

PABLO DUTRA DA SILVA

**MODELO COMPACTO DE NÃO-LINEARIDADES
EM TRANSISTORES MOS**

**Florianópolis
2006**

UNIVERSIDADE FEDERAL DE SANTA CATARINA

**PROGRAMA DE PÓS-GRADUAÇÃO
EM ENGENHARIA ELÉTRICA**

**MODELO COMPACTO DE NÃO-LINEARIDADES
EM TRANSISTORES MOS**

Dissertação submetida à
Universidade Federal de Santa Catarina
como parte do requisitos para a
obtenção do grau de Mestre em Engenharia Elétrica.

PABLO DUTRA DA SILVA

Florianópolis, Setembro de 2006.

MODELO COMPACTO DE NÃO-LINEARIDADES EM TRANSISTORES MOS

Pablo Dutra da Silva

‘Esta dissertação foi julgada adequada para a obtenção do Título de Mestre em Engenharia Elétrica, Área de Concentração em *Circuitos e Sistemas Integrados*, e aprovada em sua forma final pelo Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal de Santa Catarina.’

Prof. Carlos Galup-Montoro, Dr.
Orientador

Prof. Fernando Rangel de Sousa, Dr.
Co-orientador

Prof. Nelson Sadowski, Dr.
Coordenador do Programa de Pós-Graduação em Engenharia Elétrica

Banca Examinadora:

Prof. Carlos Galup-Montoro, Dr.
Presidente

Prof. Marcio Cherem Schneider, Dr.

Prof. Wilhelmus A. M. van Noije, Dr.

Prof. Walter Pereira Carpes Junior, Dr.

*Aos meus pais e
à Alessandra.*

AGRADECIMENTOS

Expresso minha sincera gratidão:

Primeiramente a Deus, que é a força maior que nos impulsiona para superar o que nos parece um obstáculo intransponível.

Agradeço ao CNPQ e ao Laboratório de Circuitos Integrados pelo fundamental suporte financeiro e estrutural utilizado para a realização deste trabalho.

Aos orientadores Professores Carlos Galup-Montoro e Fernando Rangel de Sousa.

Ao Professor Marcio Cherem Schneider pelas discussões técnicas e as preciosas opiniões dispensadas a este trabalho.

Aos meus pais e ao Filipe, meu irmão, pelo apoio e amizade demonstrados durante este tempo e que souberam relevar meu mau humor nos momentos mais difíceis pelos quais passei.

Um agradecimento especial à Alessandra, minha noiva, que além de ser amante, foi sempre muito compreensiva com meus horários, amiga e às vezes psicóloga.

Aos amigos de turma Alessandro, Gustavo, Marcelo e Marcio que dividiram as horas de estudo e stress durante a fase dos cursos. Aos amigos Edgar Maurício Camacho, Halmilton klimach e Luis Henrique Spiller pelas discussões técnicas proveitosas e pela paciência em ouvir desabafos.

Ao grande amigo Willian Prodanov e sua esposa que me receberam de forma amiga e acolhedora na sua casa em Genova-Itália durante os três inesquecíveis meses em que passei por lá.

Resumo da Dissertação apresentada à UFSC como parte dos requisitos necessários para a obtenção do grau de Mestre em Engenharia Elétrica.

MODELO COMPACTO DE NÃO-LINEARIDADES EM TRANSISTORES MOS

Pablo Dutra da Silva

Setembro/2006

Orientador: Carlos Galup-Montoro, Dr.

Área de Conhecimento: Circuitos e Sistemas Integrados

Palavras Chave: Modelo Compacto, Regime de Inversão Moderada, distorção de intermodulação, Circuitos Integrados aplicados a Radiofrequência (RF), Saturação da Velocidade dos Portadores, “Sweet Spot”.

Número de páginas: 59

Neste trabalho, é proposto um modelo compacto para não-linearidades em transistores MOS desenvolvido com base nas equações de canal curto do modelo “Advanced Compact MOSFET” (ACM). As maiores vantagens deste modelo são a simplicidade de suas equações e a forma explícita com que se determinam as não-linearidades do transistor MOS em função do nível de inversão. Além disso, são discutidas as causas físicas de um aumento de linearidade observado em inversão moderada, chamado “Sweet Spot”. Através de medidas, concluiu-se que efeitos de segunda ordem, principalmente a saturação da velocidade dos portadores em transistores de canal curto, são as principais causas do aumento de linearidade observado.

Abstract of Dissertation presented to UFSC as a partial fulfillment of the requirements for the degree of Master in Electrical Engineering.

COMPACT MODELING OF NONLINEARITIES IN MOS TRANSISTORS

Pablo Dutra da Silva

September/2006

Advisor: Carlos Galup-Montoro, PhD.

Area of Concentration: Integrated Circuits and systems

Keywords: Compact Model, Moderate Inversion Level, Intermodulation Distortion, Radio Frequency (RF) Integrated Circuits, Carrier Velocity Saturation, “Sweet Spot”.

Number of Pages: 59

In this work, a compact model for nonlinearities in MOS transistors derived from the short-channel equations of the Advanced Compact MOSFET (ACM) is proposed. The main advantages of the referred model are simplicity of the equations and the explicit determination of the nonlinearities of the MOSFET with respect to the inversion level. In addition, the physical causes of a linearity improvement observed in moderate inversion level, called “Sweet Spot”, are discussed. The measurements shown that second order effects, principally the carrier velocity saturation in short-channel transistors, are the main causes of the linearity improvement observed.

Sumário

LISTA DE FIGURAS	ix
LISTA DE TABELAS	x
LISTA DE SÍMBOLOS	xi
1 INTRODUÇÃO	1
2 PROBLEMAS CAUSADOS PELAS NÃO-LINEARIDADES EM CIRCUITOS DE RADIOFREQÜÊNCIA (RF) E FIGURAS DE MÉRITO PARA AMPLIFICADORES.....	3
2.1 Introdução	3
2.2 Sistemas não-lineares	3
2.3 Problemas causados pela não-linearidade em circuitos de RF.....	5
2.3.1 Compressão de ganho	5
2.3.2 Dessensibilização e bloqueio	7
2.3.3 Modulação cruzada.	8
2.3.4 Intermodulação.....	9
2.3.5 Ponto de intersecção de 3ª ordem (IP3)	11
2.4 Low Noise Amplifiers (LNA) e principais figuras de mérito.	15
2.4.1 Fator de ruído e figura de ruído.....	18
2.4.2 Sensibilidade e Faixa Dinâmica	20
2.5 Conclusão	22
3 MODELO DO TRANSISTOR MOS.....	23
3.1 Introdução	23
3.2 Considerações Gerais Sobre o Funcionamento do Transistor MOS.....	23
3.3 Modelo de Canal Longo do MOSFET	25
3.3.1 Modelo Para a Corrente de Dreno.....	26
3.3.2 Modelo da Transcondutância	28
3.4 Efeitos de Segunda Ordem do Transistor MOS	28
3.4.1 Modulação do Comprimento do Canal.	29
3.4.2 Degradação da Mobilidade com o Campo Elétrico Transversal.....	30
3.4.3 Saturação da Velocidade dos Portadores.	30
3.5 Conclusão	32
4 MODELAGEM DE IIP3 DE UM AMPLIFICADOR FONTE COMUM	33
4.1 Introdução	33

4.2	Análise não-linear de um amplificador fonte comum.....	33
4.3	Equações de Transcondutância e suas Derivadas	35
4.4	IIP3 em função da Polarização do Transistor	39
4.5	Estudo das Causas do “Sweet Spot”	42
4.6	Conclusão.....	44
5	RESULTADOS EXPERIMENTAIS.....	45
5.1	Introdução	45
5.2	Metodologia de medição	45
5.3	Extração de parâmetros tecnológicos relacionados com os efeitos de segunda ordem	48
5.4	Resultados obtidos e comparação com as equações teóricas.....	48
5.5	Conclusão.....	52
6	CONCLUSÃO	54
	ANEXO 1 – DESENVOLVIMENTO DETALHADO DO MODELO DE NÃO- LINEARIDADE.	56
	Referências Bibliográficas	58

LISTA DE FIGURAS

Figura 2-1: Exemplo de sistema não-linear e da aproximação de linearização.	4
Figura 2-2: Definição do ponto de compressão de 1dB.	7
Figura 2-3: Problema da intermodulação em forma esquemática.	9
Figura 2-4: Saída de um teste de dois tons genérico.	12
Figura 2-5: Representação gráfica do método.	15
Figura 2-6: Amplificador porta comum.	16
Figura 2-7: Amplificador “cascode” com degeneração de fonte.	17
Figura 2-8: Figura de ruído em circuitos em cascata.	19
Figura 3-1: Estrutura física de um transistor MOS.	24
Figura 4-1: Amplificador fonte comum.	34
Figura 4-2: Transcondutância e suas derivadas.	38
Figura 4-3: V_{IP3}^2 em função da densidade de cargas normalizada no lado da fonte.	40
Figura 4-4: V_{IP3}^2 em função da corrente de dreno normalizada.	41
Figura 4-5: Comparação entre modelos e medida.	43
Figura 5-1: Esquemático representando as ligações feitas para a medição.	46
Figura 5-2: (A) g_{mg} e (B) g_{mg}'' para os transistores da tecnologia TSMC 0.18 μm	49
Figura 5-3: (A) g_{mg} e (B) g_{mg}'' para os transistores da tecnologia TSMC 0.35 μm	50
Figura 5-4: V_{IP3}^2 em função da polarização para os transistores M1 e M3.	51

LISTA DE TABELAS

Tabela 2-1: Resumo das componentes de distorção harmônica.....	5
Tabela 2-2: Resumo das amplitudes das componentes da função de transferência não linear.....	10
Tabela 2-3: Amplitudes de saída das componentes fundamentais e de intermodulação. ...	12
Tabela 4-1: Transistores utilizados.	42
Tabela 5-1 Propriedades dos transistores e condições de medida utilizadas.	47
Tabela 5-2: Parâmetros tecnológicos.	48

LISTA DE SÍMBOLOS

$A_{in,out}$	-	Amplitude de entrada ou de saída do sinal.
α_θ	-	Constante de degradação da mobilidade transversal.
α_j	-	Derivadas de ordem j da expansão em séries de potências.
C'_{ox}	-	Capacitância de óxido por unidade de área.
ε	-	Parâmetro de inclusão da saturação da velocidade dos portadores.
ε_S	-	Permissividade elétrica do silício.
ε_{SiO_2}	-	Permissividade elétrica do óxido de silício.
ϕ_t	-	Tensão termodinâmica.
ϕ_S	-	Potencial de superfície.
F	-	Fator de ruído.
G	-	Ganho de pequenos sinais.
g_{md}	-	Transcondutância de dreno.
g_{mg}	-	Transcondutância de porta.
g_{ms}	-	Transcondutância de fonte.
g'_{mg}	-	Derivada primeira de g_{mg} em relação à tensão de porta.
g''_{mg}	-	Derivada segunda de g_{mg} em relação à tensão de porta.
I_D	-	Corrente de dreno.
id	-	Corrente de dreno normalizada.
i_{out}	-	Corrente de saída AC.
i_f	-	Corrente direta normalizada.
i_r	-	Corrente reversa normalizada.
I_F	-	Corrente direta.
I_R	-	Corrente reversa.
I_S	-	Corrente específica.
$IIP3$	-	Ponto de intersecção de terceira ordem referenciado à entrada.
$IM2$	-	Produto de intermodulação de segunda ordem.

$IM3$	-	Produto de intermodulação de terceira ordem
L	-	Comprimento de canal nominal.
L_{eff}	-	Comprimento de canal efetivo.
L_1, L_2	-	Indutores.
n	-	Fator de rampa.
NF	-	Figura de ruído.
$N_{i(fonte)}$	-	Ruído na entrada.
$N_{o(total)}$	-	Ruído total na saída do sistema.
$N_{o(fonte)}$	-	$N_{i(fonte)}G$.
μ_0	-	Mobilidade dos portadores na ausência de campo.
μ_S	-	Mobilidade efetiva.
μ	-	Mobilidade dependente do campo transversal.
O_{IP3}	-	Ponto de intersecção de terceira ordem referenciado à saída.
P_{IIP3}	-	Potência de entrada na qual IO3 é alcançado.
$P_{sig,in}$	-	Potência de sinal na entrada.
P_{RS}	-	Potência de ruído gerado pela resistência da fonte de sinal.
$P_{IM,in}$	-	Potência dos produtos de intermodulação referida à entrada.
Q'_I	-	Densidade de cargas de inversão por unidade de área.
Q'_B	-	Densidade de cargas de depleção por unidade de área.
Q'_{IS}	-	Densidade de cargas de inversão por unidade de área associada à fonte.
Q'_{ID}	-	Densidade de cargas de inversão por unidade de área associada ao dreno.
Q'_{IP}	-	Densidade de cargas de “pinch-off” por unidade de área.
q'_{IS}	-	Densidade de cargas de inversão normalizada associada à fonte.
q'_{ID}	-	Densidade de cargas de inversão normalizada associada ao dreno.
q'_{IDSAT}	-	Densidade de cargas de inversão máxima associada ao dreno quando a velocidade dos portadores satura.

R_L	-	Resistência de carga.
R_S	-	Resistência de fonte.
U_{CRIT}	-	Campo elétrico crítico.
V_{in}	-	Tensão de entrada.
V_{out}	-	Tensão de saída.
V_b	-	Tensão de polarização.
V_{DD}	-	Tensão de alimentação.
V_{DB}	-	Tensão dreno-corpo.
V_{SB}	-	Tensão fonte-corpo.
V_{GB}	-	Tensão porta-corpo.
V_{DS}	-	Tensão dreno-fonte.
v_{lim}	-	Velocidade limite.
V_{IIP3}	-	Amplitude de entrada onde se encontra IP3.
V_P	-	Tensão de “pinch-off”.
V_{T0}	-	Tensão de limiar do transistor.
W	-	Largura nominal do canal.
W_{eff}	-	Largura efetiva do canal.
ω	-	Frequência (rad/s).

1 INTRODUÇÃO

Os sistemas de comunicação sem fio estão tomando uma importância crucial na vida cotidiana, principalmente com o uso de telefones celulares. Os sistemas de comunicação de segunda geração (padrões TDMA, CDMA e GSM), além de sistemas de terceira geração (WCDMA e outros), que estão em operação em todo o mundo, prometem uma gama enorme de serviços, dentre eles, acesso à internet com transmissão de dados em alta velocidade, transmissão de TV e conectividade ao redor do mundo. Por essa convergência de tecnologias e serviços facilitarem a vida dos usuários é que o mercado de sistemas de comunicação sem fio vem crescendo. Como a grande vantagem desses sistemas é oferecer vários serviços, qualquer que seja o local onde esteja o usuário, um grande desafio da indústria de sistemas de comunicação sem fio é aumentar o tempo entre cargas da bateria, que é o momento no qual o sistema não pode oferecer mobilidade ao usuário. Conseqüentemente muitos dos esforços para o desenvolvimento de novas tecnologias pela indústria e pesquisadores da área de telecomunicações estão ligados a produzir sistemas com o menor consumo de potência possível.

Várias tecnologias de fabricação são utilizadas para a construção de circuitos para telecomunicações, mas a tecnologia “Complementary Metal Oxide Semiconductor” (CMOS) é muito atrativa pelo fato de que é um processo barato em relação aos outros. Além disso, o escalamento do comprimento mínimo de canal nas tecnologias mais avançadas faz com que o transistor “Metal Oxide Semiconductor Field Effect Transistor” (MOSFET) alcance frequências de transição da ordem de GHz no regime de inversão fraca ou moderada. Por outro lado, até há pouco tempo os níveis de distorção gerados por transistores MOS, operando no regime de inversão fraca ou moderada, tornavam problemática a utilização destes em circuitos que exigiam baixo consumo de potência e alta linearidade. Contudo, estudos recentes [13]-[17] têm mostrado a existência de um aumento de linearidade de transistores MOS de canal curto operando em regime de inversão moderada, em que se pode obter um bom compromisso entre frequência de transição, consumo de potência e, agora, linearidade. Esse fenômeno foi denominado de “Sweet Spot” por alguns pesquisadores.

Para que projetistas possam utilizar este ponto de alta linearidade como um requisito de projeto e extrair do circuito o melhor desempenho possível, o modelo do transistor MOS, utilizado como ferramenta de projeto, deve modelar de forma confiável

esse fenômeno. Além disso, é muito aconselhável que esse modelo seja constituído de equações simples, de forma que possibilite sua utilização em projetos à mão. Assim sendo, um dos objetivos deste trabalho é o de desenvolver um modelo não-linear do MOSFET, baseado no “Advanced Compact MOSFET” (ACM), contendo equações simples da transcondutância de porta, suas derivadas e de uma figura de mérito chamada de ponto de intermodulação de terceira ordem em função da polarização do transistor MOS.

A literatura apresentada aqui, como referência sobre o assunto [13]-[17], comprova a existência do “Sweet Spot” com o desenvolvimento de modelos complexos e medidas experimentais, mas não apresenta de forma clara uma discussão acerca dos fenômenos físicos que provocam esse aumento de linearidade no regime de inversão moderada. Neste trabalho, discutem-se os efeitos de segunda ordem envolvidos no aparecimento do “Sweet Spot” além de serem identificadas as variáveis de projeto mais importantes para o controle da posição deste ponto em relação à polarização do transistor utilizado.

Para se cumprir os objetivos acima propostos, uma revisão sobre os problemas que o comportamento não-linear de circuitos causa em sistemas de comunicação, as figuras de mérito relacionadas com esses problemas, e algumas topologias clássicas de amplificadores de baixo ruído, que são mais utilizadas em sistemas de comunicação reais, serão apresentados no capítulo 2. Uma apresentação do modelo ACM de canal longo e dos principais efeitos de segunda ordem presentes na operação de transistores será feita no capítulo 3. O desenvolvimento das equações da transcondutância, suas derivadas e do ponto de intersecção de terceira ordem em função da polarização do transistor, além de uma discussão sobre quais efeitos de segunda ordem estão relacionados com o fenômeno chamado “Sweet Spot”, serão apresentados no capítulo 4 deste trabalho. Finalmente, para a validação do modelo desenvolvido no capítulo 4, serão apresentadas no capítulo 5 comparações entre o cálculo analítico feito através das equações desenvolvidas e curvas obtidas de forma experimental.

2 PROBLEMAS CAUSADOS PELAS NÃO-LINEARIDADES EM CIRCUITOS DE RADIOFREQÜÊNCIA (RF) E FIGURAS DE MÉRITO PARA AMPLIFICADORES

2.1 Introdução

No caso dos sistemas de comunicação sem fio, a consideração das não-linearidades faz parte do estudo básico, já que estes operam em ambientes onde o sistema receptor está sujeito a captar uma quantidade enorme de sinais com níveis de potência que os projetistas não podem, a priori, determinar. Por essa razão, os sistemas de recepção estão sujeitos a problemas causados pelas não-linearidades dos sistemas físicos e esses devem ser minimizados, utilizando-se metodologias de projeto que considerem as não-linearidades dos circuitos. Para o desenvolvimento de uma metodologia de projeto RF, uma modelagem matemática, acompanhada de conceitos bem definidos dos problemas que as não-linearidades podem causar, torna-se fundamental. Este capítulo terá como objetivo principal mostrar esta modelagem matemática, acompanhada da análise dos problemas causados pela não-linearidade desses sistemas, quando na presença de ruído e sinais provenientes de outras transmissões em canais adjacentes.

Este capítulo está organizado da seguinte maneira: na seção 2.2 se apresentará como os sistemas físicos não-lineares são modelados, já na seção 2.3 serão mostrados os problemas causados em sistemas de comunicação devido à não-linearidade, na seção 2.4 serão apresentadas as figuras de mérito padrão para amplificadores de baixo ruído, além de alguns métodos de medida para as mesmas e na seção 2.5 será apresentada uma breve conclusão.

2.2 Sistemas não-lineares

Os sistemas físicos reais são não-lineares por natureza, sendo que a linearidade é uma abstração, muitas vezes útil para projeto de circuitos analógicos, porém limitada.

Na Figura 2-1 fica evidente, no esquema mostrado, que a função de transferência do circuito amplificador, representado pelo bloco “A”, não é linear para todos os valores possíveis de amplitude do sinal de entrada.

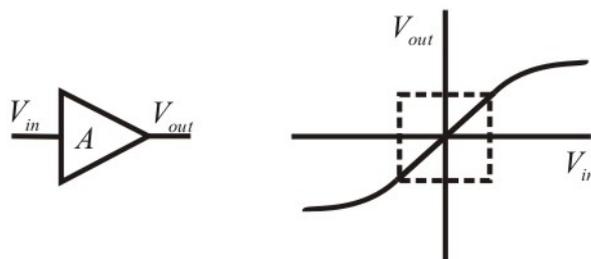


Figura 2-1: Exemplo de sistema não-linear e da aproximação de linearização.

Para linearizar esse sistema, de forma que se possa obter uma metodologia de projeto simples, limita-se a amplitude de sinal de entrada a certos valores, em sistemas em que isso pode ser controlado, e utiliza-se um modelo matemático válido para a região linear da característica de transferência (evidenciada pelo quadrado tracejado).

Considerando o sistema mostrado na Figura 2-1 como sendo um sistema sem memória, este pode ter sua característica de transferência modelada através de série de potências:

$$V_{out} = \alpha_0 + \alpha_1 v_{in} + \alpha_2 v_{in}^2 + \alpha_3 v_{in}^3 + \dots \quad (2.1)$$

Em (2.1), α_j representa a derivada de ordem j da função e são termos que dão peso a cada parcela da expansão.

Para se definir com exatidão um sistema físico, são necessários infinitos termos para essa expansão, mas para uma boa aproximação de sistemas com distorção fraca, pode-se truncar a expansão no termo de terceira ordem [1][2]. Assim se pode verificar o desempenho do sistema para distorções de até terceira ordem que, como será mostrado, são as mais importantes, além de simplificar a análise dos sistemas. Esse truncamento da característica de transferência será utilizado em toda a análise subsequente.

Para linearizar (2.1), deve-se considerar $\alpha_1 v_{in} \ll \alpha_2 v_{in}^2$ e $\alpha_3 v_{in}^3$ e escolher um $v_{in,max}$ tal que os termos de ordem superior sejam bastante menores em relação ao termo linear. Neste caso, o sistema tem uma função de transferência praticamente linear.

A expansão mostrada em (2.1) será base para o desenvolvimento e compreensão dos fenômenos relacionados com a não-linearidade em sistemas de comunicação. Deve-se lembrar que a expansão mostrada aqui serve para sistemas sem memória e que, para sistemas que possuem memória, deve-se utilizar uma expansão mais complexa.

2.3 Problemas causados pela não-linearidade em circuitos de RF

A seção anterior mostrou de forma simples como os sistemas não-lineares podem ser modelados através de uma expansão em série de potências. Nesta seção, utilizar-se-á como base a expressão (2.1) para demonstrar e compreender os problemas causados pelas não-linearidades dos sistemas eletrônicos.

Esses problemas causados pela não-linearidade dos circuitos podem ser inerentes aos amplificadores, como por exemplo, a compressão de ganho que gera distorção harmônica; ou podem ser causados apenas em amplificadores utilizados em sistemas de comunicação, pois se devem a interferências, como por exemplo, a dessensibilização ou bloqueio e a geração de produtos de intermodulação de segunda e terceira ordens. Todos esses problemas acarretam algum tipo de degradação do desempenho dos sistemas de transmissão e, principalmente na recepção, influenciando no processamento do sinal de interesse.

2.3.1 Compressão de ganho

O fenômeno de compressão do ganho é um fenômeno geral em todos os tipos de amplificadores, pois está relacionado com a limitação dada pelo barramento de alimentação do circuito e pelas tensões mínimas que devem ser aplicadas para manter o circuito em certa condição de operação. Na Figura 2-1 pode-se verificar a saturação da característica de transferência de amplificadores.

Se um sinal do tipo $v_{in} = A \cos(\omega t)$ for aplicado à entrada do circuito amplificador da Figura 2-1, obter-se-á o seguinte resultado após substituir v_{in} em (2.1) e algumas manipulações matemáticas:

$$V_{out} = \alpha_0 + \frac{\alpha_2 A^2}{2} + \left(\alpha_1 A + \frac{3\alpha_3 A^3}{4} \right) \cos(\omega t) + \frac{\alpha_2 A^2}{2} \cos(2\omega t) + \frac{3\alpha_3 A^3}{4} \cos(3\omega t) \quad (2.2)$$

Tabela 2-1: Resumo das componentes de distorção harmônica.

Frequência	DC	ωt	$2\omega t$	$3\omega t$
Amplitude	$\alpha_0 + \frac{\alpha_2 A^2}{2}$	$\alpha_1 A + \frac{3\alpha_3 A^3}{4}$	$\frac{\alpha_2 A^2}{2}$	$\frac{3\alpha_3 A^3}{4}$

Com a Tabela 2-1, fica mais fácil analisar separadamente cada parcela, que corresponde a uma determinada componente em frequência de (2.2). Nota-se que, além de uma componente DC gerada pela expansão em séries (α_0), as parcelas que representam a distorção harmônica de segunda ordem ($\alpha_2 A^2/2$) contribuem também com o nível DC. A amplitude da componente fundamental tem um fator somado ao ganho de pequeno sinal (α_1), no caso do amplificador, ocasionado pela parcela de terceira ordem da expansão ($3\alpha_3 A^3/4$). Esse resultado é o foco da discussão sobre compressão de ganho que será feita a seguir. E finalmente fica evidente a geração de distorção harmônica através das parcelas $2\omega t$ e $3\omega t$.

O ganho linearizado, como já foi comentado, é constante para qualquer amplitude de entrada. Mas o ganho do sistema real não tem um valor constante e varia segundo a amplitude do sinal de entrada aplicado. Para o caso de um amplificador, pode-se considerar um comportamento compressivo, assim $\alpha_3 < 0$. Além disso, pode-se também considerar a aproximação de distorção fraca, sendo $|\alpha_3| A^2 \ll \alpha_1$. Assim sendo, o ganho diminui com o aumento da amplitude de sinal na entrada do sistema. Isso pode ser constatado em (2.2) e com as condições apresentadas: se A for suficientemente pequeno, a parcela referente a α_3 pode ser desprezada, mas se A aumentar, a parcela referente a α_3 se tornará importante, pois essa é proporcional a A^3 (enquanto que a parcela referente a α_1 , que é proporcional à A). Esse fenômeno é chamado de compressão do ganho. Na Figura 2-1, o ganho tende a ser igual a zero com o aumento da amplitude e não menor que zero, como seria lógico pensar após o que foi dito anteriormente sobre α_3 . Mas deve-se lembrar que a expansão está truncada até a parcela de terceira ordem, e isso significa que ela não é válida para infinitos valores de tensão de entrada, pois, após certo valor de amplitude, as parcelas de mais altas ordens apresentam magnitudes significativas, não podendo ser desprezadas.

A compressão de ganho pode ser quantificada através de uma figura de mérito, chamada de ponto de compressão de 1dB. Este é definido, como o nível de amplitude do sinal de entrada que causa uma queda de 1dB no ganho em relação ao ganho linearizado teórico, representado pela linha tracejada da Figura 2-2.

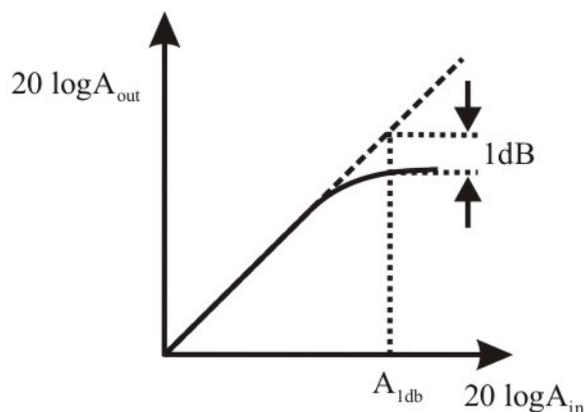


Figura 2-2: Definição do ponto de compressão de 1dB.

Segundo a definição gráfica acima, pode-se chegar a uma definição analítica, utilizando a Tabela 2-1 e substituindo A_{in} por V_{1-dB} que é a amplitude de entrada na qual o ponto de compressão é atingido, da seguinte forma:

$$20 \log \left| \alpha_1 + \frac{3}{4} \alpha_3 V_{1-dB}^2 \right| = 20 \log |\alpha_1| - 1dB \quad (2.3)$$

$$V_{1-dB} = \sqrt{0,145 \left| \frac{\alpha_1}{\alpha_3} \right|} \quad (2.4)$$

O ponto de compressão de 1dB (A_{1-dB}) pode ser utilizado como limite máximo de amplitude de entrada de um sistema para limitar distorção. Existem outras figuras de mérito nesse sentido e uma das mais clássicas é a distorção harmônica total (THD – sigla em Inglês), onde se especifica THD máxima na saída do circuito e mede-se a amplitude de entrada máxima a que essa THD é alcançada na saída. Essa figura de mérito é muito utilizada em amplificadores de uso geral. Por outro lado, o ponto de compressão de 1dB é mais utilizado em circuitos de comunicação.

2.3.2 Dessensibilização e bloqueio

Os sistemas de comunicação estão sujeitos a interferências causadas por sinais com diferentes frequências e amplitudes diversas, que variam com a distância a que o receptor se encontra da fonte de interferência. Esclarece-se que, chama-se interferência, aqui, qualquer sinal que não seja o sinal com a informação de interesse. Os problemas causados pelas interferências se tornam muito importantes para sistemas de comunicação, requerendo assim, figuras de mérito referentes às não-linearidades e que levem em consideração os efeitos causados por essa, devido às interferências.

Então, a situação em que o sinal contido no canal de transmissão tem nível de potência menor em relação às interferências é bastante comum. Isto porque não se sabe a priori a localização do receptor, que pode ser móvel. Esse poderá estar mais distante da fonte do sinal de interesse do que da fonte de interferência, que pode ser uma outra transmissão fora do canal de transmissão. Quando isso acontece, circuitos com características compressivas apresentam uma diminuição do ganho de pequenos sinais, fazendo com que o sinal que está dentro do canal de transmissão não seja amplificado. Esse fenômeno se chama dessensibilização. Quando a amplitude da interferência é suficientemente alta para levar o ganho da componente fundamental a zero, o fenômeno é conhecido como bloqueio.

Esses dois fenômenos podem ser verificados analiticamente substituindo $v_{in} = A_1 \cos(\omega_1 t) + A_2 \cos(\omega_2 t)$ em (2.1), onde ω_1 e ω_2 são as frequências dos sinais da informação e da interferência, respectivamente, resultando em:

$$v_{out} = \left(\alpha_1 A_1 + \frac{3}{4} \alpha_3 A_1^3 + \frac{3}{2} \alpha_3 A_1 A_2^2 \right) \cos(\omega_1 t) + \dots \quad (2.5)$$

Em (2.5), mostra-se somente o resultado referente ao ganho da componente fundamental, pois este é o resultado que interessa para a análise matemática do fenômeno de dessensibilização, devido ao fato de a interferência alterar o valor do ganho da componente fundamental do sinal. Se $A_1 \ll A_2$ pode-se simplificar (2.5), resultando em:

$$v_{out} = \left(\alpha_1 + \frac{3}{2} \alpha_3 A_2^2 \right) A_1 \cos(\omega_1 t) \quad (2.6)$$

Para sistemas compressivos como, por exemplo, alguns tipos de amplificadores, tem-se $\alpha_3 < 0$. Pode-se então verificar em (2.6) que o ganho para o sinal diminui com o aumento da amplitude da interferência, pois, o ganho é diminuído de um fator proporcional ao quadrado da amplitude da interferência. Se a amplitude da interferência é menor do que a do sinal de informação, então o fenômeno em questão não aparece, pois nos sistemas compressivos $\alpha_3 A_2^2 \ll \alpha_1$. Se a interferência é forte o suficiente para levar o ganho a zero, então o sinal é bloqueado completamente.

2.3.3 Modulação cruzada.

Na seção anterior, considerou-se A_2 em (2.5) como sendo um único tom de interferência para caracterizar o fenômeno de bloqueio. Considerando agora que a

amplitude A_2 de interferência seja modulada em amplitude por um outro sinal modulante, variando assim o seu valor de amplitude com o tempo. Surge assim um fenômeno relacionado com sinais de interferência que se caracteriza pela transferência de modulação ou de ruído presentes na amplitude do sinal de interferência para a amplitude de saída do sinal de informação. Esse fenômeno é conhecido na literatura como modulação cruzada.

Já que a interferência é modulada em amplitude, ou seja, a amplitude da interferência A_2 varia de acordo com o sinal modulante; o ganho do circuito também varia, transmitindo essa modulação para o sinal de informação na saída do circuito. O mesmo caso ocorre para sinais de interferência como relação sinal ruído baixa, ou seja, a amplitude do sinal de interferência varia significativamente de acordo com as variações do ruído.

2.3.4 Intermodulação

Outro fenômeno decorrente da aplicação de mais de um sinal à entrada de um sistema não-linear é o de intermodulação. No sinal de saída, aparecem componentes que não são harmônicas do sinal de informação, que são chamados de produtos de intermodulação.

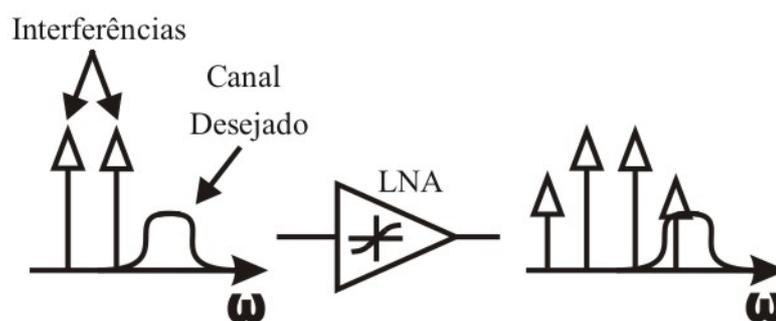


Figura 2-3: Problema da intermodulação em forma esquemática.

Através da Figura 2-3, é possível facilmente entender a problemática envolvida com o fenômeno de intermodulação em sistemas de comunicação. Na entrada do amplificador da Figura 2-3, observa-se, no domínio da frequência, o canal de informação e dois sinais interferentes, próximos a esse, com níveis de potência maiores em relação ao sinal dentro do canal de informação. Na saída do amplificador, nota-se um dos produtos de intermodulação gerados pelas não-linearidades dentro do canal de informação, causando degradação no sinal que será processado por outros blocos do receptor.

Pode-se verificar analiticamente o problema da intermodulação substituindo em (2.1) o sinal de entrada: $v_{in} = A_1 \cos(\omega_1 t) + A_2 \cos(\omega_2 t)$ sendo A_1 e A_2 as respectivas amplitudes dos dois sinais interferentes mostrados na Figura 2-3. Assim:

$$v_{out} = \alpha_1 [A_1 \cos(\omega_1 t) + A_2 \cos(\omega_2 t)] + \alpha_2 [A_1 \cos(\omega_1 t) + A_2 \cos(\omega_2 t)]^2 + \alpha_3 [A_1 \cos(\omega_1 t) + A_2 \cos(\omega_2 t)]^3 \quad (2.7)$$

Utilizando algumas identidades trigonométricas e feitas algumas manipulações algébricas, separam-se os termos de forma que cada parcela do polinômio represente uma componente em frequência. Dessa forma, pode-se analisar separadamente cada componente. Nessa análise, as componentes harmônicas foram descartadas, pois o objetivo é estudar a distorção de intermodulação.

$$\omega = \omega_1, \omega_2 : \left(\alpha_1 A_1 + \frac{3}{4} \alpha_3 A_1^3 + \frac{3}{2} \alpha_3 A_1 A_2^2 \right) \cos(\omega_1 t) + \left(\alpha_1 A_2 + \frac{3}{4} \alpha_3 A_2^3 + \frac{3}{2} \alpha_3 A_2 A_1^2 \right) \cos(\omega_2 t) \quad (2.8)$$

$$\omega = \omega_1 \pm \omega_2 : \alpha_2 A_1 A_2 \cos(\omega_1 + \omega_2)t + \alpha_2 A_1 A_2 \cos(\omega_1 - \omega_2)t \quad (2.9)$$

$$\omega = 2\omega_1 \pm \omega_2 : \frac{3\alpha_3 A_1^2 A_2}{4} \cos(2\omega_1 + \omega_2)t + \frac{3\alpha_3 A_1^2 A_2}{4} \cos(2\omega_1 - \omega_2)t \quad (2.10)$$

$$\omega = 2\omega_2 \pm \omega_1 : \frac{3\alpha_3 A_2^2 A_1}{4} \cos(2\omega_2 + \omega_1)t + \frac{3\alpha_3 A_2^2 A_1}{4} \cos(2\omega_2 - \omega_1)t \quad (2.11)$$

Tabela 2-2: Resumo das amplitudes das componentes da função de transferência não linear.

Frequência	Fundamentais (ω_1, ω_2)	IM2 ($\omega_1 \pm \omega_2$)	IM3 ($2\omega_1 \pm \omega_2, 2\omega_2 \pm \omega_1$)
Amplitude	$\alpha_1 A + \frac{9}{4} \alpha_3 A^3$	$\alpha_2 A^2$	$\frac{3\alpha_3 A^3}{4}$

Em (2.8), pode-se verificar um resultado já mostrado em (2.5) apenas para ω_1 ; mas agora, vêm-se as parcelas para as componentes fundamentais (ω_1, ω_2). Os produtos de intermodulação são mostrados por sua vez em (2.9), (2.10) e (2.11); sendo que (2.9) mostra os produtos de intermodulação de segunda ordem (IM2) e (2.10) e (2.11) mostram os produtos de intermodulação de terceira ordem (IM3).

Para que a análise se torne mais simples, considerou-se $A_1 = A_2 = A$, ou seja, as amplitudes dos sinais de interferência são iguais. A Tabela 2-2 resume os resultados dessa

simplificação e ainda mostra, de forma organizada, as amplitudes referentes a cada componente em frequência.

Os produtos de intermodulação de segunda ordem podem ser desconsiderados para análise de amplificadores, pois suas componentes em frequência não afetam o canal de comunicação. Isso porque suas componentes estão na soma e na subtração das frequências das componentes de interferência. Se $\omega_1 \approx \omega_2$, então $\omega_1 + \omega_2 \approx 2\omega_1$ e $\omega_1 - \omega_2 \approx 0$. A componente aproximadamente DC dos produtos de segunda ordem é muito importante para a análise de não-linearidades em circuitos multiplicadores em sistemas de conversão direta, pois pode saturar o próximo estágio do receptor [1]-[3].

Já os produtos de intermodulação de terceira ordem são muito importantes para a análise de amplificadores de RF pelos motivos já comentados e mostrados na Figura 2-3. Se $\omega_1 \approx \omega_2$, então as componentes dos produtos de terceira ordem que são importantes são: $2\omega_1 - \omega_2$ e $2\omega_2 - \omega_1$, pois essas aparecem próximas das componentes fundamentais, enquanto $2\omega_1 + \omega_2$ e $2\omega_2 + \omega_1$ aparecem em frequências mais altas em relação ao canal de comunicação.

A degradação de sinais, devido aos produtos de intermodulação de terceira ordem gerados por sinais interferentes próximos, é freqüente em sistemas de comunicação, tanto que uma figura de mérito foi criada para caracterizar esse fenômeno. Esta é chamada de ponto de intersecção de terceira ordem (IP3 – sigla em Inglês) e será tratado na próxima subsecção por ser a figura de mérito alvo do presente trabalho.

2.3.5 Ponto de intersecção de 3ª ordem (IP3)

Como já comentado anteriormente neste mesmo capítulo, o parâmetro IP3 é uma medida da linearidade do sistema. Essa figura de mérito é medida utilizando um método chamado de teste de dois tons (“two tone test” - em Inglês) com as seguintes condições: as amplitudes $A_1 = A_2 = A$, sendo que A é suficientemente pequeno de forma que a aproximação de pequenos sinais seja válida (ganho de pequenos sinais seja igual a α_1), e os termos de ordens mais elevadas da expansão sejam desprezíveis, ou seja, pretende-se evitar a compressão de ganho. As frequências ω_1 e ω_2 devem apresentar valores próximos um do outro (pode-se utilizar, por exemplo, uma diferença de dez por cento entre os tons aplicados).

Utilizando as considerações acima mencionadas com relação à amplitude dos tons em (2.8), (2.10) e (2.11) resultam nas seguintes amplitudes de saída mostradas em seguida.

Tabela 2-3: Amplitudes de saída das componentes fundamentais e de intermodulação.

Freqüência	ω_1, ω_2	$2\omega_1 \pm \omega_2, 2\omega_2 \pm \omega_1$
Amplitude	$\alpha_1 A$	$\frac{3\alpha_3 A^3}{4}$

Segundo os resultados apresentados na Tabela 2-3, a amplitude de saída referente às freqüências fundamentais aumenta proporcionalmente com A , enquanto as componentes de intermodulação de terceira ordem aumentam proporcionalmente com A^3 . Assim sendo, sabe-se que, em algum valor de amplitude de entrada, as amplitudes de saída, tanto das freqüências fundamentais quanto dos produtos de intermodulação de terceira ordem, se igualarão. A amplitude de entrada para a qual as amplitudes de saída mencionadas são iguais é chamada de ponto de intersecção de terceira ordem (IP3 – sigla em Inglês para “Third-Order Intercept Point”).

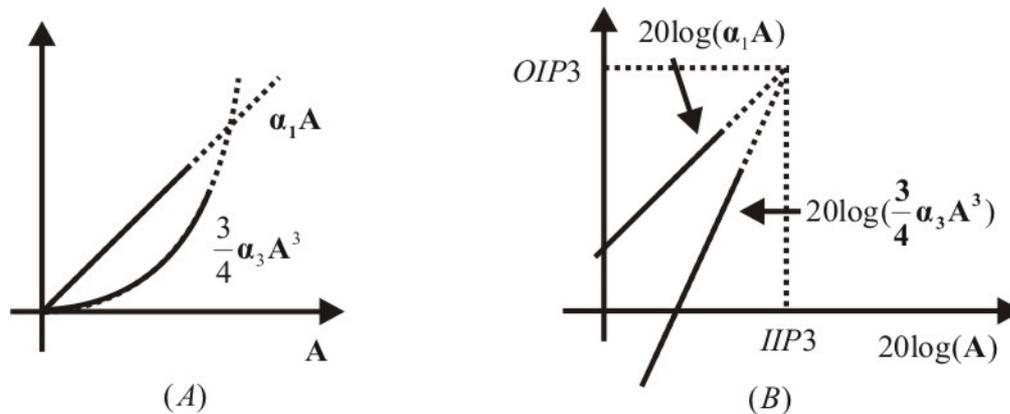


Figura 2-4: Saída de um teste de dois tons genérico.

A Figura 2-4 mostra de forma gráfica os resultados mostrados na Tabela 2-3, sendo que o gráfico da Figura 2-4 (A) mostra, na mesma escala linear, a evolução da amplitude de saída das freqüências fundamentais e da amplitude de IM3 em função da amplitude de entrada dos tons aplicados. Enquanto a Figura 2-4 (B) mostra as mesmas grandezas anteriormente mencionadas em escala log-log. Em ambas, pode-se verificar o ponto onde as curvas se interceptam.

Na Figura 2-4 (B), as curvas são pontilhadas para indicar que, na prática, IP3 é fruto de uma extrapolação, pois como já foi comentado, se a amplitude A passar de certo

valor, onde as parcelas de maior ordem que compõe o ganho das frequências fundamentais, certamente interferirão na medida de IP3. Além disso, na prática, o ponto IP3 é geralmente maior do que a amplitude necessária para alcançar o ponto onde a compressão de ganho se inicia, ou seja, o IP3 é uma medida indireta, pois não pode ser atingido na prática. A Figura 2-4 (B) também mostra dois pontos de IP3, um ponto no eixo das abscissas e outro no eixo das ordenadas, que são respectivamente: IP3 referenciado à entrada do circuito (IIP3 – Sigla em Inglês) e IP3 referenciado à saída do circuito (OIP3 – Sigla em Inglês). Aqui se trabalhará sempre com IIP3, pois parece ser o parâmetro mais apropriado e mais utilizado na literatura [1]-[5] e com relação ao cálculo da faixa dinâmica do circuito.

A grande vantagem do parâmetro IIP3 é que ele representa uma única quantidade que, por si só, serve como padrão para a comparação de linearidade de diferentes circuitos. Ou seja, não é preciso especificar, por exemplo, a amplitude de sinal aplicada para a realização dos testes ou até mesmo a carga dos circuitos [1].

É possível deduzir uma expressão simplificada para IP3 para que se possa ter uma estimativa de seu valor. Considerando a definição já mencionada, podem-se igualar as duas amplitudes mostradas na Tabela 2-3. Lembrando que, pela definição apresentada, $A = V_{IIP3}$ quando as amplitudes de saída das fundamentais e dos produtos de intermodulação de terceira ordem são iguais. Assim:

$$|\alpha_1|V_{IIP3} = \frac{3}{4}|\alpha_3|V_{IIP3}^3 \quad (2.12)$$

$$V_{IIP3} = \sqrt[3]{\frac{4|\alpha_1|}{3|\alpha_3|}} \quad (2.13)$$

Em (2.13), é apresentada uma expressão que tem como objetivo obter uma estimativa teórica do valor de V_{IIP3} que é a amplitude de entrada na qual se alcança o ponto de IIP3. Apesar de esta estimativa já ser muito importante para o projetista, ainda se faz necessária uma medida experimental para a determinação deste valor em situações práticas. Como já mencionado, a medida de V_{IIP3} deve ser feita de forma indireta pelo fato de este ponto ser impossível de se alcançar em uma situação de medida prática. Alguns métodos podem ser utilizados para medir esta figura de mérito. Dentro desses métodos, o mais preciso é o de extrapolação, que consiste em fazer uma varredura em A , respeitando às condições já mencionadas anteriormente, para que o teste de dois tons seja válido, e

medir na saída do circuito as amplitudes das componentes fundamentais e dos produtos de intermodulação de terceira ordem. Os resultados obtidos podem ser apresentados em um mesmo gráfico de escala log-log, como mostrado na Figura 2-4 (B). Para medir o V_{IIP3} faz-se a extrapolação de ambas as curvas até que elas se encontrem, como mostrado nas linhas pontilhadas da Figura 2-4 (B). Se a extrapolação for precisa, a determinação também o será.

Um método mais rápido é mostrado a seguir e pode ser importante para uma primeira avaliação de IIP3. Considerando a amplitude de entrada como sendo A_{in} , a amplitude de saída das componentes ω_1 e ω_2 como sendo A_{ω_1, ω_2} e a amplitude de IM3 como sendo A_{IM3} obtém-se:

$$\frac{A_{\omega_1, \omega_2}}{A_{IM3}} = \frac{|\alpha_1| A_{in}}{3 |\alpha_3| A_{in}^3 / 4} \quad (2.14)$$

$$\frac{A_{\omega_1, \omega_2}}{A_{IM3}} = \frac{4 |\alpha_1|}{3 |\alpha_3|} \frac{1}{A_{in}^2} \quad (2.15)$$

que em comparação com (2.13), permite escrever:

$$\frac{A_{\omega_1, \omega_2}}{A_{IM3}} = \frac{V_{IIP3}^2}{A_{in}^2} \quad (2.16)$$

Expressando (2.16) em dB, pode-se escrever da seguinte forma:

$$\underbrace{20 \log V_{IIP3}}_{IIP3_{dB}} = \frac{1}{2} \underbrace{(20 \log A_{\omega_1, \omega_2} - 20 \log A_{IM3})}_{\Delta A_{dB}} + \underbrace{20 \log A_{in}}_{A_{in}_{dB}} \quad (2.17)$$

Então, se as amplitudes de sinal são expressas em dB, o IIP3 é igual à metade da diferença entre as magnitudes de saída das componentes fundamentais e dos produtos de intermodulação somados à magnitude de entrada das componentes fundamentais. A grande vantagem desse método em relação à extrapolação é que basta medir com apenas uma amplitude de entrada. As parcelas de (2.17) podem ser expressas em potência sem nenhum prejuízo para o método, pois a definição de IIP3 é a mesma para as grandezas medidas com unidades de tensão ou potência, como mostrado na Figura 2-5.

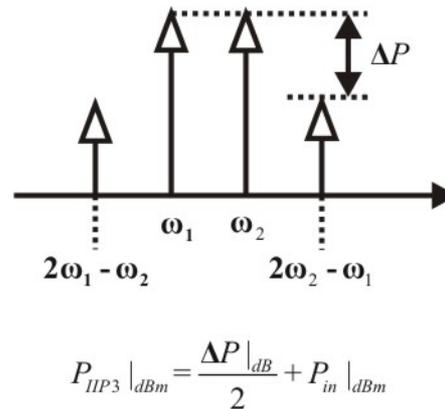


Figura 2-5: Representação gráfica do método.

A Figura 2-5 apresenta uma representação gráfica de como o método de medida de IIP3 funciona onde as potências expressas em dBm são definidas como:

$$P |_{dBm} = 10 \log \left(\frac{P}{1mW} \right).$$

2.4 Low Noise Amplifiers (LNA) e principais figuras de mérito.

Os amplificadores de baixo ruído (Low Noise Amplifiers – LNA) estão presentes em praticamente todas as estruturas de comunicação utilizadas. São blocos extremamente importantes na recepção de sinais de rádio frequência. Esse bloco de circuito deve amplificar o sinal recebido pela antena, já degradado devido ao ambiente de transmissão, sem degradar a relação sinal ruído e minimizar problemas devido à não-linearidade, de forma que, os circuitos subsequentes possam processar o sinal de informação de modo confiável. Em muitas ocasiões esses devem apresentar uma impedância característica de entrada de 50Ω , principalmente quando o sistema receptor utiliza um filtro passivo antes do LNA. Este detalhe é muito importante porque muitos filtros são sensíveis à carga de terminação [1]-[5].

O objetivo dos projetistas de LNAs é sempre alcançar o melhor compromisso entre casamento de impedâncias na entrada do amplificador, figura de ruído mínima e baixo consumo de potência. As metodologias de projeto clássicas tentam alcançar os mesmos objetivos, mas consideram que o transistor (componente utilizado para a construção de amplificadores) é um dispositivo dado por um fabricante e que tem características bem definidas; assim, as metodologias voltadas para circuitos integrados devem levar em consideração um grau de liberdade importantíssimo em microeletrônica: as dimensões do

transistor para se alcançar os objetivos de forma simples e eficiente [3]. Para isso, é necessário que os modelos de ruído e de linearidade do transistor levem em consideração esses graus de liberdade que definem seu comportamento. Com isso, será possível definir uma metodologia de projeto que leva a um bom compromisso entre ganho, casamento de impedância, figura de ruído e consumo de potência, escolhendo as dimensões e a corrente de polarização dos transistores.

Existem várias topologias de circuito utilizadas na implementação de LNAs. Para uma avaliação prévia de desempenho destas, a literatura especializada costuma utilizar dois aspectos: casamento de impedâncias e figura de ruído. Aqui serão mostradas as topologias mais utilizadas pela literatura, incluindo uma análise de algumas de suas vantagens e desvantagens mostrando também duas das figuras de mérito mais importantes no projeto de LNAs.

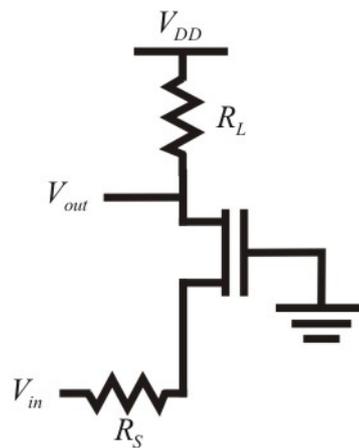


Figura 2-6: Amplificador porta comum.

Uma das mais populares topologias utilizadas é o amplificador porta comum, mostrado na Figura 2-6. O casamento de impedâncias é facilmente alcançado escolhendo a corrente de polarização e as dimensões do transistor tais que a impedância de entrada do circuito (dada por $1/g_{ms}$) seja igual à impedância da fonte, com o intuito de proporcionar a máxima transferência de potência. Mesmo assim, o excesso de ruído fica por conta da resistência de canal [3]. A topologia mostrada na Figura 2-6 não alcançará uma impedância de entrada de 50Ω sem consumir níveis de corrente proibitivos para uma operação em baixo consumo [3], lembrando que g_{ms} aumenta com o aumento do nível de inversão (para um transistor com dimensões pré-fixadas).

É possível utilizar níveis de corrente compatíveis com a operação em baixo consumo e, para superar o problema do casamento de impedâncias, pode-se utilizar uma rede de casamento reativa. Isso faz com que a aplicação do amplificador mude, pois com um casamento feito através de uma rede reativa perde-se a vantagem do casamento de impedâncias constante para aplicações em banda larga. Outro problema para o consumo é o fator de qualidade dos componentes reativos, assunto que é discutido em [4] e não será foco dessa discussão informativa.

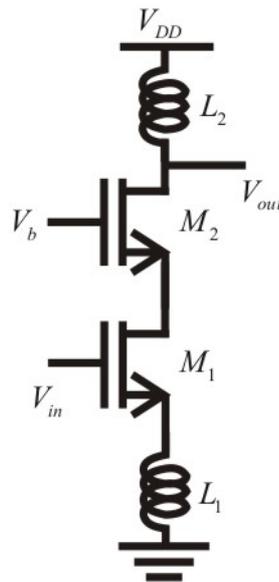


Figura 2-7: Amplificador “cascode” com degeneração de fonte.

Para a utilização em transmissões de banda estreita, que tem um grande número de aplicações, o amplificador mostrado na Figura 2-7 é largamente utilizado. A grande vantagem desse circuito é utilizar um indutor de degeneração para obter o casamento de impedâncias na entrada sem a adição de ruído térmico, pois idealmente os elementos reativos (capacitores e indutores) não apresentam ruído desse tipo. Como esse casamento de impedâncias, feito com elementos reativos, da-se apenas na frequência de ressonância, o amplificador é de banda estreita.

A indústria de sistemas de comunicação requer algumas características já citadas dos LNAs, em seus padrões de comunicação. Esses requisitos se transformam em especificações para os projetistas de circuitos integrados. Pode-se apontar as mais importantes como sendo: nível de ruído adicionado pelo circuito (caracterizado pela figura de ruído), capacidade do circuito amplificar sinais com mínima distorção (ponto de intersecção de terceira ordem), casamento de impedância com a antena ou filtro

precedente e consumo de potência. Entre essas, a mais requisitada será definida pela aplicação a que o sistema de comunicação será exposto.

Para se fazer uma medida de comparação padrão, também se criou figuras de mérito que são definições, na maioria das vezes, intuitivas para se utilizar como parâmetros de comparação de desempenho para os circuitos. Aqui serão apresentadas as mais importantes para o projeto de um LNA.

2.4.1 Fator de ruído e figura de ruído

A agitação térmica é a causa do mais importante tipo de ruído observado em sistemas eletrônicos, o ruído térmico. A energia térmica agita os elétrons de forma aleatória em diferentes direções, causando desvios aleatórios de corrente elétrica, caracterizando assim o ruído térmico. O ruído térmico tem como característica principal a densidade espectral de potência constante e por isso também é conhecido como ruído branco.

Existem outras fontes de ruído além da agitação térmica. Pode-se citar o ruído chamado de “flicker” ou $1/f$ como o principal tipo de ruído não térmico. A denominação $1/f$ deve-se à densidade espectral de potência deste tipo de ruído, que é inversamente proporcional à frequência. Uma explanação mais detalhada sobre esse assunto pode ser encontrada em [1]-[3].

Uma das preocupações que o ruído causa aos projetistas é o fato de que a potência de ruído mínima, que é detectada na entrada do circuito (chão de ruído), define a sensibilidade do circuito. Isso significa que o circuito será hábil para processar sinais com potência maior do que o chão de ruído mais o ruído adicionado pelo circuito [1]-[3].

O ruído adicionado pelo circuito eletrônico ao sinal é medido através do fator de ruído, que é definido como sendo a medida de como a relação sinal ruído é degradada pelo próprio sistema. A relação sinal ruído (SNR – sigla em Inglês) de um sinal é um fator que mostra quão mais forte é o sinal em relação ao ruído somado a ele. Através dessa definição, pode-se deduzir uma equação para o fator de ruído:

$$F = \frac{SNR_i}{SNR_o} = \frac{S_i/N_{i(fonte)}}{S_o/N_{o(total)}} = \frac{S_i/N_{i(fonte)}}{(S_i \cdot G)/N_{o(total)}} = \frac{N_{o(total)}}{N_{o(fonte)}} \quad (2.18)$$

Em (2.18), $N_{o(fonte)}$ é a potência de ruído recebida na entrada do circuito multiplicada pelo ganho de pequenos sinais, enquanto $N_{o(total)}$ representa o ruído

mencionado anteriormente ($N_{o(fonte)}$), somado à contribuição de ruído do circuito, e G é o ganho de pequenos sinais do circuito. Para interpretar essa definição, pode-se imaginar que o circuito em questão não adiciona ruído ao sinal, então $N_{o(fonte)} = N_{o(total)}$ resultando em um fator de ruído $F = 1$. Essa última afirmação leva à conclusão que o menor fator de ruído possível é igual a um e significa que o circuito não degrada a relação sinal ruído do sinal de informação recebido, ou que o circuito não adiciona ruído ao sinal recebido.

A figura de ruído é relacionada com o fator de ruído através da relação abaixo:

$$NF = 10 \log_{10} F \quad (2.19)$$

Na grande maioria das aplicações, a figura de mérito mais importante para os LNAs é o fator de ruído ou a figura de ruído, pois a figura de ruído do primeiro bloco do sistema domina a figura de ruído total do sistema, como será verificado.

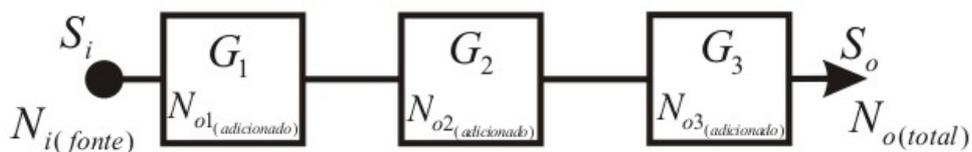


Figura 2-8: Figura de ruído em circuitos em cascata.

Considere o diagrama de blocos da Figura 2-8, onde se pode verificar três blocos de circuito em cascata, representando o caminho do sinal de informação recebido na antena de um receptor de RF. Como visto anteriormente, para calcular o fator de ruído do sistema da Figura 2-8, basta calcular o ruído de saída devido à fonte ruído na entrada e o ruído total na saída. Assim:

$$N_{o(total)} = N_{i(fonte)}G_1G_2G_3 + N_{o1(adicionado)}G_2G_3 + N_{o2(adicionado)}G_3 + N_{o3(adicionado)} \quad (2.20)$$

O ruído de saída devido à fonte de ruído na entrada é:

$$N_{o(fonte)} = N_{i(fonte)}G_1G_2G_3 \quad (2.21)$$

Substituindo (2.20) e (2.21) em (2.18) e fazendo algumas manipulações algébricas, chega-se a:

$$F = F_1 + \frac{F_2 - 1}{G_1} + \frac{F_3 - 1}{G_1G_2} \quad (2.22)$$

Onde F_n é o fator de ruído de cada bloco n de circuito individual e G_n são os diferentes ganhos de potência dos n circuitos ligados em cascata. A conclusão que se

chega com a relação (2.22), que na literatura é chamada de equação de Friis [1]-[3], é que o fator de ruído do circuito de entrada é a parcela predominante para o cálculo do fator de ruído de todo o sistema. Por essa razão, o fator de ruído ou a figura de ruído são figuras de mérito fundamentais para o projeto de LNAs e é por isso que amplificadores que recebem e amplificam sinais vindos de antenas são chamados LNA.

2.4.2 Sensibilidade e Faixa Dinâmica

Pode-se definir a sensibilidade de um receptor como sendo a mínima magnitude de sinal que o sistema consegue detectar com uma aceitável relação sinal-ruído [1]. Para calcular a sensibilidade, P_{sig} será a potência do sinal na entrada e P_{R_s} a potência do ruído gerado pela resistência da fonte de sinal, ambos considerando uma banda unitária. Então, substituindo P_{sig} por S_i e P_{R_s} por $N_{i(fonte)}$ em (2.18) obtém-se:

$$NF = \frac{SNR_m}{SNR_{out}} = \frac{P_{sig}/P_{R_s}}{SNR_{out}} \quad (2.23)$$

Isolando a potência de sinal na entrada resulta em:

$$P_{sig} = P_{R_s} \cdot NF \cdot SNR_{out} \quad (2.24)$$

A equação (2.24) prediz a sensibilidade como a potência mínima de sinal de entrada que é necessária para atingir um certo valor para a relação sinal-ruído na saída (SNR_{out}) [1]. Considera-se que a potência do sinal está distribuída ao longo da banda do canal de comunicação. Assim, integrando (2.24) em relação à banda (B) com o objetivo de se obter a potência média, utilizando $P_{sig} = P_{in,min}$ e $SNR_{out} = SNR_{min}$ e expressando as parcelas em dB ou dBm, obtém-se:

$$P_{in,min} |_{dBm} = P_{R_s} |_{dBm/Hz} + NF |_{dB} + SNR_{min} |_{dB} + 10 \log B \quad (2.25)$$

Em (2.25), $P_{in,min}$ é a potência mínima de sinal na entrada necessária para atingir a SNR_{min} e B é a banda em Hz [1].

Se for feito um casamento de impedâncias na entrada do sistema, pode-se definir P_{R_s} como a potência de ruído que a resistência característica da fonte de sinal transfere à impedância de entrada do circuito considerando a banda igual a 1Hz. Assim:

$$P_{R_s} = \frac{4KTR_s}{4} \frac{1}{R_{in}} \quad (2.26)$$

$$P_{R_s} = KT \quad (2.27)$$

Na temperatura de 300K, (2.25) pode ser simplificada como:

$$P_{in,min} |_{dBm} = -174dBm / Hz + NF |_{dB} + 10 \log B + SNR_{min} |_{dB} \quad (2.28)$$

A soma dos três primeiros termos de (2.28) é o ruído total do sistema e, algumas vezes, é chamado de chão de ruído [1].

A faixa dinâmica é geralmente definida na literatura como a razão entre a máxima magnitude de sinal de entrada que o circuito consegue tolerar e a mínima magnitude de entrada que o circuito pode detectar, mantendo uma razoável qualidade do sinal na saída [1]. Em circuitos de RF, a magnitude máxima geralmente é determinada através do comportamento dos produtos de intermodulação, e a magnitude mínima é determinada através da sensibilidade. Essa definição é chamada em inglês de “Spurious-free dynamic range” (*SFDR*) [1]-[3].

O limite máximo da faixa dinâmica pode ser definido como a máxima magnitude aplicada à entrada do circuito, em um teste de dois tons, na qual os produtos de intermodulação se mantenham abaixo do chão de ruído[1]. Pode-se quantificar essa definição da seguinte forma, considerando a equação mostrada na Figura 2-5, onde $\Delta P = P_{out} - P_{IM,out}$, e considerando também que $P_{out} = P_{in} + G$ e $P_{IM,out} = P_{IM,in} + G$, onde G é o ganho de potência do circuito e $P_{IM,in}$ é a potência dos produtos de intermodulação referidos à entrada. Tem-se:

$$P_{in} = \frac{2P_{IP3} + P_{IM,in}}{3} \quad (2.29)$$

Pela definição apresentada, $P_{IM,in}$ deve ser igual ao chão de ruído, chamado aqui de $F = -174dBm / Hz + NF + 10 \log B$, resultando em:

$$P_{in,max} = \frac{2P_{IP3} + F}{3} \quad (2.30)$$

Segundo a definição de *SFDR*, deve-se subtrair (2.30) de (2.28), resultando em:

$$SFDR = \frac{2(P_{IP3} - F)}{3} - SNR_{min} \quad (2.31)$$

Então, o *SFDR* representa a magnitude relativa máxima de interferência que um receptor consegue tolerar enquanto apresenta um sinal com uma aceitável qualidade na saída para um sinal de pequena magnitude na entrada. Essa qualidade está relacionada com SNR_{min} .

2.5 Conclusão

Neste capítulo, mostraram-se as principais conseqüências das não-linearidades dos sistemas físicos no desempenho de amplificadores RF. Analisou-se, em particular, a compressão de ganho e também os problemas que sinais de interferência localizados em canais adjacentes causam.

O estudo em nível de sistema gera figuras de mérito utilizadas como padrões de especificação de circuitos de comunicação. As figuras de mérito para amplificadores de baixo ruído aqui apresentadas se mostraram muito importantes, pois definem em que condições o circuito será eficiente na recepção de sinais contendo informação.

Além disso, a modelagem matemática aqui mostrada servirá de base para o desenvolvimento de um modelo em nível de circuito, visando oferecer ferramentas simples para projetistas desenvolverem projetos que levem em consideração a não-linearidade dos circuitos. Mas para isso, deve-se conhecer um modelo matemático em nível de dispositivo para que seja aplicado a essa metodologia. O modelo de dispositivo pode definir o quão simples e preciso será o projeto em nível de circuito, sendo que o dispositivo em questão aqui é o transistor MOS. Esse modelo é apresentado no capítulo seguinte.

3 MODELO DO TRANSISTOR MOS

3.1 Introdução

O modelo aqui apresentado e que será utilizado como base para as equações que representarão o comportamento não-linear do transistor MOS é o modelo desenvolvido na Universidade Federal de Santa Catarina, chamado “Advanced Compact MOSFET” (ACM) [6]-[9]. Esse modelo, baseado na física, tem as seguintes características: necessita de poucos parâmetros tecnológicos, possui equações simples que permitem inclusive o projeto de circuitos a mão, é contínuo em todas as regiões de operação do transistor e também preciso desde inversão fraca até inversão forte [8].

Devido à complexidade do modelo do transistor MOS, serão aqui apresentadas apenas as equações do modelo ACM, que representam as relações elétricas que modelam os fenômenos físicos mais importantes ao desenvolvimento deste trabalho.

Este capítulo será organizado da seguinte forma: a seção 3.2 apresentará as considerações gerais do funcionamento do MOSFET, a seção 3.3 mostrará as equações de corrente de dreno e de transcondutância do modelo de canal longo, a seção 3.4 apresentará os efeitos de segunda ordem ou de canal curto e, na seção 3.5, uma breve conclusão.

3.2 Considerações Gerais Sobre o Funcionamento do Transistor MOS

Não se poderia aqui falar sobre termos e qualquer tipo de nomenclatura que fazem referência ao dispositivo chamado de transistor “Metal Oxide Semiconductor Field Effect Transistor” (MOSFET – Sigla em Inglês) sem antes tentar introduzir esses termos e a nomenclatura em um resumo muito sucinto do funcionamento do dispositivo em questão.

A figura abaixo mostra um exemplo da estrutura física de um transistor MOS e pode-se verificar que o nome dado a esta estrutura é muito conveniente, pois o material utilizado para a confecção da porta dos primeiros transistores era um metal (alumínio), apesar de que nas tecnologias mais modernas, o material seja um semicondutor policristalino (polisilício), para que fique com características metálicas de condução. Abaixo desse, há uma camada de óxido de silício (SiO_2), ou outro material de alta constante dielétrica, que faz o papel de isolante e finalmente, completando o sanduíche que dá nome ao dispositivo, encontra-se o substrato semicondutor, que normalmente é de

silício dopado. Para transistores do tipo N, o substrato é dopado com impurezas do tipo P e, para transistores do tipo P, o substrato é dopado com impurezas do tipo N.

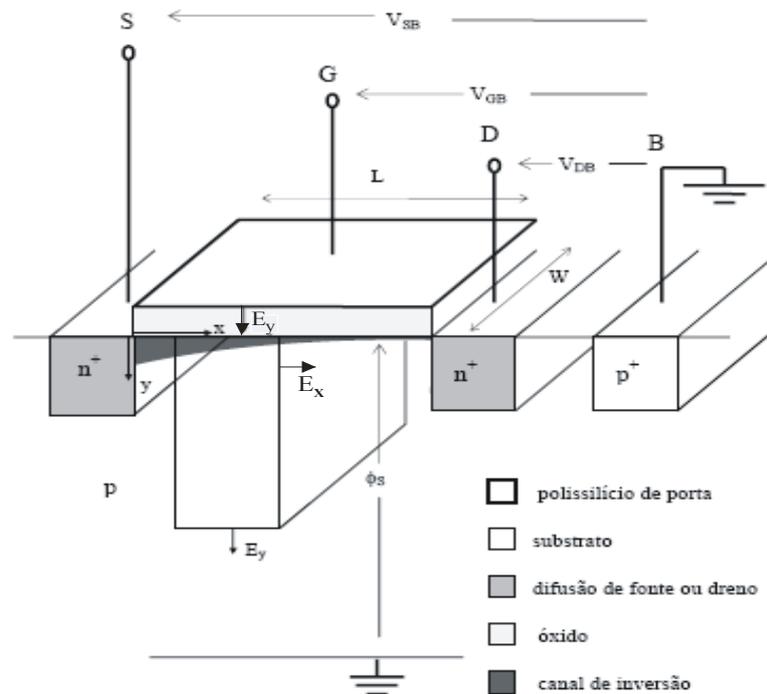


Figura 3-1: Estrutura física de um transistor MOS.

A estrutura mostrada na Figura 3-1, como exemplo, é o de um transistor canal N ou NMOS, que tem essa denominação porque os portadores de corrente no canal de inversão são elétrons e possui quatro terminais: Corpo (“Bulk”), Porta (“Gate”), Dreno (“Drain”) e Fonte (“Source”). Neste trabalho, os terminais serão designados por seus respectivos nomes em Português. Mas, na nomenclatura para o equacionamento, serão usadas as siglas dos respectivos nomes em Inglês por uma questão de compatibilidade com a nomenclatura utilizada pela maioria dos engenheiros.

As diferenças de potencial entre porta e corpo (V_{GB}), entre fonte e corpo (V_{SB}) e entre dreno e corpo (V_{DB}) definem os campos elétricos que atuam sobre a distribuição das cargas fixas e portadores no interior do substrato semiconductor, logo abaixo do óxido formado entre o terminal de porta e o substrato. No exemplo dado na Figura 3-1, quando a tensão V_{GB} é menor do que a tensão de banda plana (definida como a soma de todos os potenciais de contato entre os materiais, e o potencial induzido por cargas parasitas no interior do óxido e na interface entre este e o semiconductor [6]), a região do substrato próxima ao óxido tem sua densidade de lacunas aumentada, caracterizando-se o regime de

acumulação. Quando V_{GB} é levemente superior à tensão de banda plana, a densidade de cargas próxima à interface com o óxido é associada à depleção de lacunas, caracterizando-se o regime de depleção ou inversão fraca. Finalmente quando V_{GB} é positivo e maior do que a denominada tensão de limiar, a densidade de carga de elétrons livres na interface com o óxido aumenta de tal forma que a concentração de elétrons livres na interface supera a de lacunas no corpo do semiconductor, caracterizando-se assim o regime de inversão. Esse regime tem esse nome por ser o canal de inversão constituído por portadores que são minoritários no corpo do semiconductor. A região de inversão é a mais utilizada para polarizar o transistor MOS.

O canal de inversão é constituído por uma finíssima camada de portadores minoritários que podem se deslocar laminarmente na direção do eixo x, mostrado na Figura 3-1, desde que se aplique uma diferença de potencial entre os extremos do referido canal condutor. Considerando que a tensão aplicada ao dreno seja superior à tensão aplicada à fonte, a diferença de potencial entre as faces superior e inferior da camada de óxido isolante é menor nas vizinhanças do dreno do que na fonte. Assim, o nível de inversão do canal é maior próximo à fonte e diminui progressivamente na direção do dreno.

Com o aumento da tensão aplicada ao dreno, a região próxima a ele experimenta uma redução considerável da carga de elétrons em relação à carga na fonte, caracterizando a chamada saturação direta. A saturação direta não impede a continuação do fluxo eletrônico, mas todo o incremento adicional da tensão de dreno servirá para intensificar o campo longitudinal na zona estrangulada, por sua vez aumentando a velocidade dos portadores na região próxima ao dreno [6]. Na chamada saturação direta o dispositivo apresenta uma corrente de dreno praticamente insensível a aumentos da tensão de dreno.

3.3 Modelo de Canal Longo do MOSFET

Nesta seção, será apresentada a formulação para a corrente de dreno do modelo ACM de forma simplificada e breve, de modo que esse equacionamento sirva de base para o equacionamento alvo deste trabalho. Para um estudo mais aprofundado do modelo, sugere-se a leitura de [6]-[10].

3.3.1 Modelo Para a Corrente de Dreno

Como em qualquer componente eletrônico semiconductor, a corrente de dreno I_D possui natureza mista, podendo ser decomposta em uma corrente de deriva, proporcional ao gradiente do potencial elétrico no canal, e uma corrente de difusão, proporcional ao gradiente da concentração de elétrons ao longo do canal [6][7]:

$$I_D = \mu_S W \left(-Q'_I \frac{d\phi_S}{dx} + \phi_t \frac{dQ'_I}{dx} \right) \quad (3.1)$$

onde μ_S é a mobilidade dos portadores, W é a largura do canal, ϕ_S é o potencial de superfície, ϕ_t é o potencial térmico e Q'_I é a densidade de cargas por unidade de área.

Para se derivar uma expressão para a corrente de dreno em termos das densidades de carga de inversão na fonte (Q'_{IS}) e no dreno (Q'_{ID}), utiliza-se a aproximação fundamental segundo a qual Q'_I varia linearmente com ϕ_S . Essa aproximação fundamental é apresentada abaixo:

$$dQ'_I = nC'_{ox} d\phi_S \quad (3.2)$$

onde n é chamado de fator de rampa e $C'_{ox} = \varepsilon_{SiO_2} / t_{ox}$ é a capacitância de óxido por unidade de área, ε_{SiO_2} é a permissividade elétrica do óxido de silício e t_{ox} é a espessura de óxido.

Para obter então a equação para a corrente de dreno, substitui-se (3.2) em (3.1) e faz-se a integração ao longo do canal, resultando em [6]:

$$I_D = I_F - I_R = I(V_{GB}, V_{SB}) - I(V_{GB}, V_{DB}) \quad (3.3)$$

$$I_{F(R)} = \mu_S n C'_{ox} \frac{W}{L} \frac{\phi_t^2}{2} \left[\left(\frac{Q'_{IS(D)}}{nC'_{ox}\phi_t} \right)^2 - 2 \frac{Q'_{IS(D)}}{nC'_{ox}\phi_t} \right] \quad (3.4)$$

onde I_F e I_R são as correntes de saturação direta e reversa respectivamente, Q'_{IS} e Q'_{ID} são as densidades de carga por unidade de área do lado do terminal de fonte e de dreno respectivamente, μ_S é a mobilidade dos portadores, n é o fator de rampa, C'_{ox} é a capacitância de óxido por unidade de área, ϕ_t é o potencial térmico e W e L são a largura e o comprimento de canal equivalentes do transistor.

Os resultados apresentados por (3.3) e (3.4) permitem verificar uma importante propriedade da corrente de dreno em transistores canal longo: essa pode ser escrita como a

diferença entre duas componentes, a saber, a corrente de saturação direta (I_F), que corresponde à densidade de cargas no lado do terminal de fonte, e a corrente de saturação reversa (I_R), que corresponde à densidade de cargas do lado do terminal de dreno.

Para a maioria das aplicações, o transistor é polarizado em saturação direta, ou seja, a corrente de saturação reversa apresenta um valor muito pequeno em relação à corrente de saturação direta, resultando em $I_D \cong I_F$.

Pode-se ainda utilizar (3.4) em uma forma normalizada apresentada abaixo:

$$i_{f(r)} = \left(q'_{IS(D)} \right)^2 - 2q'_{IS(D)} \quad (3.5)$$

onde $i_{f(D)} = I_{F(D)}/I_S$ que são as correntes direta e reversa normalizadas, e $q'_{IS(D)} = Q'_{IS(D)}/Q'_{IP}$ são as densidades de carga por unidade de área normalizadas perto da fonte e do dreno respectivamente. O fator de normalização da corrente $I_S = \mu_0 n C'_{ox} \frac{W}{L} \frac{\phi_t^2}{2}$ é também chamado de corrente específica, enquanto o fator de normalização das densidades de carga $Q'_{IP} = -nC'_{ox} \phi_t$ é também chamado de densidade de cargas de “pinch-off”.

Para relacionar as densidades de carga de inversão com as tensões nos terminais do dispositivo, o modelo ACM utiliza o “Unified Charge Control Model” (UCCM), deduzido em [6][7] e mostrado abaixo por conveniência:

$$V_P - V_{SB(DB)} = \frac{Q'_{IP} - Q'_{IS(D)}}{nC'_{ox}} + \phi_t \ln \left(\frac{Q'_{IS(D)}}{Q'_{IP}} \right) \quad (3.6)$$

onde V_P é a tensão de “pinch-off”, dada no modelo ACM pela seguinte aproximação:

$$V_P = \frac{V_{GB} - V_{T0}}{n} \quad (3.7)$$

onde V_{T0} é a tensão de limiar do transistor.

A expressão apresentada em (3.6) também pode ser utilizada na sua forma normalizada como apresentado a seguir:

$$\frac{V_P - V_{SB(DB)}}{\phi_t} = q'_{IS(D)} - 1 + \ln q'_{IS(D)} \quad (3.8)$$

3.3.2 Modelo da Transcondutância

O modelo de transcondutância que será apresentado faz parte de um modelo mais completo, chamado de modelo de pequenos sinais, o qual apresenta as características dinâmicas que descrevem a operação do transistor MOS em regime quase-estático. Nesse regime, supõe-se que as variações das tensões terminais sejam suficientemente lentas para que as densidades de carga em qualquer posição do canal variem instantaneamente com as tensões terminais [10]. Um modelo de pequenos sinais quase-estático completo deve ser composto por doze parâmetros independentes, sendo que três são transcondutâncias e nove são capacitâncias [6][7].

As definições das três transcondutâncias, mencionadas acima, estão mostradas a seguir:

$$g_{mg} = \left. \frac{\partial I_D}{\partial V_G} \right|_{V_S, V_D, V_B} \quad (3.9)$$

$$g_{ms} = \left. \frac{\partial I_D}{\partial V_S} \right|_{V_G, V_D, V_B} \quad (3.10)$$

$$g_{md} = \left. \frac{\partial I_D}{\partial V_D} \right|_{V_G, V_S, V_B} \quad (3.11)$$

Assim sendo, desconsiderando a variação de n e μ_s em relação à tensão V_{GB} , a partir de (3.3) obtêm-se os seguintes resultados, segundo [6] e [7]:

$$g_{ms} = -\mu_s \frac{W}{L} Q'_{IS} \quad (3.12)$$

$$g_{md} = -\mu_s \frac{W}{L} Q'_{ID} \quad (3.13)$$

$$g_{mg} = \frac{1}{n} (g_{ms} - g_{md}) \quad (3.14)$$

O desenvolvimento completo de (3.12), (3.13) e (3.14) é apresentado em [6] e [7].

A transcondutância g_{mg} será muito importante para o desenvolvimento posterior das equações que serão utilizadas para avaliar as não-linearidades do transistor MOS.

3.4 Efeitos de Segunda Ordem do Transistor MOS

Em aplicações de RF transistores com altas frequências de transição são muito desejáveis, por isso utilizam-se transistores com comprimentos de canal mínimo. Por sua vez, nos transistores de canal mínimo, principalmente em tecnologias avançadas, os

efeitos de segunda ordem são fundamentais para a modelagem precisa do comportamento do transistor nesse tipo de aplicação. Por isso, é fundamental a revisão da modelagem destes efeitos nesta seção.

3.4.1 Modulação do Comprimento do Canal.

Esse efeito está relacionado com as contribuições do campo elétrico longitudinal ao canal para a distribuição espacial das cargas, principalmente na região próxima ao terminal de dreno. Esse problema físico é de ordem bidimensional, mas para fins de simplificação do modelo, considera-se como um efeito elétrico unidirecional no qual o comprimento efetivo do canal é reduzido com o aumento da tensão V_{DS} devido ao avanço da região de depleção.

No modelo ACM, o canal é dividido em duas seções, associadas uma à fonte e a outra ao dreno. Na seção de fonte, a aproximação de canal gradual é válida e a corrente pode ser calculada por um modelo consistente com esta hipótese, porém empregando-se em lugar do comprimento do canal, o comprimento da seção [6]. Para determinar o comprimento da seção de fonte, faz-se uma análise do campo elétrico na seção de dreno. Uma forma bem simples de modelar esse efeito é mostrada em [6][7] e é reproduzida abaixo por conveniência:

$$I_D = \frac{I_{D0}}{\left(1 - \frac{\Delta L}{L}\right)} \quad (3.15)$$

onde I_{D0} é a corrente de dreno de canal longo, calculada através de (3.3), L é o comprimento efetivo do canal e ΔL é o comprimento da seção do dreno, que pode ser modelado da seguinte forma [6][7]:

$$\Delta L = \lambda L_C \ln \left(1 + \frac{V_{DS} - V_{DS_{SAT}}}{L_C U_{CRIT}} \right) \quad (3.16)$$

onde $L_C = \sqrt{\frac{\epsilon_{Si} x_j}{C'_{ox}}}$, x_j é a profundidade da junção, ϵ_{Si} é a permissividade do silício, C'_{ox} é a capacitância de óxido, $V_{DS_{SAT}}$ é a tensão entre dreno e fonte na qual o transistor entra em saturação direta, $U_{CRIT} = \frac{v_{lim}}{\mu}$, v_{lim} é a velocidade limite, μ é a mobilidade dos portadores e λ é um parâmetro de ajuste.

Esse efeito de segunda ordem ou de canal curto é sempre desconsiderado em transistores de canais considerados longos, pois o encurtamento do canal causado por esse fenômeno é muito menor do que o comprimento total do canal. Por outro lado, é muito importante em transistores de canal curto.

3.4.2 Degradação da Mobilidade com o Campo Elétrico Transversal.

Quanto maior o campo elétrico na direção perpendicular ao canal, maior é a densidade de cargas de inversão próximas à camada de óxido de silício, aumentando assim a taxa de colisões contra a interface e entre as partículas. Assim, a mobilidade deixa de ser constante e passa a ser dependente das tensões V_{GB} , V_{SB} e V_{DB} , sendo assim variável ao longo do canal.

A dedução do modelo da mobilidade em função do campo elétrico transversal para o modelo ACM pode ser encontrada em [6], [7] e [9]. O resultado é apresentado a seguir:

$$\mu = \frac{\mu_0}{1 - \frac{\alpha_\theta}{\epsilon_{Si}} (Q'_I + nQ'_B)} \quad (3.17)$$

onde α_θ é o coeficiente de dependência entre mobilidade e campo, μ_0 é a mobilidade dos portadores na ausência de campo e $\frac{Q'_I + nQ'_B}{\epsilon_{Si}}$ é o campo elétrico médio na camada de inversão e $n = 1/2$ para elétrons e $1/3$ para lacunas.

Essa equação da mobilidade pode ser utilizada em (3.1) para inclusão da degradação da mobilidade com o campo elétrico transversal na equação da corrente. Esse efeito é muito importante, apesar de que não será considerado no desenvolvimento subsequente, a exemplo da modulação do comprimento do canal. Isso pelo fato que o efeito estudado aqui é dominante em transistores de canal longo e que para amplificadores utilizados em aplicações de RF são utilizados transistores de canal curto em que a velocidade de saturação dos portadores é o efeito de segunda ordem dominante.

3.4.3 Saturação da Velocidade dos Portadores.

Para o desenvolvimento do modelo de primeira ordem, ou de canal longo, não se considerou a relação entre a velocidade dos portadores e o campo elétrico longitudinal ao

canal. O efeito da saturação da velocidade dos portadores no modelo ACM é baseado na equação abaixo:

$$\mu_s = \frac{\mu}{1 + (\mu/v_{lim})(d\phi_s/dx)} \quad (3.18)$$

onde $\mu/v_{lim} = 1/U_{CRIT}$ e U_{CRIT} é o campo elétrico crítico que corresponde a uma velocidade de metade da velocidade limite, e μ é a mobilidade, dependente do campo transversal.

A mobilidade μ pode ser substituída por μ_0 , no caso de o modelo levar em consideração apenas o efeito da velocidade de saturação dos portadores para o desenvolvimento de equações aproximadas. No próximo capítulo, isso será considerado.

Substituindo (3.18) em (3.1), usando a aproximação fundamental do modelo ACM (3.2), e integrando esse resultado ao longo do canal, obtém-se uma equação geral para a corrente de dreno, considerando o fenômeno da saturação da velocidade dos portadores [9], que é apresentada abaixo na sua forma normalizada:

$$id = \frac{(q'_{IS} - q'_{ID})}{1 + \varepsilon(q'_{IS} - q'_{ID})} (q'_{IS} + q'_{ID} + 2) \quad (3.19)$$

onde $\varepsilon = \frac{\phi_t}{LU_{CRIT}}$ é um parâmetro referente à saturação de velocidade e L é o comprimento efetivo do canal.

Note-se que, para transistores de canal longo, tem-se $\varepsilon \ll 1$ (por exemplo, $\varepsilon = 0,0045$ para um transistor com $L = 1\mu m$) e o efeito da saturação da velocidade dos portadores tem pouca influência no cálculo da corrente de dreno. Note que o valor de ε é intimamente ligado ao comprimento de canal e que para o modelo a consideração ou não do efeito da velocidade limite se dá de forma automática.

Segundo (3.19) se q'_{IS} for fixado em determinado valor a corrente de dreno aumenta à medida que a q'_{ID} diminui. Ao mesmo tempo em que q'_{ID} diminui, a velocidade dos portadores deve aumentar, pelo fato que a diminuição de q'_{ID} implica em um aumento do campo longitudinal segundo o UCCM. No entanto essa velocidade não pode exceder a velocidade limite. Portanto a corrente máxima que pode fluir quando o transistor alcança a saturação direta é limitada pela velocidade de saturação dos portadores e a mínima densidade de carga no lado de dreno [7]:

$$I_{D_{SAT}} = -WV_{lim}Q'_{ID_{SAT}} \quad (3.20)$$

ou na sua forma normalizada:

$$i_{d_{sat}} = \frac{2}{\varepsilon} q'_{ID_{SAT}} \quad (3.21)$$

onde $i_{d_{sat}} = I_{D_{SAT}} / I_S$ é a corrente de dreno normalizada com a velocidade dos portadores saturada, e $q'_{ID_{SAT}} = Q'_{ID_{SAT}} / Q'_{IP}$ é a densidade de cargas do lado de dreno normalizada com a velocidade dos portadores saturada.

Substituindo (3.21) em (3.19) com o objetivo de se obter $q'_{ID_{SAT}}$ para qualquer regime de operação:

$$\frac{q'_{ID_{SAT}}}{q'_{IS}} = 1 + \frac{\varepsilon + 1}{\varepsilon q'_{IS}} \left(1 - \sqrt{1 + \frac{2\varepsilon q'_{IS}}{(\varepsilon + 1)^2}} \right) \quad (3.22)$$

A expressão apresentada em (3.22) terá papel fundamental no desenvolvimento do modelo de não-linearidades do MOSFET que será apresentado na próxima seção. A dedução de todas as equações apresentadas aqui estão detalhadas em [6][7][9].

3.5 Conclusão

Nesse capítulo, foi apresentado um modelo do transistor MOS, que é fundamental para o projeto de circuitos integrados em tecnologia CMOS (Sigla em Inglês para “Complementary Metal Oxide Semiconductor”). Aqui não se encontra apresentado um modelo completo, mas apenas as equações que serão utilizadas no desenvolvimento do modelo de não-linearidades. Para quem desejar conhecer mais profundamente os princípios que circundam a modelagem de transistores e até mesmo o modelo ACM, recomenda-se o estudo de [6]-[11]. Em algumas das referências apresentadas, pode-se verificar o desenvolvimento de todas as equações apresentadas aqui de forma rápida, pois este trabalho não tem por objetivo o seu desenvolvimento.

Pôde-se verificar também, nesse capítulo, a simplicidade apresentada pelo modelo ACM em suas expressões. A maioria dessas relações mostradas aqui podem ser implementadas em um simulador de circuitos em uma versão mais completa, ou seja, os cálculos de corrente devem considerar todos os efeitos de segunda ordem. As expressões como foram apresentadas de forma modular servem para o entendimento dos fenômenos físicos envolvidos e para o projeto de circuitos à mão.

4 MODELAGEM DE IIP3 DE UM AMPLIFICADOR FONTE COMUM

4.1 Introdução

Em trabalhos anteriores, [13]-[17] foi constatado, mesmo através de medições, um aumento significativo da linearidade, (utilizando como referência IIP3) na região de inversão moderada, em que se pode alcançar um bom compromisso entre consumo de potência, frequência de transição e linearidade para transistores de canal curto. Esse aumento foi chamado de “Sweet Spot” em [13] e [16] e pode ser um interessante critério de projeto para circuitos de radiofrequência de baixo consumo, em que linearidade é um fator relevante. Apesar de apresentar o “Sweet Spot” como uma realidade, a literatura o formaliza com equações complexas e, em certos casos, apresenta medidas sem apresentar uma discussão sobre os fenômenos físicos ligados a ele.

Este capítulo se propõe a apresentar o desenvolvimento de um modelo compacto para o comportamento de IIP3, em relação à polarização de transistores MOS, baseado no modelo ACM. O objetivo é obter equações simples e confiáveis para serem utilizadas no projeto de amplificadores, operando principalmente em inversão fraca e moderada. Além disso, pretende-se mostrar os fenômenos físicos envolvidos na origem do “Sweet Spot”.

Para alcançar os objetivos propostos acima, fez-se uma análise de não linearidade em um amplificador fonte comum, pois esse serve de base para outras topologias muito utilizadas em circuitos de radiofrequência, utilizando os conceitos apresentados nos capítulos anteriores.

Este capítulo está organizado desta forma: após esta breve introdução, uma análise não-linear de um amplificador fonte comum será apresentada em 4.2, as equações de transcondutância e suas derivadas em 4.3, IIP3 em função da polarização do transistor em 4.4 e, finalmente, um estudo das causas do “Sweet Spot” em 4.5.

4.2 Análise não-linear de um amplificador fonte comum

O amplificador fonte comum, como já foi mencionado anteriormente, é a topologia mais simples de amplificador. No entanto, ela permite que sua análise, tanto de ruído quanto de linearidade, possa ser estendida a outras topologias mais complexas, além de permitir que se faça uma análise do comportamento de linearidade intrínseca do transistor

MOS. Um exemplo de topologia baseada no amplificador fonte comum é mostrado na Figura 2-7, que é a configuração comumente chamada de amplificador “cascode” e que é amplamente utilizada em circuitos de comunicação como o LNA.

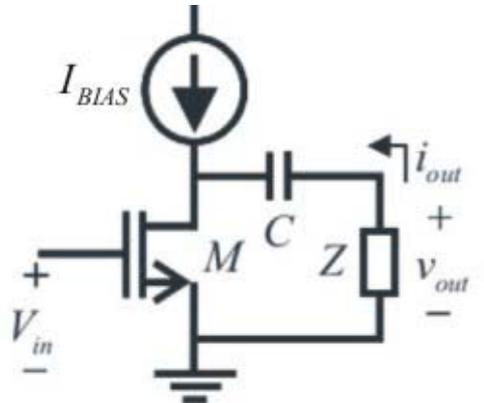


Figura 4-1: Amplificador fonte comum.

A Figura 4-1 mostra um esquemático do amplificador no qual o comportamento de IIP3 será analisado com relação à polarização do circuito. Para essa análise são feitas algumas considerações com relação ao circuito mostrado acima:

- A primeira delas é que a grandeza de saída da característica de transferência, que será deduzida mais adiante, será a corrente de dreno.
- A segunda consideração é que o transistor é polarizado na região de saturação direta.

De acordo com o que foi visto para o sistema hipotético da Figura 2-1, em que a característica de transferência foi expandida em série de potências, a característica de transferência do circuito da Figura 4-1 resulta na seguinte expressão geral para a corrente de saída:

$$i_{out} = \frac{dI_{BIAS}}{dV_{in}} v_{in} + \frac{1}{2!} \frac{d^2 I_{BIAS}}{dV_{in}^2} v_{in}^2 + \frac{1}{3!} \frac{d^3 I_{BIAS}}{dV_{in}^3} v_{in}^3 \quad (4.1)$$

onde v_{in} é a amplitude do sinal de entrada e V_{in} é a tensão de polarização DC.

Uma verificação rápida na Figura 4-1 evidencia que, para essa topologia, $V_{in} = V_G$ e $I_{BIAS} = I_D$, já que o terminal de corpo está conectado à fonte. Assim sendo, as derivadas que compõem os fatores da expansão em série de potências são obtidos através de g_{mg} , equação (3.9), e suas derivadas (g'_{mg} , g''_{mg} em relação à V_G). Lembra-se que neste capítulo se utilizará a aproximação de distorção fraca, ou seja, trunca-se a expansão a partir do

termo correspondente a g''_{mg} e esse o único termo relacionado com a distorção por intermodulação.

Substituindo os fatores que multiplicam v_{in} e v_{in}^3 em (2.13) obtém-se:

$$V_{IP3}^2 = \frac{4}{3} \frac{dI_D}{dV_G} \bigg/ \frac{1}{3!} \frac{d^3 I_D}{dV_G^3} = 8 \frac{g_{mg}}{g''_{mg}} \quad (4.2)$$

Através de (4.2), pode-se observar a relação entre V_{IP3} e a polarização do transistor representada por sua tensão de porta. Assim, é possível determinar uma expressão de V_{IP3} em função da polarização do transistor, utilizando para esse propósito o modelo de dispositivo apresentado no capítulo anterior.

4.3 Equações de Transcondutância e suas Derivadas

Como anteriormente foi dito, deve-se agora introduzir o modelo de dispositivo em (4.2), de forma que se possa desenvolver uma expressão, se possível explícita, para V_{IP3} em função da polarização do transistor.

O fato de que as novas tecnologias de fabricação de dispositivos CMOS permitem, em suas regras, comprimentos de canal mínimos cada vez menores, faz com que a frequência de transição desses dispositivos atinja valores da ordem de grandeza requeridos pelas aplicações de RF. Então a utilização de transistores com comprimentos de canal mínimos em tecnologias modernas faz com que esses sejam modelados adicionando efeitos relacionados às diminutas dimensões dos transistores atualmente utilizados no projeto de circuitos para aplicações de RF. Assim sendo, numa primeira aproximação, será considerado apenas o efeito de velocidade de saturação dos portadores, porque esse se apresenta como principal efeito de segunda ordem, relacionado com a diminuição do comprimento de canal mínimo das tecnologias utilizados atualmente.

O efeito da velocidade de saturação dos portadores foi apresentado no capítulo anterior, em que (3.21) é apresentada como a corrente de dreno normalizada em função da densidade de cargas normalizadas do lado de dreno do transistor, operando com a velocidade dos portadores saturada. Fazendo a derivada implícita de (3.21) em relação à tensão de porta, obtém-se:

$$g_{mg} = \frac{dI_D}{dV_G} = I_S \frac{2}{\varepsilon} \frac{dq'_{ID_{SAT}}}{dV_G} \quad (4.3)$$

onde $I_S = \mu_0 n C'_{ox} \frac{W \phi_t}{L 2}$ é a corrente específica já definida no capítulo anterior e repetida aqui por conveniência. Nessa dedução supõe-se que I_S e ε dependem muito pouco de V_G .

Note que (4.3) se apresenta em função da derivada da densidade de cargas normalizada no lado de dreno, quando a velocidade dos portadores está saturada, em relação à tensão de porta. Pode-se utilizar (3.22), que é a relação entre as densidades de carga normalizada do lado de dreno e do lado de fonte, para se obter uma ligação entre $q'_{ID_{SAT}}$ e V_G , pois a densidade de carga normalizada do lado de fonte está relacionada com a tensão de porta através do UCCM (3.8).

Analisando (4.3) e (3.22), pode-se prever que de (4.2) provavelmente resultará em uma equação complexa, pelo fato de haver uma raiz contendo a variável de derivação q'_{IS} . Esse fato inibe o desenvolvimento de um modelo que possa ser utilizado para prever V_{IIP3} através de cálculos a mão, além de resultar em equações complexas para a segunda derivada da transcondutância. Então, na tentativa de simplificar (3.22), multiplicou-se e dividiu-se a mesma pelo conjugado do termo entre parêntesis, resultando em:

$$\frac{q'_{ID_{SAT}}}{q'_{IS}} = \frac{2\varepsilon(1+q'_{IS}/2)}{1+\varepsilon(q'_{IS}+1)+\sqrt{1+2\varepsilon(q'_{IS}+1)+\varepsilon^2}} \quad (4.4)$$

A equação (4.4) permite algumas considerações para sua simplificação. Para essa simplificação considera-se $\varepsilon \ll 1$ e $q'_{IS} \ll 1$. Assim o termo referente à raiz quadrada resulta igual a um para as regiões de inversão fraca e início da inversão moderada. Essa simplificação foi feita considerando que se pode tolerar certa imprecisão no cálculo da carga para inversão forte, se essa não comprometer a precisão em inversão fraca e moderada, pois essas são as regiões de maior interesse neste trabalho. Assim (4.4) pode ser simplificada como segue:

$$\frac{q'_{ID_{SAT}}}{q'_{IS}} \cong \frac{2\varepsilon(1+q'_{IS}/2)}{2+\varepsilon(q'_{IS}+1)} \quad (4.5)$$

Essa aproximação feita para se obter (4.5) mantém a precisão do modelo para a inversão fraca e moderada, porém apresenta erro na região de polarização ($q'_{IS} \gg 1$) onde a parcela simplificada tem peso significativo para o cálculo da densidade de cargas.

A expressão (4.5) oferece a possibilidade de um equacionamento muito simples para g_{mg} e suas derivadas, como será verificado adiante.

Desenvolvendo a derivada de (3.8) em relação à V_G e lembrando que $dV_P/dV_G = 1/n$, obtém-se como resultado dq'_{IS}/dV_G , que por sua vez é substituído no resultado da derivação de (4.5) em relação à V_G . O resultado de tudo isso é substituído em (4.3), para se obter finalmente o resultado da primeira derivada da corrente de dreno em relação à tensão de porta, mostrado abaixo:

$$g_{mg} = \frac{dI_D}{dV_G} = \frac{2I_S}{n\phi_t} \frac{q'_{IS}}{2 + \varepsilon(1 + q'_{IS})} \left\{ \frac{2(1 + q'_{IS})[2 + \varepsilon(1 + q'_{IS})] - \varepsilon q'_{IS}(2 + q'_{IS})}{(1 + q'_{IS})[2 + \varepsilon(1 + q'_{IS})]} \right\} \quad (4.6)$$

Pode-se aproximar o termo $(2 + q'_{IS})$ mostrado em (4.6) por $(1 + q'_{IS})$ para o regime de inversão forte ($q'_{IS} \gg 10$), pois para os regimes de inversão fraca e moderada o termo em questão é desprezível. Assim obtém-se:

$$g_{mg} = \frac{dI_D}{dV_G} = \frac{2I_S}{n\phi_t} \frac{q'_{IS}}{2 + \varepsilon q'_{IS}} \left(\frac{4 + \varepsilon q'_{IS}}{2 + \varepsilon q'_{IS}} \right) \quad (4.7)$$

Verifica-se em (4.7) que a transcondutância é uma função da densidade de cargas de inversão normalizada no lado de fonte. Como a densidade de cargas normalizada do lado da fonte tem uma relação direta com a tensão de porta, então se pode utilizar a densidade de cargas normalizada como um parâmetro que se refere à polarização do circuito e nos dá uma indicação sobre o nível de inversão aplicado ao dispositivo.

Para a obtenção de g'_{mg} , deriva-se (4.7) em relação à tensão de porta e substitui-se dq'_{IS}/dV_G pelo resultado obtido com a derivação de (3.8) em relação à tensão de porta. O processo para a obtenção de g''_{mg} resume-se em derivar g'_{mg} em relação à tensão de porta e substituir novamente dq'_{IS}/dV_G pelo resultado obtido com a derivação de (3.8) em relação à tensão de porta. Os resultados dos desenvolvimentos descritos acima são os seguintes:

$$g'_{mg} = \frac{\partial^2 I_D}{\partial V_G^2} = \frac{16I_S}{(n\phi_t)^2} \frac{q'_{IS}}{(q'_{IS} + 1)} \frac{1}{(2 + \varepsilon q'_{IS})^3} \quad (4.8)$$

$$g''_{mg} = \frac{\partial^3 I_D}{\partial V_G^3} = \frac{16I_S}{(n\phi_t)^3} \frac{q'_{IS}}{(q'_{IS} + 1)^3} \frac{2 - 2\varepsilon q'_{IS} - 3\varepsilon q'^2_{IS}}{(2 + \varepsilon q'_{IS})^4} \quad (4.9)$$

O equacionamento apresentado acima resultando em (4.7), (4.8) e (4.9) é apresentado com detalhes no Anexo 1.

Agora (4.7), (4.8) e (4.9) podem ser utilizadas para compor as definições das figuras de mérito de não-linearidades, mostradas na seção 2. Além disso, a transcondutância e suas derivadas podem ser utilizadas em modelos mais complexos, onde não se utiliza a aproximação de distorção limitada. Em especial, pretende-se neste trabalho estudar o comportamento de V_{HP3} em relação à polarização do transistor, por isso se utilizará mais adiante apenas (4.7) e (4.9).

A Figura 4-2 mostra graficamente como se comportam (4.7), (4.8) e (4.9) em função de q'_{IS} para um transistor que seria fabricado na tecnologia TSMC 0.18 μm com o comprimento de canal mínimo e razão de aspecto $W/L = 100$ ($\varepsilon = 0.034$) e para o mesmo transistor considerando a aproximação de canal longo ($\varepsilon = 0$). O valor de L , mostrado na figura, corresponde ao comprimento de canal nominal do transistor, ou seja, é o comprimento mínimo permitido para o desenho do dispositivo na referida tecnologia.

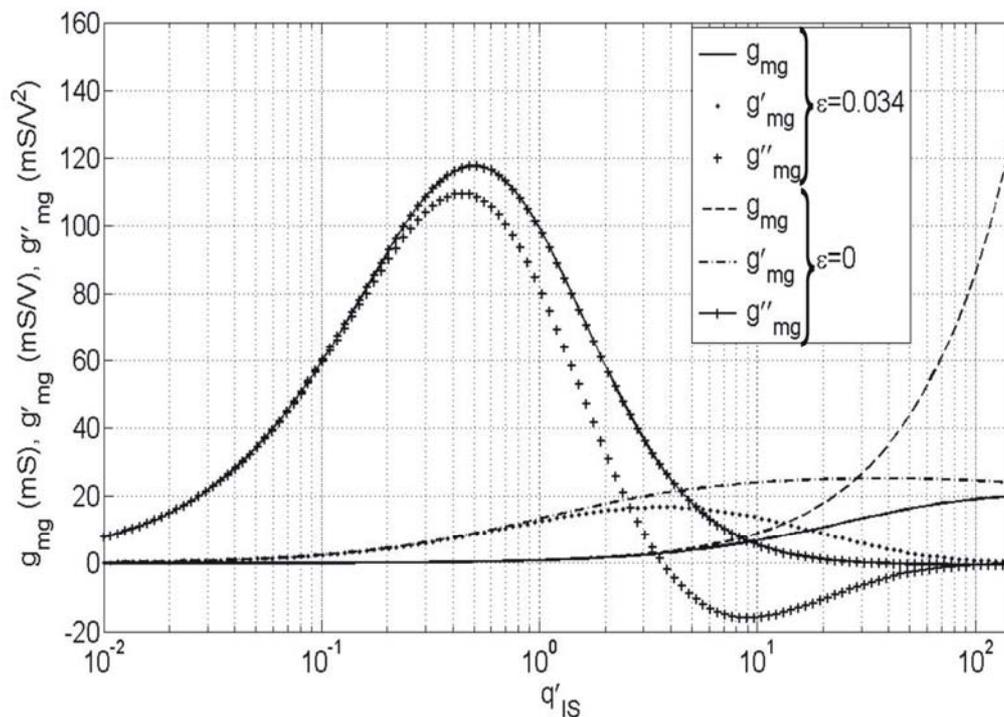


Figura 4-2: Transcondutância e suas derivadas.

Pode-se verificar na Figura 4-2 que o comportamento da transcondutância e de suas derivadas para um transistor de canal curto é bastante diferente do comportamento de um transistor de canal longo. Considerando uma curva $I_D \times V_G$ de um transistor considerado de canal longo, para tensões $V_G < V_{T0}$ (inversão fraca $q'_{IS} \ll 1$), o mecanismo

de transporte de cargas pode ser modelado como uma função exponencial; já para $V_G > V_{T0}$ (inversão moderada e forte $q'_{IS} \geq 1$), o mecanismo de transporte pode ser modelado por uma função quadrática. Considerando agora o comportamento das derivadas da corrente de dreno, pode-se dizer que qualquer derivada para $V_G < V_{T0}$, a função continuará sendo exponencial; já para $V_G > V_{T0}$, a derivada primeira tenderá a uma função de primeiro grau, a derivada segunda tenderá a uma função constante e, finalmente, a terceira derivada tenderá a zero desprezando o efeito de degradação da mobilidade com o campo transversal, como pode ser verificado na Figura 4-2. Essa é uma visão extremamente simplificada, mas que pode ser muito útil para se entender um pouco da diferença entre os comportamentos de transistores, em que se possam utilizar a aproximação de canal longo e os que devem utilizar as aproximações de canal curto. Como já foi mencionado, pode-se verificar na Figura 4-2 um comportamento bastante diferente do descrito acima para um transistor de canal longo. A maior diferença entre os comportamentos se apresenta na transição entre inversão moderada e forte, em que a primeira derivada da corrente tende a uma saturação, ou função constante. Por conseqüência a segunda derivada tende a zero, e a terceira derivada também tende a zero em inversão forte, mas com uma passagem por zero em um ponto de inversão moderada, pelo fato de haver uma inversão na taxa de variação da segunda derivada. A análise apresentada é baseada na estrutura mostrada na Figura 4-1 que tem o terminal de corpo conectado ao terminal de fonte do transistor.

4.4 IIP3 em função da Polarização do Transistor

Substituindo (4.7) e (4.9) em (4.2), obtém-se um resultado simples para V_{IIP3}^2 , mostrado abaixo:

$$V_{IIP3}^2 = (n\phi_t)^2 \left[\varepsilon q'_{IS}{}^2 + (2 + \varepsilon) q'_{IS} + 2 \right] \left| \frac{(1 + q'_{IS})(4 + \varepsilon q'_{IS})}{2 - 2\varepsilon q'_{IS} - 3\varepsilon q'_{IS}{}^2} \right| \quad (4.10)$$

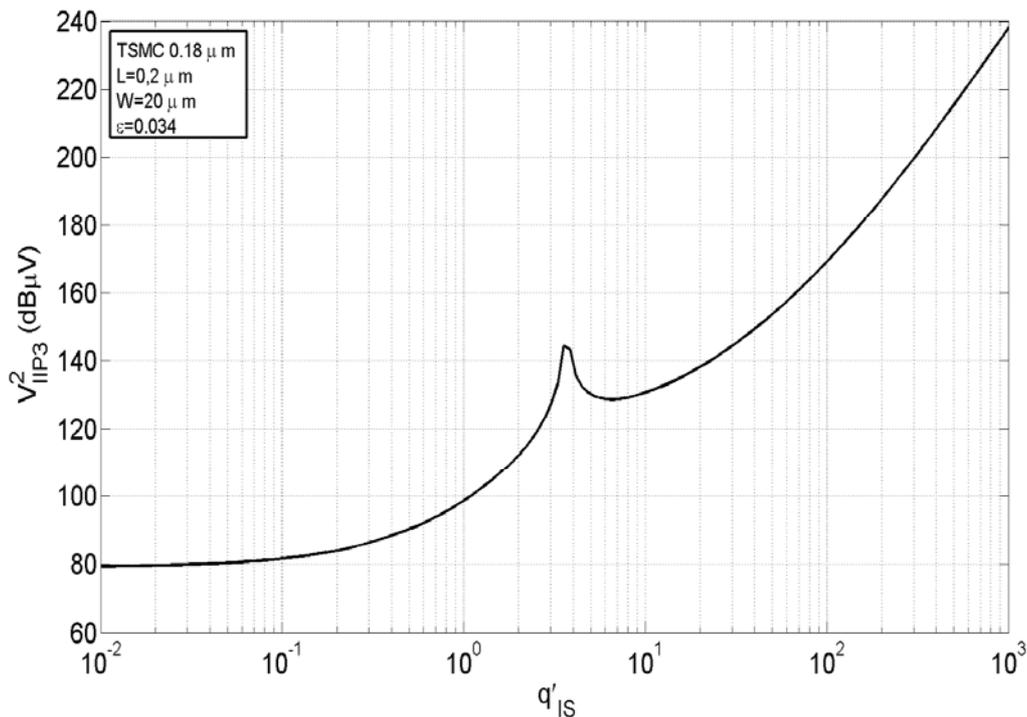


Figura 4-3: V_{IP3}^2 em função da densidade de cargas normalizada no lado da fonte.

Na Figura 4-3, o comportamento de V_{IP3}^2 em função da densidade de cargas normalizada no lado de fonte apresenta um ponto singular, chamado de “Sweet Spot” em [13], que já era esperado devido às raízes da equação de segundo grau no denominador de (4.10). Essa é uma importante vantagem da equação (4.10) em relação ao que já foi mostrado em [13]-[17], em que nenhuma equação explícita é apresentada com o objetivo de definir a polarização na qual o “Sweet Spot” é encontrado. Além disso, seria possível prever esse comportamento singular a partir da Figura 4-2, onde se verifica uma passagem por zero no mesmo ponto de polarização, que por sua vez traduz-se em um ponto de V_{IP3}^2 infinito segundo (4.2).

Uma observação muito importante sobre esse comportamento é que, para transistores fabricados com canal curto, ele aparece na região considerada de inversão moderada [13]-[15], o que faz pensar que, se for possível polarizar o circuito nesse ponto, pode-se alcançar um interessante compromisso entre consumo de potência e alta linearidade.

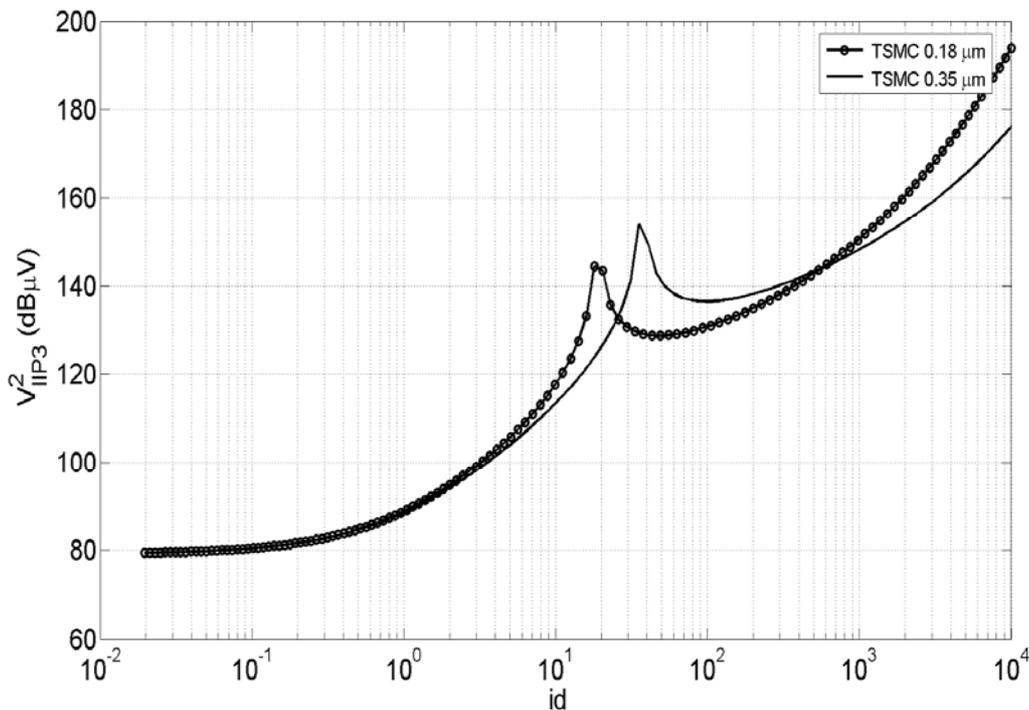


Figura 4-4: V_{IP3}^2 em função da corrente de dreno normalizada.

A Figura 4-4 apresenta o comportamento do ponto de intersecção de terceira ordem em função da corrente de dreno normalizada para transistores de duas tecnologias diferentes: TSMC 0.18μm e TSMC 0.35μm. Os dados referentes às tecnologias e aos transistores são mostrados abaixo. A corrente de dreno normalizada, que é apresentada como variável independente na Figura 4-4, é uma grandeza bem conhecida entre projetistas e corresponde à corrente de polarização do transistor.

Pode-se notar também, na Figura 4-4, que o “Sweet Spot” não se encontra no mesmo ponto de polarização para os dois transistores. O transistor com o maior comprimento de canal apresenta o “Sweet Spot” localizado em uma corrente de dreno normalizada maior em relação ao transistor com menor comprimento de canal, mas ainda dentro da região que pode ser chamada de inversão moderada. Observando (4.10), pode-se dizer que a diferença entre os dois pontos se deve à diferença entre os valores de ε para cada transistor. A diferença entre os parâmetros ε das duas tecnologias em questão mostra que o comprimento do canal é a principal variável com relação à maior ou menor influência que um transistor sofrerá, devido ao efeito de saturação da velocidade dos portadores.

Tabela 4-1: Transistores utilizados.

Tecnologia	L (μm)	W (μm)	ε	μ_0 (cm^2/Vs)
TSMC 0.18 μm	0.2	20	0.0345	264.06
TSMC 0.35 μm	0.4	40	0.0177	426.39

Os valores de ε e μ_0 mostrados na Tabela 4-1 foram extraídos da tabela de parâmetros tecnológicos do modelo BSIM fornecida em [19], utilizando a metodologia que será apresentada no próximo capítulo. Esses valores estão apresentados aqui por conveniência.

4.5 Estudo das Causas do “Sweet Spot”

Apesar de alguns trabalhos mostrarem a presença do “Sweet Spot” de forma analítica e experimental, a citar [13]-[17], alguns não discutem, de forma clara, as causas que podem levar a esse repentino aumento de linearidade na região de inversão moderada. Além disso, [13] utilizou um modelo de canal longo com o intuito de provar que o “Sweet Spot” é um fenômeno inerente ao funcionamento do transistor, mesmo sem considerar nenhum efeito de segunda ordem. Essa afirmação fez com que ela fosse verificada através de uma comparação da terceira derivada da corrente de dreno em relação à tensão de porta calculadas com o modelo ACM canal longo, com o modelo apresentado por [13] e com medidas em um transistor desenhado, utilizando a tecnologia AMIS 1.5 μm [19] com as seguintes dimensões: $L = 16\mu\text{m}$ e $W = 1600\mu\text{m}$. Além disso, essa comparação tem o objetivo de verificar se outros efeitos de segunda ordem, além da saturação da velocidade dos portadores, estão na origem do “Sweet Spot”.

O modelo da corrente de dreno utilizado em [13] é apresentado abaixo por conveniência:

$$I_D = \left(2n\mu_0 C'_{ox} \phi_t^2 \frac{W}{L} \right) \left[\ln \left(1 + e^{\frac{(V_{GS} - V_{TH})}{2n\phi_t}} \right) \right] \quad (4.11)$$

Em (4.11), I_D é a corrente de dreno, V_{GS} é a tensão entre porta e fonte, n é o fator de rampa, μ_0 é mobilidade para baixos campos elétricos, C'_{ox} é a capacitância de óxido por unidade de área, V_{TH} é a tensão de limiar do transistor e ϕ_t é o potencial térmico. É

importante afirmar que nesse modelo, assim como para o modelo ACM, μ_0 , V_{TH} , e L são considerados constantes.

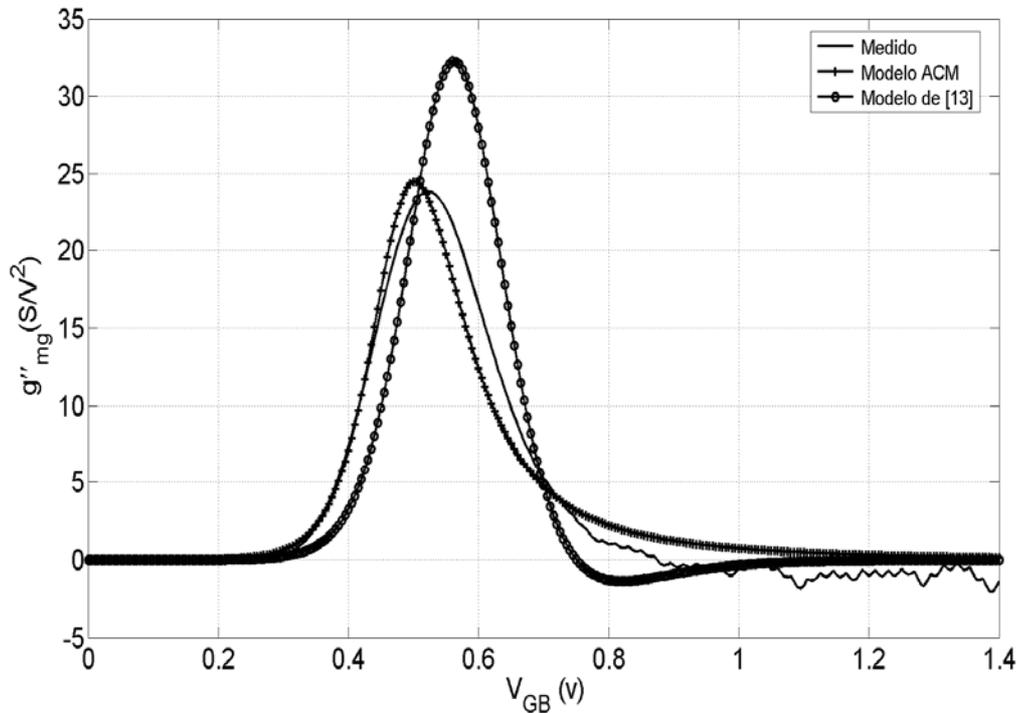


Figura 4-5: Comparação entre modelos e medida.

A Figura 4-5 mostra a comparação mencionada anteriormente. A medida (cuja as condições são mostradas na Tabela 5-1 e a metodologia apresentada no capítulo 5) foi feita de forma que o campo elétrico longitudinal não atingisse o campo crítico, dado por: $U_{CRIT} = v_{lim}/\mu_0$. Para isso, estimou-se o campo crítico através dos parâmetros tecnológicos fornecidos pelo programa MOSIS [19], resultando em $U_{CRIT} = 38,1 \times 10^3 V/cm$. Então aplicou-se uma tensão $V_{DS} = 2V$ fixa que resultou em um campo longitudinal aproximado de $1,2 \times 10^3 V/cm$, ou seja, menor do que o campo crítico estimado.

A observação mais importante a se fazer sobre a Figura 4-5 é que, assim como a medida feita, o modelo (4.11) apresenta também uma passagem por zero, apesar de estar localizada em um valor de polarização muito diferente em relação à medida. Em outras palavras, o “Sweet Spot” está presente também em transistores considerados de canal longo. Uma observação sobre o modelo ACM de canal longo é que ele apresenta uma maior precisão em relação à curva levantada experimentalmente do que o modelo apresentado em [13], apesar de não apresentar a passagem por zero da transcondutância.

Nos modelos apresentados a mobilidade dos portadores e a tensão de limiar são considerados constantes para qualquer região de operação. Além disso, o modelo ACM é um modelo baseado na física e sua aproximação de canal longo já provou ser precisa [8]. A curva que resulta de medidas experimentais, apesar da preocupação de se evitar a saturação da velocidade dos portadores, ainda provavelmente apresenta a ação de efeitos de segunda ordem como, por exemplo, degradação da mobilidade devido ao campo elétrico transversal. Pode-se concluir então, que a passagem por zero de g_{mg}'' , presente até em transistores de canal longo, deve-se a efeitos de segunda ordem, especialmente aqueles que provocam qualquer degradação na mobilidade dos portadores e que o modelo canal longo, apresentado em [13], apresenta alguma inconsistência. Aqui deve-se lembrar que o efeito de saturação da velocidade dos portadores não é caracterizado, do ponto de vista físico, como uma degradação da mobilidade. No entanto, a saturação da velocidade dos portadores é modelada como se fosse uma degradação da mobilidade com o campo elétrico longitudinal.

4.6 Conclusão

Nesse capítulo, um modelo compacto para avaliar o comportamento de IIP3 em relação à polarização de transistores MOS, que pode ser estendido para a análise não linear de amplificadores, baseado no modelo ACM, foi apresentado. Esse modelo apresenta como característica principal a simplicidade das suas equações aliada à maneira explícita e confiável com que se determina o ponto de polarização, em que ocorre o aumento de linearidade chamado de “Sweet Spot”.

Comparando a figura de mérito V_{IIP3}^2 de um amplificador fonte comum para duas tecnologias e dois comprimentos de canal diferentes, verificou-se que a principal variável, que pode ser controlada por projetistas, que determina a região de operação, em que o “Sweet Spot” ocorrerá, é o comprimento de canal, pois o principal efeito de segunda ordem na origem desse fenômeno em transistores considerados de canal curto é a saturação da velocidade dos portadores. Contudo a constatação de que o “Sweet Spot” também ocorre para transistores considerados de canal longo revelou que outros efeitos de segunda ordem, como por exemplo, a degradação da mobilidade devido ao campo elétrico transversal, tem contribuição para a ocorrência do referido fenômeno.

5 RESULTADOS EXPERIMENTAIS.

5.1 *Introdução*

No capítulo anterior, um modelo compacto para o comportamento de IIP3 em função da polarização do transistor foi apresentado com o objetivo de oferecer uma ferramenta de cálculo simplificada para projetistas. Além disso, foi verificada uma singularidade no cálculo de V_{IIP3}^2 em função da polarização do transistor, chamada de “Sweet Spot”, que consiste em um aumento repentino de linearidade na região de inversão moderada.

Com o objetivo de validar o modelo apresentado no capítulo anterior e verificar a precisão do modelo para prever o “Sweet Spot”, algumas medidas feitas em transistores integrados em tecnologias comerciais bem conhecidas foram realizadas. Os resultados destas medidas serão apresentados neste capítulo e comparados com os resultados analíticos que foram apresentados anteriormente.

Além de apresentar uma comparação entre os cálculos analíticos, obtidos através do modelo, e os resultados experimentais, este capítulo apresentará também a metodologia de medidas utilizada para obter tais resultados. O capítulo mostra também como são tratados os parâmetros tecnológicos para a determinação de ε .

Este capítulo será organizado como segue: após esta introdução, a metodologia de medição aplicada é apresentada em 5.2, a extração de parâmetros tecnológicos relacionados com os efeitos de segunda ordem em 5.3 e uma comparação entre os resultados obtidos e aqueles previstos pelas equações teóricas em 5.4.

5.2 *Metodologia de medição*

A metodologia de medida utilizada para a determinação experimental da transcondutância, suas derivadas e conseqüentemente V_{IIP3}^2 em relação à polarização do transistor é bastante simples. Descreve-se abaixo o seu funcionamento:

1º. Medição, em corrente contínua, da curva $I_D \times V_G$ com o transistor em saturação direta e com a tensão de dreno mantida constante num valor que não exceda a tensão máxima permitida pela tecnologia.

2º. Os dados obtidos no primeiro passo são processados em ferramentas computacionais que trabalham com cálculo matricial, como por exemplo Matlab[®] ou Scilab[®].

A Figura 5-1 apresenta o esquemático elétrico que representa as ligações feitas para a determinação experimental das curvas $I_D \times V_G$ mencionadas no primeiro passo para a metodologia de medida. Na figura, estão representadas as ligações elétricas feitas entre o DUT (“Device Under Test”), que indica o dispositivo a ser testado, e as fontes de alimentação, que representam as unidades de estímulo e medida (SMU – Sigla em Inglês para Source/Monitor Unit) de um analisador de parâmetros de semicondutores.

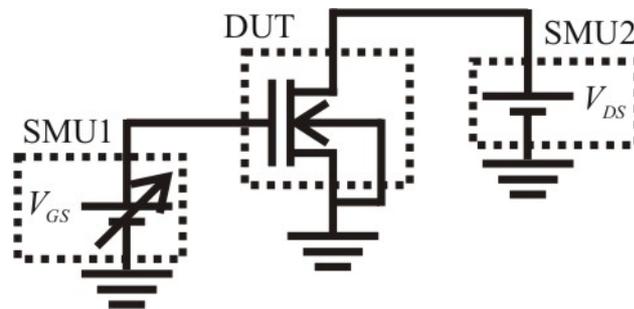


Figura 5-1: Esquemático representando as ligações feitas para a medição.

O equipamento utilizado para a extração da curva $I_D \times V_{GB}$ foi um analisador de parâmetros de semicondutores de precisão Agilent[®] 4156C. Este equipamento possui três unidades de estímulo e medida, que servem como fontes de tensão ou corrente, e de medidores de tensão e corrente. A Figura 5-1 mostra como essas unidades foram ligadas ao dispositivo a ser testado. A SMU1 faz uma varredura da tensão de porta, indicada como V_{GS} . A SMU2 tem duas funções: a primeira é a de aplicar uma tensão V_{DS} constante que dependerá da tecnologia em que o transistor sob teste foi fabricado, e a segunda função é a de medir a corrente de dreno.

Os transistores testados e as condições de medidas são mostrados na Tabela 5-1. Pode-se verificar que, para as duas tecnologias denominadas submicrométricas, foram testados dois transistores: um de canal mínimo e outro com dez vezes o canal mínimo. Já para a tecnologia de canal mínimo de $1.5\mu\text{m}$, apenas um transistor com um canal de $16\mu\text{m}$ foi testado para fazer o estudo sobre o “Sweet Spot” em transistores de canal longo analisados no capítulo anterior.

Uma das preocupações que se deve ter quando se propõe validar um modelo teórico é que, se possível, as considerações feitas para se desenvolver o equacionamento

teórico possam ser realizadas nas metodologias de medição. Um exemplo disso é que os transistores foram polarizados de forma a operarem em saturação direta e com uma tensão V_{DS} constante, pelo fato de as equações considerarem o transistor em saturação direta e que não há variação da tensão de dreno. Para o transistor da tecnologia AMIS 1.5 μm , a tensão V_{DS} teve uma restrição adicional: o campo elétrico transversal deveria ser inferior a U_{CRIT} , como já mostrado no capítulo 4 para as medidas com o transistor de canal longo.

Tabela 5-1 Propriedades dos transistores e condições de medida utilizadas.

TSMC 0.18 μm								
Transistor	L (μm)	W (μm)	N	W/L_{total}	V_{GS} (V)	V_{DS} (V)	U_{CRIT} ¹ (kV/cm)	V_{DS}/L (kV/cm)
M1	0,2	2,0	10	100	$0 < V_{GS} < 1,2$	1,2	38,63	60
M2	2,0	20	10	100	$0 < V_{GS} < 1,2$	1,2	38,63	6,0
TSMC 0.35 μm								
M3	0,4	4,0	10	100	$0 < V_{GS} < 2,0$	2,0	38,0	50
M4	4,0	40	10	100	$0 < V_{GS} < 2,0$	2,0	38,0	5,0
AMIS 1.5 μm								
M5	16	160	10	100	$0 < V_{GS} < 2,0$	2,0	38,11	1,25

Na Tabela 5-1, pode-se verificar, além das dimensões dos transistores utilizados nas medidas, as condições de medida aplicadas. Os transistores referentes às tecnologias TSMC 0.18 μm e 0.35 μm são de canal mínimo e outro de canal dez vezes o canal mínimo para cada tecnologia, com o objetivo de verificar a diferença entre o comportamento para transistores de diferentes comprimentos de canal, considerando a mesma tecnologia. Isso é importante para se comprovar a influência do comprimento de canal quanto à posição do “Sweet Spot” em relação à polarização do transistor.

Os dados obtidos com as medidas foram processados em uma ferramenta de cálculo matricial com o objetivo de se calcular g_{mg} , g'_{mg} e g''_{mg} , usando funções de alisamento, para a redução do ruído associado às derivadas computacionais.

¹ Calculados através dos parâmetros extraídos e que serão apresentados na próxima subseção.

5.3 Extração de parâmetros tecnológicos relacionados com os efeitos de segunda ordem

Como visto, o modelo desenvolvido neste trabalho possui alguns poucos parâmetros para caracterizar o comportamento de não-linearidade do MOS. Essa é uma característica vantajosa do modelo ACM, que necessita de aproximadamente uma dezena de parâmetros tecnológicos para modelar os transistores de uma determinada tecnologia.

A literatura relacionada à extração de parâmetros tecnológicos para o modelo ACM [20] apresenta três métodos de extração. O mais simples e rápido deles é o mapeamento dos parâmetros desejados em uma tabela de parâmetros já extraídos, via experimental, para outro modelo, neste caso, o modelo BSIM. Aqui apenas serão mostrados os parâmetros de interesse ao modelo apresentado.

Tabela 5-2: Parâmetros tecnológicos.

Tecnologia	V_{T0} (V)	t_{ox} (ηm)	μ_0 (cm^2/Vs)	v_{lim} (cm/s)
TSMC 0.18 μm	0,36	4,0	264,06	$1,02 \times 10^7$
TSMC 0.35 μm	0,48	7,7	426,4	$1,62 \times 10^7$
AMIS 1.5 μm	0,58	30,5	593,09	$2,26 \times 10^7$

A Tabela 5-2 mostra todos os parâmetros necessários para a determinação de outros parâmetros, como por exemplo, ε , U_{CRIT} e outros utilizados no modelo de não-linearidades desenvolvido anteriormente.

As tabelas de parâmetros tecnológicos do modelo BSIM foram fornecidos por [19] e obtidos através de medidas experimentais pelas próprios fabricantes de circuitos integrados, mais comumente chamados de “foundries”.

5.4 Resultados obtidos e comparação com as equações teóricas.

Dos resultados obtidos através das medidas feitas com os transistores apresentados na Tabela 5-1, apenas os referentes aos transistores M1, M2, M3 e M4 serão apresentados aqui, pois os resultados referentes ao transistor M5 já foram apresentados no capítulo anterior.

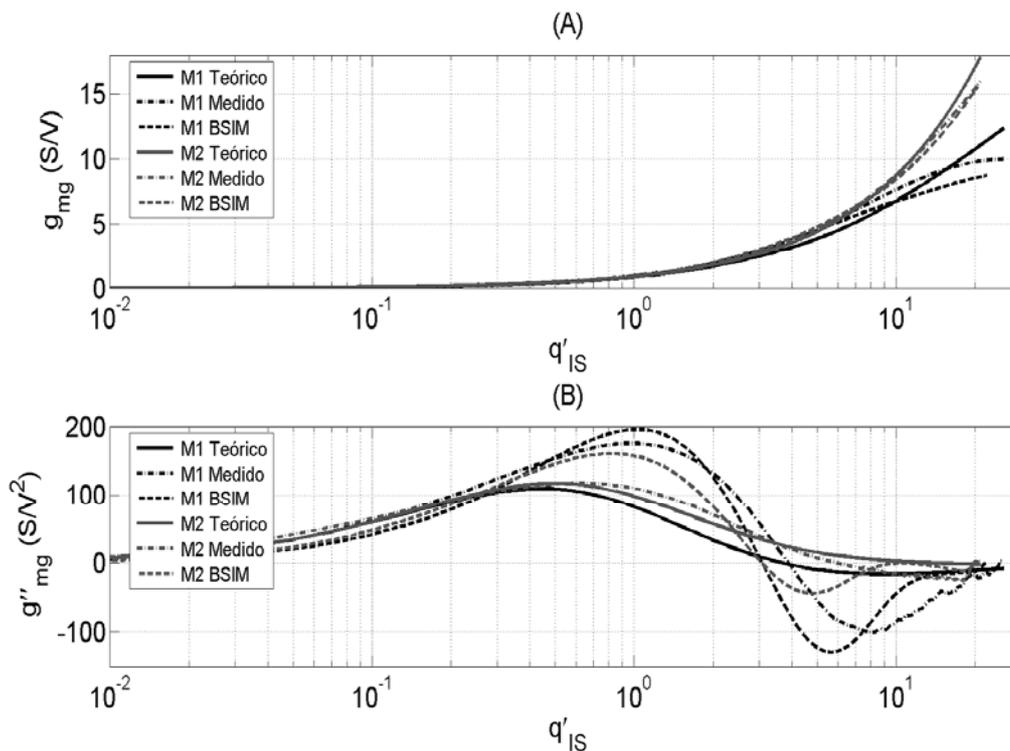


Figura 5-2: (A) g_{mg} e (B) g''_{mg} para os transistores da tecnologia TSMC 0.18 μm .

A Figura 5-2 apresenta em (A) a comparação entre g_{mg} obtido através de medidas, simulações utilizando o modelo BSIM 3 e através de (4.7) e, em (B), a comparação entre g''_{mg} obtido através de medidas, simulações utilizando o modelo BSIM 3 e através de (4.9) para os transistores M1 e M2, mostrados na Tabela 5-1. Observa-se em (B), para o transistor de canal curto (M1), que o modelo representado por (4.9) não modela g''_{mg} com precisão em toda a faixa dos níveis de inversão medidos apesar do comportamento do modelo para a terceira derivada da corrente de dreno se mostrar muito coerente em relação à medida experimental [18]. Mas, por outro lado, pode prever com bastante precisão o “Sweet Spot”, pelo fato da posição desse estar relacionada apenas com o parâmetro ε , para transistores de canal curto. Já para o transistor que pode ser considerado de canal longo (M2), pois o campo elétrico longitudinal é menor do que U_{CRIT} , o modelo apresenta uma maior precisão no cálculo de g''_{mg} para toda a faixa dos níveis de inversão medidos, apesar de não prever com precisão o “Sweet Spot” neste caso. Dessa discussão, pode-se concluir que o “Sweet Spot” depende exclusivamente de ε para transistores de canal curto, ou seja, o efeito de segunda ordem dominante, que pode ser interpretado e modelado como uma dependência da mobilidade segundo o campo elétrico longitudinal, é a

saturação da velocidade dos portadores. Diferente disso, para o transistor M2, que pode ser considerado de canal longo, o efeito dominante é a degradação da mobilidade devido ao campo elétrico transversal, que (4.9) não modela. Quanto ao modelo BSIM 3, é difícil fazer uma comparação com os resultados acima comentados, pois, como pode ser observado na Figura 5-2 (B), os resultados obtidos através da simulação utilizando o referido modelo apresenta pouca relação com os outros resultados. O ponto de polarização que propicia o “Sweet Spot” é muito diferente do ponto apresentado pelo modelo ACM e as medidas experimentais para o transistor de canal curto. Além disso, o comportamento de g_{mg}'' para o transistor de canal longo é muito parecido com o comportamento apresentado pelo transistor de canal mínimo o que é inconsistente com os resultados experimentais apresentados. Uma observação importante é que a curva de g_{mg}'' obtida através da simulação apresenta para ambos os transistores testados a passagem por zero no mesmo ponto de polarização, fato que também é inconsistente com os resultados experimentais.

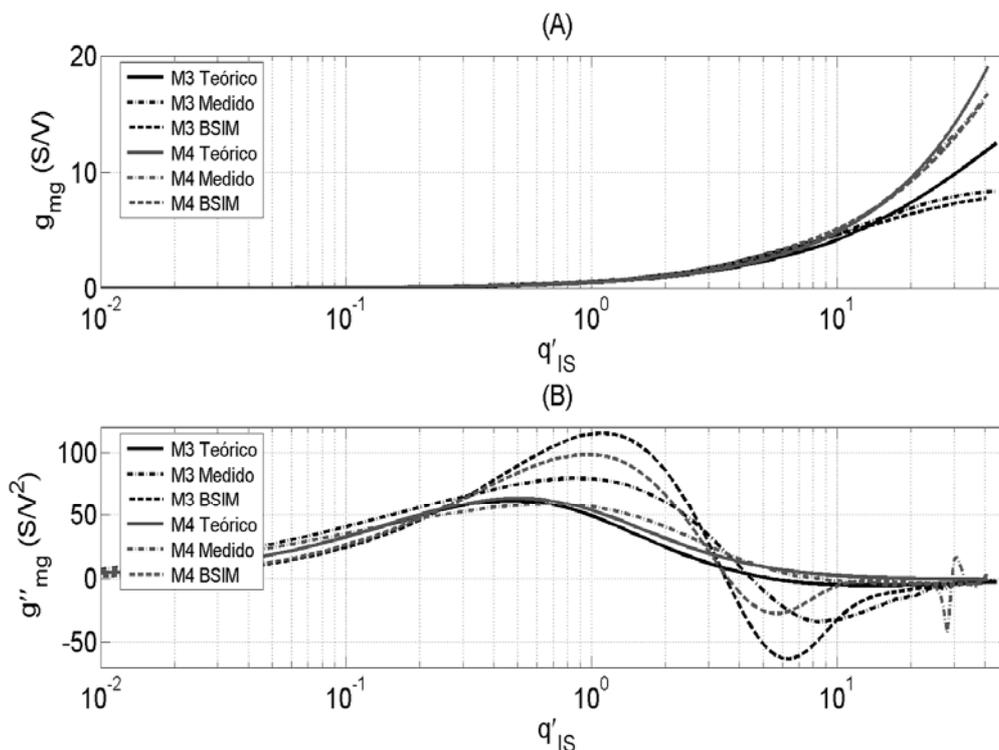


Figura 5-3: (A) g_{mg} e (B) g_{mg}'' para os transistores da tecnologia TSMC 0.35 μm .

A Figura 5-3 apresenta a comparação em (A) entre g_{mg} obtido através de medidas, simulações utilizando o modelo BSIM 3 e através de (4.7) e, em (B), a comparação entre

g''_{mg} obtido através de medidas, simulações utilizando o modelo BSIM 3 e através de (4.9) para os transistores M3 e M4 mostrados na Tabela 5-1. As observações feitas acima para os transistores da tecnologia TSMC 0.18 μm são também válidas para os transistores da tecnologia TSMC 0.35 μm . Contudo vale a pena uma comparação entre ambas com relação ao “Sweet Spot”. Devido à diferença entre os valores de ε dos dois transistores, o “Sweet Spot” pode estar localizado em níveis de inversão muito diferentes, concordando assim com (4.9) e, conseqüentemente, com (4.10). Além disso, analisando a Tabela 5-1, verifica-se que U_{CRIT} é praticamente igual para as três tecnologias testadas. Esse fato leva a crer que se houver dois transistores de mesmo comprimento de canal de tecnologias diferentes, mas com mesmo U_{CRIT} , o “Sweet Spot” se localizará no mesmo ponto de polarização para os dois transistores. Então, o comprimento de canal é a principal variável na determinação da posição do “Sweet Spot”. Quanto aos resultados obtidos com o modelo BSIM 3 para a tecnologia TSMC 0.35 μm pode-se tirar as mesmas conclusões apresentadas anteriormente. Novamente se observam as mesmas inconsistências na modelagem da terceira derivada da corrente de dreno para ambos os transistores testados.

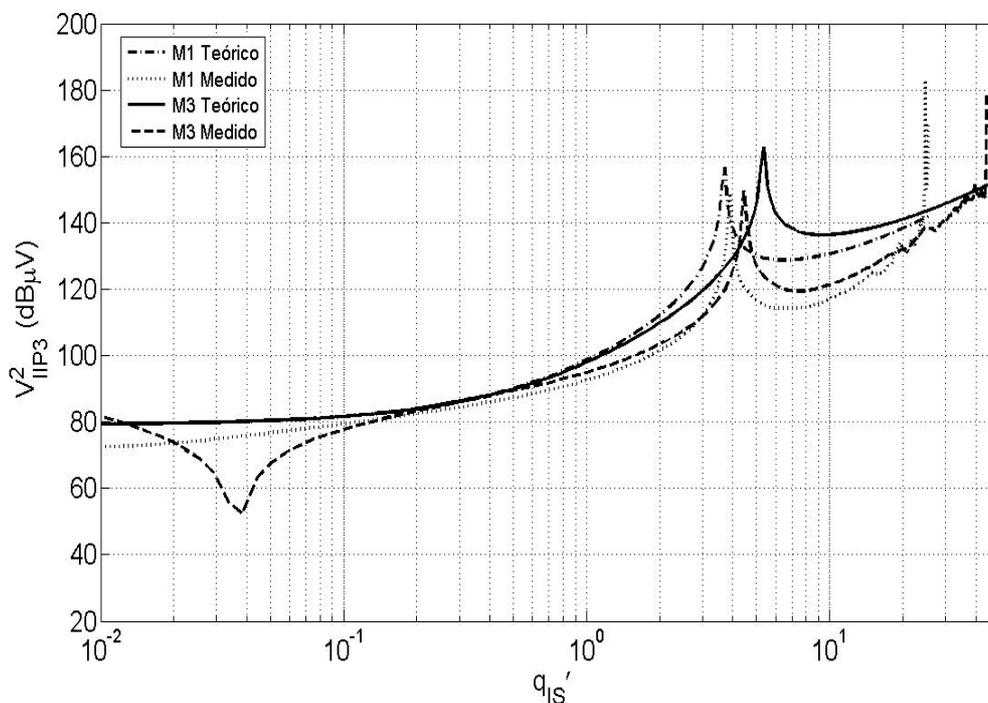


Figura 5-4: V_{IP3}^2 em função da polarização para os transistores M1 e M3.

A Tabela 5-1 mostra, através da comparação entre U_{CRIT} e a razão entre V_{DS} e o comprimento do canal, quais transistores estão com a velocidade dos portadores saturada e

os que não estão. Esse é um bom critério para definir quando um transistor é de canal curto ou não. Verificando a tensão máxima V_{DS} para a qual o campo elétrico crítico não seja ultrapassado, o transistor pode ser considerado de canal longo. Do contrário será um transistor de canal curto, pois esse está sujeito à saturação da velocidade dos portadores.

A Figura 5-4 apresenta uma comparação entre V_{HP3}^2 em função da polarização do transistor calculado através de (4.10) e medido para os transistores considerados de canal curto M1 e M3.

As comparações feitas entre as equações desenvolvidas e as medidas desenvolvidas mostram que o ponto “Sweet Spot” pode ser determinado com precisão se o parâmetro ε , relacionado com a velocidade de saturação dos portadores, também for determinado com precisão. Os valores teóricos apresentados na Tabela 5-1 utilizaram os parâmetros tecnológicos apresentados anteriormente e que foram extraídos por comparação ao modelo BSIM, cujos parâmetros podem ser encontrados em [19]. Assim a diferença entre os valores teóricos e práticos se deve ao fato de que para o cálculo de ε está se utilizando o comprimento de canal nominal e não o comprimento efetivo, que considera o encurtamento do canal devido às difusões laterais.

5.5 Conclusão

Neste capítulo, apresentou-se uma metodologia simples para avaliar o comportamento não-linear de transistores. Com essa metodologia, foram feitas medições em transistores de tecnologias e comprimentos de canal diferentes para validar as equações desenvolvidas na seção anterior e também verificar a precisão com que o “Sweet Spot” é avaliado utilizando as mesmas equações. Além disso, uma comparação com um modelo compacto BSIM 3 foi feita através de simulação.

Apresentou-se aqui também uma metodologia simples de extração de parâmetros utilizados no modelo ACM, baseada na comparação da tabela de parâmetros do modelo BSIM extraídos através de medidas experimentais [19].

Os resultados apresentados aqui deixam claro que o modelo que foi proposto consegue prever, com precisão, o ponto de polarização onde se encontra o “Sweet Spot” para transistores de canal curto. Contudo, esta precisão depende de quão preciso se determina o parâmetro ε de cada transistor, sendo que a determinação desse parâmetro está intimamente ligada ao conhecimento do comprimento de canal efetivo do transistor.

A comparação feita entre medidas experimentais, o modelo proposto por este trabalho e as simulações utilizando o modelo BSIM 3 mostraram a inconsistência do modelo BSIM 3 em modelar a terceira derivada da corrente de dreno com relação a tensão de porta.

6 CONCLUSÃO

Um modelo de não-linearidade do transistor MOS para transistores de canal curto incluindo as equações para g_{mg} , g'_{mg} , g''_{mg} e V_{IP3}^2 em função do nível de polarização do transistor, representado pela densidade de carga normalizada no lado de fonte, foi desenvolvido no trabalho aqui apresentado. As maiores vantagens do referido modelo são a simplicidade de suas equações, a forma explícita como é determinado o ponto de polarização em que o “Sweet Spot” está localizado e a flexibilidade para se calcular outras figuras de mérito relacionadas com as não-linearidades em transistores MOS além de V_{IP3}^2 .

Uma importante discussão, apresentada durante este trabalho, tem como foco as causas do aparecimento do “Sweet Spot”. Depois de uma análise comparativa feita no capítulo 4, utilizando inclusive dados experimentais, conclui-se que esse fenômeno é causado por efeitos de segunda ordem e que, para transistores de canal curto, o efeito de saturação da velocidade dos portadores é predominante. Já para transistores de canal longo, o fenômeno é causado provavelmente pela degradação da mobilidade devido ao campo elétrico transversal.

A constatação de que o parâmetro U_{CRIT} é praticamente igual para todas as tecnologias utilizadas para a validação do modelo aumenta a importância do comprimento de canal como variável principal para a determinação do parâmetro ε e, por sua vez, do nível de inversão onde o “Sweet Spot” se localiza, segundo (4.10). Isso também leva a crer que os “Sweet Spots” de dois transistores de mesmo comprimento de canal fabricados em tecnologias diferentes, mas com os valores de U_{CRIT} iguais, estarão localizados no mesmo ponto de densidade de cargas normalizada no lado de fonte, ou seja, na mesma corrente de dreno normalizada. Com o que foi exposto acima, conclui-se ainda que a precisão com que o modelo prediz o ponto de polarização em que o “Sweet Spot” se encontra está intimamente ligada à precisão com que se conhece o comprimento de canal efetivo do transistor que será projetado.

As medidas em corrente contínua realizadas, além de confirmarem algumas conclusões já apresentadas, mostram que é possível, com o modelo desenvolvido, prever o ponto de polarização ideal para o aparecimento do “Sweet Spot” de forma confiável, desde que o comprimento de canal efetivo do transistor seja bem conhecido. Além disso, o modelo se presta à predição do valor de V_{IP3}^2 para qualquer nível de inversão. As

comparações das medidas feitas com as simulações utilizando o modelo BSIM 3 mostraram algumas inconsistências do referido modelo com relação à modelagem da terceira derivada da corrente de dreno com relação à tensão de porta. A constatação destas inconsistências demonstra que nem todos os modelos conseguem modelar as não-linearidades de forma confiável.

Para trabalhos futuros, propõe-se a adição do efeito de degradação da mobilidade devido ao campo elétrico transversal, a fim de tentar predizer o “Sweet Spot”, de forma confiável, para uma maior faixa de comprimentos de canal. Novas medidas em corrente contínua com transistores com diferentes comprimentos de canal curto e diferentes tecnologias poderiam ser realizadas com o intuito de confirmar as conclusões e resultados aqui expostos. Medidas em corrente alternada, utilizando para isso sinais de teste com diferentes frequências, poderiam ser realizadas com o objetivo de verificar o comportamento do “Sweet Spot” com relação à frequência. E, finalmente, seria muito interessante utilização do modelo aqui apresentado em uma metodologia de projeto que culminaria na realização de um amplificador de baixo ruído e alta linearidade, operando com baixo consumo de potência.

ANEXO 1 – DESENVOLVIMENTO DETALHADO DO MODELO DE NÃO-LINEARIDADE.

Este anexo se propõe a mostrar detalhadamente o desenvolvimento das equações de transcondutância e suas derivadas apresentadas no capítulo 4.

Segundo o modelo ACM, a corrente de dreno com a velocidade dos portadores saturada e em regime de saturação direta é dada por:

$$I_D = I_S \frac{2}{\varepsilon} q'_{IDSAT} \quad (\text{A.1})$$

onde I_S é a corrente específica, ε o parâmetro referente à saturação da velocidade dos portadores, e q'_{IDSAT} é a densidade de cargas normalizada do lado de dreno.

Para se obter a primeira derivada da corrente de dreno em relação à tensão de porta, deriva-se implicitamente (A.1) com relação à tensão de porta, resultando em:

$$g_{mg} = \frac{dI_D}{dV_{GB}} = I_S \frac{2}{\varepsilon} \frac{dq'_{IDSAT}}{dV_{GB}} \quad (\text{A.2})$$

A derivada da densidade de cargas normalizada em relação à tensão de porta pode ser obtida através de (4.5), que por sua vez já foi devidamente aproximada para manter a simplicidade do modelo nas derivadas de mais altas ordens:

$$\frac{dq'_{IDSAT}}{dV_{GB}} = \frac{dq'_{IS}}{dV_{GB}} \varepsilon q'_{IS} \left\{ \frac{2(1+q'_{IS})[2+\varepsilon(1+q'_{IS})] - 2\varepsilon q'_{IS}(1+q'_{IS}/2)}{[2+\varepsilon(1+q'_{IS})]^2} \right\} \quad (\text{A.3})$$

A derivada da densidade de cargas do lado de fonte (q'_{IS}) em relação à tensão de porta é obtida com a derivação do UCCM (3.8) em sua versão normalizada. Lembrando que $\frac{dV_P}{dV_{GB}} = \frac{1}{n}$, obtém-se:

$$\frac{dq'_{IS}}{dV_{GB}} = \frac{1}{n\phi_t} \frac{q'_{IS}}{(1+q'_{IS})} \quad (\text{A.4})$$

Substituindo (A.4) em (A.3) e o resultado disso em (A.2), tem-se a transcondutância de porta em função da densidade de carga normalizada no lado de fonte:

$$g_{mg} = \frac{dI_D}{dV_{GB}} = \frac{2I_S}{n\phi_t} \frac{q'_{IS}}{2 + \varepsilon(1 + q'_{IS})} \left\{ \frac{2(1 + q'_{IS})[2 + \varepsilon(1 + q'_{IS})] - 2\varepsilon q'_{IS}(1 + q'_{IS}/2)}{(1 + q'_{IS})[2 + \varepsilon(1 + q'_{IS})]} \right\} \quad (\text{A.5})$$

É possível ainda fazer uma aproximação em (A.5) que resultará em uma expressão muito mais simples para a transcondutância sem que se perca a precisão do modelo. Pode-se verificar em (A.5) que em inversão fraca ($q'_{IS} \ll 1$) e moderada ($1 < q'_{IS} < 10$) a parcela negativa do numerador pode ser desconsiderada pelo fato de $\varepsilon \ll 1$. Já em inversão forte ($q'_{IS} \gg 10$) pode-se dizer que $(1 + q'_{IS}) \approx (2 + q'_{IS})$ resultando em:

$$g_{mg} = \frac{dI_D}{dV_{GB}} = \frac{2I_S}{n\phi_t} \frac{q'_{IS}}{2 + \varepsilon q'_{IS}} \left(\frac{4 + \varepsilon q'_{IS}}{2 + \varepsilon q'_{IS}} \right) \quad (\text{A.6})$$

Para se obter a segunda derivada da corrente em relação à tensão de porta basta derivar implicitamente (A.6) em relação à tensão de porta e substituir (A.4) ao resultado da derivação. Isso resulta em:

$$g'_{mg} = \frac{\partial^2 I_D}{\partial V_{GB}^2} = \frac{16I_S}{(n\phi_t)^2} \frac{q'_{IS}}{(q'_{IS} + 1)} \frac{1}{(2 + \varepsilon q'_{IS})^3} \quad (\text{A.7})$$

A terceira derivada da corrente de dreno em relação à tensão de porta é obtida de forma parecida com a segunda derivada. Deriva-se implicitamente (A.7) em relação à tensão de porta e substitui-se (A.4) ao resultado desta derivação, resultando em:

$$g''_{mg} = \frac{\partial^3 I_D}{\partial V_{GB}^3} = \frac{16I_S}{(n\phi_t)^3} \frac{q'_{IS}}{(q'_{IS} + 1)^3} \frac{2 - 2\varepsilon q'_{IS} - 3\varepsilon q'^2_{IS}}{(2 + \varepsilon q'_{IS})^4} \quad (\text{A.8})$$

As equações aqui desenvolvidas servem para o cálculo de várias figuras de mérito relacionadas à não-linearidade de circuitos eletrônicos revisadas no capítulo 2 do presente trabalho.

Referências Bibliográficas

- [1] Razavi, Behzad. “*RF Microelectronics*”. 1ª Edição. Prentice Hall Communications Engineering and Emerging Technologies Series, 1998.
- [2] Rogers, J. e Plett, C. “*Radio Frequency Integrated Circuit Design*”. 1ª Edição. Artech House, 2003.
- [3] Lee, Thomas H. “*The Design of CMOS Radio-Frequency Integrated Circuits*”. 1ª Edição. Cambridge University Press. New York, 1998.
- [4] Abidi, A. A.; Pottie, G. J. e Kaiser, W. J. “Power-Conscious Design of Wireless Circuits and Systems”. *Proceedings of IEEE*. v. 88, Outubro de 2000. pp 1528 – 1545.
- [5] Shaeffer, D. K. e Lee, T. H. “A 1.5-V, 1.5-GHz CMOS Low Noise Amplifier”. *IEEE Journal of Solid-State Circuits*. v. 32, Maio de 1997. pp 745-759.
- [6] Cunha, A. I. A. “*Um Modelo do Transistor MOS para Projeto de Circuitos Integrados*”. Brasil. Dezembro de 1996. Tese Doutorado em Engenharia Elétrica. Universidade Federal de Santa Catarina.
- [7] Galup-Montoro, C.; Schneider, M. C. e Cunha, A. I. A. “A Current-Based MOSFET Model for Integrated Circuit Design”. In: Sanchez-Sinencio, E. e Andreu, A. G. “*Low-Voltage/Low-Power Integrated Circuits and System*”. 1ª Edição. IEEE Press, August, 1998.
- [8] Galup-Montoro, C.; Schneider, M. C. e Pahim, V. C. “Fundamentals of Next Generation Compact MOSFET Models”. In: SBCCI 2005. *Proceedings of 18th Symposium on Integrated Circuits and Systems Design*. Florianópolis, Brazil, September 2005. pp 32 – 37.
- [9] Bork, B. C. “*Efeitos da Saturação de Velocidade em Aplicações de Alta Frequência do MOSFET*”. Brasil. Agosto de 2003. Dissertação Mestrado em Engenharia Elétrica. Universidade Federal de Santa Catarina.
- [10] Tsividis, Y. “*Operation and modeling of the MOS transistor.*” McGraw-Hill. New York, 1987.
- [11] Langevelde, Ronald van. “*A Compact MOSFET Model for Distortion Analysis in Analog Circuit Design*”. Tese de Doutorado. Technische Universiteit Eindhoven. Eindhoven, 1998.

- [12] Arnaud Maceira, A. “*Very Large Constant GM-C Filters*”. Uruguay, Abril de 2005. Tese de Doutorado. Universidad de la Republica.
- [13] Toole, B.; Plett, C. e Cloutier, M. “RF Circuit Implications of Moderate Inversion Enhanced Linear Region in MOSFETs”. *IEEE Transactions on Circuits and Systems – I*. v. 51, Fevereiro de 2004, pp 319 – 328.
- [14] Woerlee, P. H.; Knitel, M. J.; van Langevelde, R.; Klaassen, D. B. M.; Tiemeijer, L. F.; Scholten, A. J. e Zegers-van Duijnhoven, A. T. A. “RF-CMOS Performance Trends”. *IEEE Transactions on Electron Devices*. V. 48, August, 2001, pp 1776 – 1782.
- [15] Ramos, J.; Mercha, A.; Jeamsaksiri, W.; Jenei, D.; Rooyackers, R.; Verbeeck, R. e Decoutere, S. “90nm RF CMOS Technology for low-power 900 MHz applications”. In: ESSDERC 2004. *Proceedings of the 34th European*. Setembro de 2004. pp 329 – 332.
- [16] Niu, G.; Pan, J.; Wei, X.; Taylor, S. S.; Sheridan, D. “Intermodulation Linearity Characteristics of CMOS Transistors in a 0.13 μm Process”. In: IEEE Digest of Papers 2005. *Radio Frequency Integrated Circuits (RFIC) Symposium*. June 2005. pp. 65-68.
- [17] Aparin, V.; Brown, G.; Larson, L. E. “Linearization of CMOS LNA’s via Optimum Gate Biasing”. In: ISCAS 2004. *Proceedings of the 2004 International Symposium on Circuits and Systems*. V. 4, May 2004. pp. 748-751.
- [18] Da Silva, P.D.; Sousa, F. R.; Galup-Montoro, C. and Schneider, M. C. “Compact Modeling of Nonlinearities in Submicron MOSFETs” In Nanotech 2006. *Proceedings of Nanotech 2006*. Boston, USA, Maio de 2006. pp 753-756.
- [19] Programa MOSIS; In: <http://www.mosis.org>.
- [20] Coitinho, R. M.; Spiller, L. H.; Schneider, M. C.; Galup-Montoro, C. “A Simplified Methodology for the Extraction of the ACM MOSFET Model Parameters”. In: SBCCI 2001. *Proceedings of 14th Symposium on Integrated Circuits and Systems Design*. Pirenópolis, Brazil, September 2001. pp 136 – 141.