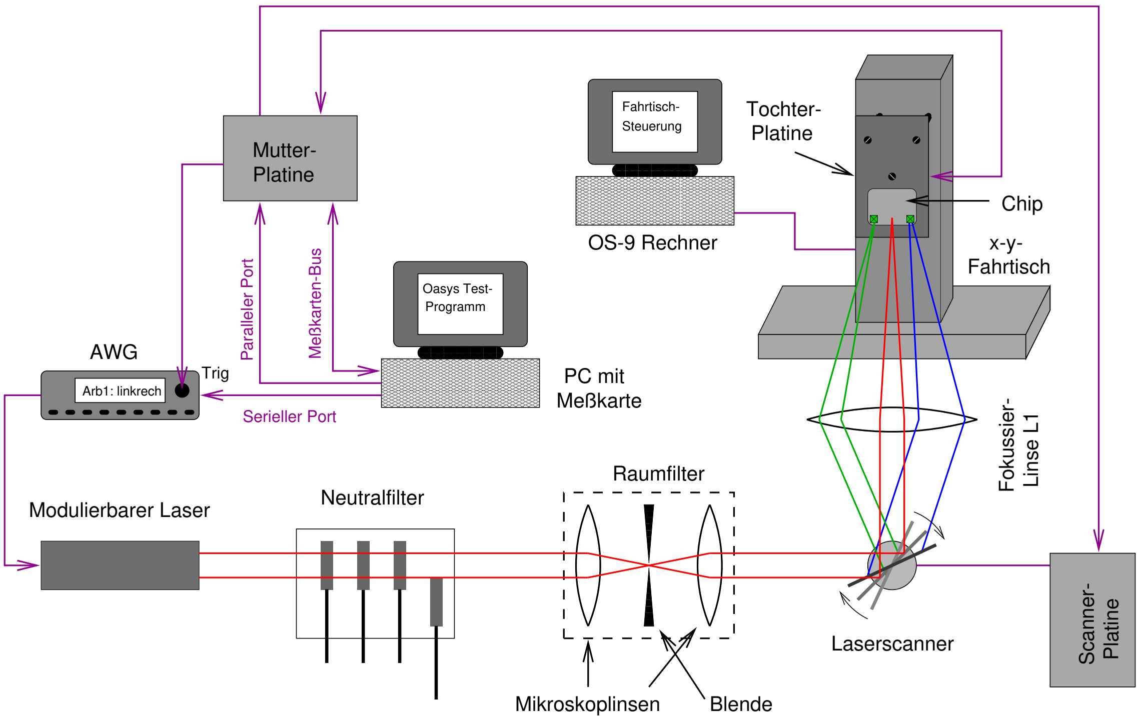

Abbildung 5.1: Meßaufbau für den Test der optischen Programmierung des Widerstandsnetzwerkes (schematisch).

Plastiklinse mit einer Brennweite von 26 mm und einer Apertur von 35 mm benutzt[2].

5.1.1 Optimale Lichtfleckgröße

Sowohl die Wellennatur des Lichtes als auch die Abbildungsfehler realer (also ausgedehnter) Linsen verhindern eine perfekte Fokussierung des Strahls auf einen Punkt. Die untere Grenze für den Durchmesser d der Strahltaille eines mit einer Linse der Brennweite f fokussierten parallelen Gaußschen Strahls ergibt sich nach [Klein 86] zu:

$$d = \frac{4\lambda f}{\pi D} \tag{5.1}$$

Dabei bezeichnet D den Durchmesser des Strahls vor der Fokussierung und λ die Wellenlänge des Laserlichtes. d ist hier als die Entfernung vom Zentrum der gaußförmigen Intensitätsverteilung definiert, für die diese auf $1/e$ des Maximalwertes abgefallen ist. Mit $f = 26\,\mathrm{mm}$, $\lambda = 675\,\mathrm{nm}$ und einem Strahldurchmesser vor der Fokussierung $D \approx 3\,\mathrm{mm}$ erhält man für den Durchmesser der Strahltaille demnach eine Mindestgröße von $7.5\,\mu\mathrm{m}$.

Die tatsächlich einzustellende Größe des Laser-Spots ist ein Kompromiß zwischen dem Übersprechen auf benachbarte Zellen bei der optischen Programmierung und der Genauigkeit dieser Programmierung. Der Einfluß der Spotgröße auf diese beiden Fehlerquellen ist

[2] Bei der verwendeten Linse handelt es sich um ein Gratisexemplar einer bi-asphärischen PXM-Leichtlinse der Firma *Eschenbach*.

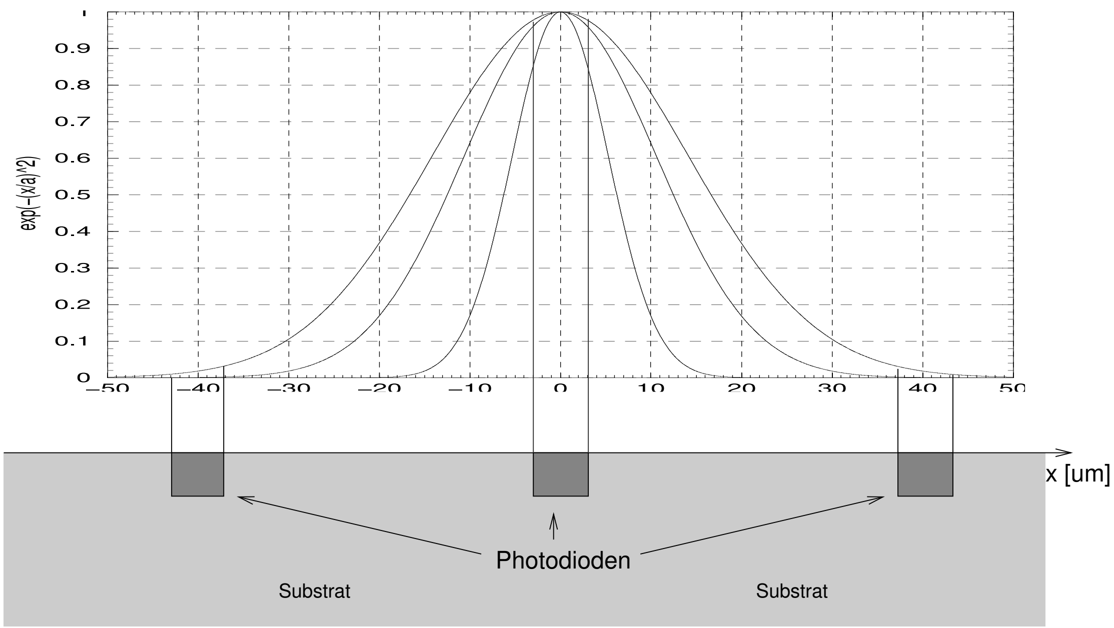

Abbildung 5.2: Zusammenhang zwischen Lichtfleckgröße und dem Übersprechen bei der optischen Programmierung einerseits und dem Fehler bei derselben verursacht durch die Ungenauigkeit der Positionierung des Spots andererseits. Das gaußförmige Intensitätsprofil aus Gl. (5.2) ist für die drei Strahlradius $d/2 = 7.5$, 15 und $20\,\mu$m über einem Schnitt durch den Chip gezeigt, von dem nur die Kontrolldioden eingezeichnet sind.

in Abb. 5.2 veranschaulicht. Ist der Lichtfleck zu weit ausgedehnt, fällt ein Teil des Lichtes auf die benachbarten Kontrolldioden und erzeugt dort einen unerwünschten Photostrom. Dagegen bedeutet ein kleinerer Lichtfleck einen steileren Verlauf der Lichtintensität, so daß sich eine Abweichung von der Soll-Position des Laserspots in einem größeren Fehler für den Photostrom niederschlägt. Der optimale Kompromiß hängt dabei entscheidend von der erreichbaren Positioniergenauigkeit des Lichtflecks ab. Fordert man, daß die Lichtintensität am Mittelpunkt der nächsten beiden Kontrolldioden auf 1% des Maximalwertes abgefallen ist, so ergibt sich für einen gaußförmigen Intensitätsverlauf gegeben durch

$$\frac{J}{J_0} = e^{-\frac{x^2}{d^2}} \tag{5.2}$$

der maximal zulässige Strahlradius $d/2 = 18.6\,\mu$m. Die Brennweite der Linse ist also insofern vernünftig gewählt, als daß die gewünschte Spotgröße einen gewissen Sicherheitsabstand zu der theoretisch möglichen besitzt. Im Übrigen macht die obenstehende Abschätzung auch deutlich, warum eine Fokussierung *vor* dem Scanner nicht in Frage kommt.

5.1.2 Berechnung des Strahlengangs

Da die Genauigkeit, mit der die Kontrolldioden programmiert werden können, ganz eng mit der Abbildung des Fokuspunktes auf dem Chip verknüpft ist, soll diese hier für den einfachsten Fall geometrischer Optik mit der Linsendicke Null durchgerechnet werden. Abb. 5.3 zeigt einen Laserstrahl mit Durchmesser $2r$, der um den Winkel α gegenüber der optischen

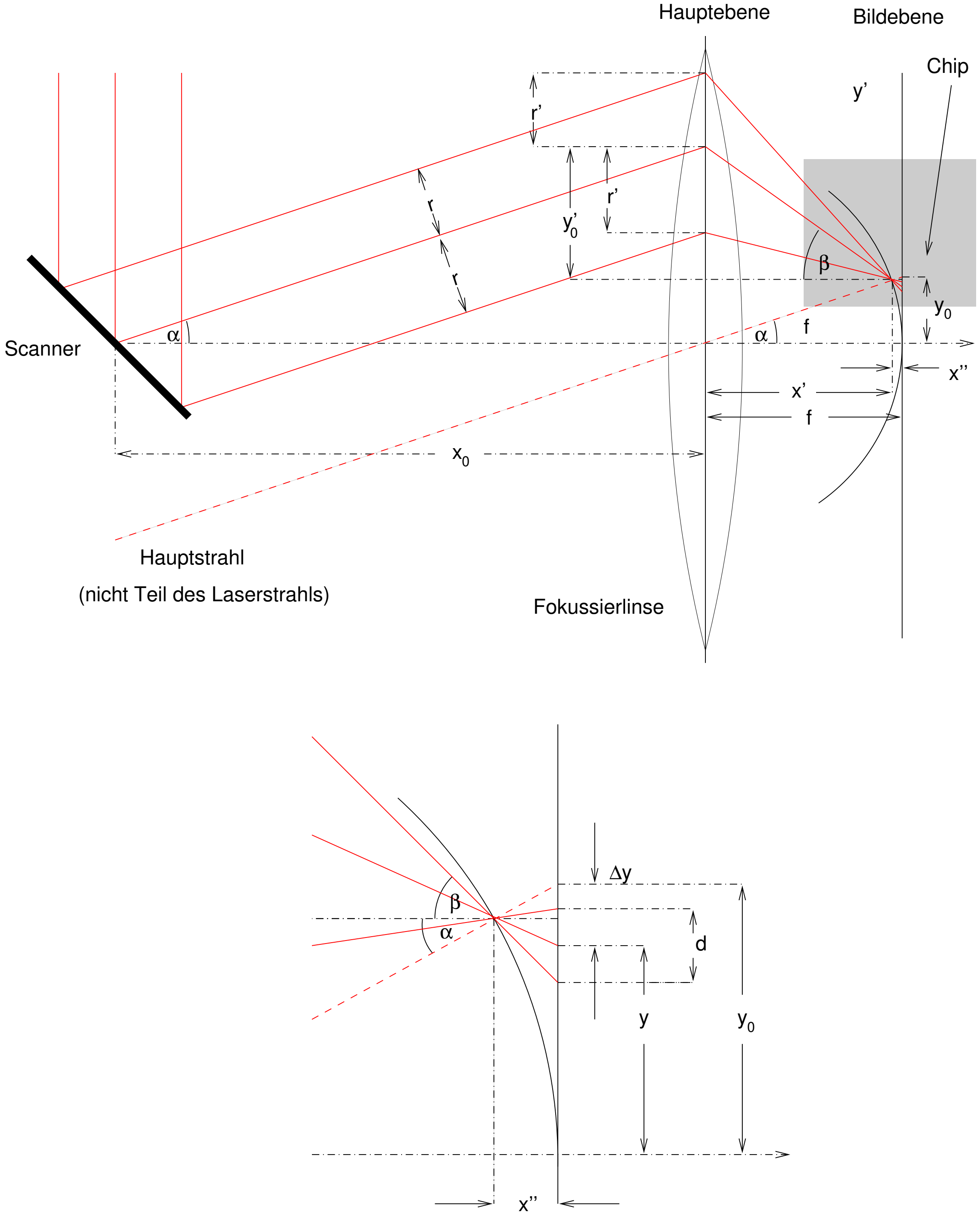

Abbildung 5.3: Strahlengang des ausgelenkten Laserstrahls (rot). Der untere Teil zeigt eine Vergrößerung des grau unterlegten Ausschnitts der oberen Zeichnung.

Achse der Linse ausgelenkt ist. Der durch die Linse erzeugte Fokuspunkt liegt dann für alle Auslenkungen auf einem Kreis (3-dimensional auf einer Sphäre), dessen Radius durch die Brennweite f der Linse gegeben ist und dessen Mittelpunkt mit dem der Linse zusammenfällt. Das ist darin begründet, daß die hier betrachtete ideale Linse paralleles Licht im Abstand f hinter der Linse auf einen Punkt in der Weise fokussiert, daß der die Linsenmitte durchquerende Hauptstrahl, der allerdings hier gar nicht zum betrachteten Strahlenbüschel gehört und nur zu Hilfszwecken eingezeichet ist, nicht abgelenkt wird.

Die Tatsache, daß sich die Fokuspunkte des Laserstrahls auf einem Kreis befinden, hat einerseits zur Folge, daß der Mittelpunkt des auf die Bildebene treffenden Lichtflecks um Δy gegenüber dem Auftreffort des Hauptstrahls zur optischen Achse verschoben ist und andererseits, daß der auf die Bildebene treffende Strahl die endliche Ausdehnung d besitzt (dargestellt im unteren Teil der Abb. 5.3).

Für ein unter dem Winkel α auf die Linse treffendes paralleles Strahlenbüschel, das Rotationssymmetrie bzgl. des Haupstrahls besitzt, erwartet man das Maximum der Lichtintensität bei:

$$y_0 = f \tan\alpha \tag{5.3}$$

Um nun den Mittelpunkt und die Ausdehnung des eingezeichneten Strahlenbüschels in der Bildebene zu berechnen, werden zuerst die Abstände des Fokuspunktes von der Hauptebene und der Bildebene bestimmt. Aus der Abbildung ergibt sich:

$$\begin{aligned} x' &= f\cos\alpha && (5.4)\\ x'' &= f(1-\cos\alpha) && (5.5) \end{aligned}$$

Für den vertikalen Abstand zwischen dem Durchstoßpunkt des mittleren Laserstrahls durch die Hauptebene und dem Fokuspunkt läßt sich schreiben:

$$y_0' = x_0\tan\alpha - x'\tan\alpha = \tan\alpha(x_0 - f\cos\alpha) \tag{5.6}$$

Nun kann man den Winkel β zwischen dem mittleren Strahl des fokussierten Laserstrahls und der optischen Achse durch α und x_0 ausdrücken.

$$\beta = \arctan\frac{y_0'}{x'} = \arctan\left(\frac{\tan\alpha(x_0 - f\cos\alpha)}{f\cos\alpha}\right) \tag{5.7}$$

Damit erhält man zusammen mit Gl. (5.5) für Δy schließlich:

$$\Delta y = (\tan\beta + \tan\alpha)x'' = \left(\frac{(x_0 - f\cos\alpha)\tan\alpha}{f\cos\alpha} + \tan\alpha\right) f(1-\cos\alpha) \tag{5.8}$$

Der Mittelpunkt des fokussierten Strahlenbüschels trifft die Bildebene also im Abstand

$$y = y_0 - \Delta y = \left[1 - \left(\frac{x_0 - f\cos\alpha}{f\cos\alpha} + 1\right)(1-\cos\alpha)\right] f\tan\alpha \tag{5.9}$$

von der optischen Achse. Anstatt zu versuchen, diese Gleichung zu lösen, soll hier nur für einen gegebenen Winkesl α die Abweichung ΔY von dem sich für den Hauptstrahl ergebenden Abstand berechnet werden, da der genaue Wert wegen der Fülle anderer Ungenauigkeiten experimentell durch Iteration gewonnen werden muß. Für y_0 wird die Hälfte des Abstandes der Positionsdetektordioden, also 1.45 mm, eingesetzt. Der maximal benötigte Auslenkwinkel ergibt sich damit unter Vernachlässigung von Δy zu:

$$\alpha_{max} = \arctan \frac{y_0}{f} = 3.19^\circ = 55.7\,\mathrm{mrad} \tag{5.10}$$

Setzt man das, sowie den Abstand des Scanners von der Linse $x_0 \approx 70\,\mathrm{mm}$ in Gl. (5.8) ein, so ergibt sich:

$$\Delta y = 6.1\,\mu\mathrm{m} \tag{5.11}$$

Damit ergibt sich folgende Abschätzung für die Abweichung der Spotposition bei dem maximalen Auslenkwinkel α_{max} von der für einen linearen Zusammenhang zwischen Ablenkwinkel α und Veränderung der Spotposition auf dem Chip erwarteten:

$$y - y_{lin} = f \tan\alpha - f\alpha - \Delta y \approx 1.5\,\mu\mathrm{m} - 6.1\,\mu\mathrm{m} = -4.6\,\mu\mathrm{m} \tag{5.12}$$

Da eine Positioniergenauigkeit von etwa $1\,\mu$m angestrebt wird, muß dieser Abweichung bei der Ansteuerung von Scanner und AWG Rechnung getragen werden.

Für den Durchmesser des Strahlenbüschels in der Bildebene erhält man unter Verwendung des geeigneten Strahlensatzes:

$$d = 2r' \frac{x''}{x'} = 2r' \frac{1-\cos\alpha}{\cos\alpha} = 2r(1-\cos\alpha) \tag{5.13}$$

Numerisch ergibt sich für den maximalen Auslenkwinkel α_{max} mit $r \approx 4\,\mathrm{mm}$ eine Aufweitung des Fokuspunktes auf:

$$d \approx 4\,\mu\mathrm{m} \tag{5.14}$$

Da in der obigen strahlenoptischen Behandlung weder durch die endliche Ausdehnung der Linse verursachte Fehler, noch die Wellennatur des Lichtes berücksichtigt wurden, ist der in Gl. (5.14) angegebene Strahldurchmesser in Wirklichkeit als Abschätzung für eine Vergrößerung des ohnehin schon ausgedehnten Lichtflecks zu verstehen. Auch dieser Effekt muß, zusammen mit der Veränderung des Strahldurchmessers durch die Linsenfehler bei der Ansteuerung von Scanner und AWG mit einbezogen werden, wenn die optische Programmierung genau sein soll.

Zum Schluß soll noch die Auflösung der Lichtfleckpositionierung beschrieben werden. Eine Auslenkung des Laserstrahls um $25^\circ \approx 0.44\,\mathrm{rad}$ entspricht gemäß Herstellerangaben[3] einer Analogspannung von 9.6 V, der maximale Auslenkwinkel $\alpha_{max} \approx 3.19^\circ$ einer Ablenkung des Strahls um 1.45 mm. Daraus ergibt sich die dem ganzen Dynamikbereich von $2\alpha_{max}$ entsprechende Spannungsdifferenz zu 2.45 V. Da für die Steuerung des Scanners einer der beiden 12-Bit-DAC's der Meßkarte mit einer externen Referenzspannung von 4 V verwendet wurde ergibt sich eine Auflösung von $1.16\,\mu$m.

5.2 Messungen mit den Positionsdetektoren

5.2.1 Reaktion des Positionsdetektors auf Eingangssignaländerungen

Bei der Diskussion der für die Positionsdetektoren entwickelten Ladungsverstärker in Kapitel 3.7 wurde davon ausgegangen, daß sich die gewünschte Ausgangsspannung binnen zweier

[3] Benutzt wurde der Scanner *S10* der Firma *Scanlab*

Taktzyklen am Ausgang der *sample and hold* Stufe (siehe Abb. 3.27) einstellt. Um das zu testen, wurde anstatt einer Lichtsignaländerung das Taktsignal für die Positionsdetektoren vor der Messung etwa 20 ms ausgesetzt und der linke Positionsdetektor mit einem Laserstrahl kontinuierlich beleuchtet. Die Taktrate betrug 5 kHz. Abb. 5.4 zeigt die mit dem Oszilloskop aufgenommene Messung. Vor dem Beginn des ersten Taktzyklus beträgt die Ausgangsspannung fast 5 V und nähert sich nach etwa 10 Taktzyklen ähnlich einer Exponentialfunktion einem Wert von etwa 2 V an.

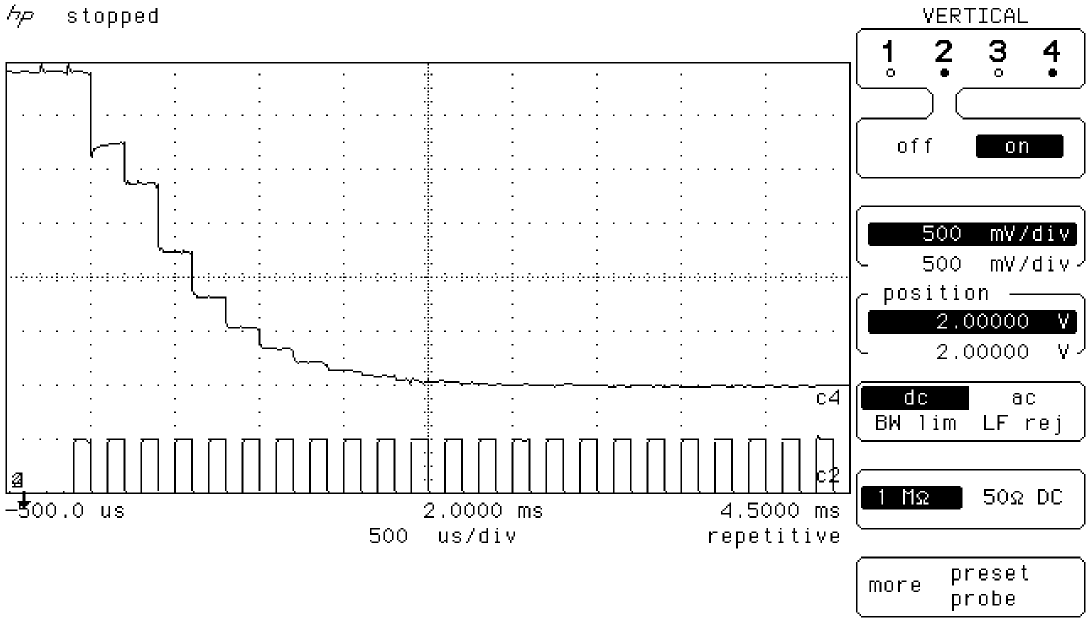

Abbildung 5.4: Oszilloskopbild des Ausgangssignals eines der Positionsdetektoren. Am unteren Bildausschnitt ist das Taktsignal für die Posititionsdetektoren zu sehen, die obere Kurve entspricht dem Ausgangssignal des Positionsdetektors. Die x-Achse entspricht einem Bereich von 3 V. Der Abstand zwischen zwei horizontalen Gitterlinien beträgt 500 mV, der zwischen zwei vertikalen 500 μs.

Wie läßt sich die Diskrepanz zwischen vorhergesagtem und gemessenem Verhalten erklären? In der Beschreibung der Ladungsverstärker wurden die Photodioden als durch die einfallende Lichtintensität einstellbare Stromquellen aufgefaßt. Tatsächlich ergibt sich für jede der 4 dreieckigen Photodioden aus den Prozeßparametern eine Kapazität von etwa 2.3 pF. Abb. 5.5 zeigt noch einmal den invertierenden Ladungsverstärker und ein Ersatzschalbild von dem linken Teil der Schaltung, das die Kapazität der Photodiode C_{photo} beinhaltet. Da der Operationsverstärker dafür sorgt, daß der rechte Anschluß von $S2$ nach dem Schließen des Schalters so schnell wie möglich das an seinem positiven Eingang anliegende Potential V_{ref} annimmt, kann er zur Betrachtung der Ladungsübertragung durch eine Spannungsquelle mit Potential V_{ref} ersetzt werden.

Gegenüber der Betrachtung aus Kapitel 3.7 ergibt die Einführung von C_{photo} zwei Änderungen: Erstens entlädt der Photostrom bei geöffnetem Schalter $S1$ die Kapazität C_{photo}, wird also während der ganzen Zeit eines Taktes auf einer der beiden Kapazitäten integriert. Das Ausgangssignal fällt daher bei gleicher Taktfrequenz um den Faktor zwei größer aus. Zwei-

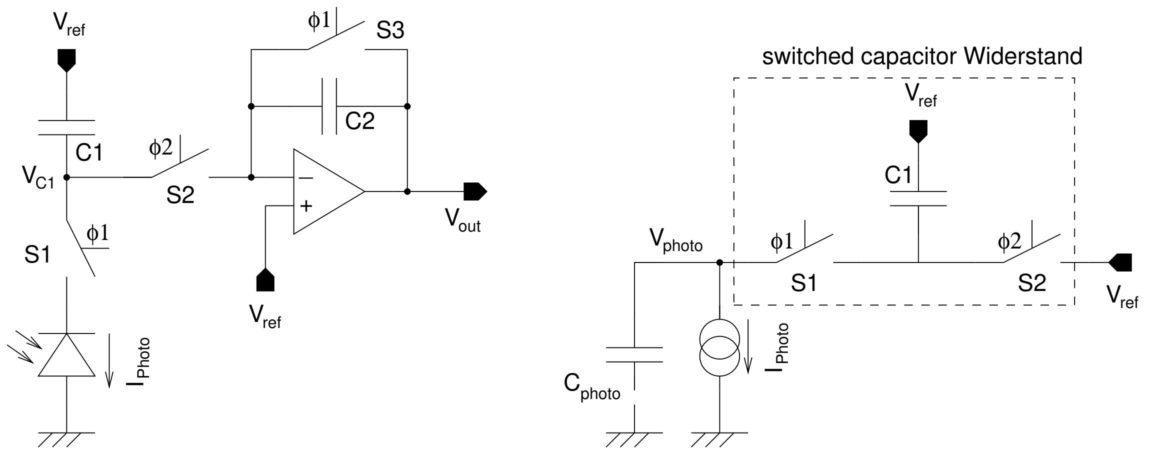

Abbildung 5.5: Links: Schaltbild des invertierenden Ladungsverstärkers. Zusätzlich zu Abb. 3.24 ist noch die Kapazität der Photodiode eingezeichnet. **Rechts:** Ersatzschaltbild des Ladungsverstärkers auf der linken Seite.

tens wirkt die Kombination von $S1$, $S2$ und $C1$ für die Spannungsdifferenz zwischen V_{photo} und V_{ref} wie ein durch geschaltete Kondensatoren realisierter Widerstand, dessen Wert durch ([Geiger 90]):

$$R = \frac{1}{C_1 f_{clock}} \tag{5.15}$$

gegeben ist. Dabei bezeichnet f_{clock} die Frequenz, mit der $\phi1$ bzw. $\phi2$ getaktet werden. Bei konstantem Photostrom und Taktsignal stellt sich ein Gleichgewichtszustand ein, dergestalt, daß für die pro Taktzyklus übertragene Ladungsmenge q gilt:

$$q = TI_{photo} \tag{5.16}$$

Dieser Gleichgewichtszustand stellt sich exponentiell mit der Zeitkonstanten

$$\tau = RC_{photo} = \frac{C_{photo}}{C_1} f_{clock} \tag{5.17}$$

ein. Ist die Zeitkonstante groß gegen die Zeit T eines Taktzyklus, kann man für die Ausgangsspannung des Ladungsverstärkers, die proportional zu der pro Taktzyklus übertragenen Ladung ist, näherungsweise schreiben:

$$V_{out} = V_{gl} + (V_0 - V_{gl})e^{-\frac{t}{\tau}} = V_{gl} + (V_0 - V_{gl})e^{-\frac{t}{T}\cdot\frac{C_1}{C_{photo}}} = V_{gl} + (V_0 - V_{gl})e^{-k\frac{C_1}{C_{photo}}} \tag{5.18}$$

Dabei bezeichnet V_{gl} die Gleichgewichtsspannung, die sich für $t \to \infty$ einstellt, V_0 die Spannung $V_{out}(t = 0)$ und k die Anzahl der vergangenen Taktzyklen. Fordert man nun, daß die Differenz zwischen dem Ausgangssignal und der zu messenden Gleichgewichsspannung kleiner oder gleich 1% der Differenz zum Zeitpunkt $t = 0$ ist, daß also gilt

$$\frac{V_{out} - V_{gl}}{V_0 - V_{gl}} = e^{-k\frac{C_1}{C_{photo}}} \leq 0.01 \tag{5.19}$$

so ergibt sich die dafür nötige Anzahl an Taktzyklen zu:

$$k = -\frac{C_{photo}}{C_1} \ln(0.01) \approx 20.77 \tag{5.20}$$

Im letzten Schritt wurde für $C_{photo} = 4.6\,\mathrm{pF}$, die Summe der Kapazitäten von zwei der dreieckigen Photodioden des Positionsdetektors, und für $C_1 = 1.02\,\mathrm{pF}$ eingesetzt (vgl. Kapitel 3.7). Das Ergebnis stimmt qualitativ und größenordnungsmäßig mit der in Abb. 5.5 gezeigten Messung überein.

Um bei den Messungen mit den Positionsdetektoren die entsprechende Genauigkeit zu erreichen, wurde zur Befriedigung von Gl. 5.20 sichergestellt, daß die Ausgangsspannung der Positionsdetektoren frühestens 20 Taktzyklen nach jeder zu erwartenden Beleuchtungsänderung oder Taktpause digitalisiert wurden.

5.2.2 Bestimmung der Lichtfleckgröße

In Kapitel 5.1.1 wurde festgestellt, daß die Größe des Lichtflecks sowohl die Genauigkeit der optischen Konfiguration der Transkonduktanz-Verstärker als auch das Übersprechen bei dieser beeinflußt. Weiterhin wurde in Kapitel 5.1.2 angedeutet, daß die Fokussierung des Laserstrahls durch eine Fülle von Effekten beschränkt ist. Insofern ist eine Abschätzung der Lichtfleckgröße einerseits für die optische Einstellung der Widerstände dringend erforderlich und ein erster Schritt zur Charakterisierung des optischen Aufbaus andererseits.

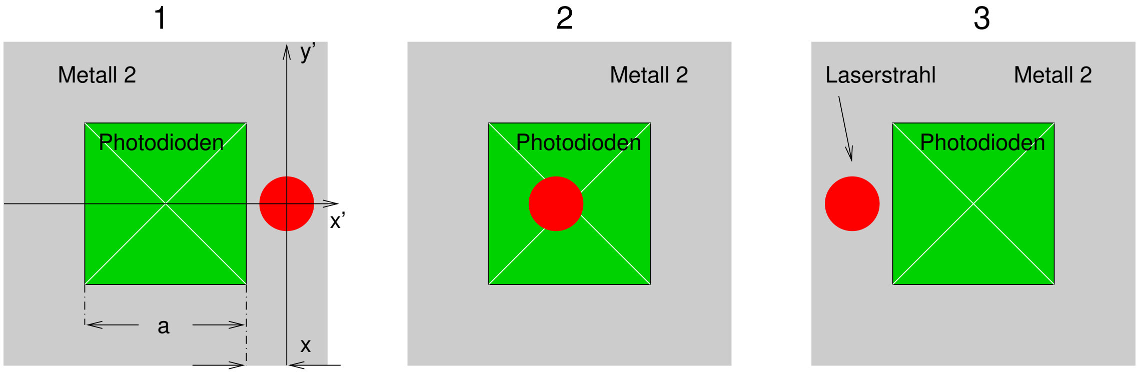

Abbildung 5.6: Position des Laserspots zu drei verschiedenen Zeitpunkten bei der Messung der Lichtfleckgröße. Der Lichtfleck ist im Vergleich zur Messung etwas vergrößert dargestellt. Im linken Bild ist noch das Koordinatensystem des Laserspots, der Abstand der rechten Kante des Quadrates zum Ursprung dieses Koordinatensystem und die Kantenlänge a des Quadrates eingezeichnet.

Die Vorgehensweise zur Bestimmung der Lichtfleckgröße ist in Abb. 5.6 angedeutet: Der Chip wird mit Hilfe des Fahrtisches so in x- oder y-Richtung verfahren, daß der Laserstrahl anfangs nur die Photodioden umgebende Metallschicht trifft, dann das ganze durch die 4 Photodioden bedeckte Quadrat durchkreuzt, und schließlich wieder ganz durch die Metallschicht auf der anderen Seite abgeschirmt wird. Dabei wird der entsprechende Positionsdetektor so betrieben, daß die Photoströme der zwei sich gegenüberliegenden Photodioden addiert werden. Die beiden Ausgangssignale für die zwei Diodenpaare werden nach der Digitalisierung addiert, so daß sich ein zur insgesamt die quadratische Fläche der Photodioden treffenden Lichtleistung proportionales Signal ergibt.

Sei nun die Lichtintensität des Laserstrahls in der Chipebene bzgl. des in Abb. 5.6 eingetragenen Koordinatensystems durch J(x',y') und der Abstand dieses Koordinatensystems von der rechten Kante des von den Photodioden aufgespannten Dreiecks gegeben. Dann gilt für die auf das Quadrat fallende Lichtleistung P_{ph}:

$$P_{ph}(x) = \int_{-(a-x)}^{x} dx' \int_{-\frac{a}{2}}^{\frac{a}{2}} dy' J(x', y') \tag{5.21}$$

Da die Ausdehnung des Lichtflecks klein ist im Vergleich zu der Fläche des Quadrates, kann man ohne großen Fehler a in den Integralgrenzen durch ∞ ersetzen. Differenziert man die Lichtleistung aus Gl. 5.21 nach x, erhält man:

$$\frac{d}{dx} P_{ph}(x) = \int_{-\infty}^{\infty} dy' J(x, y') \tag{5.22}$$

Falls sich die Lichtintensität $J(x, y)$ in einen x- und einen y-abhängigen Teil separieren läßt, was für das erwartete Gaußsche Strahlprofil möglich ist, kann man die Integration über y' ausführen und erhält für den x-abhängigen Teil der Lichtintensität:

$$J_x(x) = \frac{1}{C} \frac{d}{dx} P_{ph}(x) \tag{5.23}$$

Die Konstante C ist dabei durch das Integral

$$C = \int_{-\infty}^{\infty} dy' J_y(y') \tag{5.24}$$

gegeben und hat die Einheit $\sqrt{\mathrm{W}}$, womit sich die Einheit von J_x zu $\frac{\sqrt{\mathrm{W}}}{\mathrm{m}}$ zu ergibt.

Für die andere Seite des Quadrats gilt eine analoge Rechnung mit anderem Vorzeichen. Der Verlauf des Strahlprofils in x- oder y-Richtung kann also durch Differenzieren der auf das Quadrat fallenden Lichtleistung bzgl. x oder y gewonnen werden.

Die oben geschilderte Vorgehensweise soll im Folgenden für je einen Abtastvorgang in x- und in y-Richtung angewandt werden.

Für beide Messungen wurde der Fahrtisch mit einer Geschwindigkeit von 10 $\frac{\mu\mathrm{m}}{\mathrm{s}}$ bewegt. Die Ausgangssignale der Positionsdetektoren wurden währenddessen 10 mal pro Sekunde ausgelesen. werden. Die Rohdaten der Messung sind in Abb. 5.7 gezeigt. Beide Kurven weisen Datenpunkte auf, die von dem übrigen Verlauf der Kurve deutlich abweichen. Allerdings ist die Anzahl für die im linken Teil gezeigte Messung deutlich höher als für die im rechten Teil dargestellte. Die ausreissenden Datenpunkte lassen sich durch die fehlende Synchronisiation zwischen dem Taktsignal für die Positionsdetektoren und der Digitalisierung der Ausgangsspannungen der Posititionsdetektoren erklären. Auf Grund der endlichen Anstiegsgeschwindigkeit und Einschwingzeit des für die Ladungsverstärker verwendeten Operationsverstärkers stimmt das Ausgangssignal der Positionsdetektoren zu Beginn der *sample*-Phase ($\phi 2$ aktiv in Abb. 3.27) nicht mit dem auszulesenden Wert überein. Die Wahrscheinlichkeit dafür, den falschen Wert auszulesen, ergibt sich grob aus dem Produkt der Taktfrequenz mit der Zeit, die

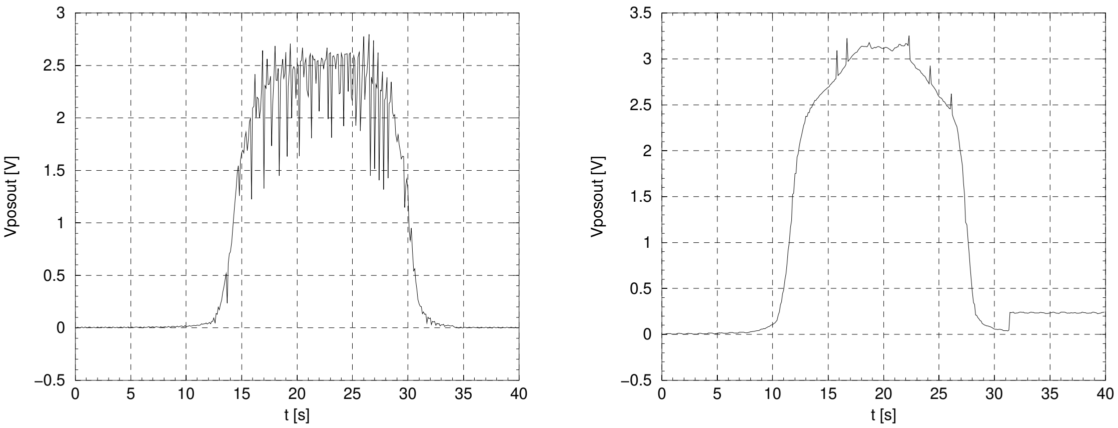

Abbildung 5.7: Summe der Ausgangssignale der Positionsdetektoren für eine Bewegung des Fahrtisches in x- (links) und y-Richtung (rechts).

der Operationsverstärker benötigt um die richtige Ausgangsspannung einzustellen. Letztere läßt sich grob durch

$$t_{falsch} = (V_{out} - V_{ref}) \;\; \mathrm{SR} \;\; + t_s \tag{5.25}$$

abschätzen. Mit den Werten aus Tabelle 3.1 ergibt sich für eine Ausgangsspannung von 4 V eine Zeit von 4 μs, in der die Auslese des Positionsdetektors den falschen Wert liefert. Die unterschiedliche Anzahl an ausreissenden Datenpunkten in den beiden Messungen ist entsprechend der obigen Betrachtung durch die unterschiedliche benutzte Taktfrequenz verursacht. Weiterhin fällt auf, daß Fehlmessungen für große Ausgangsspannungen häufiger auftreten. Das ist dadurch zu erklären, daß die benutzten Operationsverstärker für Ausgangsspannungen ab etwa 4 V wesentlich mehr Zeit benötigen, um der Eingangsspannung zu folgen.

Schließlich sollte sich für genügend kleine Ausdehnungen des Laserspots in der Mitte der gezeigten Kurven ein Plateau ausbilden. Daß das nicht der Fall ist, ist folgendermaßen zu verstehen: Wenn der Lichtfleck sich in der Nähe einer der Innenseiten des Photodiodenquadrates befindet, ergibt sich nur an einem Ausgang des Positionsdetektors eine hohe Ausgangsspannung. In den vorliegenden Messungen war die Lichtintensität im Verhältnis zur gewählten Taktfrequenz so groß, daß die Ausgangsverstärker in diesem Fall bereits einen Sättigungszustand erreichten. Befindet sich der Lichtfleck dagegen in der Mitte des Photodiodenquadrates, so teilt sich der erzeugte Photostrom auf zwei Ladungsverstärkerschaltungen auf, von denen dann keine die Sättigungsspannung erreicht.

Um trotz der schlechten Qualität der Meßdaten eine Abschätzung der Lichtfleckgröße zu erhalten, wurde wie folgt vorgegangen: Für die Auswertung der im linken Teil der Abb. 5.7 gezeigten Messung wurden zuerst alle offensichtlich nicht dem eigentlichen Kurvenverlauf entsprechenden Datenpunkte entfernt. Anschließend wurde ein neuer Datensatz erstellt, indem der neue Datenpunkt jeweils aus dem Mittelwert des Datenpunktes und der 20 benachbarten Datenpunkte errechnet wurde, um einen fortlaufenden Mittelwert zu bilden. Der resultierende Datensatz wurde schließlich differenziert und entsprechend der Gleichsetzung 1 sec = 10 μm skaliert. Um die Daten zur Bestimmung der Spotgröße in y-Richtung auszuwerten wurde ein fortlaufender Mittelwert über 10 benachbarte Datenpunkte verwendet und das Ergebnis der Differentiation wie im vorhergenden Fall skaliert. Abb. 5.8 zeigt das Ergebnis für die x- und y-Richtung. Die schwarzen Flanken stammen jeweils aus dem Ergebnis der Differentiation der

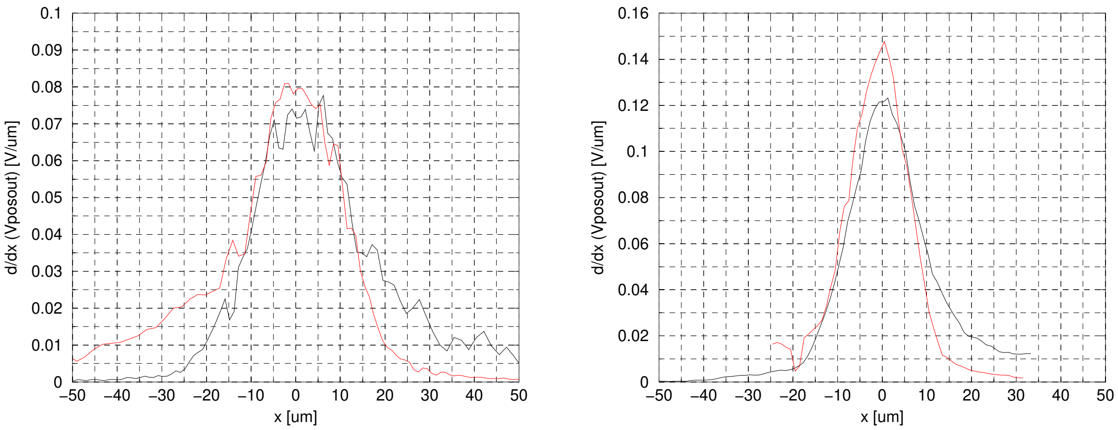

Abbildung 5.8: Aus den Kurven aus Abb. 5.7 extrahierte Verläufe der Lichtintensität in x- (links) und y-Richtung (rechts). Weitere Erklärungen im Text.

linken Flanke der dazugehörigen Kurve aus Abb. 5.7, die roten ergaben sich für die rechten Flanken. Für letztere wurde noch das Vorzeichen geändert und beide Kurven wurden jeweils so ausgerichtet, daß sich ihr Maximum auf der x-Achse befindet.

Wegen des oben diskutierten Sättigungseffektes geben die linken Hälften der schwarzen Kurven und die rechten Hälften der roten in Abb. 5.8 den wirklichen Intensitätsverlauf besser wieder als die jeweils anderen Hälften. Die so betrachteten extrahierten Intensitätprofile sind einem gaußförmigen Verlauf zumindest ähnlich. Aus den Kurven kann man auf einen Spotradius von ungefähr $15\,\mu$m in x-Richtung und $10\,\mu$m in y-Richtung schließen. Der größere Wert für die x-Richtung kann einerseits durch einen asymetrischen Laserstrahl, andererseits auch durch eine stärkere Aufweitung des Strahls in x-Richtung durch die in Kapitel 5.1.2 angedeuteten Effekte verursacht sein. Die gemessenen Lichtfleckgrößen sind nach den Überlegungen aus Kapitel 5.1.1 für eine optische Programmierung des Netzwerkes geeignet, sofern sie angesichts der Fehlerhaftigkeit der Daten sowie der durchgeführten Datenverarbeitung mit der Realität übereinstimmen.

5.2.3 Test der Modulierbarkeit der Laserdiode

Für die optische Programmierung des Bias-Stromes der Transkonduktanz-Verstärker muß die auf die einzelnen Kontrolldioden fallende Lichtmenge

$$E_{ph} = \int_0^t dt' \int_{A_{kont}} dA' J(x', y', t') \tag{5.26}$$

variiert werden (A_{kont} bezeichnet wieder die optisch aktive Fläche der Kontrolldioden). Dazu kann einerseits die Beleuchtungszeit t, andererseits die Lichtintensität $J(x, y, t)$ benutzt werden.

Die Lichtleistung des benutzten Diodenlasers[4] kann durch Anlegen einer Modulationsspannung moduliert werden. Soll die optische Konfiguration des Netzwerkes durch eine Variation

[4]Es handelt sich um das Modell *DLS 15-675-10 mod* der Firma *Spindler & Hoyer*.

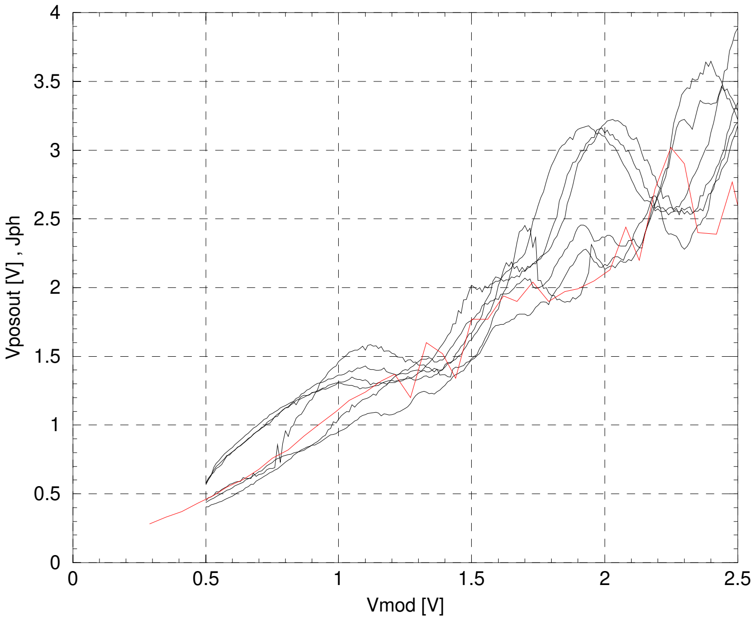

Abbildung 5.9: Intensität des von dem Diodenlaser emittierten Lichtes in Abhängigkeit von der Modulationsspannung. Die schwarzen Kurven wurden mit den Positionsdetektoren aufgenommen, die rote mit einem Photometer. Um die beiden Kurven vergleichen zu können, wurde die rote Kurve entsprechend skaliert.

der Lichtintensität erreicht werden, muß man den Zusammenhang zwischen der Modulationsspannung und der vom Laser emittierten Lichtleistung kennen. Dieser wurde daher mehrmals mit Hilfe der Positionsdetektoren und einmal mit Hilfe eines Photometers[5] vermessen. Die Messungen wurden mit dem in Abb. 5.1 gezeigten Aufbau durchgeführt. Für die Messung mit dem Photometer wurde dieses zwischen den Neutralfiltern und dem Raumfilter in den Strahlengang gebracht. Das Ergebnis ist in Abb. 5.9 dargestellt.

Der Zusammenhang zwischen der Modulationsspannung und der gemessenen Lichtintensität ist nicht monoton und nur mit einer Genauigkeit von etwa 40% reproduzierbar. Da der mit dem Photometer aufgenommene Zusammenhang einen ähnlichen Verlauf aufweist wie der mit den Positionsdetektoren gemessene, sind die Abweichungen von dem erwarteten linearen Verlauf mit großer Wahrscheinlichkeit auf das vom Laser emittierte Licht zurückzuführen.

5.3 Optische Konfiguration des Widerstandsnetzwerkes

Zur Einstellung der Widerstandswerte G des Netzwerkes muß auf jede Photodiode eine definierte Lichtmenge aufgebracht werden. Dazu wird der Laserstrahl mit Hilfe des Galvanometer-Scanners in x-Richtung abgelenkt. Aus Gl. 5.26 geht hervor, daß zur Variation der Lichtmenge zwei Parameter zur Verfügung stehen: Die Lichtintensität und die Zeit, in der die Photodiode beleuchtet wird.

Prinzipiell bieten sich für die Bewegung des Scanners nun zwei Bewegungsformen an:

[5]Benutzt wurde eine Kombination aus einem *J18 LumaColor II Photometer* und dem Tastkopf *J1812* der Fimra *Tektronix*. Die relative Genauigkeit wird vom Hersteller mit 5% angegeben.

Entweder der Scanner verharrt auf jeder der 64 Kontrolldioden eine bestimmte Zeit, die zusammen mit der während dieser Zeit gewählten Lichtintensität die aufgebrachte Lichtmenge bestimmt, oder er wird so bewegt, daß der Laserstrahl die 64 Photodioden mit konstanter Geschwindigkeit überstreicht. Bei Benutzung der letzteren Methode kann die aufzubringende Lichtmenge nur durch die Lichtintensität eingestellt werden. Ein Scan, bei dem jede Photodiode einzeln angesteuert wird, ist gegenüber dem mit kontinuierlicher Ablenkgeschwindigkeit zeitaufwendiger, da der Spiegel vor jeder Photodiode abgebremst, und dahinter wieder beschleunigt werden muß.

Da die für einen Scan benötigte Zeit für eine Übertragung des Prinzips auf eine größere Anzahl zu programmierender Bauelemente möglichst klein sein sollte, wurden die im verbleibenden Teil des Kapitels geschilderten Messungen ausschließlich mit konstanter Scangeschwindigkeit durchgeführt.

5.3.1 Meßablauf

Der Ablauf der Messungen mit optischer Konfiguration ist in Abb. 5.10 dargestellt. Damit

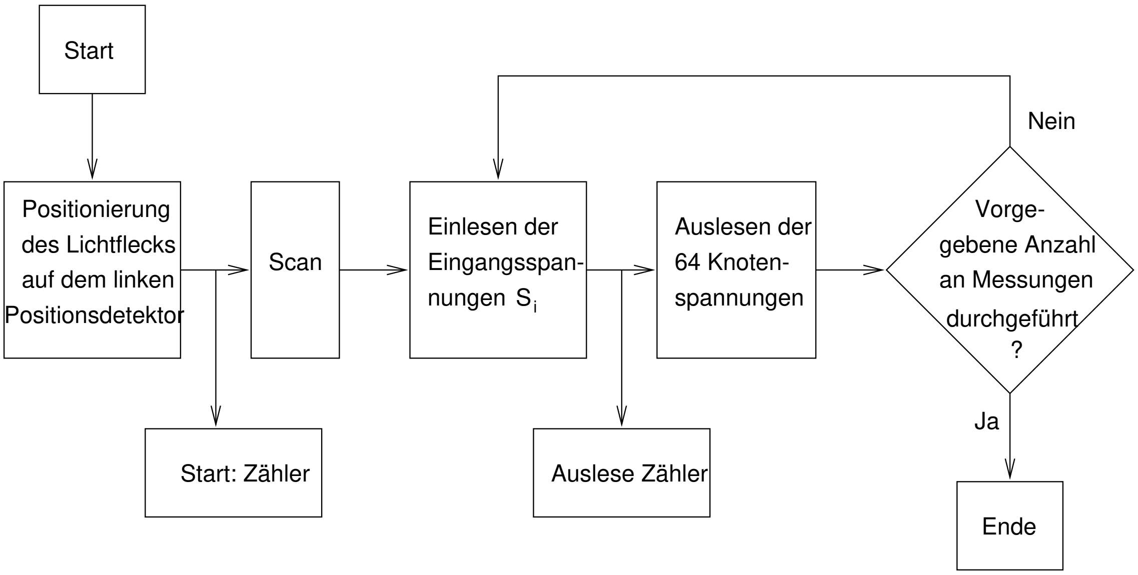

Abbildung 5.10: Ablauf der Messungen mit optischer Konfiguration des Netzwerkes.

die x-Position zu Beginn eines Scanvorgangs genau definiert ist, wird der Laserstrahl zuerst in x-Richtung auf der kleeblattförmigen Photodiodenanordung des Positionsdetektors zentriert. Dazu wird die Spannung zur Steuerung des Scanners mit einem modifizierten Newton-Verfahren so lange verändert, bis das der Differenz der Photoströme in x-Richtung proportionale Ausgangssignal einen vorgegebenen Wert unterschreitet. Gleichzeitig mit dem Beginn des Scans wird ein Zähler gestartet, der vor jeder Messung der Ausgangsspannungen ausgelesen wird. Der Zähler ist Teil der Meßkarte und wird von dieser getaktet. Dadurch kann die Zeit, die bis zu einer bestimmten Messung der Knotenspannungen seit dem Beginn des Scans vergangen ist, gemessen werden. Nach der Beendigung des Scanvorgangs werden die Eingangsspannungen S_i in die Analogspeicherzellen geschrieben. Vor der Auslese der Knotenspannungen wird durch die Auslese des Zählers die Zeit, die seit dem Beginn des Scans bzw. seit der letzten Auslese der Knotenspannungen vergangen ist, gemessen. Die Anzahl der Schreib- und Lesezyklen kann im Meßprogramm vorgegeben werden.

Um die optische Konfigurierbarkeit des Netzwerkes zu testen, sollte der Leitwert des Transkonduktanzverstärkers, der zu Knoten 31 gehört, größer eingestellt werden als der aller anderen. Dazu wird die Spannung, die die Auslenkung des Scanners kontrolliert wie im oberen Teil von Abb. 5.11 gezeigt mit konstanter Änderungsgeschwindigkeit erhöht, um den Laserstrahl mit konstanter Geschwindigkeit über die Kontrolldioden zu bewegen. Die Spannungen V_{links}

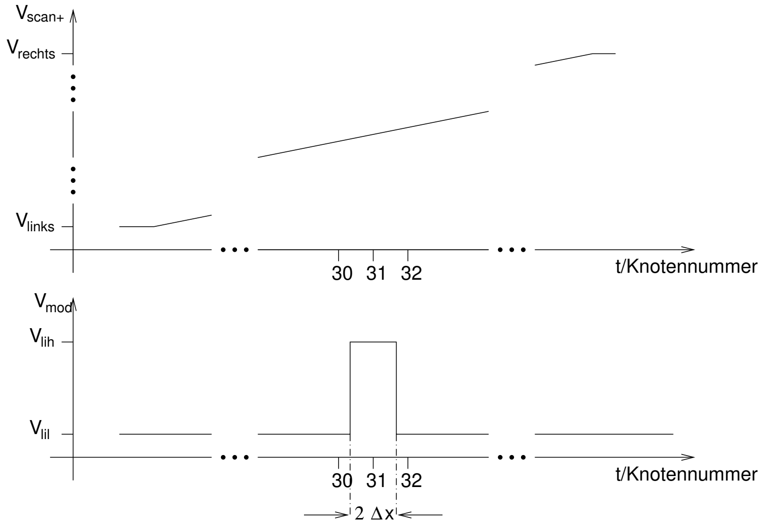

Abbildung 5.11: Zeitlicher Verlauf der Spannungen V_{scan+} (oben) und V_{int} (unten), die die Position des Scanners und die Lichtintensität bestimmen, bei der Durchführung eines Scans.

und V_{rechts} sind die Spannungen, für die sich der Lichtfleck in der Mitte der kleeblattfömig angeordneten Photodioden des linken bzw. rechten Positionsdetektors befindet. Die Spannung V_{mod}, die die Lichtintensität des Lasers kontrolliert, wird dabei wie im unteren Teil von Abb. 5.11 dargestellt variiert, ist also fast für den ganzen Scan mit V_{lil} identisch und nimmt nur in einem kleinen Zeitfenster, das der Distanz $2\Delta x$ auf dem Chip entspricht, den Wert V_{lih} an.

Der Zusammenhang zwischen der Position des Laserspots auf dem Chip und der Zeit, in der die Laserintensität moduliert wird, wird durch die Synchronisation der durch den AWG erzeugten Spannung V_{mod} und der Spannung V_{scan+}, die durch einen DAC der Meßkarte geliefert wird, gestiftet. Dazu werden die Zeitachsen für die beiden zu erzeugenden Spannungsverläufe geeignet skaliert und die Erzeugung dieser bei jedem Scan gleichzeitig gestartet. Der Einfachheit halber wurde dabei von einem linearen Zusammenhang zwischen der Spannung V_{scan+} und damit dem Ablenkwinkel des Scanners und der Position des Laserspots auf dem Chip ausgegangen, also alle in Kapitel 5.1.2 erwähnten unerwünschten Effekte der Fokussierlinse vernachlässigt.

Für das die Position der Kontrolldiode 31 einschließende Intervall $2\Delta x$, für das $V_{mod} = V_{lih}$ sein sollte, wurde der Wert $30\,\mu$m gewählt. Mit dem in Kapitel 5.2.2 ermittelten Spotradius von $15\,\mu$m sollte also ein Übersprechen des für den Knoten 31 bestimmten Intensitätssignals festzustellen sein.

Die Sprungantwortzeit auf einen Sprung der Eingangssignalspannung um 10% des Signalshubes, der zum Vollauschlag führte, wird vom Hersteller mit 2.4 ms angeben. Die für einen Scan benötigte Auslenkung beträgt $2\alpha_{max} \approx 7.4°$, also etwa 15% vom Vollausschlag des Scanners. Um nun während des Scans eine möglichst gleichmäßige Geschwindigkeit des Laserspots zu erreichen, wurde die Zeit für einen Scan mit 20 ms deutlich größer als die vom Hersteller angegebene Mindestzeit gewählt.

5.3.2 Ergebnis

Der oben geschilderte Meßablauf wurde für $V_{lih} = 2.5\,\mathrm{V}$ und $V_{lil} = 0.8$ und $1\,\mathrm{V}$ durchgeführt. Der Bias-Strom I_{biasR} für die Hres-Widerstände betrug dabei 198 nA. Als Eingangsmuster dienten 0.5 V große *räumliche* Spannungsspitzen. Die zwei Kurven mit dem großen Dynamikbereich in Abb. 5.12 gehören zu Messungen bei denen der Eingang 30 des Netzwerkes auf einem um 0.5 V höheren Potential lag als alle anderen Eingänge. Für die Messung

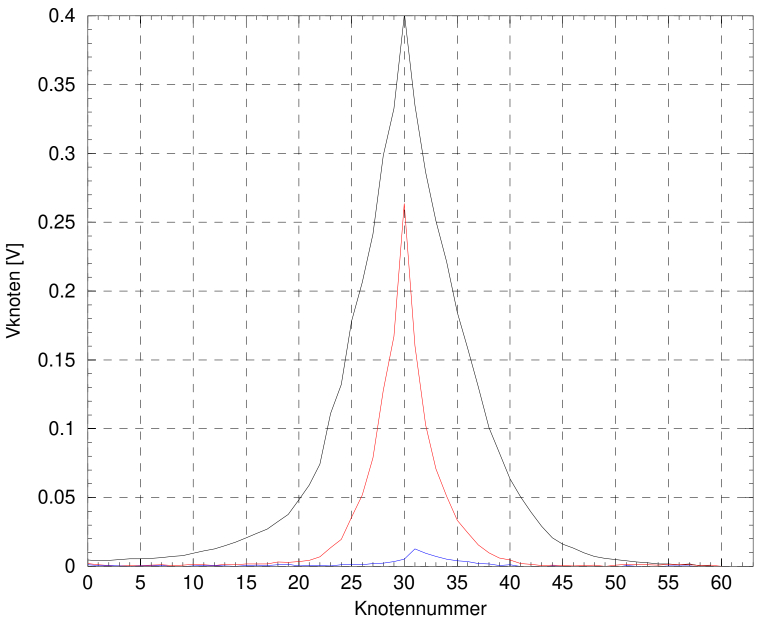

Abbildung 5.12: Reaktion des Netzwerkes auf 0.5 V große *räumliche Spannungsspitzen* nach einer optischen Konfiguration der in Kapitel 5.3.1 beschriebenen Form für $V_{lih} = 2.5\,\mathrm{V}$ und $V_{lil} = 0.8\,\mathrm{V}$ (schwarze Kurve) bzw. 1 V (farbige Kurven). Bei der blauen Kurve zu Grunde liegenden Messung wurde die Spannungsspitze an Eingang 31, bei der für die Messung der anderen Kurven an Eingang 30 angelegt.

der blauen Kurve wurde nur der Knoten 31 durch eine 0.5 V große Spannungsspitze erregt.

Die Tatsache, daß der höher gewichtete Eingang nicht wie beabsichtigt der des Knotens 31, sondern derjenige des Knotens 30 war, läßt sich mit dem Schleppfehler des Laserscanners erklären, der laut Herstellerangabe bei einer konstanten Ablenkgeschwindigkeit von 5 rad weniger als 3 mrad betragen sollte. Die durchschnittliche Geschwindigkeit für die durchgeführten Scans ergibt sich zu:

$$\omega = \frac{2\alpha_{max}}{20\,\mathrm{ms}} \approx 5.6\,\frac{\mathrm{rad}}{\mathrm{s}} \tag{5.27}$$

Wenn man die 3 mrad, die der Schleppfehler maximal betragen darf, gemäß Gl. 5.3 auf eine

Strecke auf dem Chip umrechnet, erhält man 78 μm. Daß der Schleppfehler bei den gemachten Messungen ziemlich genau die Hälfte des erlaubten Wertes betrug, also genau den Abstand zweier Kontrolldioden ausmachte, ist als glücklicher Umstand zu werten.

Die schwarze und die rote Kurve aus Abb. 5.12 weisen einen ähnlichen Verlauf auf wie die in Abb. 3.16 dargestellten, nur daß die rote Kurve ein deutlich niedrigeres Maximum besitzt. Die Messung liefert also qualitativ genau das gewünschte Ergebnis, namentlich eine deutlich höhere Gewichtung eines Eingangssignales gegenüber den restlichen.

Um zu untersuchen, inwieweit das Netzwerk auf die beschriebene Stimulation mit einem exponentiellen Abfall reagiert, wurden die Meßdaten in Abb. 5.13 auch noch logarithmisch aufgetragen. Um die Linearität der sich ergebenden Kurven besser beurteilen zu können,

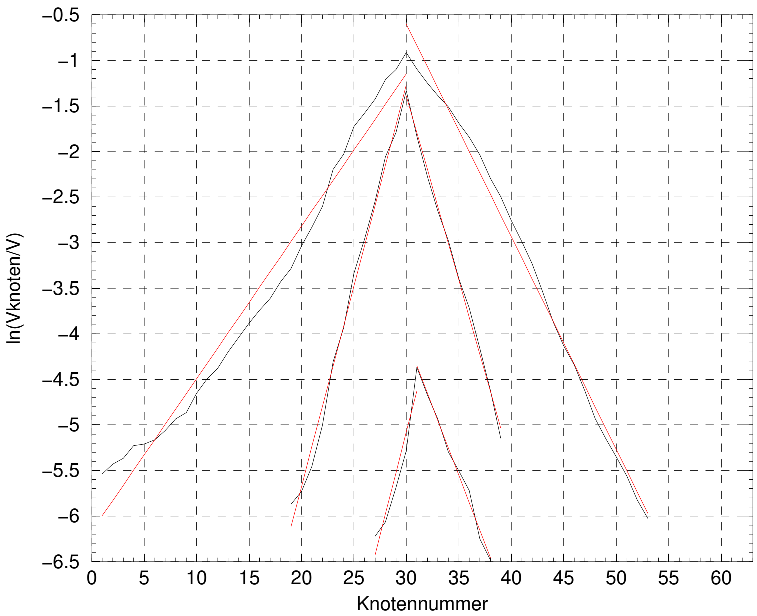

Abbildung 5.13: Logarithmische Darstellung der in Abb. 5.12 dargestellten Daten (schwarz). Zur Einschätzung der Linearität sind auch noch die Regressionsgeraden eingetragen.

wurde für alle Kurvenhälften eine lineare Regression durchgeführt. Dazu wurden nur die in der Abbildung gezeigten Datenpunkte benutzt. Die Kurven weisen im Mittel einen linearen Verlauf auf, schwanken aber um die Regressionsgeraden. Inwieweit diese Schwankungen das *fixed pattern noise* der Widerstände des Netzwerks widerspiegeln oder auf Ungenauigkeiten bei der optischen Konfiguration zurückzufüren sind, kann anhand der vorliegenden Daten nicht entschieden werden.

5.3.3 Zeitliches Verhalten nach der optischen Konfiguration

Für eine mögliche Anwendung des vorgestellten Konzeptes der optischen Konfiguration ist die Zeitspanne, in der die optisch eingestellten Spannungen (V_C in Abb. 3.11) und damit im untersuchten System die Leitwerte G konstant sind, von entscheidender Bedeutung. Daher wurde unter denselben Bedingungen, die der Messung der schwarzen Kurve aus Abb. 5.12 zu Grunde liegen, das Netzwerk nach der optischen Konfiguration für 500 s immer wieder mit dem Eingangsmuster beschrieben und anschließend ausgelesen. In Abb. 5.14 sind die ausgelesenen Ausgangsspannungsmuster in einer 3-d Graphik gegen die Zeit aufgetragen.

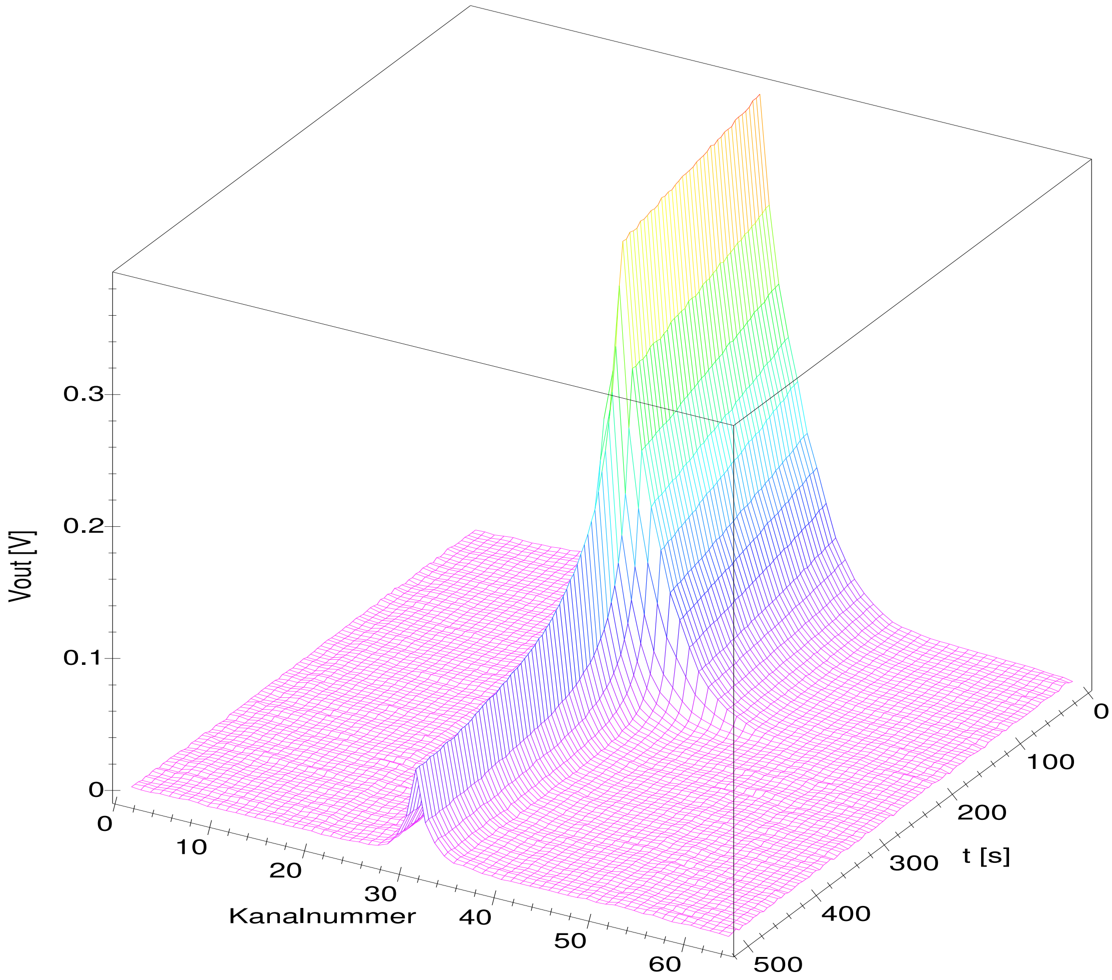

Abbildung 5.14: Dreidimensionale Darstellung des Ausgangsmusters des Netzwerkes aufgetragen gegen die Zeit.

In den ersten 200 s ist das Ausgangsmuster nahezu konstant. Da die Auslese des Zählers, der die Zeit zwischen den Messungen wie in Abb. 5.10 illustriert messen sollte, einige Probleme bereitete, kann ausgeschlossen werden, daß die als 200 s aufgetragene Zeitspanne in Wirklichkeit nicht nur ein Zehntel also, 20 s, betrug. Die Zeitangaben für die letzten 300 s sollten aber zumindest der Größenordnung nach stimmen.

Die Antwort des Netzwerkes entspricht also zu Beginn der Messung der aus Abb. 5.12. Am Ende ist die Antwort schmaler und kleiner geworden und entspricht qualitativ eher der Situation für eine kleine Resetspannung bei elektrischer Einstellung der Leitfähigkeit G der Transkonduktanzverstärker. Das ist damit zu erklären, daß die Kontrolldioden Kraft ihres Dunkelstromes und des durch Restlicht erzeugten Photostromes auch die Resetspannungen der anderen Knoten verringern und damit die Leitfähigkeit der Transkonduktanzverstärker erhöhen. Da sich die Resetspannung von Knoten 30 nach der optischen Konfiguration vermutlich schon in der Nähe des Massepotentials befindet, nimmt der Unterschied zwischen der Leitfähigkeit des Transkonduktanz-Verstärkers von Knoten 30 und der der restlichen Transkonduktanz-Verstärker mit der Zeit ab.

Aus dem zeitlichen Verhalten des Netzwerkes auf dasjenige der einzelnen Spannungen der Kondensatoren zur Speicherung der optischen Infomation zu schließen, dürfte schwierig sein. Die Zeitskala, auf der sich die Veränderungen der Ausgangssignale des Netzwerkes abspielen, ist aber mit den Messungen aus Kapitel 2, zumindest was die Größenordnung betrifft, grob in Einklang zu bringen. Außerdem legt die Messung nahe, daß das Netzwerk zumindest noch für Zeiträume in der Größenordnung von Sekunden die optisch eingestellte Konfiguration beibehält und also Eingangsmuster während dieser Zeit entsprechend der gewählten Konfiguration verarbeitet werden können.

Zusammenfassung und Ausblick

Im Rahmen der Diplomarbeit wurde ein optisch-elektronisches System zur optischen Konfiguration eines in CMOS-Technologie implementierten Widerstandsnetzwerkes entworfen. Das System wurde bis zur Testreife entwickelt und die prinzipielle Funktionalität durch erste Messungen nachgewiesen.

Um das Konzept der optischen Speicherung zu testen, wurden Teststrukturen auf einem ersten Testchip *VisionTest96* vermessen. Das Konzept erwies sich insofern als tragfähig, als daß Kondensatoren durch den Laserstrahl aufgeladen werden können und die Spannung für mindestens einige 100 ms bis auf ein Prozent konstant bleibt. Eine optische Konfiguration des Netzwerkes im in der Arbeit diskutierten Sinn und eine anschließende Datenverarbeitung durch das konfigurierte Netzwerk sollte in dieser Zeitspanne möglich sein. Allerdings ergab sich nur für eine der untersuchten Schaltungen der theoretisch vorhergesagte zeitliche Verlauf der Kondensatorspannung. Die großen Unterschiede im Verhalten der verschiedenen Schaltungen können nicht erklärt werden. Außerdem konnte die Diffusionslänge der freien Ladungsträger im Substrat des vermessenen Chips auf etwa 80 μm experimentell abgeschätzt werden.

Das realisierte optisch konfigurierbare Netzwerk besitzt 64 Knoten und wurde in einem 0.6 μm-Prozeß der Firma *AMS* implementiert. Als Widerstände wurden Transkonduktanz-Verstärker und die von C. Mead ([Mead 89]) entwickelte Hres-Schaltung verwendet. Abgesehen von den Transistoren in der Schaltung, die den Bias-Strom für die Transkonduktanz-Verstärker erzeugen, arbeiten alle Transistoren des Widerstandsnetzwerkes im Subthreshold-Bereich oder im Bereich zwischen starker und schwacher Inversion. Das sorgt dafür, daß sich die Schaltung für einen großen Bereich von Bias-Strömen ähnlich verhält, schränkt aber den Dynamikbereich ein und erhöht den *transistor mismatch* zwischen einzelnen Komponenten des Netzwerkes, führt also insgesamt zu einem schlechteren *'signal to fixed pattern noise ratio'*.

Im Rahmen der Meßgenauigkeit und der erreichten nachträglichen Offsetkorrektur zeigte das Netzwerk bei Stimulation mit räumlichen Spannungsspitzen die erwartete lokale Mittelung des Eingangssignals, also namentlich eine mit dem Abstand vom erregten Knoten exponentiell abnehmende Knotenspannung, in guter Übereinstimmung mit den Ergebnissen der Simulation.

Um einen Zusammenhang zwischen dem Koordinatensystem des Laserscanners und demjenigen des Chips herzustellen, wurde ein Positionsdetektor entwickelt, mit dem der Laserspot iterativ auf einen ausgezeichneten Punkt auf dem Chip ausgerichtet werden kann. Dazu werden Photoströme nach dem Prinzip der Ladungsverstärkung wahlweise subtrahiert oder addiert. Mit Hilfe der Positionsdetektoren konnte der Intensitätsverlauf des Laserstrahl in der Chipebene in x- und y-Richtung gemessen werden. Die sich ergebenden Intensitätsprofile sind – im Rahmen der Verläßlichkeit der genommenen Daten – sehr gut mit der theoretischen Vorhersage und den Resultaten der Messungen zur optischen Konfiguration des Netzwerkes vereinbar. Eine Messung des der Differenz der Photoströme proportionalen Ausgangssignals wurde nicht explizit in der Arbeit vorgestellt. Die erfolgreich durchgeführte optische Konfiguration des Netzwerkes impliziert aber, daß die Positionsdetektoren ihre Aufgabe gut erfüllen. Außerdem deuten die bei der Justierung des Systems gemachten Erfahrungen darauf hin, daß die Position des Lichtflecks mindestens auf einen Mikrometer genau justiert werden kann, die Positioniergenauigkeit in dem benutzten Aufbau also durch den Fahrtisch bzw. den Laserscanner beschränkt ist.

Die Messungen an dem Positionsdetektor haben weiterhin gezeigt, daß beim Entwurf der Ladungsverstärker die Kapazität der Photodioden unzulässigerweise vernachlässigt wurde. Deswegen darf der Positionsdetektor nach jeder Änderung der Eingangsströme erst nach etwa 20 Taktzyklen ausgelesen werden. Für das entwickelte System stellt das aber keine Einschränkung dar, da die Zeit zur Auslese der Positionsdetektoren klein ist im Verhältnis zu der Zeit für einen Scanvorgang.

Das Widerstandsnetzwerk wurde schließlich mit einem einfachen Testmuster optisch konfiguriert. Qualitativ reagierte das Netzwerk wieder wie von den Simulationen vorausgesagt, die prinzipielle Realisierbarkeit der optischen Konfiguration ist damit gezeigt. Das Verhalten des Netzwerkes nach der optischen Konfiguration ist für Zeitspannen in der Größenordnung einer Sekunde in guter Näherung stabil, so daß das Netzwerk nach der optischen Konfiguration für einen solchen Zeitraum zur Verarbeitung von Eingangsmustern benutzt werden kann.

Software, externe Elektronik und optischer Aufbau wurden soweit entwickelt, daß das System jetzt quantitativ untersucht werden könnte. Das Gesamtsystem ist sowohl durch die genaue optische Justierung (oder Dejustierung) und die Fehler bei der Abbildung des Laserstrahls als auch durch die Offsetvariationen des Widerstandsnetzwerkes charakterisiert. Nur wenn alle diese Einflüsse korrigiert werden können, läßt sich das Netzwerk mit hoher Genauigkeit konfigurieren. Wie sich bei dem Versuch, die Offsetvariationen der einzelnen Widerstände des Netzwerkes zu ermitteln, gezeigt hat, ist eine genaue Kenntnis dieser Variationen nicht immer leicht zu erlangen. Weiterhin wird es im allgemeinen schwierig sein, eine einmal gewählte optische Justierung beizubehalten. Aus diesen Gründen scheint die Implementierung einer automatischen, rechnergesteuerten Kalibration des Systems sowohl für eine genaue quantitative Analyse des Netzwerkes selbst als auch im Hinblick auf eine mögliche Anwendung des Konzeptes dringend erforderlich. Eine solche automatische Kalibration könnte möglicherweise durch die Verwendung neuronaler Netze oder genetischer Algorithmen in der Meßsoftware erreicht werden.

Hinsichtlich der Anwendbarkeit des vorgestellten Systems ist einzuwenden, daß die Programmierung von 64 Spannungen elektrisch mit weit weniger Aufwand um Größenordnungen schneller durchgeführt werden kann, als durch die gewählte Methode der Ablenkung des Laserstrahls mittels mechanischer Bewegung. Insofern müßte der parallelen Verarbeitung der Eingangsspannungen auch eine parallele Konfiguration des Netzwerkes gegenüberstehen, um mit der Performance herkömmlicher Mikroelektronik konkurrieren zu können.

Literaturverzeichnis

[Allen 90] P. E. Allen, D. R. Holberg, *„CMOS Analog ciruit design"*, Oxford University Press, New York, Heidelberg 1987

[AMS 0.8] *„0.8 μm CMOS Process Parameters"*, Document#: 9933006, Revision#: A, Austria Mikro Systeme International AG, Unterpremstätten, Östereich, 1994

[Cad] *„Cadence Online Library Openbook: Spectre Reference manual 4.4.1"*, Cadence Design Systems, San Jose, USA

[Geiger 90] R. L. Geiger, P. E. Allen, N. R. Strader, *„VLSI Design techniques for analog and digital circuits"*, Addison-Wesley, McGraw-Hill, Inc. 1990

[Hertz 91] J. Hertz, A. Krogh, R. G. Palmer, *„Introduction to the theory of neural computation"*, Addison-Wesley, 1991

[Klein 86] M. V. Klein, T. E. Furtak *„Optik"*, Springer Verlag Berlin, Heidelberg, New York, 1986

[Kobayashi 91] H. Kobayashi, J. L. White, A. A. Abidi, *„An Active Resistor Network for Gaussian Filtering of Images"*, IEEE J. Solid-State Circuits, vol. 26, no. 5, pp.738-748, May 1991

[Laker 94] K. R. Laker, W. M. C. Sansen *„Design of analog integrated circuits and systems"*, McGraw-Hill, Inc. 1994

[Loose 96] M. Loose, *„Layout und Test eines Systems adaptiver Photorezeeptoren in analoger CMOS-Technologie"*, Diplomarbeit im Institut für Hochenergiephysik, 1996, Universität Heidelberg

[Mead 89] C. A. Mead, *„Analog VLSI and neural systems"*, Addison Wesley, 1991

[Moini 97] A. Moini, *„Vision chips or seeing silicon"*, WWW page at URL: <http://www.eleceng.adelaide.edu.au/Groups/GAAS/Bugeye/visionchips/>, Last update April 1997 [Accessed 6 Sep 1998]

[Rojas 93] R. Rojas, *„Theorie der neuronalen Netze: Eine systematische Einführung"*, Springer Verlag Berlin, Heidelberg, New York, 1993

[Ritter 90] H. Ritter, T. Martinetz, K. Schulten, *„Neuronale Netze: Eine Einführung in die Neuroinformatik selbstorganisierter Netzwerke"*, Addison Wesley, 1991

[Sze 85] S. M. Sze, *„Semiconductor Devices, physics and technology“*, John Wiley & Sons, 1985

[Theodori 97] M. Theodori, *„Design und Messung optoelektronischer Strukturen in CMOS Technologie“*, Diplomarbeit im Institut für Hochenergiephysik, 1997, Universität Heidelberg

Danksagung

Herzlichen Dank allen, die zur Durchführung meiner Diplomarbeit beigetragen haben, insbesondere

Herrn Prof. K. Meier dafür, daß er diese Arbeit ermöglicht hat und für seine vorbildliche Betreuung,

Herrn Prof. U. Straumann für die Übernahme der Zweitkorrektur,

Herrn M. Keller für die Betreuung der Arbeit, vor allem in technischen Fragen, für die kritische Korrektur des Manuskriptes und insbesondere für seine Unterstützung in der letzten Nacht vor der Submisison des Chips,

Herrn D. Droste für die ausführliche und bereitwillige Beantwortung zahlreicher Fragen, die kritische Durchsicht des Manuskriptes und insbesondere für die hervorragende Einführung in die Programmierung der seriellen Schnittstelle,

Herrn M. Loose für seine tatkräftige Hilfe bei vielen Problemen, die kritische Durchsicht des Manuskriptes, sowie die Hilfe beim Editieren einiger Postscript Dateien,

Herrn T. Maucher für zahlreiche praktische Hilfestellungen, seine enorme Hilfsbereitschaft und für seine hervorragende Latex-Einführung mit Beispieldiplomarbeit,

Herrn J. Schemmel für die sehr effektive Hilfe sowohl beim Debuggen der Software als auch beim Entwurf der externen Elektronik und die Beantwortung vieler Fragen zum Chipentwurf,

Herrn L. Schmidt-Mende für die sorgfältige Korrektur des Manuskripts und seine aufopfernde Hilfe beim Drucken der Arbeit,

Herrn J. Boelsems und Herrn A. Achenbach für ihre praktische Unterstützung im Meßlabor,

allen weiteren Mitgliedern des ASIC-Labors für die Hilfsbereitschaft und die stets angenehme Arbeitsatmospähre,

allen Institutsmitarbeitern, die einen reibungsfreien technischen und organisatorischen Ablauf ermöglicht haben,

meinen Eltern für die umfassende Unterstützung während meines ganzen Studiums.