

Inhaltsverzeichnis

Inhaltsverzeichnis.....	1
1. Fotodioden.....	1
1.1 Aufbau und Arbeitsweise einer Si Fotodiode	2
1.2 Das Kennlinienfeld der Fotodiode	4
1.3 Ersatzschaltungen zur Verhaltensbeschreibung	7
1.4 Simulation zur Bandbreite und zum Rauschen	8
1.4.1 Rauschstrom und Signalstrom und deren Überlagerung.....	8
1.4.2 Beispiel für sehr kleine Bandbreite	9
1.4.3 Beispiel für sehr hohe Bandbreite	9
1.5 Spektrale Empfindlichkeit und NEP	10

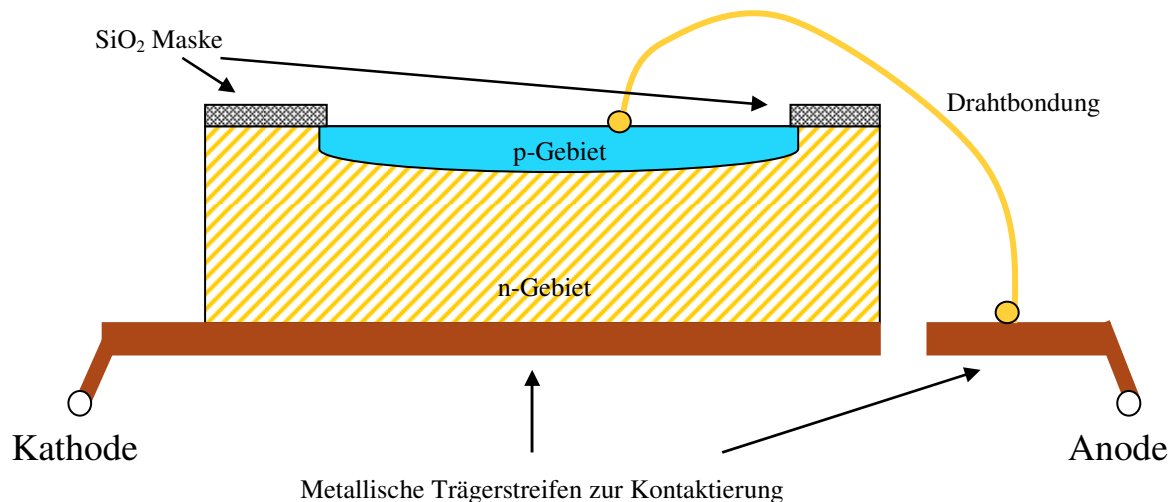
1. Fotodioden

Es wird der Aufbau von Si Fotodioden und deren Funktion beschrieben. Es folgen ein Ersatzschaltbild zur Verhaltensbeschreibung und eine Simulation zur Veranschaulichung des Signal-Rauschverhältnisses bei verschiedenen Bandbreiten der Anordnung. Ein Rechenbeispiel zur NEP (Noise Equivalent Power) schließt die kurze Einführung ab.

1.1 Aufbau und Arbeitsweise einer Si Fotodiode

Der prinzipielle Aufbau ist im Querschnitt in der folgenden Abbildung dargestellt:

Abb.1



Das p-Gebiet ist verglichen mit dem Aufbau einfacher Si-Schaltioden sehr dünn ausgeführt um ein gutes Eindringen des Lichtes bis hinab zum pn-Übergang zu ermöglichen. Eine hohe Absorption der Strahlung wird angestrebt.

Im reinen Elementhalbleiter (z.B. reines Silizium ohne Dotierung) kommt es bei Raumtemperatur durch den geringen Bandabstand von kleiner 2eV zu sporadischen Wechseln des Energieniveaus einzelner Valenzelektronen ins Leiterband. Diesen Vorgang nennt man Kombination. Es entsteht dabei ein Überschusselektron im Leiterband und ein fehlendes Elektron genannt Defektelektron am Spenderatom. Durch die stochastischen Bewegungen aller Leiterband-Elektronen im Kristallgitter wird die fehlende Besetzung am Spenderatom irgendwann auch wieder durch ein anderes Elektron ergänzt und damit ein Potentialausgleich angestrebt. Diesen Vorgang nennt man Rekombination.

Es stellt sich am undotierten Halbleiter ein statistisches Gleichgewicht von Rekombinationen und Kombinationen ein. Die Häufigkeit dieser beiden Vorgänge steigt mit der Temperatur oder auch mit der Zuführung von Lichtenergie. Messbar ist es am ohmschen Widerstand des betrachteten Halbleiterblocks. Er sinkt mit steigender Temperatur.

Durch die zusätzlich eindotierten Atome dritter Hauptgruppe im p-Gebiet wird das beschriebene Gleichgewicht zugunsten der Defektelektronen gezielt geändert. Diese Atome nennt man Akzeptoren, da sie in der Lage sind noch weitere Elektronen aufzunehmen, zu akzeptieren. Im n-Gebiet wird das Gleichgewicht zugunsten der Elektronen geändert durch eindotierte Atome der fünften Hauptgruppe. Diese nennt man Donatoren, da sie zusätzliche Elektronen abgeben können (lateinisch: donare = geben). Durch die Donatorelektronen werden Energieniveaus besetzt, die dem reinen Halbleiter nicht zugänglich sind.

Werden nun ein p- und ein n-Gebiet aneinandergebracht, wie in der Schnittzeichnung dargestellt entsteht sofort ein Ladungsträgerausgleich an der Nahtstelle des Kontaktes p zu n. Dabei ziehen die Akzeptoratomkerne des p-Gebietes die Überschusselektronen aus dem n-Gebiet zu sich herüber. Umgekehrt werden Defektelektronenstellen in der Nähe der Kontaktstelle p-n weggezogen vom p- ins n-Gebiet. Diese Ladungsträgerbewegung nennt man Diffusionsstrom. Er entsteht durch das Ladungsträgerkonzentrationsgefälle vom p zum n Gebiet. Im Ergebnis entsteht eine Verarmung von Ladungsträgern an der Kontaktfläche p-Gebiet zu n-Gebiet. Da sich die Defektelektronen im n-Gebiet anreichern und die Elektronen im p-Gebiet und zwischen Ihnen eine Ladungsträgerarme Zone entsteht bildet sich durch die gegenüberliegenden unterschiedlichen Ladungsträger hier ein elektrisches Feld aus. Die zugehörige verarmte Zone nennt man Raumladungszone (RLZ). Durch den Ladungsträgerabfluss verbleiben im Kristallgitter ortsfeste Akzeptorionenrümpfe und Donatorionenrümpfe und es baut sich durch deren gegenüberliegende Anordnung wieder ein entgegengesetztes Feld auf, das Diffusionsfeld. Letzteres treibt dann einen Strom in entgegengesetzter Richtung den Driftstrom. In der Folge stellt sich am unbelasteten pn-Übergang ein Gleichgewicht zwischen Driftstrom und Diffusionsstrom ein. Es gilt:

$$I_{Drift} = I_{Diff} \quad (I)$$

Wichtig ist, dass durch die Diffusionsfeldstärke und die Ladungsträgerverarmung an der Kontaktfläche p zu n ein Potentialwall entsteht, der immer erst durch eine angelegte zusätzliche Spannung, die Schleusen- oder Flussspannung, überwunden werden muss, bevor es zu einem nennenswerten Stromfluss in Durchlassrichtung an der Diode kommen kann. Die Durchlassrichtung ist dann geschaltet, wenn der Minuspol am n-Gebiet (Kathode) liegt und der Pluspol am p-Gebiet (Anode).

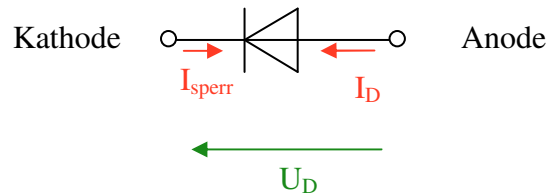
Eine Sperrpolung hingegen trägt zur Verbreiterung der anfangs beschriebenen RLZ bei, da dann die äußere zusätzliche Feldstärke die Diffusionsfeldstärke erhöht. Mit steigender Sperrspannung verbreitert sich die Verarmungszone immer mehr, bis sie irgendwann den gesamten Kristall durchsetzt und bis an die Anschlussmetallisierungen heranreicht. In diesem Fall erreicht die Sperrschichtkapazität den minimal möglichen Wert der dann von der Gehäuse- und Anschlussgeometrie und dem Dotierungsgrad von p und n bestimmt wird.

Die Funktion der PN- und auch der PIN- Fotodiode basiert auf dem inneren Fotoeffekt. Durch die dünne Auslegung des p-Gebietes kann das Licht, also der Photonenstrom tief in die Kristallgitterschichten der Fotodiode eindringen. Bei zur Bandlücke passender Wellenlänge des einfallenden Lichtes werden so zusätzliche Kombinationen verursacht. Da die RLZ arm an Ladungsträgern war und somit kein nennenswerter Strom floss, führen die durch Lichteinfall generierten neuen Ladungsträger sofort zu einem nennenswert höheren Strom. Das ist der Fotostrom, der über mehrere Dekaden proportional mit der optischen einfallenden Leistung ansteigt.

1.2 Das Kennlinienfeld der Fotodiode

Am Zweipol Fotodiode werden zu Beginn folgende Definitionen gemacht:

Abb. 2:



Der nichtlineare Zusammenhang zwischen dem Zweipolstrom und der Zweipolspannung sei wie folgt gegeben:

$$I_D = I_d \left(e^{\frac{U_D}{U_T}} - 1 \right) - I_{Ph} \quad (\text{II})$$

Das ist die Gleichung von William B. Shockley zur Beschreibung des Verhaltens von Halbleiterdioden ergänzt um den lichtabhängigen Fotostrom I_{Ph} .

I_d ist der Dunkelstrom (Sättigungsstrom).

Der Ausdruck U_T ist die Temperaturspannung und definiert sich wie folgt:

$$U_T = \frac{k \cdot T}{e} \quad (\text{III})$$

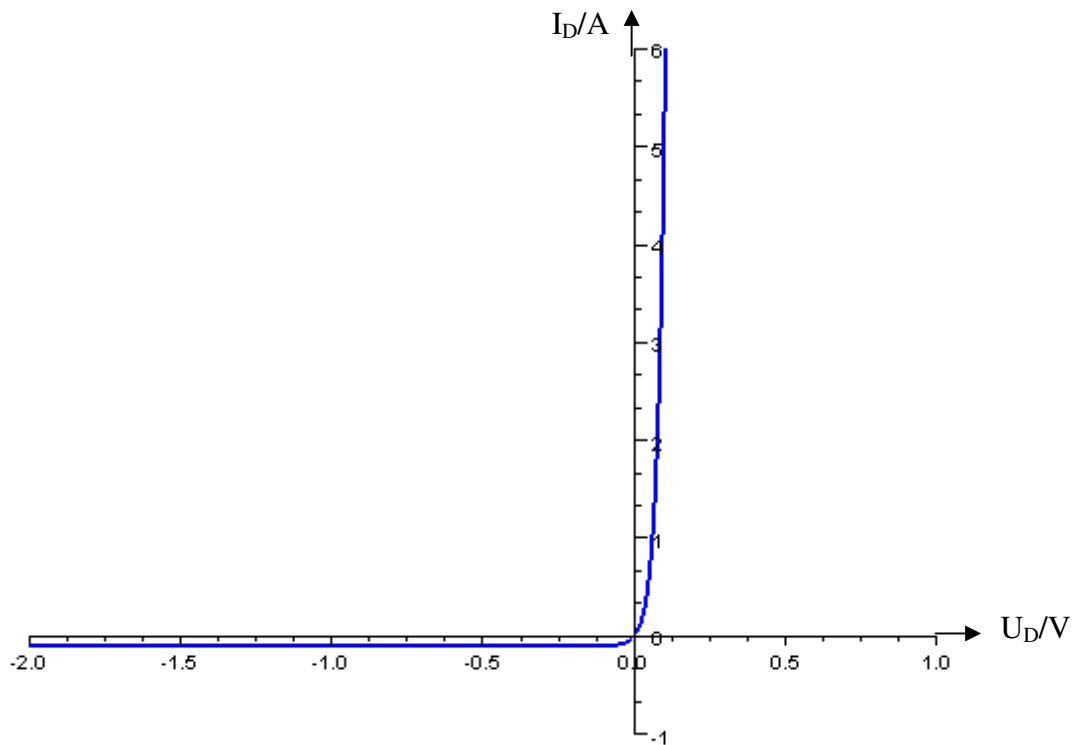
In (3) sind T die absolute Temperatur in Kelvin, e die Elementarladung und k ist die Boltzmannkonstante.

Für eine Raumtemperatur von $T=25^\circ\text{C}=298\text{K}$ und $k=8,617343 \cdot 10^{-5} \text{ eV/K}$ erhält man:

$$U_T = \frac{k \cdot T}{e} = \frac{8,617343 \cdot 10^{-5} \text{ eV} \cdot 298\text{K}}{e \cdot K} = \underline{\underline{25,68\text{mV}}} \quad (\text{III})$$

Mit II und III lässt sich jetzt die Funktion $I_D = f(U_D)$ beispielhaft für einen Dunkelstrom von 1nA und T=298K grafisch darstellen:

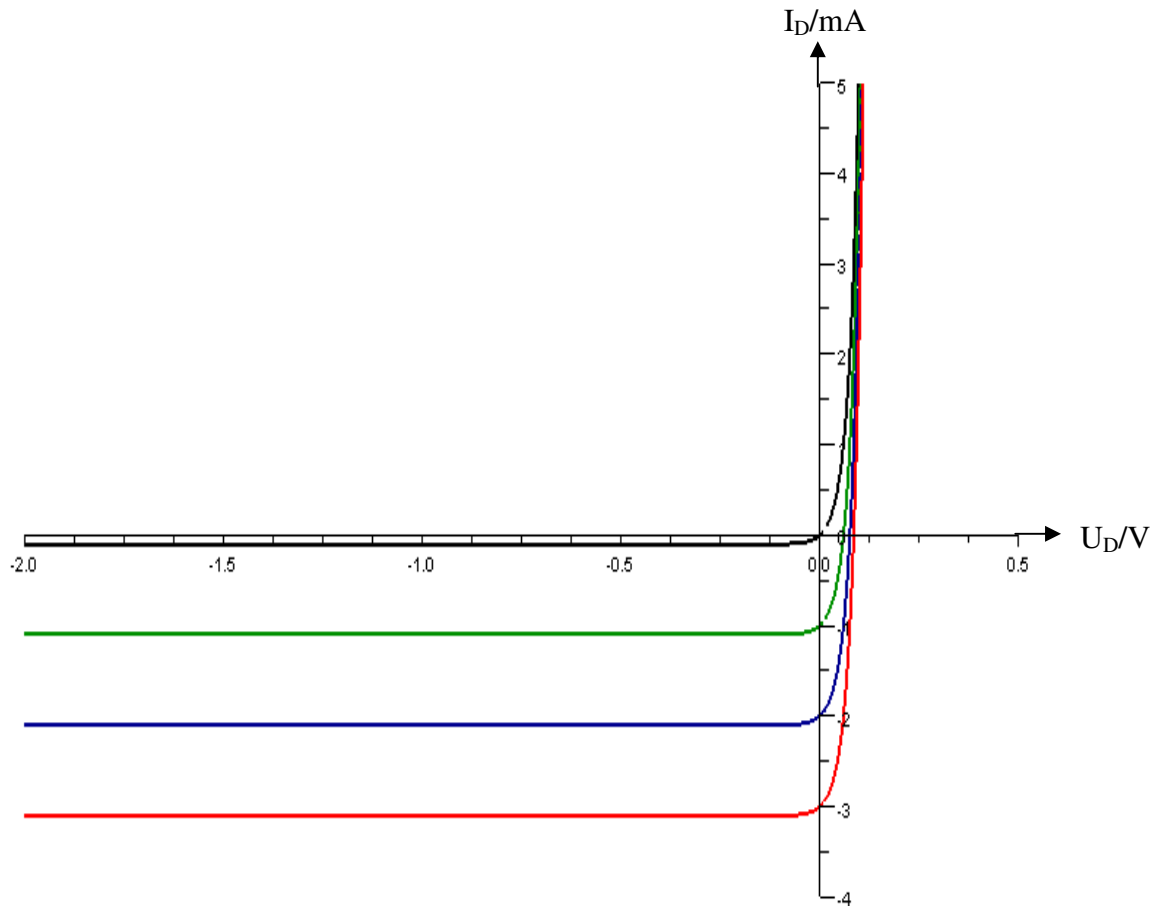
Abb.3:



Aus (II) ist weiterhin ersichtlich, dass der Diodenstrom im Dunkelfall 0 ist, wenn die Diodenspannung genauso groß wie die Temperaturspannung ist. Durch die begrenzte Auflösung des Graphen könnte man annehmen, dass die Spannung 0V beträgt bei $I=0A$.

Für drei verschiedene Fotoströme größer 0A erhalten wir folgende Funktionsverläufe:

Abb. 4:



Der Lichteinfall steuert den Sperrstrom und führt mit steigendem Fotostrom zu einer senkrechten Verschiebung des durch die Shockley Gleichung beschriebenen Verlaufes nach unten.

Variiert zum Beispiel in einer optischen Übertragungsstrecke die optische Leistung einer sendenden Leuchtdiode im Bitstrom zwischen zwei Werten, dann würde das den zwei Fotostromverläufen grün und rot entsprechen.

Das gilt für den Fall, dass die Fotodiode negativ vorgespannt ist also in Sperrrichtung betrieben wird. Der nahezu waagerechte Verlauf dieser Kennlinien legt die Modellierung dieses Verhaltens in Form einer idealen Stromquelle nahe. Wie man im Diagramm sieht, bleibt der Strom über einen weiten Sperrspannungsbereich nahezu konstant. Diese Spannungsunabhängigkeit ist eine Eigenschaft der idealen Stromquelle.

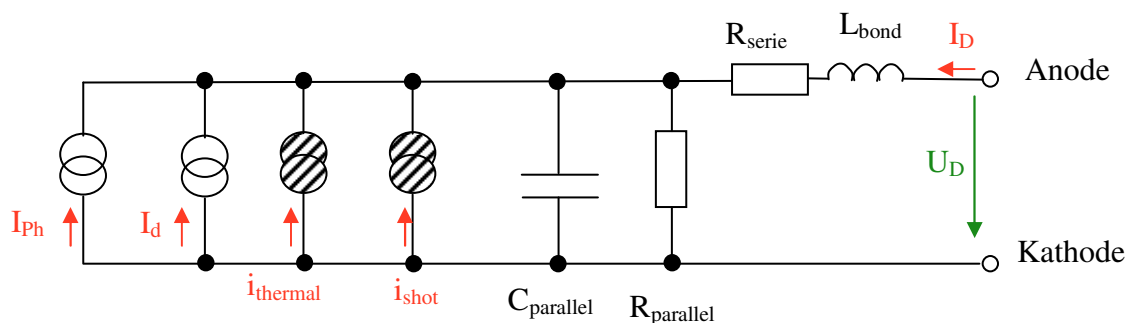
1.3 Ersatzschaltungen zur Verhaltensbeschreibung

Eine Ersatzschaltung bestehend aus idealen passiven Grundzweipolen, die da wären R, L, C, Stromquelle und Spannungsquelle beschreibt die verschiedenen Verhaltensaspekte für die meisten Anforderungen genau genug. Durch die Beschränkung auf diese 5 Grundelemente hält sich außerdem der mathematische Aufwand in Grenzen.

In der Ersatzschaltung beginnen wir mit einer idealen Stromquelle zur Modellierung des variierenden Fotostromes. Eine weitere Stromquelle kann den nahezu konstanten Dunkelstromverlauf, wie er in Abb. 2 ersichtlich ist nachbilden.

Außerdem gibt es zwei Ursachen für Stromrauschen an den Klemmen der Fotodiode. Diese werden mit zwei weiteren Rauschstromquellen getrennt beschrieben. Ein Serienwiderstand modelliert die Bahnwiderstände der realen Fotodiode. Der Bonddraht lässt sich in erster Näherung gut mit einer kleinen Längsinduktivität ausreichend genau nachbilden und spielt bei hohen Datenraten eine wichtige Rolle.

Die in 1.1. erwähnten gegenüberliegenden Ladungsträgeransammlungen und die dazwischen befindliche Verarmung an Ladungsträgern wird mit einer parallel liegenden Kapazität abgebildet:



Die mathematische Formulierung des thermisch verursachten Stromrauschens ist gegeben mit:

$$\overline{i_{thermal}^2} = 4 \cdot k \cdot T \cdot \frac{1}{R_{parallel_effektiv}} \cdot B \quad (IV)$$

Die Formel stammt von Henry Nyquist (1928).

In (IV) ist B die Bandbreite der betrachteten Anordnung und $R_{Parallel_effektiv}$ die Zusammenfassung aller vorhandenen parallelen Wirkwiderstände die am Zweipol anliegen. Dies sind im Normalfall der Lastwiderstand in Reihe zu R_{Serie} und der $R_{Parallel}$ der Fotodiode.

Das Schrotrauschen (engl. shot noise) ist gegeben mit:

$$\overline{i_{shot}^2} = 2 \cdot e \cdot I_{Sperr} \cdot B = 2 \cdot e \cdot (I_{ph} + I_d) \cdot B \quad (V)$$

Es ist Sperrstromabhängig und entsteht durch Vorgänge am pn-Übergang. In (V) steht e wieder für die Elementarladung.

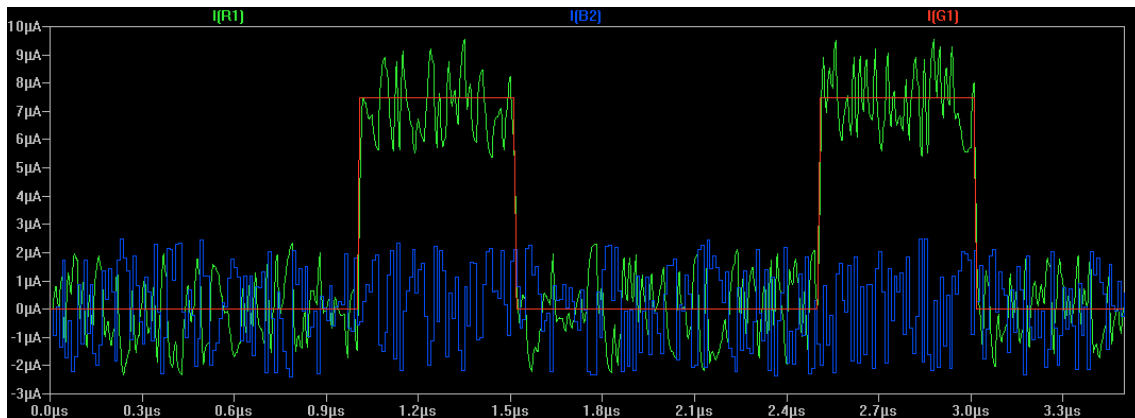
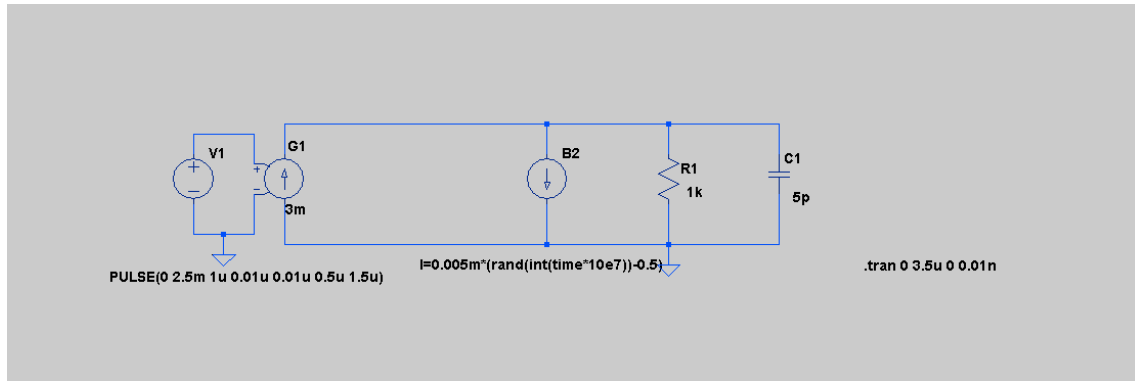
Mit dem nun vorliegenden Modell ist die Fotodiode und ihr Verhalten für unsere Zwecke ausreichend genau beschrieben.

Anmerkung: Für viele Betrachtungen, die nur der Funktion dienen reicht oft auch das rauschfreie Modell aus oder es können ohne viele Einbußen einige Elemente entfallen. Ist der Dunkelstrom beispielsweise sehr klein gegenüber dem Fotostrom, dann kann dessen Quelle entfallen da ihr Beitrag vernachlässigbar ist usw.

1.4 Simulation zur Bandbreite und zum Rauschen

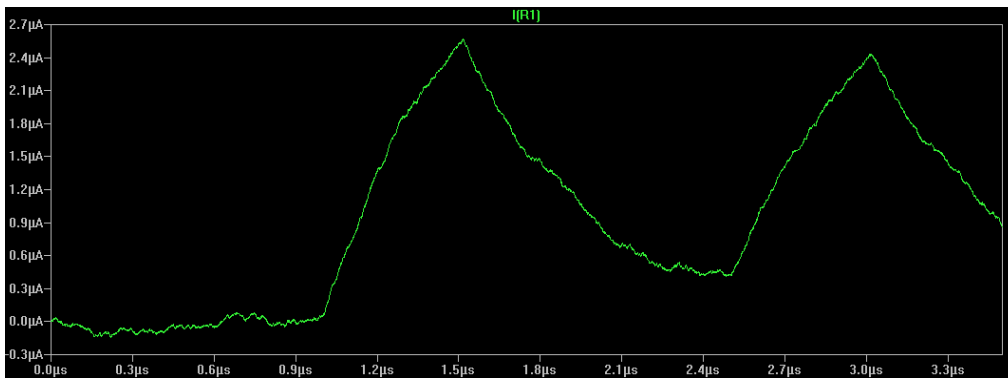
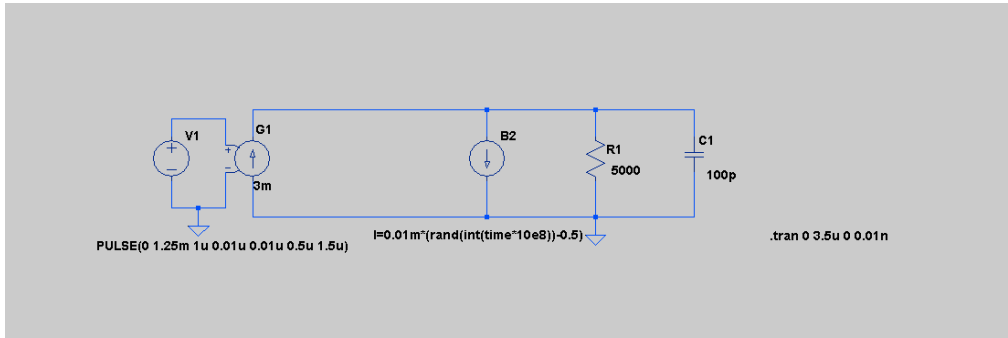
Die folgenden 3 Fälle für die Lastwiderstandsauslegung und Sperrschichtkapazitätsvariation durch zum Beispiel Sperrspannungsänderung sollen aufzeigen wie wichtig ein guter Kompromiß zwischen der hohen Bandbreite für viele Daten einerseits und einer kleinen Bandbreite zur Rauschbegrenzung andererseits ist. Zur Simulation wurde der frei verfügbare Schaltungssimulator LT Spice verwendet. Experimentieren Sie mal damit!

1.4.1 Rauschstrom und Signalstrom und deren Überlagerung



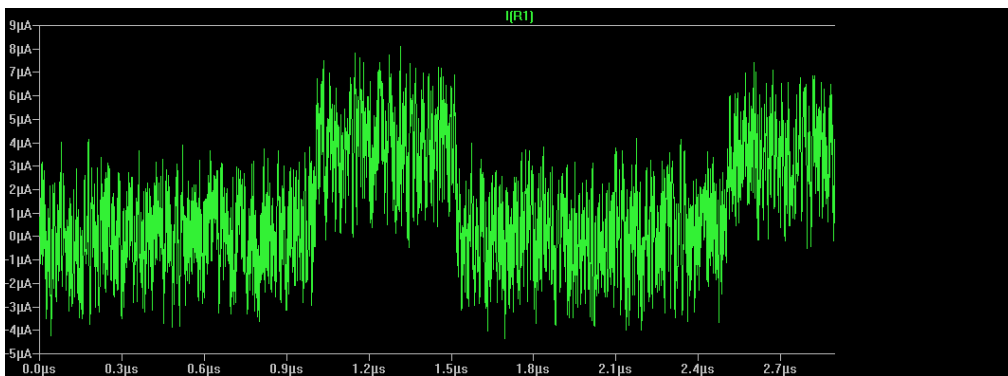
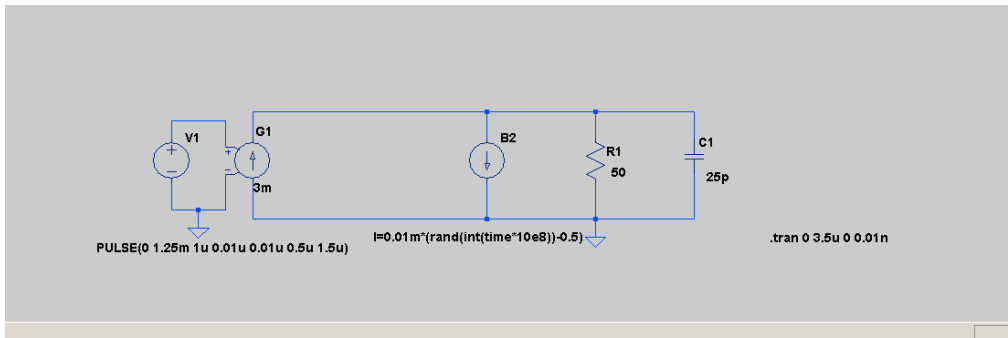
Die Modellierung des Rauschens wurde durch eine Zufallszahlengeneration vereinfacht realisiert. Die Signalquelle ist ein Pulsgeneratormodell, das eine spannungsgesteuerte Stromquelle treibt. Dies bildet den durch optische Leistung variierten Fotostrom nach.

1.4.2 Beispiel für sehr kleine Bandbreite



Ergebnis: kein Rauschen aber zu langsam für den Datenstrom durch viel zu kleine Bandbreite.

1.4.3 Beispiel für sehr hohe Bandbreite



Ergebnis: Steile Flanken und dadurch sehr schnell aber das Rauschen kommt in die Größenordnung des Signals! Bald kommt die Dynamikgrenze....

1.5 Spektrale Empfindlichkeit und NEP

Zum Abschluss der Einführung noch ein Rechenbeispiel zur Ermittlung der Grenzemfindlichkeit und damit der unteren Dynamikgrenze einer Fotoempfängeranwendung.

Aus dem Herstellerdatenblatt der Fotodiode werden dafür zwei wichtige Größen benötigt.

Eine ist die NEP und die andere die spektrale Empfindlichkeit R_{SP} .

NEP steht für Noise Equivalent Power. Damit ist die optische Leistung gemeint, mit der die Diode bestrahlt werden muss um an Ihren Klemmen genauso viel Signalleistung wie Rauschleistung zu erzeugen. Es ist quasi der Punkt, oberhalb dessen die Signalleistung anfängt größer als die Rauschleistung zu werden. Da die Rauschleistung proportional mit der Bandbreite ansteigt und letztere vom System abhängt in dem die Diode zur Anwendung kommt wird die NEP auf Wurzel 1 Hertz normiert angegeben.

$$NEP = \frac{P_{opt\ min}}{\sqrt{B}} = \left[\frac{W}{\sqrt{Hz}} \right] \quad (VI)$$

Die spektrale Empfindlichkeit sagt etwas über den Wirkungsgrad aus. Es ist das Verhältnis von erzeugtem Strom und einfallender aufgewendeter optischer Leistung:

$$R_{SP} = \frac{I_{ph}}{P_{opt}} = \left[\frac{A}{W} \right] \quad (VII)$$

Die Berechnung beginnt mit einem Leistungsvergleich. Es werden eine PIN Fotodiode mit Sperrspannung und ihr Lastwiderstand R_L betrachtet.

Die Signalleistung ergibt sich zu:

$$P_{Sig} = I_{ph}^2 \cdot R_L \quad (VIII)$$

Die Rauschleistung ergibt sich zu:

$$P_{Noise} = i_N^2 \cdot R_L \quad (IX)$$

Das Verhältnis der Leistungen soll 1 sein, so wie es für die NEP definiert ist:

$$\frac{P_{Sig}}{P_{Noise}} = \frac{I_{ph}^2 \cdot R_L}{i_N^2 \cdot R_L} = 1 \quad (X)$$

Gesucht ist jetzt die Größe der notwendigen optischen Leistung für diesen Grenzfall. (VII) wird umgestellt und in (X) eingesetzt. R_L wird gekürzt und das Stromrauschen wird durch seine beiden verschiedenen Ursachen beschrieben, die zueinander unkorreliert angenommen werden:

$$\frac{P_{Sig}}{P_{Noise}} = \frac{I_{ph}^2}{i_N^2} = \frac{(R_{SP} \cdot P_{opt\ min})^2}{i_{thermal}^2 + i_{shot}^2} = 1 \quad (X)$$

Umstellung nach $P_{opt\ min}$ und (IV) und (V) eingesetzt in (X) ergibt:

$$(P_{opt\ min})^2 = \frac{1}{R_{SP}^2} \cdot (i_{thermal}^2 + i_{shot}^2) = \frac{1}{R_{SP}^2} \cdot \left(4kTB \cdot \frac{1}{R_L} + 2 \cdot e \cdot I_{Sperr} \cdot B \right) \quad (X)$$

Die minimale optische Leistung ergibt sich damit wie folgt:

$$P_{opt\ min} = \frac{1}{R_{SP}} \cdot \sqrt{B} \cdot \sqrt{4kT \cdot \frac{1}{R_L} + 2 \cdot e \cdot I_{Sperr}} \quad (X)$$

Für den Fall, dass das thermische Rauschen um vieles größer als das Schrotrauschen ist kann man folgern, dass eine Vervielfachung des Lastwiderstandes ungefähr eine Halbierung der benötigten optischen Leistung verursacht.

Zum Abschluss der Betrachtungen ein Beispiel für eine Dimensionierung:

Gegeben seien:

Gesucht: $P_{opt\ min}$

$$R_{SP} = 0,6 \cdot \frac{A}{W}$$

$$I_{Sperr} = I_d = 1nA$$

$$R_L = 50\Omega$$

$$T = 298K \quad \text{und} \quad C_{Sperr} = 15pF$$

$$P_{opt\ min} = \frac{1}{0,6} \frac{W}{A} \cdot \sqrt{1Hz} \cdot \sqrt{4 \cdot 1,3806504 \cdot 10^{-23} \cdot \frac{Ws}{K} \cdot 298K \cdot \frac{1 \cdot A}{50 \cdot V} + 2 \cdot 1,6021765 \cdot 10^{-19} \cdot As \cdot 1nA} \quad (X)$$

$$P_{opt\ min} = \frac{1}{0,6} \frac{W}{A} \cdot \sqrt{1Hz} \cdot \sqrt{329,147 \cdot 10^{-24} A^2s + 320,435 \cdot 10^{-30} \cdot A^2s} \quad (X)$$

Das Ergebnis für 1 Hertz Bandbreite ist:

$$\frac{P_{opt\ min}}{\sqrt{1Hz}} = 30,23pW \cdot \sqrt{s} \quad (X)$$

Für eine angenommene Sperrkapazität von 15pF erhält man unter der Annahme, dass ein Tiefpaß erster Ordnung vorliegt eine Bandbreite von:

$$B = \frac{1}{2\pi \cdot R_L \cdot C_{Sperr}} = \frac{1}{2\pi \cdot 50\Omega \cdot 15pF} = 212MHz \quad (XI)$$

Dadurch benötigt man wegen der größeren Rauschbandbreite folgende minimale optische Leistung:

$$P_{opt\ min} = 30,23pW \cdot \sqrt{212 \cdot 10^6 Hz} \cdot \sqrt{s} = \underline{\underline{440nW}} \quad (XI) \text{ in } (X)$$

In (XI) ist zu erkennen, dass ein kleiner Lastwiderstand gut für eine große Bandbreite ist. In (X) wiederum war zu sehen, dass er schlecht für das Rauschen ist. Hier ist für das System der beste Kompromiss zu finden. Siehe dazu auch die Beispielsimulation in 1.4.