

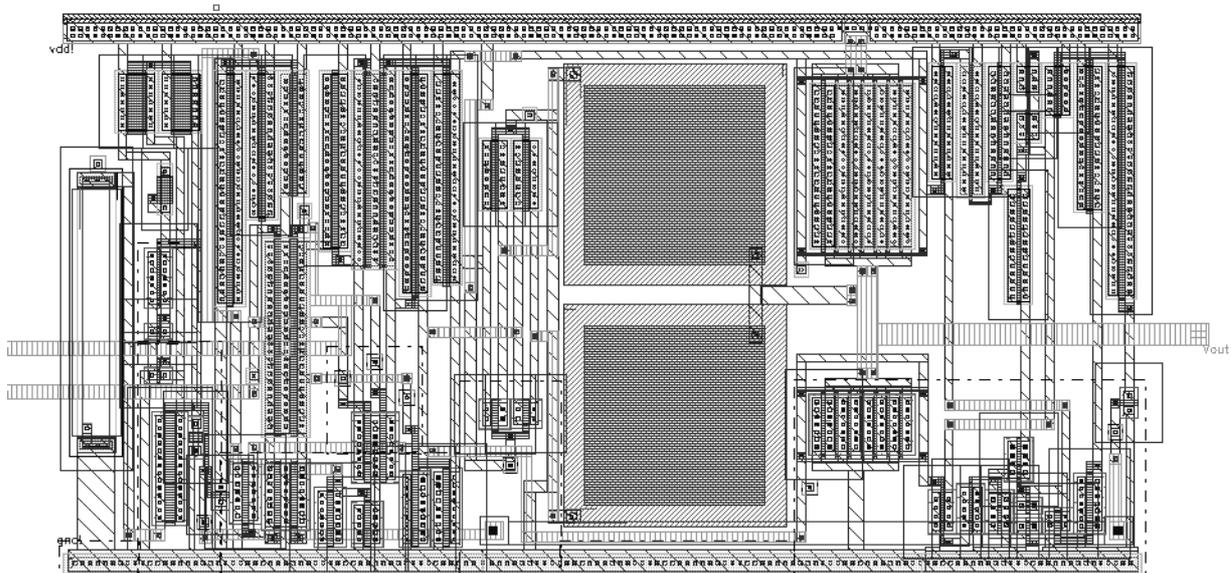
UNIVERSITÀ DEGLI STUDI DI PISA
Corso di Laurea Magistrale in Ingegneria Elettronica

P. Bruschi

DISPENSE DEL CORSO
PSM

(Progettazione di Sistemi Microelettronici)

MODULO MICROELETTRONICA ANALOGICA



2013

Indice

Premessa	3
Capitolo 1. Introduzione alla progettazione full-custom.....	4
1.1 Concetti generali.....	4
1.2 Flusso di progetto	5
1.3 Caratteristiche generali di uno Schematic Editor	7
1.4 Caratteristiche generali di un layout editor	10
1.5 Elementi fondamentali di un processo CMOS n-well standard a due livelli di metal.	13
1.6 Componenti passivi per circuiti integrati.	15
Capitolo 2. Considerazioni generali relative ai dispositivi per circuiti integrati	22
2.1 Transistori bipolari: dimensionamento.....	22
2.2 Nozioni utili per il dimensionamento di transistori MOS in circuiti analogici.....	25
2.3 Errori di fabbricazione e uguaglianza tra dispositivi nominalmente identici.....	28
Capitolo 3. Specchi di corrente	36
3.1 Definizioni generali.....	36
3.2 Specchi di corrente a MOSFET:.....	39
3.3 Specchi di corrente a transistori bipolari.....	44
3.4 Riepilogo delle caratteristiche degli specchi di corrente.....	56
3.5 Effetto degli errori di matching sugli specchi di corrente.....	57
Capitolo 4. Amplificatori differenziali: concetti e strutture di base.	60
4.1 Definizioni generali riguardanti gli amplificatori differenziali.....	60
4.2 Coppia di transistori MOSFET accoppiati di source (source coupled pair).....	63
4.3 Coppia di transistori bipolari connessi di emettitore.....	68
4.4 Amplificatori differenziali a carichi resistivi.	70
4.5 Considerazioni generali sugli amplificatori a carichi resistivi.	77
Capitolo 5. Amplificatori differenziali: circuiti.....	78
5.1 Amplificatore differenziale semplice a MOSFET con uscita single-ended	78
5.2 Amplificatori differenziali cascode a MOSFET.....	86
5.3 Amplificatori a MOSFET a larga dinamica (OTA e OTA rail-to-rail)	92

5.4	Amplificatore differenziale a BJT, carichi attivi e uscita single ended.....	96
Capitolo 6.	Circuiti analogici di largo impiego	100
6.1	Riferimenti di tensione.	100
6.2	Traslatori di tensione.	105
6.3	Stadi di uscita.	107
6.4	Cenni agli amplificatori operazionali CMOS.....	115
Appendice A.	Breve guida al simulatore elettrico SPICE	129
Appendice B.	Cenni ai modelli usati da SPICE per gli elementi a semiconduttore	145
Appendice C.	Casi salienti di calcolo della resistenza vista	155

Premessa

Questa dispensa costituisce il principale riferimento bibliografico per il primo modulo del corso di “Progettazione di Sistemi Microelettronici”, del Corso di Laurea Specialistica in Ingegneria Elettronica. Questo modulo didattico, equivalente a 6 CFU, comprende le basi della progettazione *full-custom* di circuiti integrati e la descrizione delle configurazioni elementari maggiormente utilizzate per la sintesi di circuiti integrati analogici. Gli argomenti relativi alla progettazione di circuiti integrati digitali, non inclusi in questa dispensa, vengono trattati nel secondo modulo del suddetto corso. Il corso “Progettazione di Sistemi Microelettronici” sostituisce per l’anno accademico 2002-03 il corso “Microelettronica II” della laurea in Ingegneria Elettronica, vecchio ordinamento.

Nella prima parte di questo trattato, corrispondente ai capitoli 1 e 2, vengono presentati i concetti fondamentali della progettazione di circuiti integrati in modalità “*full-custom*”, descrivendone le varie fasi e, in parallelo, i relativi ausili CAD, ormai assolutamente indispensabili per lo sviluppo di un circuito integrato. Viene inoltre affrontato il problema del dimensionamento dei dispositivi attivi principali, i transistori MOS e bipolari, evidenziando i parametri sui quali ha effettivamente controllo il progettista e gli effetti che tali parametri hanno sulle caratteristiche elettriche dei dispositivi stessi. Attenzione particolare viene dedicata al matching dei dispositivi, che costituisce uno degli aspetti più importanti per la progettazione di circuiti analogici integrati, proponendo una tecnica di analisi avente massima generalità di applicazione.

Nella seconda parte del trattato vengono introdotte quelle configurazioni circuitali che costituiscono la base per la comprensione e la sintesi di circuiti analogici di qualsiasi complessità. Saranno quindi presi in esame gli specchi di corrente (capitolo 3), gli amplificatori differenziali (capitoli 4 e 5) e altri blocchi funzionali quali i riferimenti di tensione e un semplice amplificatore operativo (capitolo 6). In questa sezione della dispensa, che riflette l’approccio seguito nel corso, si è cercato di evidenziare i vantaggi e gli svantaggi delle varie soluzioni esaminate, allo scopo di fornire al lettore la capacità di scegliere la topologia che di volta in volta consente di soddisfare meglio le specifiche. Nello stesso tempo è stata dedicata particolare cura alla comprensione del ruolo dei vari elementi circuitali, allo scopo di indicare al lettore come modificare il comportamento dei circuiti proposti mediante il dimensionamento dei parametri più importanti e, nello stesso tempo, favorire la capacità di sintetizzare nuove topologie partendo dal livello più basso, ovvero dal singolo transistor.

Completano la presente dispensa tre appendici, le prime due dedicate al programma di simulazione più diffuso, SPICE, e l’ultima alla raccolta dei risultati relativi al calcolo delle resistenze viste in vari casi salienti richiamati frequentemente nel testo.

Per un approfondimento degli argomenti trattati in questa dispensa si consiglia la consultazione dei testi seguenti:

1. P. R. Gray, R. G. Meyer, “Circuiti integrati analogici”, McGraw-Hill Libri Italia srl, Milano 1994.
2. R. J. Baker, H. W. Li, D. E. Boyce, “CMOS - Circuit design, layout and simulation”, IEEE Press, New York 1998.
3. R. Gregorian, G. C. Themes, “Analog MOS integrated circuits for signal processing”, John Wiley & Sons, Singapore 1986.

Capitolo 1. Introduzione alla progettazione full-custom

1.1 Concetti generali

La progettazione full-custom di un circuito integrato (o di una sua parte) richiede che il progettista effettui la definizione della topologia del circuito elettrico e il suo dimensionamento fino ad arrivare al livello dei singoli componenti elementari (transistori, resistori etc.). La modalità di progettazione full-custom prevede inoltre che venga disegnato il **layout** del circuito integrato. Anche in questa fase è compito del progettista effettuare il posizionamento e l'interconnessione dei singoli componenti elementari. In alcuni casi, specie lavorando con processi CMOS, può essere richiesto di disegnare anche la struttura dei singoli componenti. In ogni caso nella progettazione full-custom il progettista ha la possibilità di specificare tutte le maschere che definiscono i vari passi di processo necessari alla fabbricazione di un circuito integrato. Questo costituisce una differenza sostanziale con la progettazione semi-custom nella quale alcune strutture sono predefinite (es. diffusioni p^+ e n^+ , polisilicio) e il progettista interviene solo sullo schema delle metallizzazioni che effettuano le interconnessioni (routing).

La progettazione di circuiti integrati viene svolta utilizzando ambienti CAD che vengono spesso indicati con la sigla EDA (electronic design automation). Come accennato le due fasi principali del progetto di un circuito integrato full custom sono le seguenti:

- 1) Progetto del circuito elettrico (livello elettrico di descrizione del circuito).
- 2) Progetto del layout (livello fisico di descrizione del circuito).

La prima fase viene svolta utilizzando un programma denominato **schematic editor**. Questo programma consente di posizionare i componenti sul foglio di lavoro che rappresenta il circuito e di specificarne le interconnessioni. A questo livello nulla viene specificato riguardo al posizionamento reale dei componenti nel chip.

Il progetto del layout, svolto mediante il **layout editor**, consiste nel definire le maschere da utilizzare nella fabbricazione del circuito integrato. Ogni maschera è associata a un passo di processo e specifica le aree del chip nelle quali va applicato quel determinato passo di processo. Per esempio, la maschera della "metal1", specifica dove deve essere presente il primo livello di metallizzazione. Il disegno del layout è spesso un lavoro di routine affidato a un tecnico diplomato. Tuttavia è necessario che il progettista conosca tutti i dettagli che caratterizzano il passaggio dal circuito elettrico al layout poichè dovrà essere in grado di dare indicazioni sui punti critici del circuito in modo da evitare che interconnessioni troppo lunghe o la prossimità di dispositivi introducano elementi parassiti non tollerabili. Altri esempi di fenomeni indesiderati riguardano l'accensione di mosfet parassiti sotto interconnessioni sottoposte a tensioni elevate, accoppiamenti termici dovuti al posizionamento errato di dispositivi sensibili in prossimità di elementi di potenza o il dimensionamento errato della larghezza delle metallizzazioni relativamente alla corrente massima che sono destinate a condurre.

1.2 Flusso di progetto

La figura 1.1 mostra il tipico flusso di progetto di un circuito full-custom in un generico ambiente EDA. L'ambiente di progetto (ovvero il pacchetto software) deve essere preparato mediante l'installazione delle librerie costituenti il cosiddetto **design-kit**. Questo è un insieme di files (technology files) che adattano l'ambiente di lavoro al dato processo con cui si vuole far fabbricare il circuito. Il design kit deve quindi essere fornito dalla silicon foundry ("fonderia", ovvero l'industria presso la quale verrà fabbricato il circuito).

AMBIENTE DI PROGETTAZIONE FULL-CUSTOM EDA (Electronic Design Automation)

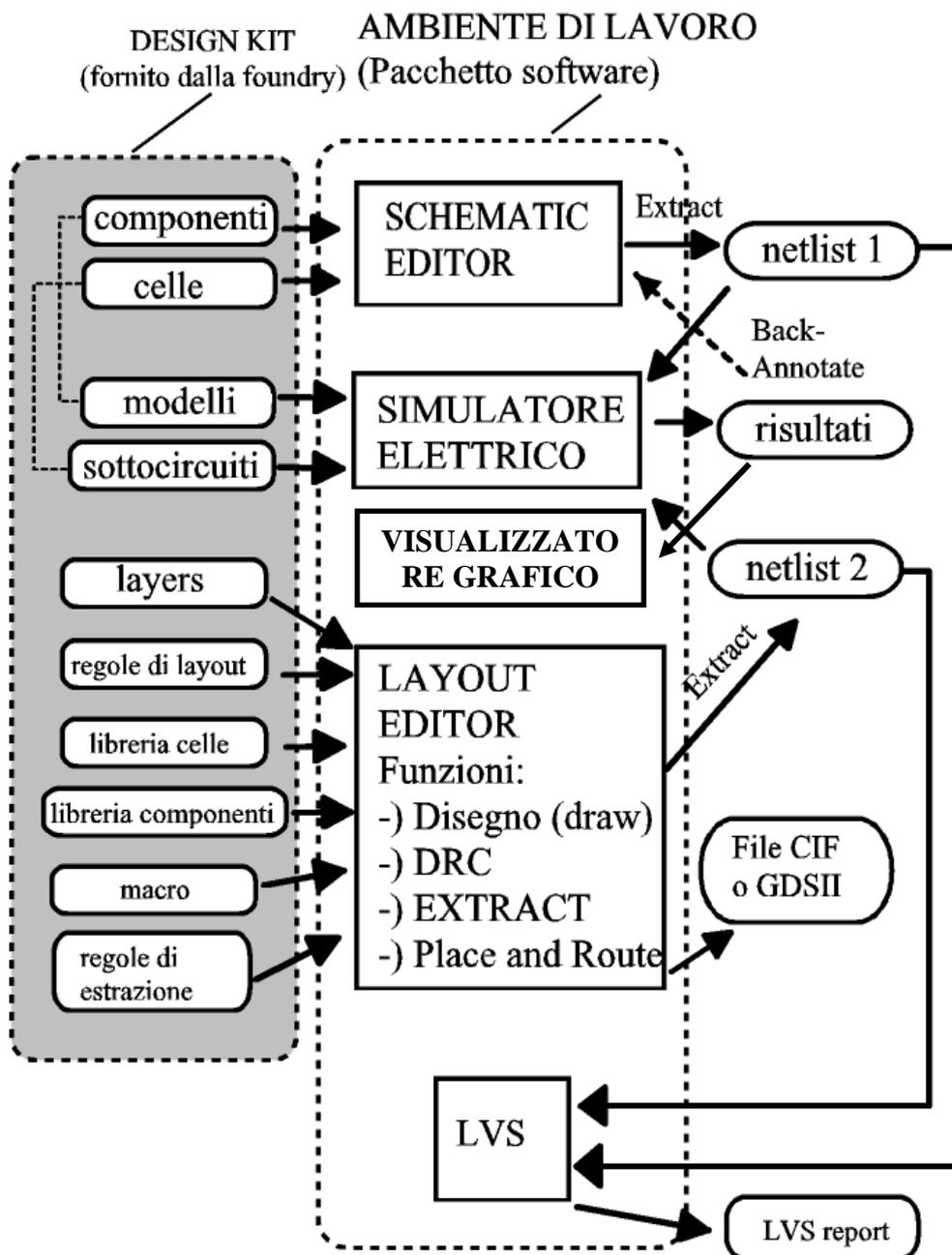


Figura 1.2.1

Come accennato si inizia disegnando il circuito elettrico (schematic capture). Il design editor è personalizzato con i technology files contenenti i componenti disponibili con quella tecnologia. Per esempio, se si lavora con un processo bipolare si avranno a disposizione BJT pnp e npn, resistori e condensatori ma non si potranno utilizzare transistori MOSFET. Lo schematic editor consente in genere di utilizzare anche componenti denominati “standard” che non sono propri del processo. Quindi, facendo riferimento all’esempio precedente di processo bipolare, pur non trovando il transistor MOS nella libreria di processo (che appartiene al design-kit), posso sempre introdurre un MOSFET della libreria standard. Devo però tener presente che tale componente non potrà essere incluso nel circuito integrato. A che servono quindi i componenti standard? Essi possono essere usati per simulare componenti esterni al chip, per esempio carichi (resistori, condensatori di carico) o circuiti esterni che forniscono segnale al nostro circuito integrato o ne prelevano il segnale di uscita. Il circuito integrato può quindi essere simulato ricostruendo il più possibile le condizioni di funzionamento reali.

Il design kit comprende inoltre in genere una serie di celle (porte logiche, flip-flops, amplificatori operazionali etc.) che possono essere utilizzate nel modulo che si sta progettando. Ovviamente queste celle utilizzano unicamente componenti del processo con cui si sta lavorando.

La fase di progettazione elettrica viene assistita da un simulatore elettrico con il quale si verifica se il circuito (o una sua versione parziale) si comporta effettivamente come richiesto. Il simulatore elettrico viene inoltre usato attivamente, specie per circuiti analogici, per dimensionare i parametri dei componenti (es. resistenze, capacità, lunghezze e larghezze di MOSFET). Nella stragrande maggioranza dei casi, infatti, per via analitica si riesce solo a fissare dei valori di massima per i parametri. In pratica si eseguono serie di simulazioni effettuate variando il parametro in esame e controllando le prestazioni del circuito (o, nel caso più semplice il punto di riposo) fino ad ottenere il valore opportuno. Alcuni simulatori elettrici consentono di effettuare questa operazione automaticamente.

Per dare risultati consistenti con la realtà, il simulatore deve contenere una descrizione accurata del comportamento fisico dei componenti. Ciò si ottiene mediante files opportuni detti “**modelli**” o “**models**”. Per accedere al simulatore occorre far sì che dal disegno del circuito (schematic view) si passi ad una rappresentazione del circuito stesso in un formato compatibile con il simulatore. Questa rappresentazione consiste in un file, quasi sempre di testo, detto **netlist**. Se nel circuito si fa uso di celle di libreria, di solito la netlist contiene soltanto un riferimento alla data cella la quale deve essere presente anche in una libreria di sottocircuiti nel formato compatibile con il simulatore.

I risultati della simulazione possono essere letti mediante un post-processore grafico, o se sono semplicemente informazioni sul punto di riposo possono essere riportati direttamente sul circuito con un’operazione di **back-annotation**.

Quando il circuito elettrico soddisfa ai requisiti che ci siamo posti, possiamo passare alla fase di disegno del **layout**. Il layout editor è personalizzato con alcuni technology files. **I layers** consentono, al progettista di indicare, con le modalità che vedremo più avanti, tutte e sole le lavorazioni che sono concesse in quel processo e sulle quali può intervenire. Per esempio, in un processo CMOS standard, il progettista non potrà indicare aree nelle quali egli intende far posizionare uno strato sepolto (buried layer) in quanto questo è disponibile solo in altri tipi di processo (es. bipolare o BiCMOS). **Le regole di layout** indicano al programma DRC (design rule checker) quelle che sono le regole che devono essere rispettate affinché il layout sia realizzabile (es. distanza minima tra due piste dello stesso livello di metallizzazione). Il DRC evidenzia nel layout tutti i punti che violano le regole del processo. Le librerie dei

componenti e delle celle contengono i layout di tutti i corrispondenti componenti o celle presenti nelle librerie dello schematic editor. **Le macro** sono programmi che consentono di disegnare più agevolmente componenti o oggetti vari nel layout. Per esempio è possibile che il design kit di un processo bipolare contenga una macro che disegni matrici di transistori in numero specificabile in un'opportuna finestra e poi colleghi gli stessi transistori in parallelo.

Quando il layout è pronto e non presenta più errori durante il DRC, occorre verificare che il layout stesso corrisponda al circuito elettrico che abbiamo progettato. Si effettua quindi un confronto tra il layout e lo schema elettrico. Per fare ciò occorre estrarre dal layout il suo schema elettrico equivalente. L'operazione che si compie si dice appunto estrazione di parametri (**extract**). Dovrà essere presente un opportuno file tecnologico che "insegna" al programma che compie l'estrazione come associare a certe figure geometriche componenti elettronici (es. resistori, transistori...) e interconnessioni. Il circuito così estratto (di solito può essere una netlist o un formato intermedio) viene confrontato con il circuito estratto dallo schema elettrico. Questo tipo di confronto viene denominato **LVS (layout versus schematic)**. È possibile confrontare i due circuiti solo a livello topologico oppure confrontare anche i valori dei parametri dei dispositivi (resistenze, dimensioni dei transistori etc.). L'operazione LVS produce un report che indica quali differenze esistono tra layout e schema elettrico ovvero, se il test viene passato con successo, indica l'assenza di differenze.

È inoltre possibile estrarre una netlist che contiene alcuni tipi di **componenti parassiti** che sono dipendenti dal tipo di layout e che quindi non si potevano prevedere a livello di schema elettrico: per esempio, la resistenza di una pista di collegamento (supposta nulla nello schema elettrico) dipende dalla lunghezza della pista stessa, la quale può essere stabilita solo dopo aver disegnato il layout. Altri tipi di parametri parassiti che si possono estrarre sono le capacità tra due linee di interconnessione o tra le linee stesse e il substrato. I parametri parassiti interni ai dispositivi (per esempio la capacità tra base e emettitore di un BJT) sono invece già inclusi nei modelli del simulatore e non devono quindi essere estratti. La rete estratta in questo modo può essere simulata per avere una stima più precisa delle prestazioni del circuito (**simulazione post-layout**). Questa è particolarmente utile per verificare gli effettivi tempi di ritardo delle celle digitali in quanto in questo caso le capacità e le resistenze parassite possono avere un'influenza determinante. Si osservi che ai fini dell'operazione di LVS non si devono assolutamente estrarre i componenti parassiti in quanto la verifica va fatta solo sui componenti di libreria e sulla topologia del circuito.

Se il circuito che abbiamo progettato è una cella che dovrà essere utilizzata in un modulo più complesso il lavoro termina qui. Se invece il circuito è un intero chip, possiamo procedere ad alcune operazioni finali richieste dal particolare processo (operazioni di die finishing) e poi produrre un formato finale del layout (GDSII o CIF) che può essere accettato dalla fonderia per la produzione delle maschere richieste per la realizzazione del chip.

1.3 Caratteristiche generali di uno Schematic Editor

Ogni circuito che viene creato o importato costituisce un modulo o cella. Ad ogni cella possono essere associate più viste, ovvero diverse rappresentazioni, ciascuna delle quali serve ad uno scopo particolare. Per quanto riguarda lo schematic editor, le viste di uso più comune sono la vista "schematic" e la vista "symbol". La vista schematic descrive lo schema elettrico vero e proprio della cella, ovvero come la cella è costituita da celle più semplici connesse tra loro da fili elettrici (wires). La vista symbol consente l'inclusione della cella in circuiti più complessi. Nella vista symbol compaiono solo i terminali (pin) di connessione (ingresso,

uscita o generici) e una rappresentazione grafica mnemonica della cella che ne ricordi la funzione (es. un triangolo per un amplificatore). Ovviamente nella vista schematic di ciascuna cella vengono inserite celle più semplici sotto forma di viste symbol.

Non tutte le celle possiedono una vista schematic: i componenti elementari, quali i resistori, i transistori, i condensatori ed altri non vengono generalmente descritti come connessione di elementi ancora più semplici. Per queste celle, dette **primitive**, esiste solo la vista symbol.

Esempio di struttura gerarchica dello schema elettrico di una porta AND a 3 ingressi

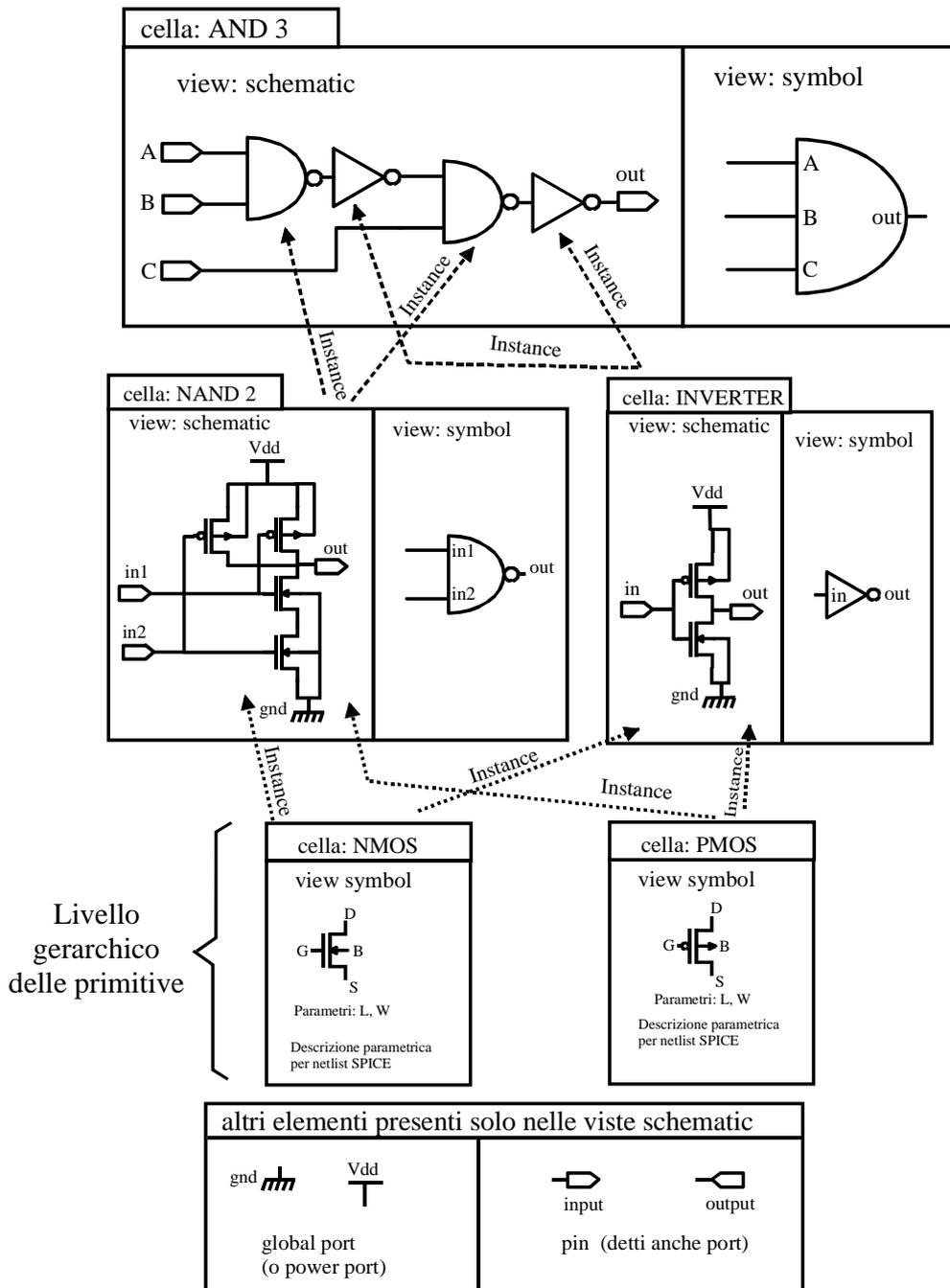


Figura 1.3.1

Gerarchia (Hierarchy). È possibile definire un criterio di classificazione gerarchica dei circuiti costituenti un progetto. Se il modulo A utilizza al suo interno un modulo B allora il modulo A si trova ad un livello gerarchico più elevato rispetto a B. Non è possibile creare situazioni cicliche ovvero far sì che A utilizzi B e contemporaneamente che B utilizzi A.

1 Se stiamo creando un circuito che utilizza un dato simbolo, è spesso utile esaminare il sottocircuito, ovvero la vista schematic, che “si trova sotto” quel simbolo. Tale operazione si indica come “discesa gerarchica” (*descend hierarchy*). Ovviamente non è possibile operare una discesa gerarchica su una primitiva.

Istanza (Instance). Ogni volta che inseriamo in un modulo un simbolo viene effettuata una cosiddetta “istanza” del dispositivo o sottocircuito che corrisponde al simbolo stesso. Naturalmente in un dato modulo si possono creare più istanze dello stesso oggetto. Ciascuna istanza costituisce un oggetto a sé stante, proprio come risultano essere, per esempio, tutti i transistori dello stesso modello montati in un circuito stampato. Le istanze di un simbolo provvisto di parametri possono poi essere “personalizzate”, ovvero quelle caratteristiche associate ai parametri possono cambiare da istanza a istanza. **Importante!** Se si modifica un oggetto (modulo o primitiva), tutte le istanze di tale oggetto presenti in tutti i moduli del progetto (che saranno ovviamente tutti a livelli gerarchici superiori) vengono modificate.

Parametri. Come accennato, ad un simbolo possono essere associati parametri che ne modificano le caratteristiche. Per esempio, se il simbolo è una primitiva che rappresenta un transistor MOS, è possibile che a tale simbolo siano associati due parametri che consentano di specificare la lunghezza e la larghezza di canale. È importante notare che i parametri consentono di cambiare individualmente le caratteristiche delle varie istanze di una stessa cella. In questo modo, per esempio, si possono inserire MOSFET con caratteristiche diverse nello stesso circuito utilizzando sempre la stessa cella primitiva. Se il simbolo rappresenta un sottocircuito è possibile utilizzare i parametri per cambiare alcune delle proprietà delle primitive in esso contenute. Per esempio è possibile associare un parametro alla larghezza di canale di tutti i MOSFET di una cella. In questo modo, associando ad una data istanza della cella un valore al parametro, tutti i transistori di quell’istanza assumono come lunghezza di canale il valore assegnato al parametro.

Fili di collegamento (wires). I fili consentono di collegare tra loro i terminali delle varie istanze presenti nel modulo (ovvero di collegare tra loro i vari elementi del circuito, siano esse istanze di primitive o di sottocircuiti) Attenzione: i fili devono necessariamente iniziare e terminare su un terminale di una istanza o su un punto di un altro filo.

Nodi del circuito: vengono in genere identificati con la parola “net”. Per esempio, il nodo 35 sarà identificato con la sigla NET35. Si può forzare un nome a piacere ad un nodo introducendo una etichetta (**label**). Associare a due nodi distinti la stessa label equivale a connetterli con un filo e quindi a fonderli in uno stesso nodo. Ciò può essere utile per evitare di portare connessioni intricate e poco leggibili in un circuito complesso.

Port (o pin): costituiscono i terminali di collegamento tra il modulo e l’esterno. Se al modulo in esame vogliamo associare un simbolo per poterlo utilizzare in schemi più complessi (ovvero gerarchicamente superiori), si seleziona un sottoinsieme delle port introdotte nel modulo affinché esse compaiano in ogni istanza del nostro modulo consentendone il collegamento con il resto del circuito utilizzatore. A ciascuna port deve essere associato un nome mnemonico (es. in, out...). Un nodo a cui è associata una port assume automaticamente il nome della port stessa.

Simboli globali Le etichette hanno solitamente visibilità limitata solo al sottocircuito in cui compaiono. Se è necessario che un nodo sia accessibile all’interno di tutti i moduli a

prescindere dal livello gerarchico, occorre associare ad esso un simbolo globale. Tipicamente si associano simboli globali alle alimentazioni (Vdd, gnd, Vss etc. etc.).

1.4 Caratteristiche generali di un layout editor

Foglio di lavoro. Il foglio di lavoro rappresenta la porzione superficie del substrato di silicio sulla quale verrà realizzato il circuito. Come per lo schema elettrico, il layout può riguardare un intero circuito integrato (chip) o più semplicemente una cella da includere successivamente in un chip. Generalmente, comunque, vengono fissati dei vincoli riguardo all'area che dovrà occupare il layout del circuito in esame. Per permettere il rispetto di questi vincoli è possibile introdurre una cornice ausiliare che identifica le dimensioni dell'area a disposizione. I punti del foglio di lavoro sono identificati da un sistema di coordinate interne. A queste possono essere fatte corrispondere delle coordinate reali (per esempio micron) e un'origine (tipicamente posizionata nel centro dell'area di lavoro. Poiché tutti i processi tecnologici hanno una minima dimensione fabbricabile (risoluzione), tutti gli oggetti dovranno avere dimensioni e coordinate coincidenti con multipli interi della risoluzione. Per questo motivo il layout editor viene predisposto (dal design_kit) in modo tale che l'operatore possa posizionarsi solo su una griglia di punti aventi come passo proprio la risoluzione.

Oggetti. Gli oggetti che un progettista può introdurre sono fondamentalmente di tre tipi:

- **Strutture fisiche:** sono figure geometriche bidimensionali (rettangoli e/o poligoni generici) che indicano un'area nella quale il progettista intende effettuare un'operazione particolare. Il tipo di operazione viene specificato stabilendo il tipo di layer (si veda più avanti per il concetto di layer) a cui l'oggetto si riferisce.
- **Informazioni per l'estrazione e per il DRC.** Queste non si traducono in nessuna struttura fisica nel chip ma servono per stabilire come deve comportarsi l'estrattore o il controllore di regole di layout. Per esempio possono essere introdotte delle port su determinati oggetti usati come connessioni (es. diffusioni, strisce di metal o di polisilicio). Siccome a ciascuna port è associato un nome simbolico, questa informazione viene riportata nella netlist, consentendo di rintracciare agevolmente il nodo corrispondente. Altri oggetti sono per esempio poligoni che istruiscono il DRC a non controllare le regole di layout nelle aree corrispondenti.
- **Commenti.** Sono disegni geometrici (cornici, guide, righelli) o stringhe di testo che facilitano la lettura o la stesura del layout ma non si traducono in oggetti fisici o in informazioni per operazioni da effettuarsi sul layout (estrazione etc.)

Layers: Come accennato sono proprietà che possono essere associate agli oggetti presenti nel layout per specificarne la funzione. I layer che indicano a quale operazione fisica si riferisce un oggetto geometrico si indicano come **layer tecnologici**. In pratica, prima di introdurre un oggetto fisico nel layout (figura geometrica, tipicamente un rettangolo) occorre selezionare un layer tra quelli disponibili. Per esempio, se il layer è quello che rappresenta il polisilicio, tutti i rettangoli (o altre figure geometriche se permesse dal processo) che vengono tracciati dopo tale selezione si tradurranno in corrispondenti oggetti di polisilicio durante la fabbricazione. Idealmente a ciascun layer corrisponde una maschera di processo. In realtà le maschere possono essere in numero superiore ai layer. Certamente i layer sono in numero sufficiente a far sì che, mediante essi, il progettista possa determinare in modo univoco la struttura di tutte le maschere del processo. I layer tecnologici possono avere due tipi di relazione con la maschera corrispondente:

-) Clear field: le figure disegnate con quel layer si traducono in aree non trasparenti nella maschera.

-) Dark field: le figure si traducono in aree chiare (trasparenti).

Questo tipo di suddivisione è in genere ininfluenza per il lavoro del progettista al quale basta sapere che dove lui posiziona una figura geometrica, lì verrà fatta l'operazione specificata convenzionalmente dal layer. La figura 1.2 mostra come due differenti layer, metal 1 e contacts, si traducano diversamente nelle corrispondenti maschere. Questo è necessario perchè le aree scure delle maschere sono quelle in cui il fotoresist non viene rimosso. Nel caso di metal 1, infatti, dove rimane il fotoresist rimane anche la metallizzazione (protetta dal fotoresist). I contacts invece, essendo buchi nell'ossido CVD (chemical vapour deposition), vengono individuati da aree chiare dove appunto il fotoresist non rimane e consente all'attacco chimico di rimuovere l'ossido.

L'uso della convenzione dark field (sostanzialmente la maschera è il negativo del layer) facilita il disegno del layout in quanto permette di disegnare layout più semplici e più leggibili.

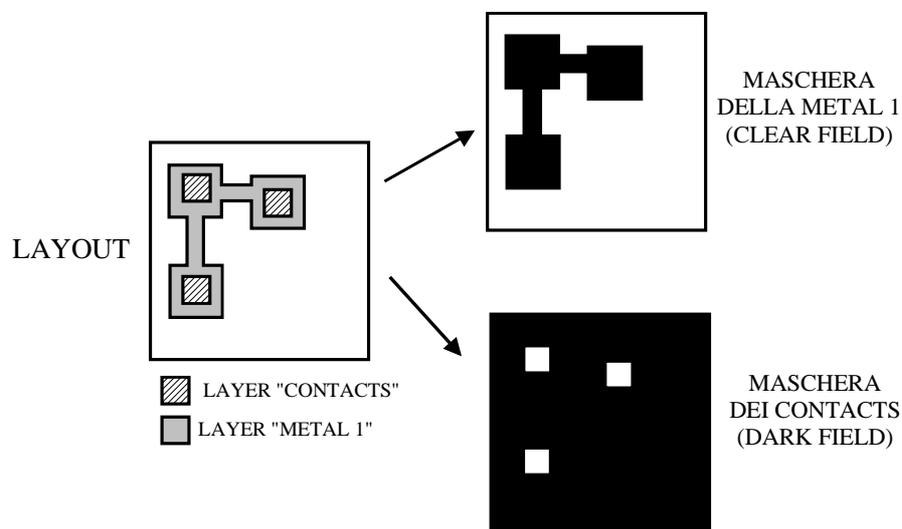


Figura 1.4.1

I layer tecnologici sono elencati all'inizio del manuale di processo e sono associati ad un numero che identifica la maschera a cui sono associati.

Regole di layout. Le regole di layout (design rules) costituiscono un insieme di regole di tipo geometrico che devono essere rispettate affinché il progetto risulti fabbricabile con una resa sufficiente. Si possono distinguere le seguenti regole di layout:

- **Larghezza (width).** Questo tipo di regola riguarda le dimensioni di un oggetto. Di solito la regola fissa la **minima** dimensione che può assumere un oggetto. Se si viola una regola di larghezza minima, l'oggetto (per esempio una pista di collegamento) rischia di risultare di larghezza nulla. Naturalmente oggetti di layer diversi avranno diverse larghezze minime.
- **Distanza (spacing).** In questo caso viene fissata la minima distanza tra due oggetti distinti. Se questa regola viene violata si rischia che i due oggetti possano risultare in contatto.

- **Sovrapposizione (overlap** oppure **margin** oppure **surround**) Misura di quanto un oggetto A deve estendersi oltre il bordo di un altro oggetto B affinché vi sia una completa sovrapposizione di A rispetto a B. Se non si rispetta questa regola si rischia che vi sia un tratto di B non incluso nella sagoma di A.
- **Coincidenza (coincidence** oppure **intersection**). Misura di quanto lungo deve essere il tratto in cui due oggetti sono sovrapposti. Se non si rispetta questa regola i due oggetti rischiano di risultare disgiunti.

Alcuni esempi di regole di layout sono mostrate in Figura 1.4.2. Le regole di layout possono essere sia in unità di lunghezza (es. micron) o adimensionali. Nel secondo caso si parla di “**regole in lambda**”. Le regole in lambda sono espresse in multipli di un’unità base detta appunto lambda (λ).

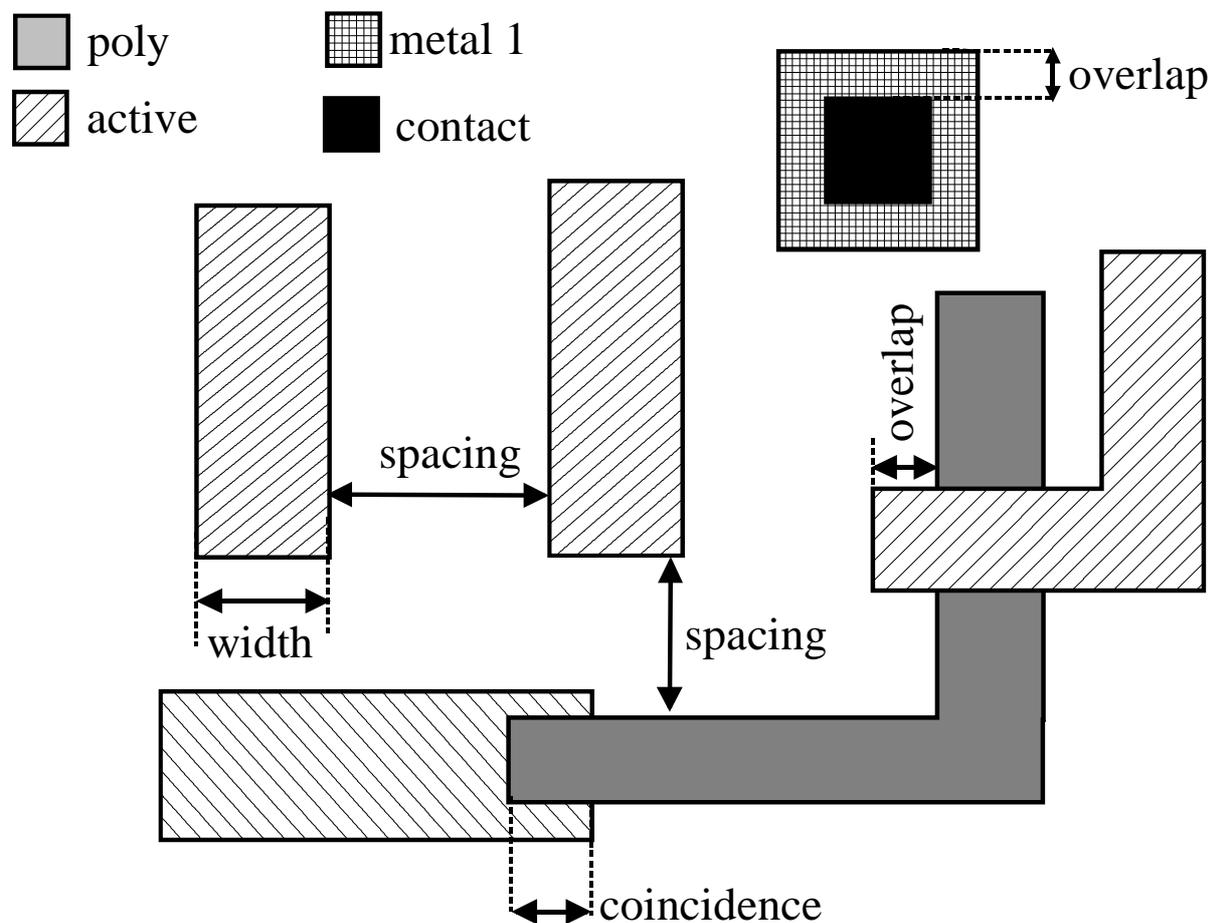


Figura 1.4.2

L’unico riferimento a unità di misura reali è effettuato specificando il valore di lambda in micron. Queste regole avevano un’utilità pratica quando le minime dimensioni degli oggetti erano fissate dalla risoluzione ottica del processo litografico che era proprio posta pari a lambda. In queste condizioni al progredire della risoluzione del processo, tutte le regole scalano della stessa quantità, derivando tutte dai limiti del processo fotolitografico. Questo a prescindere dal fatto che il layer interessato fosse, per esempio, una metal o una n-well. Il

vantaggio evidente è quello di poter adattare un layout ad un nuovo processo con una semplice operazione di scalatura, ovvero variando soltanto il valore di λ . Attualmente molti dei limiti sono posti dai processi fisici (diffusioni laterali, resistenza all'elettromigrazione, non planarità delle superfici, etc.). Pertanto, per esempio, si può avere un progresso nella minima dimensione del canale dei MOSFET senza che migliori la minima dimensione delle metal.

Pertanto per adattare un layout ad un nuovo processo, occorre ridisegnare gran parte del layout stesso. Infatti, facendo riferimento all'esempio, se si scalasse tutto per beneficiare delle dimensioni ridotte dei MOSFET, si ridurrebbero anche le metal, violando la relativa regola di layout. Per questo motivo le regole dei moderni processi sono espresse in **micron**.

Layer derivati. Oltre ai layer tecnologici possono essere presenti alcuni layer derivati (talvolta in numero molto maggiore ai layer tecnologici stessi). I layer derivati sono ottenuti con operazioni logiche tra layer tecnologici o tra altri layer derivati.

Per esempio, si può definire un layer n -active (area attiva drogata n) come l'*and* logico tra i layer corrispondenti all'area attiva e al drogaggio $n+$. In questo caso verrà considerata n -active l'intersezione tra aree attive e drogaggio $n+$. Un altro esempio è la definizione di un layer "gate" ottenibile come intersezione (and) tra polisilicio e area attiva. Questi layer non hanno un significato fisico e non vengono quindi tradotti in maschere. Essi sono utilizzati dal DRC per compiere un controllo più agevole e dall'estrattore. Per esempio l'estrattore può utilizzare il layer "gate" per riconoscere i transistori MOSFET presenti nel layout.

1.5 Elementi fondamentali di un processo CMOS n -well standard a due livelli di metal.

Nella Figura 1.5.1 è mostrato in modo molto sintetico il flusso di processo che porta a realizzare un circuito integrato in tecnologia CMOS n -well. Lo scopo è comprendere quelli che sono gli elementi che il progettista del layout può posizionare sul chip. Per questo motivo sono state eliminate tutte quelle peculiarità del processo che non aggiungono informazioni utili al progettista microelettronico. Per una rappresentazione del processo più aderente alla realtà si rimanda ad un testo di tecnologie microelettroniche. A ciascuna delle operazioni descritte in seguito è associato un layer mediante il quale il disegnatore del layout può indicare le zone dove vuole che sia applicata l'operazione in questione.

1) Il substrato di partenza è di tipo p . Vengono inizialmente realizzate, in aree selezionate del substrato, le n -well, ovvero delle zone drogate n . Le n -well ospiteranno i p -MOS, mentre gli n -MOS saranno realizzati nelle aree di substrato dove non sono presenti le well.

2) Sul substrato viene cresciuto un ossido termico (ossido di campo) che funge da isolamento. Il progettista può decidere che in alcune zone (aree attive) l'ossido di campo non sia presente. Le aree attive serviranno per costruire i MOSFET e per contattare il substrato e le well.

3) Successivamente viene cresciuto l'ossido di gate nelle aree attive e dappertutto viene depositato uno strato di polisilicio. Il polisilicio viene rimosso e con esso l'ossido di gate. Il progettista ha a disposizione un layer (poly) con il quale specifica le zone dove il polisilicio deve rimanere. Se si mantiene il polisilicio all'interno di una parte di un'area attiva, sotto di esso vi rimane sempre anche l'ossido di gate. In questo modo il polisilicio e l'ossido di gate

sottostante costituiscono rispettivamente il gate e l'isolante di un MOSFET.

4) Con opportune maschere vengono drogate le aree attive $n+$ e $p+$. Il drogaggio avviene solo nelle zone di intersezione del corrispondente layer di drogaggio con un'area attiva. Le aree attive devono essere necessariamente drogate (altrimenti si viola una regola di layout).

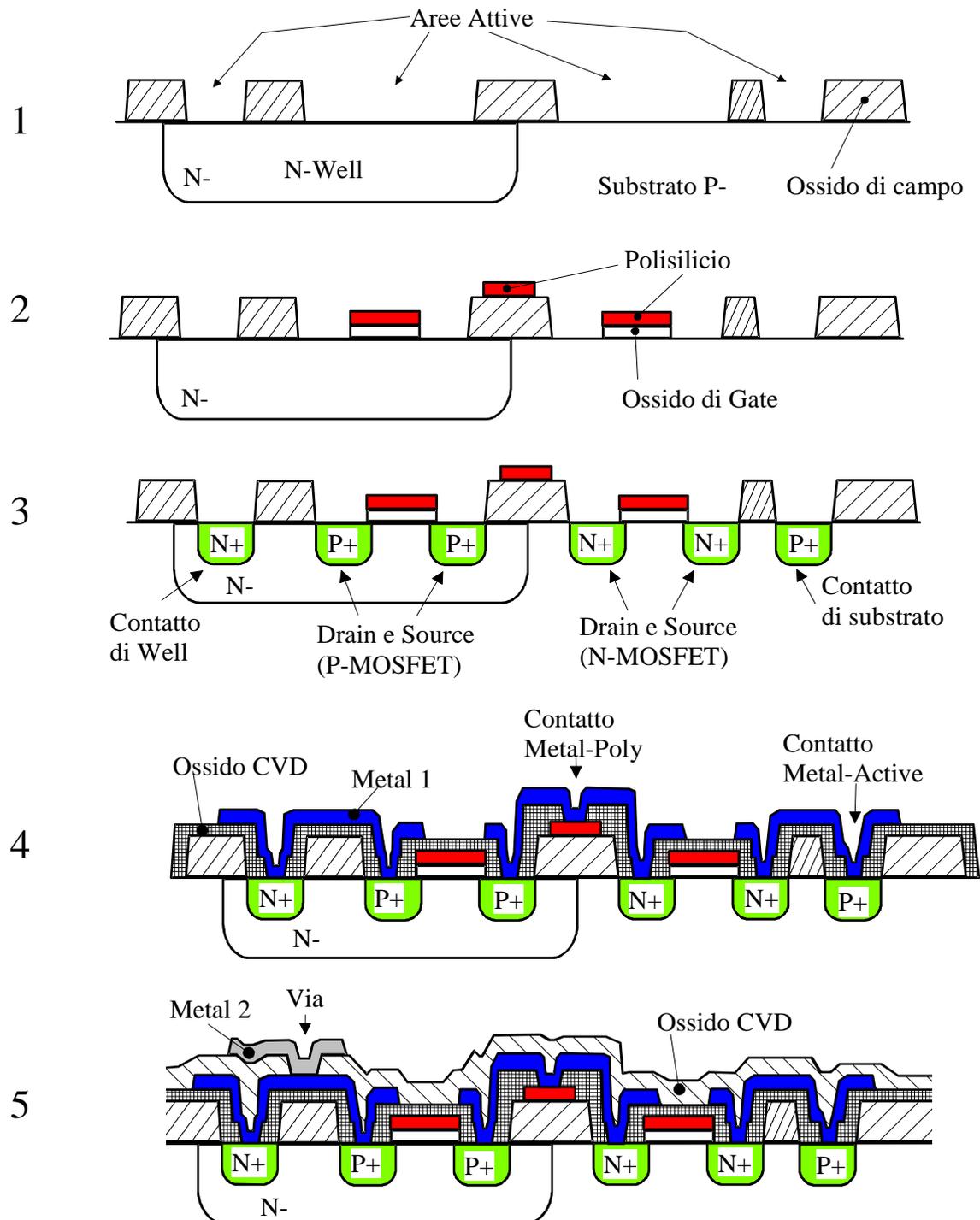


Figura 1.5.1

Si possono distinguere quattro casi diversi, come mostrato in figura:

- aree attive n^+ sul substrato con funzione di drain e source degli n -MOS;
- aree attive p^+ nella n -well con funzione di drain e source dei transistori p -MOS;
- aree attive p^+ nel substrato con funzione di contatto di substrato (substrate tap);
- aree attive n^+ nel substrato con funzione di contatto di well (well tap).

5) Viene depositato un isolante (ossido di silicio) CVD (ovvero da fase vapore). Il progettista può definire dove rimuovere selettivamente questo ossido mediante un layer di “contact”. La metallizzazione (metal1) che viene depositata successivamente sopra l’ossido CVD, e definita mediante un suo proprio layer, potrà entrare in contatto con il polisilicio o le aree attive proprio attraverso queste aperture nell’ossido CVD. In questo modo si possono, per esempio, connettere i drain/sources dei MOSFET ad una pista di metal1.

6) Viene depositato un ulteriore strato di ossido CVD sopra il quale viene depositato un secondo livello di metal (metal2). Il progettista ha a disposizione due layer, “metal2” e “via”, mediante i quali, rispettivamente, definire le piste di metal2 e decidere dove l’ossido CVD deve essere rimosso per consentire il contatto tra la metal2 e la metal1, realizzando appunto una “via”.

Naturalmente il passo 6 può essere ripetuto più volte per realizzare ulteriori livelli di metal. Il chip viene infine ricoperto da uno strato di passivazione (ossido o nitruro di silicio). Il progettista ha un layer opportuno a disposizione per decidere dove aprire questo strato di passivazione per esporre aree di metal (pad) che serviranno a realizzare le interconnessioni con i terminali del contenitore in cui il chip verrà incapsulato.

Altri processi tecnologici:

Processo	Dispositivi attivi
Bipolare	BJT npn verticali e pnp laterali.
BiFET	BJT npn verticali e pnp laterali, J-FET (tipicamente a canale p)
BiCMOS	n -MOS, p -MOS, BJT npn verticali e pnp laterali
BCD	n -MOS, p -MOS, BJT npn verticali e pnp laterali, D-MOS.

1.6 Componenti passivi per circuiti integrati.

In questo paragrafo verranno descritti i vari modi con cui si possono realizzare resistori e condensatori integrati. Va subito detto che i componenti passivi non si prestano bene ad essere integrati. Per quanto riguarda i resistori, per ottenere componenti con ridotti elementi parassiti e valori di resistenza elevati occorre aggiungere passi di processo addizionali (maschere aggiuntive) rispetto a quelli necessari per la fabbricazione dei componenti attivi. Ciò è vero anche per i condensatori, i quali, in più, presentano forti limitazioni nei valori ottenibili (di fatto limitati ad alcune decine di pF) e grandi ingombri. Tuttavia, la completa integrazione di sistemi analogici su un unico chip richiede quasi sempre l’uso di componenti passivi. Si pensi semplicemente al fatto che quando si richiede un amplificatore con guadagno preciso occorre chiudere in reazione uno o più amplificatori operazionali. Il filtraggio analogico sia a tempo continuo che a tempo discreto (condensatori commutati) richiede l’uso di condensatori. Tutti i processi

Resistori

Il tipico layout di un resistore integrato è mostrato nella Figura 1.6.1.

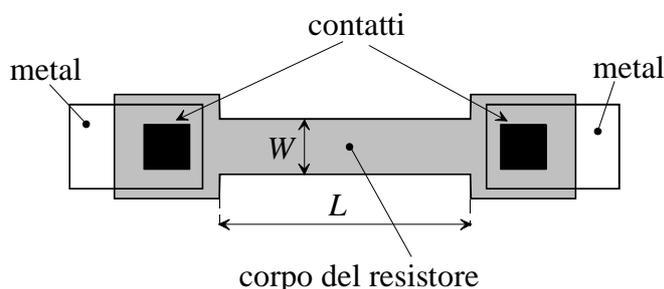


Figura 1.6.1

La resistenza nominale risulta data dalla seguente espressione:

$$R = R_s \frac{L}{W}$$

Dove R_s indica la resistenza di strato del materiale usato per il corpo del resistore mentre L e W sono le dimensioni indicate in figura. A questo valore di resistenza vanno aggiunte le resistenze di contatto spesso non trascurabili (di solito alcune decine di Ω). In base al materiale usato per il corpo del resistore si distinguono quattro tipi di resistori integrati:

- Resistori a film sottile
- Resistori in polisilicio a bassa o alta resistività
- Resistori diffusi.
- Resistori a transistori MOS.

Nei resistori a film sottile si utilizza uno strato metallico aggiuntivo, di solito Nichel-Cromo o Tantalio, caratterizzato da elevata resistenza di strato ($10\text{-}1000 \Omega/\square$) e basso coefficiente di temperatura ($\pm 100 \times 10^{-6} \text{ }^\circ\text{C}^{-1}$). I resistori a film sottile sono di elevata qualità e si prestano ad essere aggiustati in valore alla fine del processo mediante “laser trimming”. In pratica, a chip completamente fabbricato e funzionante, si distruggono sezioni del resistore mediante un fascio laser fino ad ottenere le prestazioni desiderate dal circuito (per esempio offset nullo in un amplificatore operazionale).

I resistori in polisilicio presentano qualità analoga a quella dei resistori a film sottile. Il polisilicio è presente in tutti quei processi che comprendono transistori MOS (CMOS, BiCMOS, BCD) come materiale per il gate. Sfortunatamente il polisilicio utilizzato per i gate è fortemente drogato e la sua resistenza di strato è conseguentemente bassa ($20\text{-}100 \Omega/\square$). A meno di non tollerare grandi ingombri di area, i resistori in polisilicio a bassa resistività sono dunque limitati a resistenze di valore medio-basso (alcuni $\text{k}\Omega$). In alcuni processi è possibile limitare il drogaggio di aree selezionate di polisilicio. In tale modo si può disporre di polisilicio ad elevata resistenza di strato per la realizzazione di resistori di valore elevato (fino a $\text{M}\Omega$). Naturalmente ciò comporta una complicazione del processo e l’aggiunta di almeno una maschera. Il progettista ha a disposizione un layer che posizionato su aree selezionate di polisilicio lo trasforma in polisilicio ad alta resistività.

Nei resistori diffusi il corpo del resistore è costituito da uno strato diffuso di drogaggio opposto a quello del substrato in cui è posizionato. Teoricamente ogni strato diffuso ($n+$, $p+$, n -well etc.) può essere usato per realizzare resistori. Un esempio di resistore diffuso, utilizzato in processi bipolari e BiCMOS è mostrato in Figura 1.6.2.

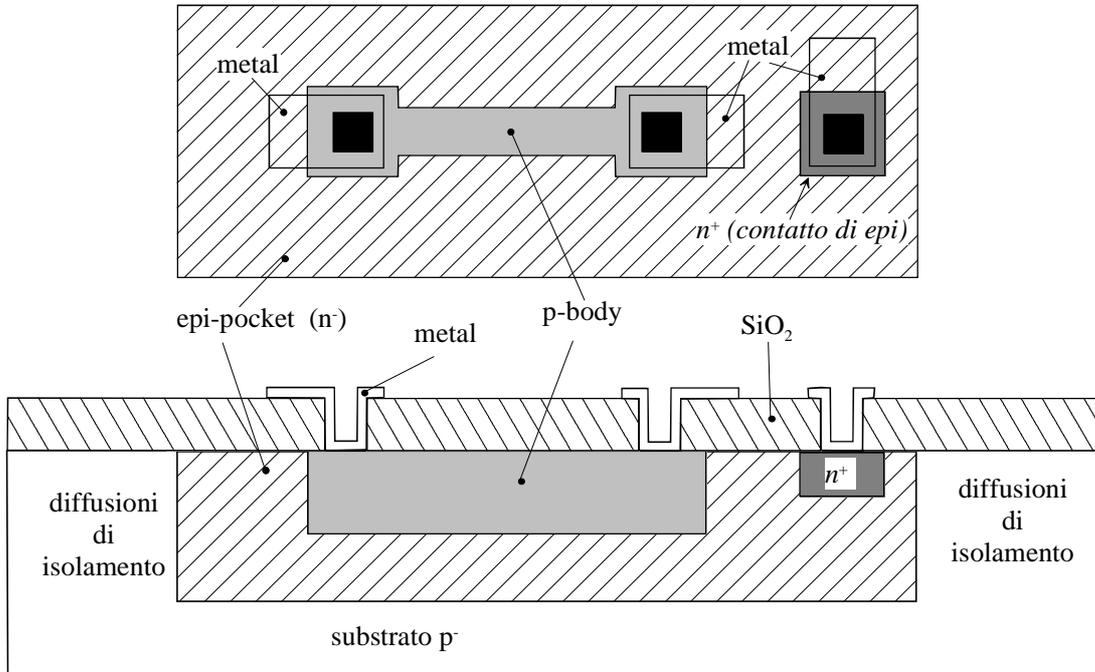


Figura 1.6.2

In questo resistore il corpo è costituito dalla diffusione p usata per le basi dei transistori bipolari nnp . Tale strato si indica spesso con p -body. Si noti la presenza di un terminale aggiuntivo per polarizzare la “sacca” di strato epitassiale (epi pocket). Lo scopo è mantenere in inversa la giunzione tra p -body e strato epitassiale onde evitare che consistenti correnti parassite sfuggano verso lo strato epitassiale medesimo. Alcuni strati comunemente usati per i resistori diffusi sono raccolti nella tabella seguente:

Strato	R_S (Ω/\square)
p -body (base degli npn)	100-200
$n+$ (emettitori)	2-10
$n+$ (source/drain)	10-50
$p+$ (source / drain)	30-100
n -well	0.5k-2k
Strato epitassiale	2k-5k

In alcuni casi è possibile aumentare la resistenza di strato riducendo la sezione del resistore mediante una diffusione di segno opposto a quella del corpo effettuata sopra al corpo del resistore stesso. Nel caso in esempio (resistore di p -body) si può effettuare una diffusione di tipo $n+$ sopra a parte del p -body. Occorre usare una diffusione meno profonda del p -body, quale quella usata per gli emettitori. I resistori ottenuti in questo modo si indicano “resistori strozzati”.

Problemi inerenti l'uso di resistori diffusi. I resistori diffusi presentano alcuni inconvenienti rispetto ai resistori a film sottile e a polisilicio. I problemi derivano dal fatto che il corpo del resistore è sempre in contatto con un substrato attraverso una giunzione $p-n$. Pertanto:

- Occorre sempre garantire che la giunzione in questione abbia polarizzazione inversa (o al limite nulla)
- Anche se la condizione precedente è rispettata esistono sempre delle correnti di perdita verso il substrato nel quale il resistore è alloggiato. Occorre verificare che queste correnti siano trascurabili rispetto alla corrente che mediamente attraversa il resistore.
- Il confine del corpo del resistore è in gran parte determinato da zone di svuotamento. Siccome l'ampiezza di queste ultime varia con la polarizzazione, anche la sezione del resistore e quindi la sua resistenza dipende dalle tensioni applicate ai terminali. Questa dipendenza della resistenza dalla tensione genera distorsione nel caso di ampi segnali e comunque limita la precisione del resistore. L'effetto è tanto più grande quanto meno drogato è lo strato usato.
- È presente una notevole capacità tra il corpo del resistore e il substrato nel quale è alloggiato.

Nei resistori ottenuti da transistori MOS, il corpo del resistore coincide con il canale del transistor stesso. Il gate del transistor viene polarizzato in modo che il transistor sia acceso. Inoltre, per ottenere un comportamento sufficientemente lineare occorre che le V_{DS} (ovvero le tensioni applicate ai capi del resistore) siano molto ridotte. I resistori ottenibili con transistori MOS sono mostrati nella Figura 1.6.3.

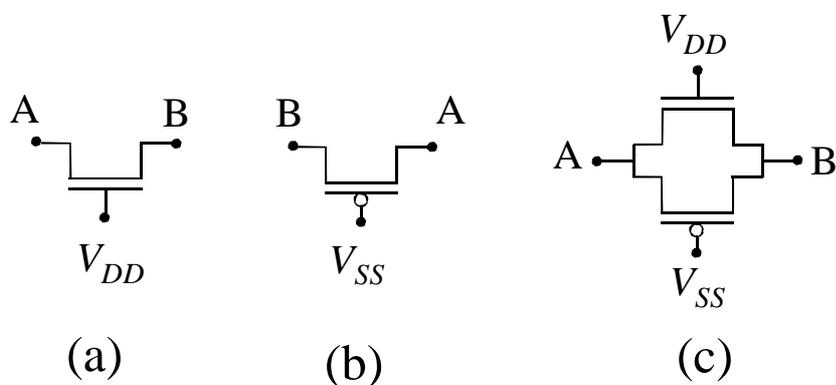


Figura 1.6.3

In tutti e tre i casi si considera che la tensione applicata al resistore (V_{AB}) sia piccola, per cui si possa considerare $V_A \cong V_B = V$. La V_{GS} sarà pertanto identificabile con $V_G - V$. In queste condizioni si verifica semplicemente che le resistenze equivalenti dei MOS sono date da:

$$\text{Caso (a): } R_n = \frac{1}{\beta_n (V_{GS} - V_m)} = \frac{1}{\beta_n (V_{DD} - V - V_m)} \quad \text{dove: } \beta_n = \mu_n C_{ox} \frac{W_n}{L_n}$$

$$\text{Caso (b): } R_p = \frac{1}{\beta_p |V_{GS} - V_{tp}|} = \frac{1}{\beta_p (V - V_{SS} - |V_{tp}|)} \quad \text{dove: } \beta_p = \mu_p C_{ox} \frac{W_p}{L_p}$$

Agendo opportunamente su W e L dei transistori si può ottenere la resistenza voluta. Con questo metodo si possono realizzare resistori di valore elevato anche in quei processi CMOS che non prevedono layer dedicati. Naturalmente i resistori a transistore MOS presentano vari difetti, tra cui la limitazione sulla tensione applicata che deve rimanere piccola (poche centinaia di mV) per mantenere il dispositivo in zona lineare. Inoltre, come per i resistori diffusi, sono presenti correnti di perdita verso il substrato a causa delle giunzioni di drain e source. Un ulteriore difetto è la dipendenza della resistenza dalla differenza di potenziale tra i terminali e il ground (indicata con V nelle formule precedenti). Per ovviare a questo inconveniente si può usare il parallelo tra un p -MOS e un n -MOS, come mostrato in (c). In questo caso la resistenza diventa:

$$R_{pn} = R_p \parallel R_n = \frac{1}{V(\beta_p - \beta_n) + (\beta_n V_{DD} - \beta_p V_{SS} - \beta_n V_m - \beta_p |V_{tp}|)}$$

La situazione più vantaggiosa si ottiene quando si fissa $\beta_n = \beta_p$, cancellando così la dipendenza dalla V . Naturalmente ciò accade solo per i valori di V in cui entrambi i transistori sono accesi. Quindi:

$$\left\{ \begin{array}{l} \beta_n = \beta_p = \beta \\ V_{SS} + |V_{tp}| < V < V_{DD} - V_m \end{array} \right. \Rightarrow R_{pn} = \frac{1}{\beta(V_{DD} - V_{SS} - V_m - |V_{tp}|)}$$

Condensatori.

I condensatori sono elementi che si prestano poco ad essere integrati in quanto, per ottenere capacità anche di pochi pF, occorre impegnare grandi aree sul chip. I valori massimi di capacità sono limitati al centinaio di pF. Tuttavia essi costituiscono l'elemento base di alcuni circuiti di estrema importanza, quali i filtri a condensatori commutati, i campionatori, gli amplificatori con annullamento dinamico dell'offset (auto-zero amplifiers) e i convertitori analogico-digitali e digitale-analogici dinamici. I condensatori sono inoltre indispensabili per effettuare la compensazione della risposta in frequenza dei circuiti reazionati.

I tipi più comuni di condensatori integrati sono i seguenti:

- Condensatori Metallo-Isolante-Metallo (MIM).
- Condensatori polisilicio-polisilicio.
- Condensatori polisilicio-diffusione n^+ .
- Condensatori a giunzione.
- Condensatori a transistore MOS.

Come per i resistori integrati, i vari processi possono supportare tipi diversi di condensatori. I condensatori MIM e polisilicio-polisilicio sono quelli che presentano caratteristiche più vicine ai condensatori ideali. Pertanto, se si deve eseguire un progetto che fa largo uso di condensatori e si richiede che questi ultimi presentino buone caratteristiche (precisione, indipendenza dalla tensione e dalla temperatura e limitati elementi parassiti), occorre assicurarsi che il processo preveda condensatori di uno di questi due primi tipi.

I condensatori MIM sono ottenuti sovrapponendo due armature di metal e assottigliando l'ossido tra le due. Infatti, l'isolante che separa due livelli di metallizzazione ha un elevato spessore (dell'ordine di 1 μm) proprio per mantenere bassa la capacità negli incroci casuali tra

le piste di interconnessione. Il progettista deve specificare dove vuole ridurre lo spessore dell'isolante per realizzare intenzionalmente una capacità. A livello di processo è necessaria quindi una maschera aggiuntiva e, a livello di design-kit, sarà presente un layer opportuno che indica dove l'ossido deve essere assottigliato. La struttura di un condensatore MIM è mostrata in Figura 1.6.4(a) dove il layer di assottigliamento dell'ossido è stato indicato con "cap".

Nel caso di condensatori polisilicio-polisilicio è necessario che il processo preveda due livelli di poly. Come per i condensatori MIM, normalmente i due livelli sono separati da un ossido relativamente spesso. Per realizzare una capacità si rimuove localmente l'ossido e al suo posto si cresce per ossidazione (come per il gate dei MOSFET) un sottile strato di SiO_2 . In questo modo, sovrapponendo nell'area selezionata i due livelli di polisilicio questi ultimi risultano separati da uno strato molto sottile e ben controllabile di isolante. Le capacità per unità di area che si possono realizzare sono in genere più elevate di quelle dei MIM.

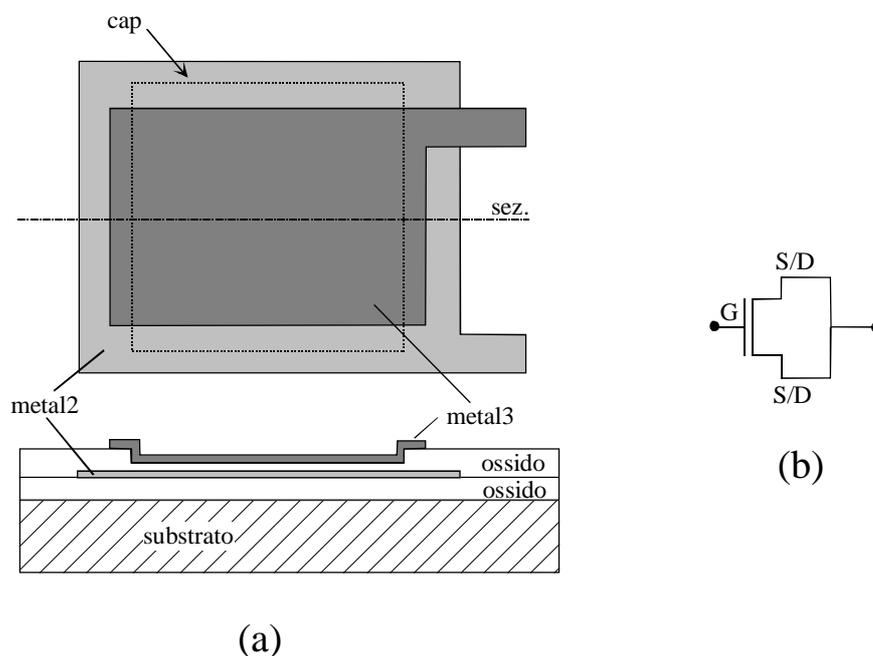


Figura 1.6.4

Nei condensatori polisilicio- diffusione n^+ il condensatore viene realizzato tra una porzione di substrato fortemente drogato, ricoperto dello stesso ossido di gate con cui si realizzano i MOSFET, e uno strato di polisilicio. È necessario drogare fortemente il substrato che funge da armatura per far sì che esso si comporti come un metallo, impedendo fenomeni di svuotamento e inversione. Le capacità che si possono ottenere sono elevate ma si ha lo svantaggio di avere un terminale del condensatore (la diffusione n^+) che presenta una giunzione parassita verso il substrato che la circonda. Rispetto ad un processo CMOS standard è richiesta una maschera in più per drogare il substrato sotto l'ossido di gate.

I condensatori a giunzione si affidano alla capacità di una giunzione polarizzata in inversa. Ovviamente le controindicazioni all'uso di questi condensatori sono molteplici, a cominciare dalla dipendenza della capacità dalla tensione applicata (utile per variare la frequenza di risonanza in circuiti accordati RF), la necessità di mantenere una polarizzazione inversa (quindi la tensione ai capi non può invertirsi) e la presenza di una giunzione parassita verso il

substrato. Tuttavia questo tipo di condensatore è spesso l'unico a disposizione in processi semplici quali il processo bipolare standard in quanto non richiede maschere aggiuntive. Le caratteristiche di questi condensatori sono accettabili per l'uso come elementi di compensazione in frequenza.

I condensatori a transistori MOSFET sfruttano la capacità gate-canale in condizioni di forte inversione e zona lineare di funzionamento. Per realizzare un condensatore con questa tecnica si connettono source e drain in cortocircuito, formando la prima armatura, come mostrato in Figura 1.6.4(b). La seconda armatura è costituita dal gate. Ovviamente è necessario garantire non solo che la tensione tra le due armature, ovvero la V_{GS} , sia sempre maggiore di zero, ma anche maggiore della tensione di soglia per garantire che sotto il gate sia presente il canale. Ovviamente anche per questi condensatori è presente una giunzione parassita tra un'armatura (source/drain) e il substrato.

Capitolo 2. Considerazioni generali relative ai dispositivi per circuiti integrati

2.1 Transistori bipolari: dimensionamento

Il parametro *area*. In un processo che comprenda transistori bipolari (esempi: processo bipolare, BiCMOS, BCD) il dimensionamento dei transistori si riduce alla scelta dell'area di emettitore del dispositivo stesso. In realtà, di solito, si prende come riferimento il transistoro di quel processo avente area di emettitore minima che di qui in poi indicheremo come "transistore elementare". Per ogni transistoro del circuito integrato viene quindi specificato un parametro denominato "*area*" che indica quante volte l'area del transistoro a cui è associato è più grande di quella del transistoro elementare. Per esempio, un transistoro che ha $area=10$ avrà un'area di emettitore pari a 10 volte quella del transistoro elementare. Il parametro *area* è quindi, per i transistori bipolari, un numero puro (cioè adimensionale). Non bisogna fare confusione con le aree (e i perimetri) di drain e source dei MOSFET che sono specificati come valori effettivi (ovvero in m^2).

Ovviamente, in un processo bipolare, esisterà come minimo almeno un transistoro elementare npn e un transistoro elementare npn. Alcuni processi prevedono più tipi di transistori npn (o pnp) e pertanto esisterà per ciascun tipo un transistoro elementare. Un processo versatile potrebbe per esempio prevedere un tipo di npn a bassa tensione, che indicheremo con npnBV, e un tipo ad alta tensione (nnpHV). Esisteranno quindi un transistoro elementare npnBV e uno di tipo npnHV. In questo caso, se inseriamo un transistoro npnHV e scegliamo per esso $area=25$, quel transistoro avrà un'area (di emettitore!) pari a 25 volte quella del transistoro elementare di tipo npnHV. L'assegnazione dell'*area* deve avvenire già in fase di disegno del circuito elettrico, in modo che le simulazioni possano tenere conto del dimensionamento. In fase di layout, l'assegnazione del parametro *area* genera effettivamente una cella di transistoro che ha un'area di emettitore che è più grande di quella del transistoro elementare di quanto specificato dal parametro *area*. Le celle di layout che in fase di "instance" possono essere personalizzate geometricamente sono denominate "celle parametriche". Solitamente l'area è un numero intero.

A seconda di quanto previsto dal *design kit*, se specifichiamo per un transistoro un'area >1 la cella effettivamente inserita nel layout può essere ottenuta in uno dei seguenti modi:

- Viene introdotto un transistoro che ha un emettitore con un lato *area* volte il corrispondente lato del transistoro elementare.
- Viene introdotto un numero pari ad *area* di transistori elementari affiancati (rispettando le regole di layout) che vengono poi automaticamente connessi in parallelo. Questa soluzione è solitamente scelta per i transistori laterali (pnp) per i quali è difficile riscaldare le dimensioni mantenendo inalterate le proprietà.
- Vengono costruiti transistori di forma compatta (strutture ripiegate, interdigitate etc. etc.) che presentano un'area effettiva di emettitore pari ad *area* volte quella del transistoro elementare. Questa soluzione è da preferirsi per i bjt di potenza, caratterizzati da elevati valori di *area*.

In ogni caso, tutte le strutture componenti in transistoro (base, collettore, giunzioni di isolamento etc. etc.) vengono riscaldate automaticamente in accordo con le accresciute dimensioni dell'emettitore per mantenere il rispetto delle regole di layout.

Scelta dell'area. La scelta di un transistoro dell'area viene effettuata seguendo vari criteri.

Innanzitutto va detto che se i transistori fossero ideali, ovvero l'equazione che specifica le correnti di collettore in zona attiva diretta:

$$I_C = J_S A_E e^{\frac{V_{BE}}{V_T}} \left(1 + \frac{V_{CB}}{V_A} \right) \quad I_B = \frac{I_C}{\beta} \quad (2.1.1)$$

fosse rispettata per qualsiasi valore di I_C e I_B allora ai fini del dimensionamento di un circuito occorrerebbe fissare solo i rapporti tra le aree dei vari transistori e non i valori effettivi delle aree. Se prendiamo come esempio uno specchio di corrente semplice, si ha, ritenendo valide le equazioni (2.1.1), che la funzione che lega la corrente di uscita a quella di ingresso dipende solo dal rapporto tra le aree del transistoro di ingresso e quello di uscita. Pertanto, moltiplicando o dividendo le aree dei transistori dello specchio per lo stesso fattore, il funzionamento dello specchio dovrebbe risultare inalterato. In realtà, per un dato transistoro, le equazioni (2.1.1) sono valide solo in un intervallo di correnti di collettore che, seppur ampio varie decadi, non è illimitato.

Solitamente nel manuale di processo è riportato il grafico di Gummel (o Gummel-plot), ovvero il grafico in scala semilogaritmica delle correnti di collettore e di base in funzione della tensione base-emettitore (V_{BE}). Tale grafico è riferito ad una tensione V_{CE} fissata ad un valore tale da garantire che il transistoro sia in piena zona attiva (es. 5 V). Il grafico di Gummel viene in genere riportato solo per il transistoro elementare di tipo di BJT. Nella figura seguente è riportato un esempio di Gummel-plot.

GUMMEL PLOT

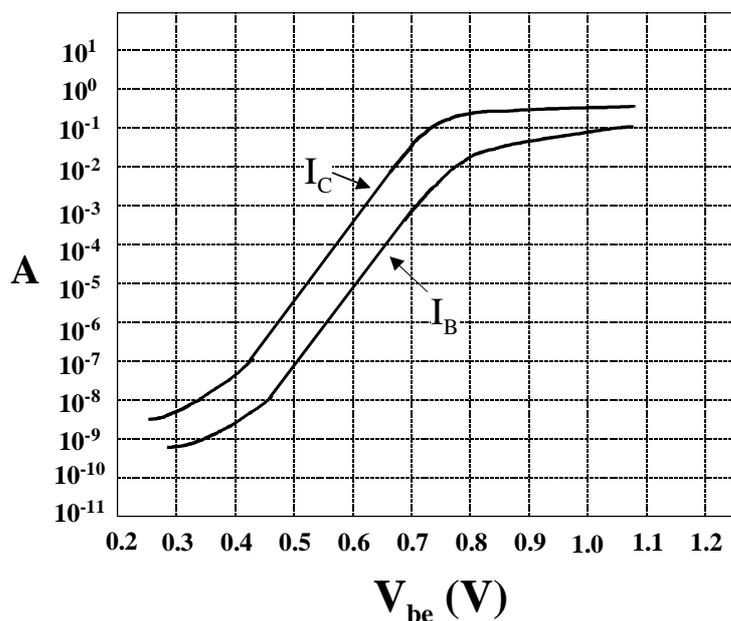


Figura 2.1.1

Come si vede il grafico della I_C è rettilineo in un ampio intervallo di correnti di collettore ($0.1 \mu\text{A}$ - 10 mA). In questo range di correnti (e quindi di tensioni V_{BE}) l'equazione (2.1.1) è soddisfatta. In basso la deviazione è dovuta alle correnti di ricombinazione nella zona di svuotamento della giunzione base-emettitore e alla corrente inversa di saturazione della giunzione base-collettore. In alto intervengono effetti di alta iniezione e la caduta di tensione sulla resistenza serie di base. Se si esclude il caso dei transistori destinati ad erogare potenze elevate (stadi di uscita) il parametro *area* deve essere scelto in modo da garantire che la

corrente di collettore vada a lavorare nella zona rettilinea. Infatti un transistor avente $area \neq 1$ può essere sempre rappresentato come il parallelo di più transistori elementari, ciascuno dei quali riceverà quindi una corrente pari a $I_C/area$.

Quindi, se un transistor deve portare da progetto una corrente che eccede il limite superiore della zona rettilinea del Gummel plot, basta assegnargli un'area sufficiente a far sì che la corrente che riceve ciascun transistor elementare che lo compone (ovvero $I_C/area$) cada nella zona corretta del Gummel-plot del transistor elementare. Come controindicazioni, un valore di area elevato comporta elevati ingombri e grandi capacità parassite (es. collettore-substrato e base-collettore) Nella figura seguente è mostrata simbolicamente l'equivalenza tra un transistor di area pari a 3 e il parallelo di tre transistori elementari.

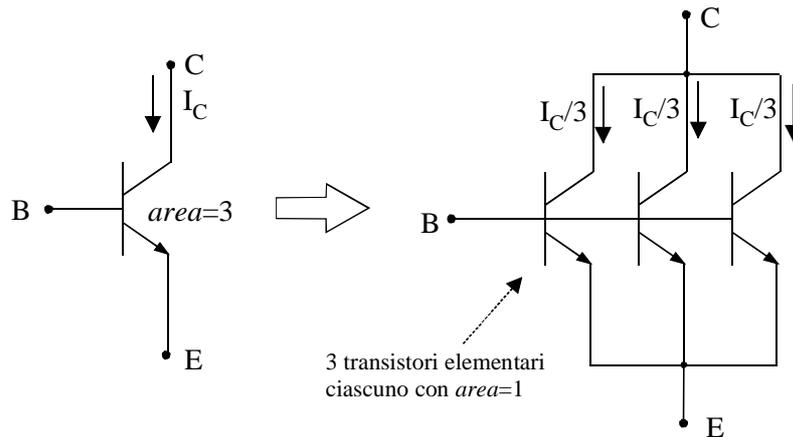


Figura 2.1.2

Nel caso di transistori di potenza è spesso necessario sfruttare anche il tratto superiore della caratteristica, dove inizia la deviazione dall'andamento rettilineo. Appare ovvio che non conviene spingersi molto oltre il ginocchio in quanto a fronte di modesti guadagni in termini di I_C occorre applicare V_{BE} che vanno ben oltre il tipico valore 0.6-0.7 V.

Sempre con riferimento alle equazioni (2.1.1), rammentiamo che anche il beta non è un parametro costante ma varia con la I_C . Il grafico seguente mostra un andamento tipico del beta in funzione della I_C per il transistor elementare. Al solito, per valutare il beta di un transistor con $area \neq 1$ occorre riportare nel grafico non la I_C ma il valore del rapporto $I_C/area$.

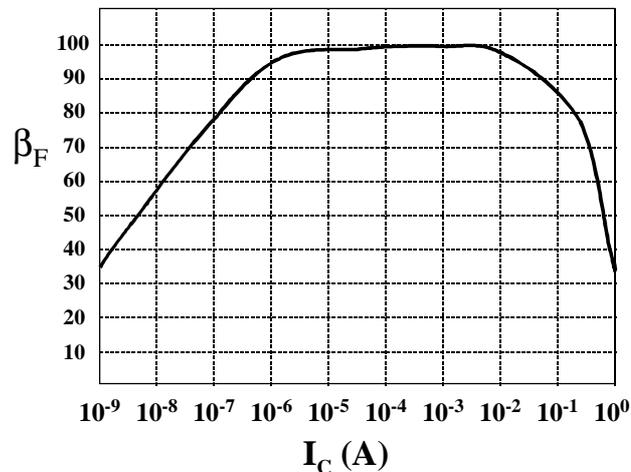


Figura 2.1.3

2.2 Nozioni utili per il dimensionamento di transistori MOS in circuiti analogici.

Il dimensionamento dei transistori MOS (ovvero la scelta di L e W) è un'operazione che consente di variare un grande numero di parametri tra i quali, limitatamente al funzionamento in continua e a basse frequenze, il g_m , la r_d , il valore della V_{DSAT} (ovvero il valore di V_{DS} che separa la zona triodo da quella di saturazione, pari a $V_{GS}-V_t$). Come vedremo trattando gli amplificatori differenziali a MOSFET, il dimensionamento dei transistori modifica il guadagno dello stadio, cosa che non accade per i transistori bipolari. Pertanto non è possibile stabilire dei criteri semplici e universali per il dimensionamento dei transistori MOS in un circuito analogico. Tuttavia si possono fare alcune considerazioni riguardo alle proprietà dei MOS che possono semplificare questa operazione. Fisseremo la nostra attenzione sui due seguenti aspetti delle proprietà dei MOSFET:

- La variazione del g_m e di r_d al variare della V_{DS} .
- La variazione del prodotto $g_m r_d$ in saturazione al variare del punto di riposo.

Per quanto riguarda il primo punto, partiamo da un'espressione generica della I_D in funzione di V_{GS} e V_{DS} , trascurando per semplicità l'effetto della V_{BS} (presente tra l'altro solo se il source non è connesso al substrato). I parametri g_m e r_d possono essere espressi come derivate parziali di questa espressione:

$$g_m = \frac{\partial I_{DS}(V_{GS}, V_{DS})}{\partial V_{GS}}, \quad \frac{1}{r_d} = \frac{\partial I_{DS}(V_{GS}, V_{DS})}{\partial V_{DS}} \quad (2.2.1)$$

Se facciamo riferimento alle equazioni semplificate della I_D , relative al LEVEL 1 di SPICE:

$$\text{zona triodo: } V_{DS} < V_{GS} - V_T \Rightarrow I_D = \beta \left[(V_{GS} - V_T)V_{DS} - \frac{V_{DS}^2}{2} \right] (1 + \lambda V_{DS}) \quad (2.2.2)$$

$$\text{saturazione: } V_{DS} \geq V_{GS} - V_T \Rightarrow I_D = \frac{\beta}{2} (V_{GS} - V_T)^2 (1 + \lambda V_{DS})$$

Si ottengono le seguenti espressioni per il g_m e il parametro $1/r_d$:

$$g_m = \begin{cases} \text{in zona triodo: } \beta V_{DS} (1 + \lambda V_{DS}) \\ \text{in saturazione: } \beta (V_{GS} - V_T) (1 + \lambda V_{DS}) \end{cases} \quad (2.2.3)$$

$$\frac{1}{r_d} = \begin{cases} \text{in zona triodo: } \beta [(V_{GS} - V_T) - V_{DS}] (1 + \lambda V_{DS}) + \lambda \beta \left[(V_{GS} - V_T)V_{DS} - \frac{V_{DS}^2}{2} \right] \\ \text{in saturazione: } \lambda \frac{\beta}{2} (V_{GS} - V_T)^2 \end{cases} \quad (2.2.4)$$

Queste espressioni possono essere semplificate nelle seguenti formule:

$$g_m = \begin{cases} \text{in zona triodo: } \beta V_{DS} \\ \text{in saturazione: } \beta (V_{GS} - V_T) \end{cases} \quad \frac{1}{r_d} = \begin{cases} \text{in zona triodo: } \beta [(V_{GS} - V_T) - V_{DS}] \\ \text{in saturazione: } \lambda \frac{\beta}{2} (V_{GS} - V_T)^2 \cong \lambda I_D \end{cases} \quad (2.2.5)$$

I grafici della figura seguente mostrano simbolicamente l'andamento di g_m e $1/r_d$ al variare della V_{DS} come si ricava dalle equazioni approssimate (2.2.5) con l'unica differenza di aver fatto sì che la funzione $1/r_d$ sia continua nel passaggio tra zona triodo e zona di saturazione (l'espressione approssimata di $1/r_d$ valida in zona triodo andrebbe altrimenti a zero per $V_{DS} = V_{GS} - V_T$)

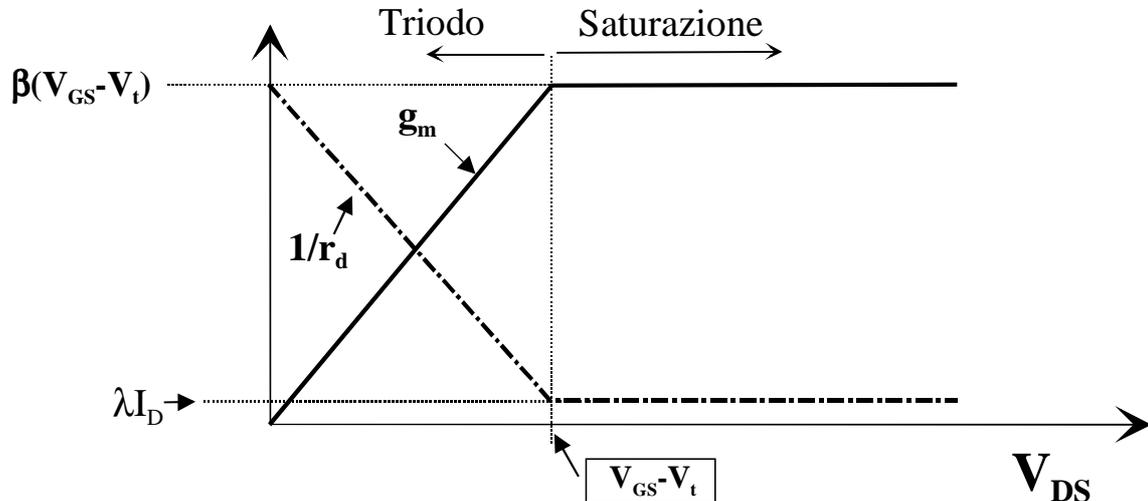


Figura 2.2.1

Dunque, quando si passa dalla saturazione alla zona triodo:

1. Il g_m si riduce fino a diventare nullo per $V_{DS}=0$.
2. La r_d diminuisce progressivamente tendendo, per $V_{DS}=0$, ad un valore pari all'inverso del g_m di saturazione. Questo valore è notoriamente molto più piccolo del valore assunto da r_d in saturazione.
3. Conseguentemente il prodotto $g_m r_d$, pari al massimo guadagno teorico di tensione realizzabile con il MOSFET per il dato punto di riposo, tende a zero per $V_{DS} \rightarrow 0$.

Dall'ultimo punto si comprende come, dal punto di vista dell'amplificazione, sia particolarmente svantaggioso far lavorare i MOSFET in zona triodo.

Per quanto riguarda prodotto $g_m r_d$ in saturazione osserviamo che, facendo riferimento alle equazioni semplificate (2.2.5) esso vale:

$$g_m r_d = \frac{g_m}{I_D} \cdot \frac{1}{\lambda}$$

Il rapporto g_m/I_D è una grandezza importante per le applicazioni analogiche dei MOSFET e vale:

$$\frac{g_m}{I_D} \cong \beta(V_{GS} - V_T) \frac{2}{\beta(V_{GS} - V_T)^2} = \frac{2}{(V_{GS} - V_T)} \quad (2.2.6)$$

Quindi $g_m r_d$ non dipende dal beta del transistoro ma solo da $V_{GS}-V_T$ e dal lambda. Quest'ultimo dipende dalla lunghezza di canale secondo una proporzionalità inversa. Pertanto, per avere transistori con elevati $g_m r_d$ occorre avere lunghezze di canale non minime.

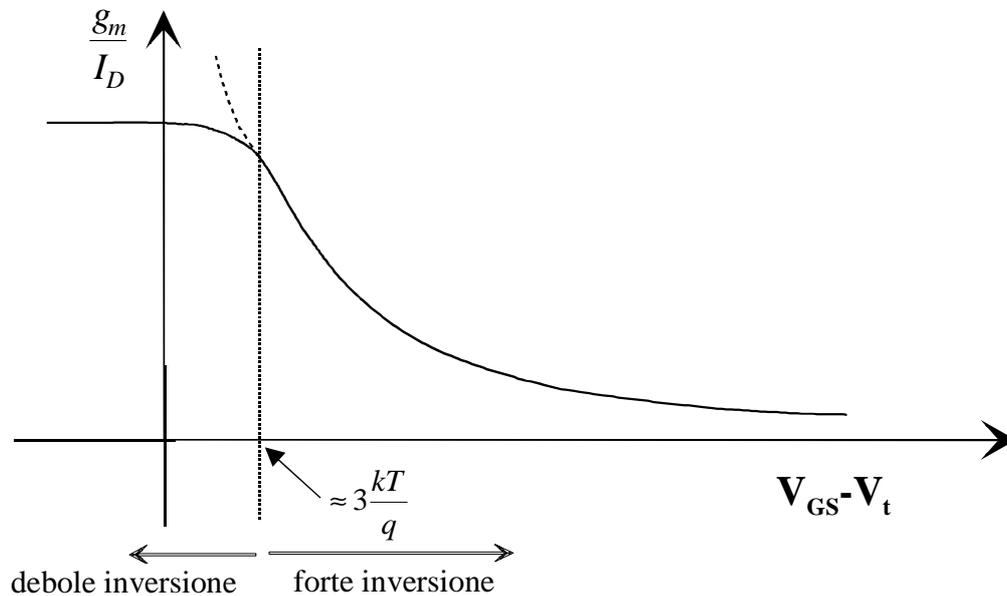


Figura 2.2.2

In pratica aumentare eccessivamente la lunghezza di canale per ricercare elevati guadagni in continua causa un degradamento della risposta in frequenza, in quanto, per mantenere lo stesso rapporto W/L occorre aumentare di pari passo anche W , con la conseguenza di aumentare le capacità parassite. È raro che per aumentare il guadagno statico si arrivi a scegliere lunghezze di canale superiori a cinque volte la lunghezza minima.

Per quanto riguarda il rapporto g_m/I_D sembrerebbe invece che si possano raggiungere valori arbitrariamente grandi (e quindi prodotti $g_m r_d$ arbitrariamente grandi) semplicemente scegliendo un punto di riposo con $V_{GS}-V_T$ infinitesima. In realtà quando $V_{GS}-V_T$ scende sotto un valore di qualche kT/q (tipicamente $3kT/q$, ovvero circa 75 mV a temperatura ambiente) il transistoro entra in zona di debole inversione (weak inversion) e le equazioni paraboliche della I_D non sono più valide. Per $V_{GS}-V_T$ che tende a zero o diventa negativa (zona di funzionamento sotto soglia), le equazioni della corrente da paraboliche diventano esponenziali. Il risultato è che il valore di $g_m r_d$ non va all'infinito per $V_{GS}-V_T$ che tende a zero ma si stabilizza ad un valore costante, come mostrato nella Figura 2.2.2. In sostanza, già attorno a $V_{GS}-V_T=100$ mV non conviene più scendere con la V_{GS} in quanto i guadagni che si ottengono sono modesti e, anzi, si riduce il g_m a spese delle prestazioni in termini di risposta in frequenza (vedi: amplificatori differenziali).

2.3 Errori di fabbricazione e uguaglianza tra dispositivi nominalmente identici.

Premessa. Spesso occorre che due dispositivi abbiano stesse caratteristiche elettriche (“*matching* dei dispositivi”). L’esempio più comune è costituito dagli amplificatori differenziali dove l’eguaglianza dei dispositivi di ingresso (transistori) e di carico (transistori o resistori) garantisce la simmetria del circuito e quindi bassi valori della tensione di offset. Un altro caso si ha negli specchi di corrente dove talvolta si richiede che la corrente di uscita sia il più possibile uguale a quella di ingresso e ciò si traduce nella necessità di avere transistori il più possibile uguali nel ramo di ingresso e di uscita.

Facendo riferimento ad un certo parametro circuitale, fisico o geometrico, occorre innanzitutto precisare alcune definizioni:

- Valore nominale: è il valore del parametro fissato da progetto. Si parlerà di dispositivi o circuiti “nominalmente identici” per indicare che essi sono stati progettati identici. Ovviamente, a causa degli errori di fabbricazione essi risulteranno differenti.
- Valore reale: è il valore del parametro che risulta effettivamente realizzato sul chip. Ovviamente il valore reale cambia da chip a chip, ovvero da realizzazione a realizzazione.
- Errore sul valore nominale: è la differenza tra il valore reale e il valore nominale.
- Errore di *matching*: è la differenza tra i valori di quel parametro per due dispositivi presenti sullo stesso chip, nominalmente identici (ovvero progettati identici).

Gli errori sul valore nominale e di *matching* vengono spesso espressi come errori relativi, ovvero come frazione del valore nominale.

L’errore sul valore nominale è di norma assai più grande degli errori di *matching*. Esso può raggiungere il 40 % del valore nominale mentre gli errori di *matching* possono essere ridotti a valori inferiori all’1% del valore nominale e si possono ulteriormente ridurre con gli accorgimenti che verranno esposti in seguito. Questo accade perché i dispositivi che si trovano su uno stesso chip, specie se spazialmente vicini, tendono ad avere valori dei parametri fisici (drogaggi, spessore dell’ossido, dimensioni effettive) molto simili, anche se assai diversi dai corrispondenti valori nominali. In altre parole, dispositivi nominalmente identici che si trovano su uno stesso chip possono essere significativamente diversi dai dispositivi progettati, ma tra di loro sono generalmente molto simili.

Un’altra distinzione da fare è quella tra errore sistematico ed errore casuale, definizione che si applica sia all’errore sul valore nominale, sia all’errore di *matching*. Prendiamo in esame un grande numero (teoricamente infinito) di dispositivi nominalmente identici e consideriamo la media dell’errore sul valore nominale che caratterizza un certo parametro. Se questa media non è nulla, ma, supponiamo per esempio, è positiva, significa che “mediamente” i dispositivi realizzati hanno un valore di quel parametro superiore a quello progettato. In questo caso siamo in presenza di un errore sistematico. È chiaro che questa è una situazione scorretta, indice di problemi nel progetto o nel processo di fabbricazione dei chip. Un errore sistematico (ovvero una media dell’errore non nulla) su un errore di *matching* significa che uno dei due dispositivi (per esempio quello “più a destra” o “più in alto”) ha un parametro sempre più grande (o più piccolo) dell’altro. Questo denota quasi sempre un errore di progetto: per esempio uno dei due dispositivi viene posizionato più vicino ad un elemento di potenza e quindi avrà una temperatura sempre più elevata dell’altro e ciò altera sempre nello stesso senso tutti i parametri influenzati dalla temperatura.

La differenza tra l'errore vero e proprio e la media dell'errore (ovvero lo scostamento dal valore effettivo dal valor medio del parametro) è il cosiddetto errore casuale. L'errore casuale ha quasi sempre distribuzione gaussiana. L'errore casuale viene caratterizzato dalla deviazione standard σ che è una stima statistica della radice quadrata della varianza e quindi la identifichiamo con essa.

La situazione corretta è quella in cui l'errore sistematico sia assente indipendentemente dal fatto che si tratti di errore di matching o errore sul valore nominale e quindi permanga solo l'inevitabile componente casuale. In altri termini ciò comporta che:

- se si parla di errore sul valore nominale, la media dei valori del parametro effettuata su un grande numero di realizzazioni coincide effettivamente con il valore nominale
- se si parla di errore di matching, i due dispositivi hanno parametri "mediamente uguali", ovvero nessuno dei due mostra una tendenza ad essere superiore (o inferiore) all'altro.

La Figura 2.3.1 mostra schematicamente la distribuzione dei valori di un generico parametro A evidenziando l'errore sistematico e casuale. Per distribuzione di A si intende la funzione densità di probabilità di A.

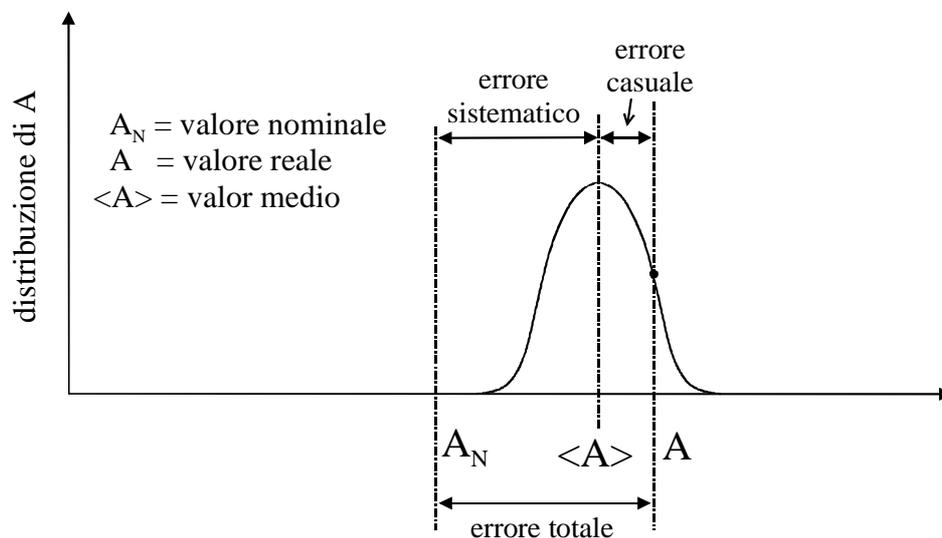


Figura 2.3.1

Si ricorda che entro un intervallo $\bar{A} \pm 3\sigma$ finisce oltre il 99 % dei valori reali.

Rappresentazione degli errori di matching.

Consideriamo due dispositivi nominalmente identici ed un parametro ad essi associato (per esempio, se sono MOSFET, la lunghezza di canale). Indichiamo con A questo parametro. A_1 sarà il valore del parametro per il dispositivo 1 e A_2 per il dispositivo 2. Generalmente si caratterizza la coppia di valori A_1 e A_2 facendo riferimento alla loro media e alla differenza, ovvero:

$$\bar{A} \equiv \frac{A_1 + A_2}{2}; \quad \Delta A \equiv A_1 - A_2 \quad (2.3.1)$$

Invertendo questa trasformazione si ottiene:

$$A_1 = \bar{A} + \frac{\Delta A}{2}; \quad A_2 = \bar{A} - \frac{\Delta A}{2} \quad (2.3.2)$$

L'errore di matching assoluto è indicato dalla differenza tra i valori dei parametri, ovvero da ΔA . Il valore medio può spesso differire significativamente dal valore nominale a causa dell'errore sul valore nominale. Si faccia attenzione al fatto che il valor medio \bar{A} definito con l'equazione (2.3.1) è in realtà una media effettuata tra due soli dispositivi che tra l'altro si trovano sullo stesso chip e quindi non coincide con il valor medio del valore nominale (indicato con $\langle A \rangle$ in Figura 2.3.1) che è una media effettuata su un grande numero di chip. In assenza di grandi errori di fabbricazione \bar{A} è comunque una stima dell'ordine di grandezza del valore nominale. Pertanto si può identificare il rapporto $\Delta A/\bar{A}$ con l'errore di matching relativo.

Errori di matching sulle grandezze derivate.

Indichiamo con G una grandezza elettrica che è funzione di alcuni parametri A, B, C, \dots secondo una legge analitica qualsiasi, ovvero $G=G(A,B,C,\dots)$. Indichiamo inoltre con 1 e 2 due dispositivi caratterizzati dalla grandezza G . Dunque per il dispositivo 1 sarà: $G_1=G(A_1,B_1,C_1,\dots)$ e per il dispositivo 2: $G_2=G(A_2,B_2,C_2,\dots)$.

Spesso è importante conoscere come dall'errore di matching sui parametri A, B, \dots si passi all'errore di matching (assoluto o relativo) su G .

Se si sviluppa G attorno al valor medio arrestandosi al primo ordine, si ottiene:

$$\begin{aligned} \Delta G = G(A_1, B_1, \dots) - G(A_2, B_2, \dots) &= \left(\bar{G} + \frac{\partial G}{\partial A} \cdot \frac{\Delta A}{2} + \dots \right) - \left(\bar{G} - \frac{\partial G}{\partial A} \cdot \frac{\Delta A}{2} + \dots \right) = \\ &= \frac{\partial G}{\partial A} \Delta A + \frac{\partial G}{\partial B} \Delta B + \dots \end{aligned} \quad (2.3.3)$$

Per quanto riguarda l'errore relativo, che come detto può essere assimilato a $\Delta G/\bar{G}$:

$$\frac{\Delta G}{\bar{G}} = \frac{\partial G}{\partial A} \frac{\Delta A}{\bar{G}} + \frac{\partial G}{\partial B} \frac{\Delta B}{\bar{G}} + \dots \quad (2.3.4)$$

È significativo il caso in cui la grandezza G risulta data da un'espressione del tipo:

$$G = \frac{a \cdot b \cdot c \cdot \dots}{A \cdot B \cdot C \cdot \dots}$$

In questo caso si verifica facilmente che l'errore relativo risulta:

$$\frac{\Delta G}{\bar{G}} = \left(\frac{\Delta a}{\bar{a}} + \frac{\Delta b}{\bar{b}} + \frac{\Delta c}{\bar{c}} + \dots \right) - \left(\frac{\Delta A}{\bar{A}} + \frac{\Delta B}{\bar{B}} + \frac{\Delta C}{\bar{C}} \right)$$

Più in generale, se α, β, γ sono costanti, si ha:

$$G = A^\alpha B^\beta C^\gamma \Rightarrow \frac{\Delta G}{G} = \alpha \frac{\Delta A}{A} + \beta \frac{\Delta B}{B} + \gamma \frac{\Delta C}{C} \quad (2.3.5)$$

Un altro caso saliente è quello che si ha quando la grandezza derivata G assume la forma:

$$G = \ln(A^\alpha B^\beta C^\gamma)$$

Applicando la (2.3.3) e svolgendo le derivate si ottiene:

$$\Delta G = \alpha \frac{\Delta A}{A} + \beta \frac{\Delta B}{B} + \gamma \frac{\Delta C}{C} \quad (2.3.6)$$

Si osservi come in questo caso (logaritmo di un prodotto) l'espressione sia simile alla (2.2.5), ma dia come risultato lo scarto assoluto (ΔG) invece che relativo.

In realtà, come abbiamo visto, quello che è importante è caratterizzare statisticamente l'errore di matching, ovvero stimare la sua deviazione standard. Si può allora fare riferimento all'espressione che dà la deviazione standard della somma H delle variabili casuali A , B , C , valida solo per A , B , C incorrelate:

$$H = A + B + C \Rightarrow \sigma_H = \sqrt{\sigma_A^2 + \sigma_B^2 + \sigma_C^2} \quad (2.3.7)$$

e a quella che dà la deviazione standard del prodotto di una variabile casuale A per una costante k :

$$\sigma_{(kA)} = k\sigma_A \quad (2.3.8)$$

Applicando le equazioni (2.3.7) e (2.3.8) alla (2.3.5) con le considerazioni fatte si ottiene un'espressione per la deviazione standard della variabile $\Delta G / \bar{G}$

$$G = A^\alpha B^\beta C^\gamma \Rightarrow \sigma_{\frac{\Delta G}{G}} = \sqrt{\alpha^2 \sigma_{\frac{\Delta A}{A}}^2 + \beta^2 \sigma_{\frac{\Delta B}{B}}^2 + \gamma^2 \sigma_{\frac{\Delta C}{C}}^2} \quad (2.3.9)$$

Esempio: applichiamo quanto visto al caso di due resistori per circuiti integrati (per esempio in polisilicio o diffusi) in cui la resistenza è data da:

$$R = R_s \frac{L}{W} \quad (2.3.10)$$

dove R_s è la resistenza di strato, L la lunghezza del resistore e W la larghezza.

Questi due resistori non avranno la stessa resistenza ma saranno caratterizzati da un errore relativo pari a:

$$\frac{\Delta R}{R} = \frac{\Delta R_s}{R_s} + \frac{\Delta L}{L} - \frac{\Delta W}{W}$$

Come esempio numerico si supponga che la resistenza di strato sia uguale per entrambi (o, per lo meno che le sue variazioni siano trascurabili), e che sia $\Delta L = 0.1 \mu\text{m}$, $\Delta W = -0.15 \mu\text{m}$ e che inoltre si abbia $\bar{W} = \bar{L} = 2 \mu\text{m}$. Si ottiene: $\Delta R / \bar{R} = -0.025$, ovvero la variazione relativa tra

le due resistenze sarà di -2.5 %. In realtà, questo che abbiamo visto è un caso particolare relativo ad una singola coppia di resistori. Come già accennato il dato importante è quello statistico, ovvero la deviazione standard della variazione di resistenza. Per calcolare quest'ultima avremmo bisogno delle deviazioni standard delle grandezze $\Delta W / \bar{W}$ e $\Delta L / \bar{L}$. In genere quello che è fissato dal processo è la deviazione standard dell'errore di matching assoluto, ΔL e ΔW , e non del valore relativo. Questo errore è legato alla risoluzione geometrica dei processi tecnologici (fotolitografia, impiantazione ionica, attacchi degli ossidi, diffusione) e, generalmente, non dipende dalla dimensione di W e L (in sostanza l'errore su una dimensione di 1 micron o 100 micron è lo stesso).

In generale, data una grandezza A , se:

- le variazioni ΔA e la media \bar{A} sono indipendenti;
- ΔA è a media nulla (come richiesto per avere errore sistematico di matching nullo);
- le variazioni ΔA sono molto più piccole di \bar{A} ;

allora si può fare l'approssimazione:

$$\sigma_{\frac{\Delta A}{A}} \cong \frac{\sigma_{\Delta A}}{\langle A \rangle} = \frac{\sigma_{\Delta A}}{A}$$

dove con A abbiamo indicato il valore nominale che, in assenza di errori sistematici coincide con la media dei valori della variabile A . Tornando al nostro resistore si potrà quindi scrivere:

$$\sigma_{\frac{\Delta R}{R}} \cong \sqrt{\left(\sigma_{\frac{\Delta R_s}{R_s}}\right)^2 + \left(\frac{\sigma_W}{W}\right)^2 + \left(\frac{\sigma_L}{L}\right)^2}$$

purché, naturalmente, si possano considerare incorrelati gli errori ΔW , ΔL e ΔR_s . Si osserva che l'errore può essere ridotto aumentando le dimensioni nominali del componente, in quanto, come già detto, le deviazioni standard delle dimensioni geometriche non dipendono dai valori nominali delle dimensioni stesse. Si noti che il progettista non può invece agire su $\sigma_{\Delta R_s / R_s}$ che risulta invece fissato dal processo.

Se la precisione del processo è identificata da $\sigma_{\Delta W} = \sigma_{\Delta L} = 0.1 \mu\text{m}$ e $\sigma_{\Delta R_s / R_s} = 0.005$, e il resistore ha dimensioni nominali $W=L=10 \mu\text{m}$, applicando le considerazioni precedenti si ottiene:

$$\sigma_{\frac{\Delta R}{R}} = 0.015$$

Pertanto l'errore di matching relativo tra le due resistenze è del 1.5 %. Più semplicemente, se facciamo riferimento a $3\sigma=0.045$, ciò significa che per più del 99 % delle coppie di resistori effettivamente fabbricate la differenza tra i due resistori sarà minore del 4.5 %.

Criteri per la riduzione degli errori di matching. Se vogliamo che due dispositivi presenti in un chip siano il più possibile uguali occorre rispettare alcuni criteri di progetto elencati qui di seguito.

1. Disegnare due dispositivi effettivamente uguali, ovvero sovrapponibili per traslazione. Per esempio, facendo riferimento alla formula (2.3.10), due resistori potrebbero avere stessa

resistenza anche con valori di L e W diversi, purché il rapporto W/L sia uguale per essi. Tuttavia, a causa degli effetti di perimetro (effetti di bordo), i due resistori sarebbero affetti da un errore sistematico sul valore della resistenza: gli effetti di bordo pesano di più nel resistore più piccolo (ovvero con W e L più piccoli) a causa del maggior rapporto perimetro/area. Stesso discorso vale per i MOSFET per i quali non è sufficiente che siano uguali i rapporti W/L ma è necessario che i dispositivi abbiano uguali valori di L e W . Anche coppie di dispositivi nominalmente identici ma con forme diverse (es. lineare e interdigitata) devono essere evitate.

2. I due dispositivi devono avere stessa orientazione. Facendo sempre riferimento ai resistori, se i loro lati più lunghi (ovvero quelli paralleli alle linee di corrente) sono disposti secondo assi aventi inclinazione diversa (es. ortogonali o formanti un angolo di 45°), risentiranno in modo diverso degli stress del substrato o, in alcuni casi, saranno caratterizzati da diverse proprietà elettriche (es. resistività, piezoresistività etc..).
3. La direzione della corrente dovrà essere la stessa nei due dispositivi. Questa regola deriva dal fatto, che correnti nei semiconduttori possono produrre effetti diversi (es. tensioni) se il loro verso è concorde o discorde con gradienti termici presenti nel materiale.
4. Scegliere ove possibile dimensioni non minime per i dispositivi. Come visto nell'esempio relativo al resistore, l'errore relativo sul valore della resistenza diminuisce al crescere delle dimensioni L e W del resistore stesso. Sempre a titolo di esempio, si può citare l'errore di matching tra le tensioni di soglia di due MOSFET: in questo caso lo scarto ΔV_t è inversamente proporzionale a $\sqrt{W \cdot L}$.
5. I due dispositivi dovranno essere posizionati il più possibile vicini (compatibilmente con le regole di layout) in modo da far sì che le differenze dei parametri fisici (drogaggio, spessore dell'ossido etc.) siano il più possibile uguali.
6. Effettuare, ove possibile, la disposizione baricentrica dei due dispositivi. Essa consiste nello sdoppiare ciascuno dei due dispositivi in due elementi identici in parallelo e nel disporre i quattro elementi risultanti come mostrato nella Figura 2.3.2(b). La denominazione "baricentrica" deriva dal fatto che i due transistori composti M1 (M1A + M1B) e M2 (M2A + M2B) hanno geometricamente lo stesso baricentro, indicato con una lettera X nella Figura 2.3.2(b). Il vantaggio di questa configurazione è mostrato nella Figura 2.3.3, dove vengono mostrate le linee di livello (supposte localmente parallele) di un qualche parametro fisico D (per esempio drogaggio del substrato). Vengono distinti tre casi a seconda dell'inclinazione delle linee di livello. La freccia indica la direzione in cui il parametro D cresce.

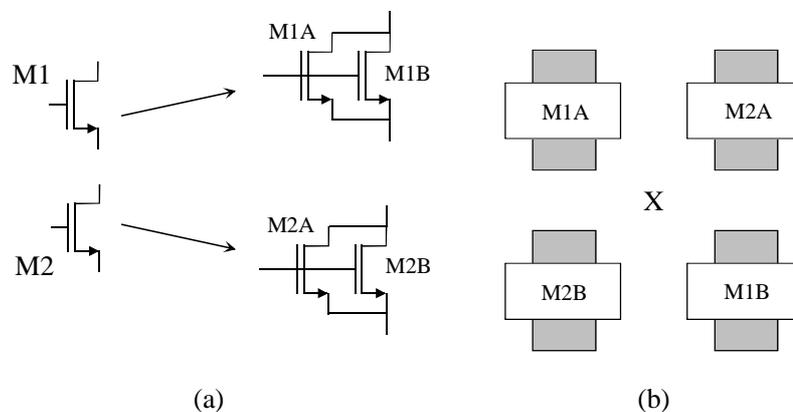


Figura 2.3.2

Osserviamo che nel caso 1 e nel caso 2 la situazione per i due transistori composti è simmetrica ovvero M1 e M2 hanno un elemento con $D=D1$ e l'altro con $D=D2$. Per cui M1 e M2 sono equivalenti. Nel terzo caso M2 ha entrambi i componenti con $D=D2$, mentre M1 ha un componente a $D=D1 < D2$ e l'altro a $D=D3 > D2$ per cui vi è, almeno al primo ordine una sorta di compensazione e le differenze tra M1 e M2, ovvero gli errori di matching, risultano ridotte.

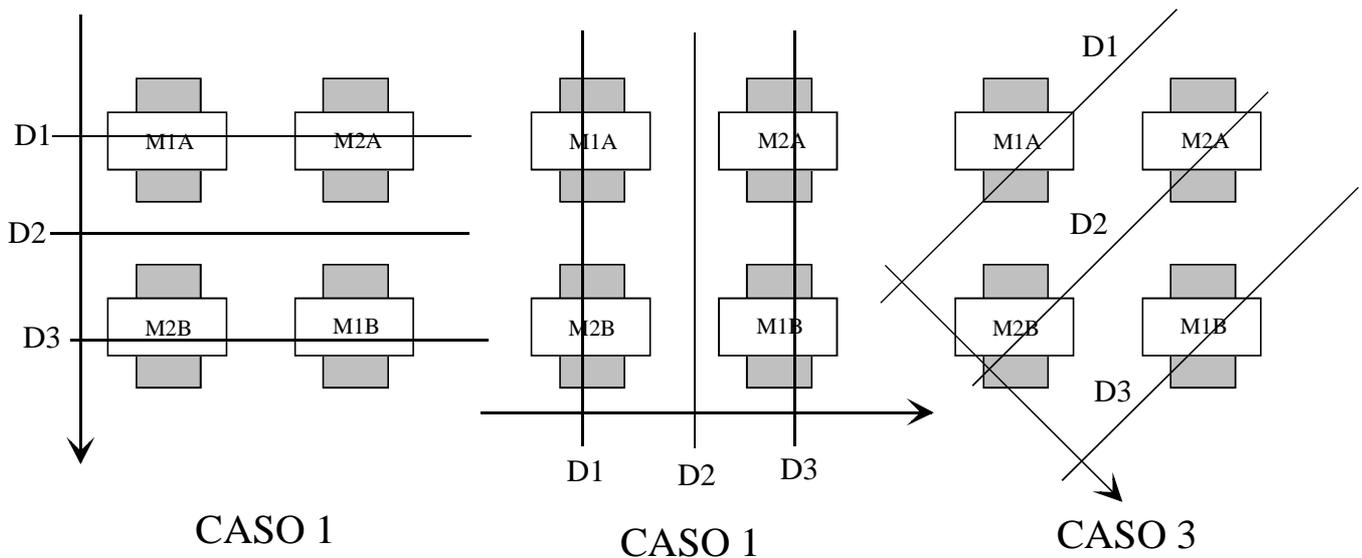


Figura 2.3.3

Cenni alla realizzazione di rapporti precisi tra le caratteristiche di due dispositivi.

Si faccia riferimento al caso in cui necessitiamo di due MOSFET aventi beta ($\beta = \mu C_{ox} W / L$) che stanno in rapporto 1:N. Questa situazione è utile, per esempio, negli specchi di corrente, per realizzare rapporti precisi tra le correnti. Nel nostro caso le correnti risulterebbero in rapporto 1:N.

Se noi semplicemente disegniamo due transistori aventi stesse lunghezze di canale e larghezze W in rapporto 1:N, il risultato cercato si ottiene solo parzialmente in quanto gli effetti di bordo peseranno di più percentualmente sul transistore a W più piccola e quindi i beta effettivi non saranno in rapporto preciso 1:N. La soluzione corretta è quella di realizzare un transistore con $W=W_1$ e il secondo transistore come il parallelo di N transistori identici al primo. In questo modo la corrente del secondo transistore sarà la somma delle correnti degli N transistori componenti e questo equivale elettricamente ad avere un beta N volte quello del singolo transistore componente. Se invece abbiamo bisogno di due transistori con beta in rapporto $M:N$ basterà realizzare un transistore con M transistori elementari e l'altro con N transistori elementari. Naturalmente la stessa cosa vale per coppie di transistori bipolari che devono avere un rapporto preciso tra le aree di emettitore.

Per quanto riguarda i resistori, se un resistore deve valere R_1 e l'altro NR_1 , non conviene realizzare il secondo semplicemente come un resistore lungo N volte il primo ma bisogna connettere in serie N resistori identici al primo. In questo modo siamo sicuri di moltiplicare per N anche le resistenze di contatto che altrimenti peserebbero percentualmente di più sul resistore più piccolo.

Infine, per realizzare rapporti precisi tra capacità non è corretto realizzare condensatori con aree diverse ma, anche qui, conviene connettere in parallelo più condensatori elementari.

Capitolo 3. Specchi di corrente

3.1 Definizioni generali.

Gli specchi di corrente sono dispositivi a tre terminali i quali sono denominati:

- **Terminale di riferimento:** è quello rispetto al quale si misurano le tensioni di ingresso e uscita e che riceve la somma delle correnti di ingresso e uscita. Solitamente coincide con uno dei terminali di alimentazione (power supply rail) ma vi sono delle applicazioni in cui questo non accade.
- **Terminale di ingresso:** la corrente che entra nel terminale di ingresso verrà indicata con I_R , mentre V_{in} sarà la tensione tra il terminale di ingresso e quello di riferimento.
- **Terminale di uscita:** la corrente che entra nel terminale di uscita verrà indicata con I_{out} , mentre V_{in} sarà la tensione tra il terminale di ingresso e quello di riferimento.

In uno specchio di corrente ideale la corrente di uscita è proporzionale alla corrente di ingresso. Ovvero la legge ideale dello specchio di corrente è:

$$I_{out} = k_S \cdot I_R$$

dove k_S è un coefficiente caratteristico dello specchio che non dipende da I_R e dalla tensione V_{out} .

Occorre subito precisare che uno stesso specchio non può trattare correnti e tensioni di segno qualsiasi. In sostanza gli specchi di corrente si possono suddividere in due gruppi:

- Specchi di tipo **n** (detti anche pozzi di corrente o *current sink*) in cui il terminale di riferimento deve avere potenziale inferiore agli altri due e le correnti di ingresso e uscita scorrono dai rispettivi terminali verso il riferimento stesso.
- Specchi di tipo **p** (detti anche sorgenti di corrente o *current source*) in cui il terminale di riferimento deve avere potenziale superiore a quello degli altri due e le correnti di ingresso e uscita scorrono dal riferimento verso i rispettivi terminali.

La Figura 3.1.1 rappresenta i due tipi di specchio di corrente specificando le grandezze elettriche a cui si è fatto riferimento.

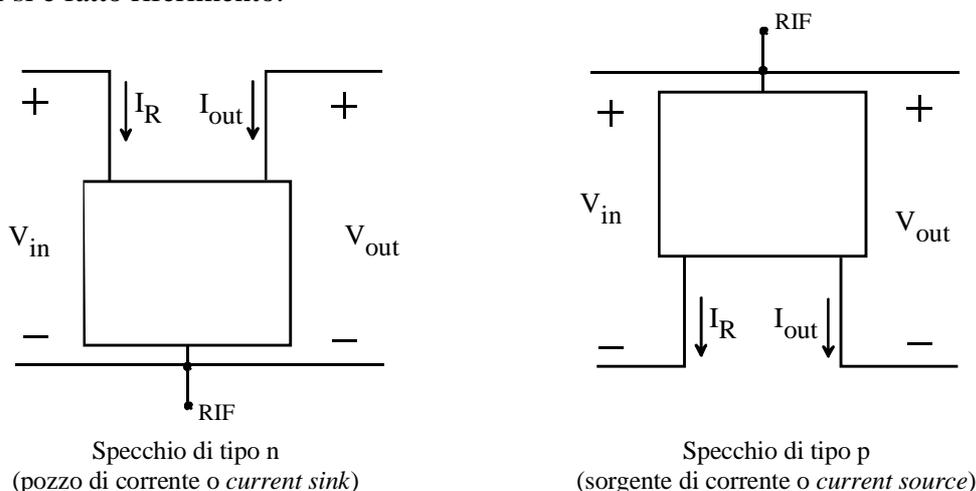


Figura 3.1.1

I requisiti che uno specchio deve soddisfare vengono elencati qui di seguito assieme ai parametri che misurano quanto uno specchio reale si avvicina allo specchio ideale.

1. Indipendenza della corrente I_{out} dalla tensione V_{out} . Questa è espressa dal parametro resistenza di uscita differenziale, definito come

$$R_{out} = \left(\frac{\Delta I_{out}}{\Delta V_{out}} \right)^{-1}$$

Più alta è la resistenza di uscita, più indipendente è la corrente I_{out} dalla V_{out} . La R_{out} è un parametro importante in molte applicazioni degli specchi. Poiché come vedremo la resistenza R_{out} non è una caratteristica della sola topologia circuitale di uno specchio ma è anche circa inversamente proporzionale alla corrente di uscita, viene introdotto il parametro “tensione di Thevenin” o “tensione a vuoto” definito come:

$$V_{Th} = I_{out} R_{out}$$

Negli specchi di corrente più diffusi la V_{Th} non dipende dalla corrente di uscita e misura quindi la qualità della topologia circuitale.

2. Ampio intervallo di tensioni di uscita in cui lo specchio funziona correttamente. In pratica esiste una tensione di uscita minima che indicheremo con V_{MIN} sotto la quale la corrente di uscita inizia a diminuire rapidamente. Per $V_{out}=0$ avremo inevitabilmente $I_{out}=0$ indipendentemente da I_R . In altre parole per $V_{out} < V_{MIN}$ la resistenza di uscita diminuisce rapidamente rispetto al valore nominale valido per $V_{out} > V_{min}$ (può diminuire di vari ordini di grandezza). La tipica caratteristica I_{out} in funzione di V_{out} di uno specchio è mostrata in Figura 3.1.2. L'intervallo di tensioni V_{out} in cui lo specchio funziona correttamente si indica dinamica di uscita.
3. Bassa tensione di ingresso (V_{IN}): questo requisito è motivato dal fatto che se la tensione di ingresso è troppo alta può accadere che il circuito deputato ad erogare la I_R non sia in grado di fornirla. Ovviamente la tensione richiesta in ingresso non deve superare la differenza tra le tensioni di alimentazioni positiva e negativa ($V_{DD}-V_{SS}$) altrimenti lo specchio non potrà mai funzionare correttamente. Per esempio, se uno specchio richiede una V_{IN} di 5 V e il circuito che lo usa deve essere alimentato con $V_{SS}=-1.5$ V, $V_{DD}=1.5$ V lo specchio non potrà mai essere polarizzato correttamente e quindi non è adatto a lavorare con quelle tensioni di alimentazione. Ovviamente la V_{IN} dipende anche dalla corrente di ingresso.
4. Tensione di ingresso V_{IN} che varia poco con la I_R , ovvero bassa resistenza differenziale di ingresso. Questa caratteristica è particolarmente utile se si vuole che lo specchio si comporti come un amplificatore ideale di corrente.
5. Ampia possibilità di variare il rapporto $k_S=I_{out}/I_R$ variando il dimensionamento dello specchio. Vi sono specchi di corrente che non soddisfano questa condizione in quanto lavorano correttamente solo per $k_S=1$.

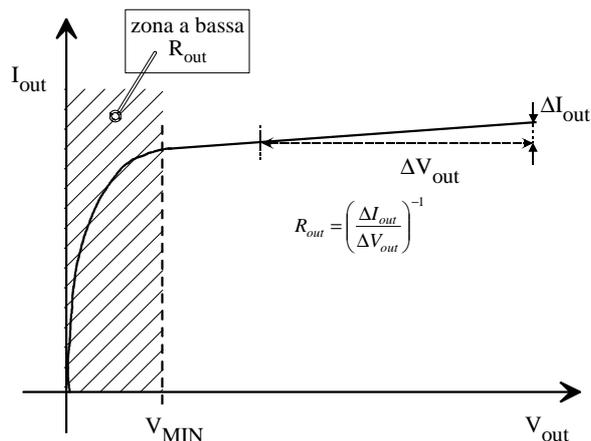


Figura 3.1.2

Circuito equivalente alle variazioni per uno specchio di corrente (in continua).

In Figura 3.1.3 si riporta il circuito alle variazioni di uno specchio. Si noti che il coefficiente che lega le variazioni della corrente di ingresso a quelle della corrente di uscita (indicato con k'_S in figura) è generalmente differente da quello statico k_S . Inoltre, la corrente $k'_S I_R$ non confluirà tutta nella I_{out} , a meno che il circuito utilizzatore non mostri verso lo specchio una resistenza molto inferiore a R_{out} .

Per il calcolo dei parametri del circuito linearizzato si supporrà sempre che il terminale di riferimento sia a massa.

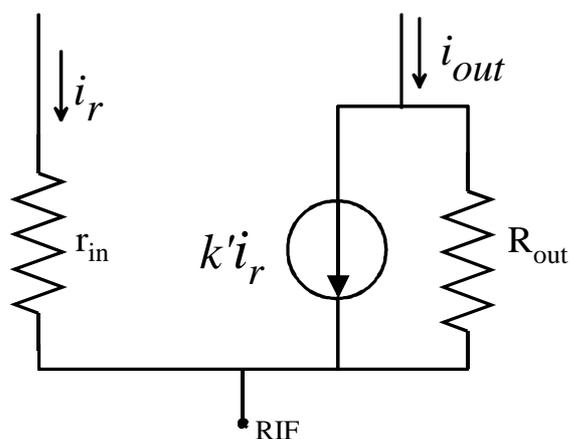


Figura 3.1.3

3.2 Specchi di corrente a MOSFET:

Specchio semplice a MOSFET.

La figura seguente mostra la struttura dello specchio di corrente a MOSFET di tipo semplice nelle due realizzazioni a n-MOS e p-MOS.

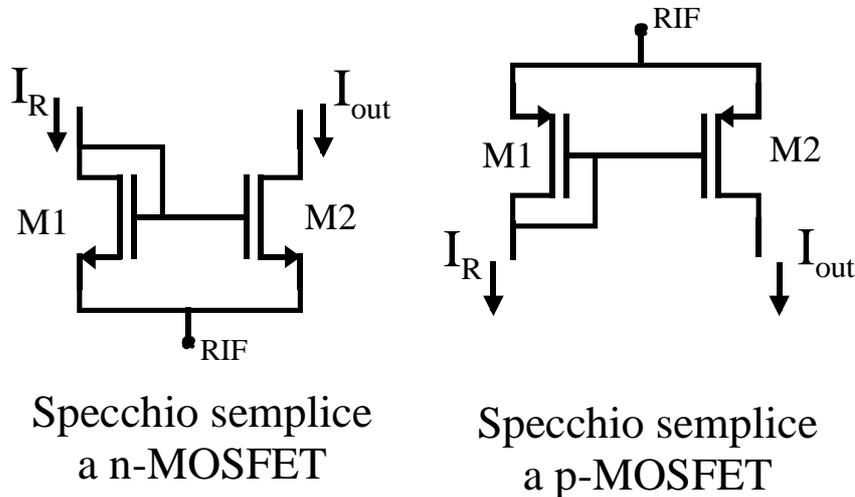


Figura 3.2.1

Il funzionamento dello specchio semplice è il seguente: I transistori M1 e M2 hanno la stessa V_{GS} e quindi, trascurando l'effetto della V_{DS} il rapporto tra le loro correnti di drain è dato da:

$$\frac{I_{D2}}{I_{D1}} = \frac{I_{OUT}}{I_R} \cong \frac{\mu_n C_{ox} \frac{W_2}{2L_2} (V_{GS} - V_T)^2}{\mu_n C_{ox} \frac{W_1}{2L_1} (V_{GS} - V_T)^2} = \frac{\beta_2}{\beta_1} = \frac{W_2 L_1}{W_1 L_2} \quad (3.2.1)$$

L'equazione precedente è valida se M1 e M2 sono in saturazione. M1 è collegato a diodo e quindi è sempre in saturazione (oppure è spento); M2 sarà in saturazione per $V_{out} > V_{GS} - V_T$. Pertanto, si può variare il rapporto tra le correnti k_s semplicemente agendo sul rapporto β_2/β_1 , ovvero sulle dimensioni L e W di M1 e M2.

Si trova immediatamente che la resistenza di uscita è pari alla resistenza differenziale di drain del MOSFET M2 ovvero $R_{out} = r_{d2}$ (ai fini del calcolo della resistenza di uscita in continua M2 ha gate e source a massa). Poiché si ha:

$$r_{d2} = \frac{\lambda_2^{-1}}{I_{D2}} = \frac{\lambda_2^{-1}}{I_{out}}$$

la tensione di Thevenin dello specchio risulta proprio: $V_{Th} = \lambda_2^{-1}$.

Questo valore della resistenza di uscita (e della tensione di Thevenin) presuppone che M2 sia in saturazione. Come già accennato, la V_{out} dovrà quindi soddisfare la condizione:

$$V_{out} > V_{GS} - V_T \equiv V_{MIN}$$

La tensione di ingresso è invece semplicemente pari a V_{GS1} e la resistenza di ingresso è pari a:

$$R_{in} = \frac{1}{g_{m1}} \parallel r_{d1} \cong \frac{1}{g_{m1}} \quad (3.2.2)$$

Per quanto lo specchio semplice sia ampiamente utilizzato e costituisca l'elemento base di molti blocchi circuitali più complessi, in molti casi si richiede l'impiego di specchi di corrente a resistenza di uscita più alta. Si può ricorrere allora allo specchio *cascode* che verrà esposto nel paragrafo seguente.

Specchio *cascode* a MOSFET.

Lo schema dello specchio *cascode* è mostrato nella figura seguente nelle due versioni n e p

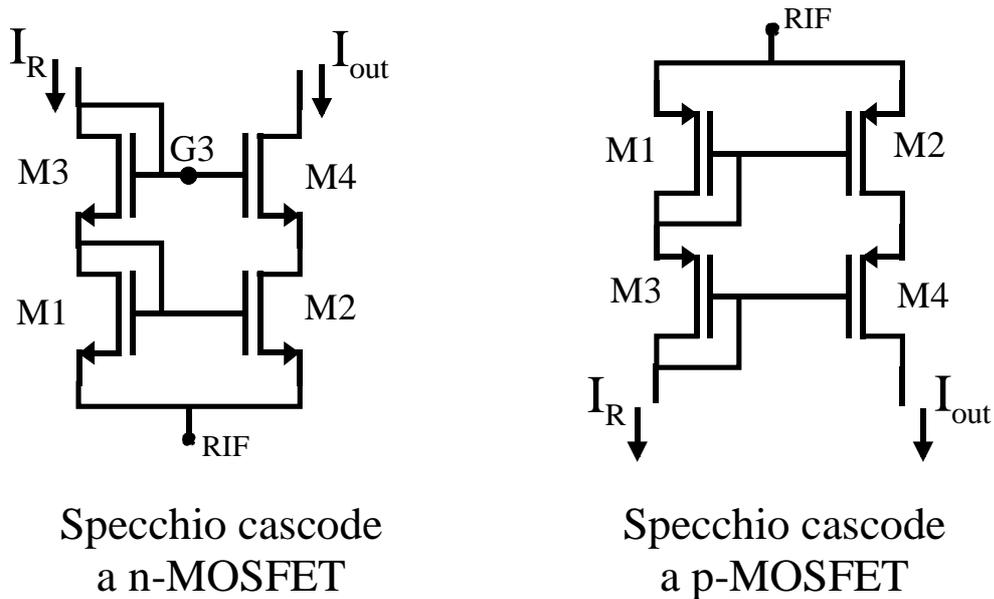


Figura 3.2.2

Funzionamento. Osserviamo innanzitutto che se M1 e M2 sono in saturazione la relazione tra le correnti I_{D1} e I_{D2} è la stessa che per lo specchio semplice, espressa dall'equazione (3.2.1). Ma, banalmente, $I_{D1}=I_{D3}=I_R$ e $I_{D2}=I_{D4}=I_{out}$, pertanto anche nello specchio *cascode* il rapporto tra le correnti di ingresso e di uscita (k_S) è fissato da M1 e M2. Il vantaggio è che nel cascode anche $V_{DS1}=V_{DS2}$, per cui la relazione (3.2.1) è vera con maggior precisione. Il ruolo di M3 e M4 è proprio quello di far sì che $V_{DS1}=V_{DS2}$, indipendentemente da V_{out} . Per dimostrare questo assunto ipotizziamo che tutti i transistori siano in saturazione (M1 e M3 sono sicuramente in saturazione per il montaggio a diodo). Osserviamo subito che, come per lo specchio semplice, $V_{GS1}=V_{GS2}=V_{GS}$. Dunque, trascurando in prima approssimazione l'effetto delle V_{DS} :

$$I_R = \frac{\beta_1}{2}(V_{GS} - V_T)^2 = \frac{\beta_3}{2}(V_{GS3} - V_{T3})^2$$

$$I_{out} = \frac{\beta_2}{2}(V_{GS} - V_T)^2 = \frac{\beta_4}{2}(V_{GS4} - V_{T4})^2$$

Si osservi che siccome i source di M3 e M4 non sono al potenziale di riferimento, (come invece accade per M1 e M2), le tensioni di soglia di M3 e M4 sono state poste diverse da quelle di M1 e M2 (indicate con V_T) per tenere conto dell'effetto body. Dividendo membro a membro le equazioni precedenti si ottiene:

$$\frac{(V_{GS3} - V_{T3})}{(V_{GS4} - V_{T4})} = \sqrt{\frac{\beta_1 \beta_4}{\beta_2 \beta_3}}$$

Da progetto si fa sì che:

$$\frac{\beta_3}{\beta_4} = \frac{\beta_1}{\beta_2} \Rightarrow (V_{GS3} - V_{T3}) = (V_{GS4} - V_{T4}) \quad (3.2.3)$$

Da questa equazione e dal fatto che $V_{G3}=V_{G4}$ (i rispettivi gate sono connessi assieme) si ottiene:

$$V_{G3} - V_{S3} - V_{T3} = V_{G3} - V_{S4} - V_{T4} \Rightarrow V_{S4} - V_{S3} = V_{T3} - V_{T4}$$

A questo punto si trova che $V_{S4}=V_{S3}$ ovvero $V_{T4}=V_{T3}$. Infatti, siccome la tensione di soglia cresce con la tensione di source, (essendo il body a potenziale fissato), se fosse $V_{S4}>V_{S3}$ dovrebbe anche essere $V_{T4}>V_{T3}$, in contrasto con l'equazione precedente. Analogamente si arriva al fatto che $V_{S4}<V_{S3}$ è assurdo e pertanto deve essere $V_{S3}=V_{S4}$.

Ma $V_{S3}=V_{D1}$ e $V_{S4}=V_{D2}$, **pertanto l'assunto $V_{DS1}=V_{DS2}$ è dimostrato**. Naturalmente ciò è vero se è rispettata la relazione sui β espressa dall'equazione (3.2.3). Con le condizioni di progetto espresse dall'equazione (3.2.3) risulta quindi:

$$V_{DS2} = V_{DS1} = V_{GS1} \quad (3.2.4)$$

Resistenza di uscita. Il calcolo della resistenza di uscita si può affrontare semplicemente osservando che per valutarla entriamo con il generatore di prova sul drain di M4 il quale ha il gate a massa (alle variazioni) e ha sul source una resistenza equivalente data dalla r_{d2} . Pertanto la resistenza di uscita sarà data da:

$$R_{out} = r_{d2} + r_{d4}(1 + g_{m4}r_{d2}) \quad (3.2.5)$$

Di solito M4 è identico a M2 e, avendo anche la stessa I_D , avranno stessi parametri differenziali, ovvero $g_{m2}=g_{m4}=g_m$, $r_{d2}=r_{d4}=r_d$. Pertanto la resistenza differenziale risulta:

$$R_{out} = r_d(2 + g_m r_d) \quad (3.2.6)$$

e, ponendo $\lambda_4=\lambda_2=\lambda$, la tensione di Thevenin risulta:

$$V_{Th} = \lambda^{-1}(2 + g_m r_d)$$

Dinamica di uscita. Il limite per la tensione V_{out} è dato dal punto in cui M4 esce dalla saturazione ed entra in zona triodo. In condizioni di funzionamento corretto (M2 e M4 in saturazione) la R_{out} è molto elevata e possiamo considerare trascurabili le variazioni della I_{out} . Per cui:

$$\Delta I_{out} \cong 0 = g_{m4} \Delta V_{GS4} + \frac{1}{r_{d4}} \Delta V_{DS4} \Rightarrow \Delta V_{GS4} = -\frac{\Delta V_{DS4}}{g_{m4} r_{d4}}$$

Ma, poiché $V_{DS2}=V_{G3}-V_{GS4}$ e V_{G3} è costante:

$$\Delta V_{DS2} = -\Delta V_{GS4} = \frac{\Delta V_{DS4}}{g_{m4} r_{d4}}$$

Quindi, finché M4 è in zona di saturazione e quindi il prodotto $g_{m4}r_{d4}$ è molto maggiore di 1, la V_{DS2} varia molto meno della V_{DS4} e le variazioni della $V_{out}=V_{DS2}+V_{DS4}$ coincidono con le variazioni della V_{DS4} . Per cui, diminuendo progressivamente la V_{out} , si arriverà ad una condizione in cui la V_{DS4} è al limite della zona triodo, ovvero $V_{DS4}=V_{GS4}-V_{T4}$. Per quanto detto, la V_{DS2} sarà variata poco rispetto al valore V_{GS2} dato dalla (3.2.4). Pertanto la minima tensione V_{out} sarà data da:

$$V_{MIN} = V_{GS2} + (V_{GS4} - V_{T4}) = V_T + (V_{GS2} - V_T) + (V_{GS4} - V_{T4})$$

Una situazione tipica è quella in cui si fa $\beta_2=\beta_4$ (e quindi, per soddisfare l'equazione (3.2.3) anche $\beta_1=\beta_3$). In queste condizioni risulta:

$$V_{GS2} - V_{T2} = V_{GS4} - V_{T4} = \sqrt{\frac{2I_{out}}{\beta_2}}$$

La V_{min} risulta allora:

$$V_{MIN} = V_T + 2(V_{GS} - V_T) = V_T + 2\sqrt{\frac{2I_{out}}{\beta_2}} \quad (3.2.7)$$

In termini di dinamica di uscita lo specchio cascode risulta assai penalizzato rispetto allo specchio semplice in quanto la V_{MIN} contiene un termine V_T che, a differenza di $V_{GS}-V_T$ non può essere ridotto con un opportuno dimensionamento.

Tensione di ingresso. La tensione di ingresso è semplicemente data da:

$$V_{in} = V_{GS1} + V_{GS3}$$

Questa tensione può essere considerevole, considerando che le V_{GS} contengono entrambe un termine V_T e che la V_{T3} è aumentata dall'effetto body.

Resistenza di ingresso: è pari alla somma delle resistenze differenziali di M1 e M3 nel montaggio a diodo. Pertanto:

$$R_{in} = \frac{1}{g_{m1}} \parallel r_{d1} + \frac{1}{g_{m3}} \parallel r_{d3}$$

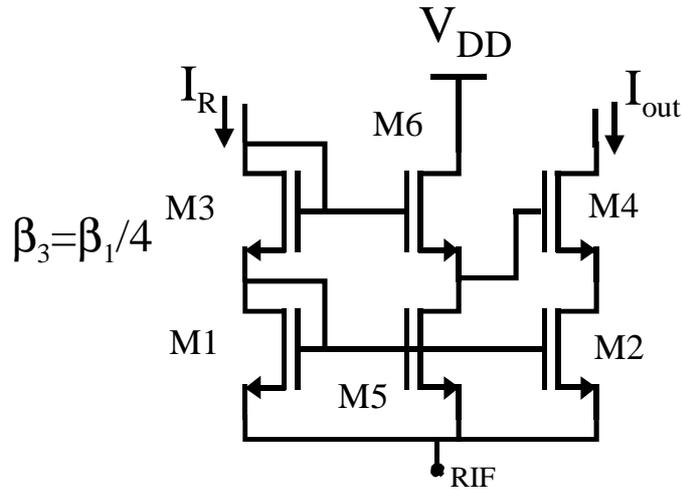
Se M1 e M3 sono identici, avendo essi anche la stessa corrente si trova semplicemente che:

$$R_{in} = 2 \left(\frac{1}{g_{m1}} \parallel r_{d1} \right) \cong \frac{2}{g_{m1}}$$

Specchio cascode a larga dinamica.

Lo specchio cascode a larga dinamica mostrato nella Figura 3.2.3 consente di ottenere una dinamica di uscita che si estende in basso fino ad una V_{MIN} pari a $2(V_{GS} - V_T)$, riducibile quindi a poche centinaia di millivolt. Questo schema trae spunto dal fatto che nel cascode classico la V_{DS2} è fissata a V_{GS} , ovvero ad un potenziale molto maggiore di quello minimo per la saturazione ($V_{GS} - V_T$). Dal punto di vista della dinamica di uscita ciò è uno svantaggio in quanto la V_{MIN} è proprio data dalla somma della V_{DS2} (che rimane costante) e della V_{DS4SAT} . Il cascode a larga dinamica fissa invece la V_{DS2} al limite della zona triodo, ovvero a $V_{GS} - V_T$. Questo valore della V_{DS2} sommato al valore limite per la V_{DS4} (ovvero V_{DS4SAT}) dà la V_{MIN} che

risulta quindi proprio pari a $2(V_{GS} - V_T)$. Ovviamente esiste un'analogia configurazione a p-MOS che realizza la versione complementare.



Specchio cascode a larga dinamica a n-MOSFET

Figura 3.2.3

Dimensionamento. Il circuito viene dimensionato nel seguente modo:

- $M5=M6$
- $M2=M4$
- $\beta_3 = \frac{\beta_1}{4}$

Queste condizioni sono fondamentali per il funzionamento del circuito.

L'analisi del circuito verrà svolta ipotizzando che l'effetto body sia trascurabile.

Al solito il rapporto tra le correnti viene fissato dai transistori in basso. Pertanto si ha:

$$\frac{I_{out}}{\beta_2} = \frac{I_R}{\beta_1} = \frac{I_{D5}}{\beta_5} \quad (3.2.8)$$

La V_{DS2} risulta data da:

$$V_{DS2} = V_{GS1} + V_{GS3} - V_{GS6} - V_{GS4}$$

Ovvero, esplicitando le V_{GS} , si ottiene:

$$V_{DS2} = \left(\sqrt{\frac{2I_R}{\beta_1}} + \sqrt{\frac{2I_R}{\beta_3}} - \sqrt{\frac{2I_{D6}}{\beta_6}} - \sqrt{\frac{2I_{out}}{\beta_4}} \right) + (V_{T1} + V_{T3} - V_{T6} - V_{T4})$$

Avendo considerato trascurabile l'effetto body, tutte le tensioni di soglia sono uguali e pertanto si annullano nell'espressione della V_{DS2} . Utilizzando l'equazione (3.2.8) e ricordando che $\beta_4=\beta_2$, $\beta_6=\beta_5$ e $\beta_3=\beta_1/4$ si ottiene infine:

$$V_{DS2} = \left(\sqrt{\frac{2I_R}{\beta_1}} + 2\sqrt{\frac{2I_R}{\beta_1}} - \sqrt{\frac{2I_R}{\beta_1}} - \sqrt{\frac{2I_R}{\beta_1}} \right) = \sqrt{\frac{2I_R}{\beta_1}} = V_{GS1} - V_T$$

Pertanto la V_{DS2} è in questo caso fissata al valore $V_{GS}-V_T$ come ipotizzato. La presenza dell'effetto body modifica un po' questo risultato poiché l'annullamento delle V_T non è più perfetto. In ogni caso si può ritoccare il parametro β_3 con l'ausilio di un simulatore e compensare questo effetto.

Resistenza di uscita. Si ripete esattamente quanto detto per lo specchio cascode convenzionale ottenendo lo stesso risultato.

Tensione di ingresso. Anche qui è la somma di due V_{GS} e vale solo la pena sottolineare il fatto che la V_{GS3} in questo schema è più grande della V_{GS1} non solo per l'effetto body ma anche per il beta ridotto.

Resistenza di ingresso. È la somma delle resistenze differenziali di M1 e M3 connessi a diodo. A causa del beta minore, la resistenza differenziale di M3 sarà doppia rispetto a quella di M1.

Svantaggi. Nello specchio cascode a larga dinamica la V_{DS2} è diversa dalla V_{DS1} . Come abbiamo visto ciò consente una maggiore dinamica ma fa sì che l'equazione (3.2.1) sia soddisfatta con minor precisione che per lo specchio cascode. In particolare se si fa $\beta_1=\beta_2$ per avere nominalmente $I_{out}=I_R$, questo risultato sarà più preciso nello specchio cascode convenzionale che nel cascode a larga dinamica.

3.3 Specchi di corrente a transistori bipolari.

Gli specchi a transistori bipolari sono caratterizzati dalla complicazione introdotta dalla corrente di base dei BJT. Questa deve essere ricavata dalle correnti I_R e/o I_{out} e pertanto costituisce un elemento di errore nel rapporto $k_S=I_R/I_{out}$. Nell'analisi che seguirà l'errore introdotto dalla I_B verrà espresso, come ordine di grandezza, in funzione di $1/\beta$. Verranno esposti esclusivamente schemi di tipo npn (current sink): di tutti gli schemi esposti esiste la versione complementare a transistori pnp (current source). Attenzione: in tutta la sezione relativa ai circuiti a transistori bipolari, con la notazione V_T si indicherà la quantità kT/q .

Specchio di corrente semplice.

Per questo specchio, mostrato in Figura 3.3.1 si può scrivere l'equazione corrispondente alla (3.2.1) degli specchi a MOSFET. Ovviamente $V_{BE1}=V_{BE2}$.

$$\frac{I_{C2}}{I_{C1}} = \frac{I_{S2} \left(e^{\frac{V_{BE2}}{V_T}} - 1 \right)}{I_{S1} \left(e^{\frac{V_{BE1}}{V_T}} - 1 \right)} = \frac{I_{S2}}{I_{S1}} = \frac{A_{E2}}{A_{E1}} = k_S \quad (3.3.1)$$

Questa equazione presuppone che si possa scrivere la corrente di saturazione come il prodotto di una densità di corrente di saturazione J_S , uguale per tutti i transistori, per l'area di emettitore A_E dei transistori stessi. Inoltre si è trascurato l'effetto delle V_{CE} (effetto Early).

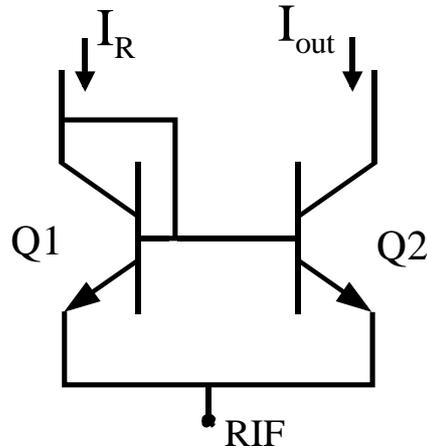


Figura 3.3.1

Questo effetto si preferisce introdurlo con la resistenza di uscita. L'equazione (3.3.1) fornisce il rapporto tra le I_C . Quello che ci interessa è però il rapporto tra I_R e I_{out} . Abbiamo:

$$I_{out} = I_{C2}$$

$$I_R = I_{C1} + I_{B1} + I_{B2} = I_{C1} + \frac{I_{C1}}{\beta} + \frac{I_{C2}}{\beta}$$

Unendo la (3.3.1) alle relazioni precedenti si ottiene:

$$\frac{I_{out}}{I_R} = k_S \frac{1}{1 + \frac{1+k_S}{\beta}} \cong k_S \left(1 - \frac{1+k_S}{\beta} \right) \quad (3.3.2)$$

L'errore dovuto alla corrente di base è quindi dell'ordine di $1/\beta$ ($2/\beta$ se si imposta un rapporto unitario tra I_{out} e I_R).

Ripetendo ragionamenti analoghi a quelli applicati allo specchio semplice a MOSFET si ottiene per lo specchio semplice a BJT:

Resistenza di uscita:

$$R_{out} = r_o = \frac{V_A}{I_{out}}$$

la corrispondente tensione di Thevenin è quindi pari alla tensione di Early, V_A .

Dinamica di uscita: la V_{MIN} coincide con la V_{CESAT} , di Q2 sotto la quale la corrente I_{out} e la resistenza di uscita diminuiscono rapidamente. La V_{MIN} è quindi dell'ordine di 100 mV.

Tensione di ingresso: è pari a V_{BE1} , ovvero dell'ordine di V_γ .

Resistenza di ingresso:

$$R_{in} = r_{o1} \left\| \frac{1}{g_{m1}} \right\| \cong \frac{1}{g_{m1}}$$

Specchi di corrente con resistenze di degenerazione.

Per aumentare la resistenza di uscita si possono adottare varie soluzioni. Un metodo semplice consiste nell'introdurre resistori di degenerazione di emettitore come negli schemi seguenti.

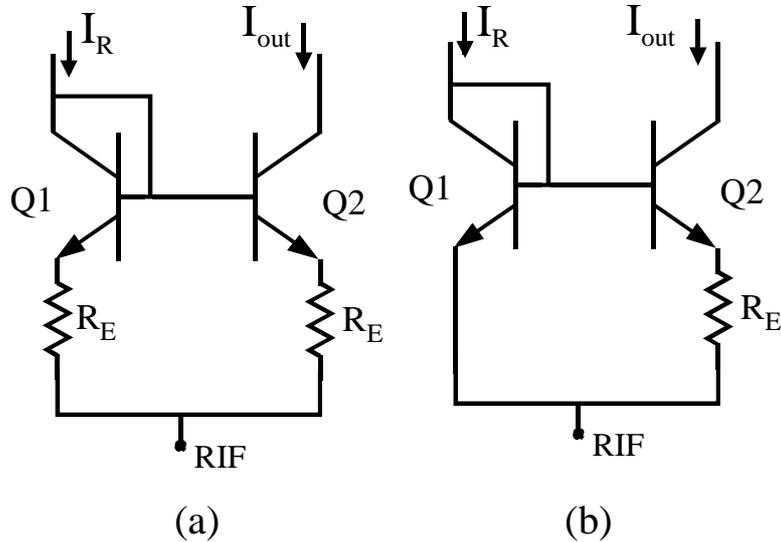


Figura 3.3.2

Lo schema di Figura 3.3.2(a) funziona effettivamente da specchio di corrente se Q1 e Q2 sono uguali e in tal caso I_R è circa uguale a I_{out} (a meno del solito errore dell'ordine di $1/\beta$ dovuto alla corrente di base). La resistenza di uscita, per $R_E \ll h_{ie2}$, è pari a:

$$R_{out} \cong r_{o2}(1 + g_{m2}R_E)$$

Se sul ramo di ingresso si omette la resistenza di degenerazione si ottiene lo schema di Figura 3.3.2(b) ovvero un "demagnificatore di corrente", denominato *generatore di corrente di Widlar*. Considerando il caso $Q1=Q2$ e esprimendo la tensione su R_E come $V_{BE1} - V_{BE2}$ si dimostra facilmente che vale la relazione:

$$\frac{I_{out}}{I_R} = e^{-\frac{R_E I_{out}}{V_T}}$$

Pertanto il rapporto tra I_{out} e I_R è minore di 1 e non è costante ma diminuisce all'aumentare di I_{out} (ovvero di I_R). Questo circuito è utilizzato tipicamente per produrre una corrente piccola (I_{out}) a partire da una corrente assai più grande (I_R). In queste condizioni l'unica incognita è la R_E che si calcola invertendo l'equazione precedente:

$$R_E = \frac{V_T}{I_{out}} \ln\left(\frac{I_R}{I_{out}}\right)$$

Naturalmente anche in questo caso il rapporto I_{out}/I_R calcolato con le equazioni precedenti è affetto da un errore dell'ordine di $1/\beta$ a causa delle correnti di base.

Specchio di corrente cascode a BJT.

Anche per gli specchi a BJT si può utilizzare la configurazione cascode per aumentare la resistenza di uscita e rendere quindi la corrente di uscita meno dipendente dalla tensione di uscita. Lo schema di uno specchio cascode a BJT npn è mostrato nella figura seguente:

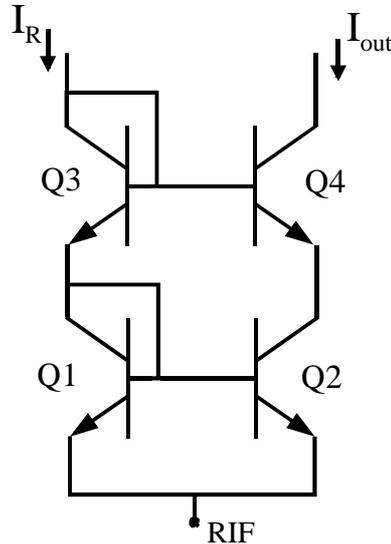


Figura 3.3.3

Funzionamento: anche qui sono i transistori in basso a stabilire il rapporto tra la corrente di uscita e quella di ingresso. In particolare vale la relazione (3.3.1). Come nello specchio cascode a MOSFET, i transistori Q3 e Q4 garantiscono che $V_{CE1} = V_{CE2}$. Perché ciò accada con esattezza occorre che il dimensionamento avvenga nel seguente modo:

$$\frac{A_{E4}}{A_{E3}} = \frac{A_{E2}}{A_{E1}} = k_S \cong \frac{I_{out}}{I_R}$$

Per quanto già affermato questo implica anche che siano valide le seguenti relazioni tra le correnti di saturazioni dei rispettivi transistori:

$$\frac{I_{S4}}{I_{S3}} = \frac{I_{S2}}{I_{S1}} \cong \frac{I_{out}}{I_R} \Rightarrow \frac{I_{out}}{I_{S4}} = \frac{I_R}{I_{S3}}$$

Da ciò deriva:

$$V_{BE3} \cong V_T \ln\left(\frac{I_R}{I_{S3}}\right) \cong V_T \ln\left(\frac{I_{out}}{I_{S4}}\right) \cong V_{BE4}$$

A questo punto basta osservare che:

$$V_{CE2} = V_{BE1} + V_{BE3} - V_{BE4} \cong V_{BE1} = V_{CE1}$$

Pertanto, essendo uguali le V_{CE} di Q1 e di Q2, e pure uguali le loro V_{BE} , la relazione tra le loro I_C seguirà l'equazione (3.3.1) con precisione. Purtroppo il rapporto tra le I_C determinabile con precisione da progetto) non coincide con il rapporto tra I_{out} e I_R a causa delle correnti di base. Si verifica facilmente che:

$$I_R = I_{C1} + I_{B1} + I_{B2} + I_{B4}$$

$$I_{out} = I_{C4} = I_{E4} - I_{B4} = I_{C2} - I_{B4}$$

Considerando che tutti i transistori hanno lo stesso beta e che I_{C4} è circa uguale a I_{C2} si ottiene:

$$I_{out} \cong I_{C2} \left(1 - \frac{1}{\beta} \right)$$

$$I_R \cong I_{C1} \left(1 + \frac{1}{\beta} + 2k_s \frac{1}{\beta} \right)$$

Per cui:

$$\frac{I_{out}}{I_R} \cong k_s \frac{\left(1 - \frac{1}{\beta} \right)}{\left[1 + \frac{1}{\beta} (1 + 2k_s) \right]} \cong k_s \left[1 - \frac{2}{\beta} (1 + k_s) \right]$$

Quindi dal punto di vista dell'errore introdotto dalle correnti di base anche lo specchio cascode si comporta come lo specchio semplice, essendo l'errore relativo dell'ordine di $1/\beta$. In particolare, se $k_s=1$ l'errore relativo è circa $4/\beta$.

Resistenza di uscita: Il calcolo della resistenza di uscita non è così immediato come nel caso del cascode a MOSFET. Infatti le correnti di base, assenti nei MOSFET, giocano qui un ruolo importante. La resistenza d'uscita del cascode a BJT risulta pari a:

$$R_{out} \cong r_{o4} \left(1 + \frac{h_{fe4}}{1 + k_s} \right)$$

Nel caso di $k_s=1$ si ha:

$$R_{out} = r_{o4} \left(1 + \frac{h_{fe4}}{2} \right)$$

Dinamica di uscita. Con argomenti analoghi a quelli usati per il cascode a MOSFET si dimostra che la V_{MIN} è data da:

$$V_{MIN} = V_{CE4SAT} + V_{BE1} = V_{CESAT} + V_\gamma$$

Quindi si passa dai circa 100 mV dello specchio semplice ai 0.8-0.9 V del cascode il quale ha quindi una dinamica di uscita più ridotta.

Tensione e resistenza di ingresso: si dimostra facilmente che:

$$V_{in} = V_{BE1} + V_{BE3} = 2V_{\gamma} \quad R_{in} \cong 2 \left(r_{o1} \parallel \frac{1}{g_{m1}} \right)$$

Specchi di corrente con amplificazione della corrente di base.

Questi specchi di corrente sono ottenuti dagli schemi visti precedentemente aggiungendo una circuiteria accessoria (uno o al massimo due transistori aggiuntivi) che ha lo scopo di fornire le correnti di base che alimentano i transistori veri e propri dello specchio. Lo specchio semplice con amplificazione di corrente è mostrato nella figura seguente:

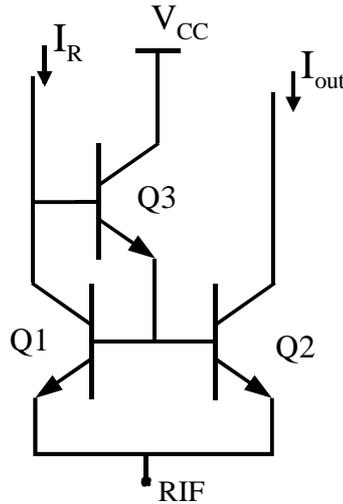


Figura 3.3.4

Per questo specchio vale sempre l'equazione (3.3.1) che fornisce il rapporto tra I_{C1} e I_{C2} . La resistenza di uscita e la dinamica di uscita sono le stesse dello specchio semplice. Un piccolo svantaggio è costituito dalla maggiore V_{IN} che qui è pari chiaramente a $2V_{\gamma}$. Quello che cambia sostanzialmente è il minor effetto delle correnti di base sul rapporto I_{out}/I_R . Si ha infatti:

$$I_R = I_{C1} + I_{B3} = I_{C1} + \frac{I_{B1} + I_{B2}}{\beta_3 + 1} \cong I_{C1} \left[1 + \frac{1 + k_S}{\beta(\beta_3 + 1)} \right]$$

$$I_{out} = I_{C2}$$

Pertanto risulta:

$$\frac{I_{out}}{I_R} = k_S \cdot \frac{1}{1 + \frac{1 + k_S}{\beta(\beta_3 + 1)}} \cong k_S \left[1 - \frac{1 + k_S}{\beta(\beta_3 + 1)} \right]$$

Si può verificare che l'errore relativo introdotto dalle correnti di base è dell'ordine di $1/\beta^2$ (considerando genericamente uguali tutti i beta). In questo modo si possono ottenere rapporti I_{out}/I_R che differiscono dal rapporto delle aree per meno dello 0.1%. Questo vantaggio può però essere completamente annullato dal fatto che, a causa della resistenza di uscita relativamente bassa, la corrente I_{out} risente molto della V_{out} e quindi questa precisione si ha in pratica solo se la V_{out} varia poco e, anzi, rimane vicina al valore V_{BE1} .

Se ciò non si verifica e richiediamo ugualmente una precisione elevata si ricorre alla soluzione cascode con amplificazione di corrente mostrato nella figura seguente

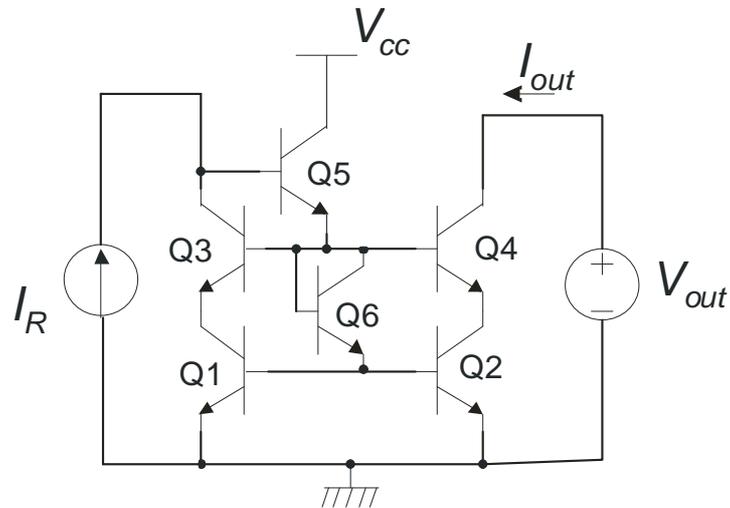


Figura 3.3.5- Specchio di corrente cascode a BJT con amplificazione della corrente di base.

In questo specchio Q5 fornisce direttamente la corrente di base a Q3 e Q4. Sempre Q5 fornisce attraverso Q6 (equivalente a un diodo) la corrente di base a Q1 e Q2.

Calcolo del contributo delle correnti di base all'errore sul valore del guadagno dello specchio (I_{out}/I_R) rispetto al valore nominale $area_2/area_1$.

In condizioni nominali, sappiamo che, con ottima approssimazione:

$$\frac{I_{C2}}{I_{C1}} = \frac{area_2}{area_1}$$

Per quanto riguarda I_R , possiamo scrivere:

$$I_R = I_{C3} + I_{B5} = I_{E3} - I_{B3} + I_{B5}$$

Quindi:

$$I_R = I_{C1} - I_{B3} + I_{B5}$$

Per quanto riguarda I_{out} :

$$I_{C2} = I_{E3} = I_{C4} + I_{B4}$$

Quindi, siccome $I_{C4} = I_{out}$:

$$I_{out} = I_{C2} - I_{B4}$$

Pertanto:

$$\frac{I_{out}}{I_R} = \frac{I_{C2} - I_{B4}}{I_{C1} - I_{B3} + I_{B5}} = \frac{I_{C2} - \frac{I_{C4}}{\beta_4}}{I_{C1} - \frac{I_{C3}}{\beta_3} + I_{B5}}$$

Introducendo le seguenti approssimazioni, valide solo al fine del calcolo delle correnti di base:

$$I_{C4} \cong I_{C2} ; I_{C3} \cong I_{C1} \quad (3.3.3)$$

si ottiene:

$$\frac{I_{out}}{I_R} = \frac{I_{C2}}{I_{C1}} \frac{1 - \frac{1}{\beta_4}}{1 - \frac{1}{\beta_3} + \frac{I_{B5}}{I_{C1}}} = k_S \frac{1 - \frac{1}{\beta_4}}{1 - \frac{1}{\beta_3} + \frac{I_{B5}}{I_{C1}}} \quad (3.3.4)$$

Ma:

$$I_{B5} = \frac{I_{E5}}{\beta_5 + 1} = \frac{I_{B1} + I_{B2} + I_{B3} + I_{B4}}{\beta_5 + 1} = \frac{\frac{I_{C1}}{\beta_1} + \frac{I_{C2}}{\beta_2} + \frac{I_{C3}}{\beta_3} + \frac{I_{C4}}{\beta_4}}{\beta_5 + 1} \quad (3.3.5)$$

Quindi, considerando le approssimazioni (3.3.3) e supponendo che $\beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta$:

$$\frac{I_{B5}}{I_{C1}} \cong 2 \frac{1 + k_S}{(\beta_5 + 1)\beta} \quad (3.3.6)$$

Il termine I_{B5}/I_{C1} è quindi dell'ordine di $1/\beta^2$. Applicando l'approssimazione di Taylor al denominatore di (3.3.4), si ottiene.

$$\frac{I_{out}}{I_R} \cong k_S \left(1 - \frac{1}{\beta_4}\right) \left(1 + \frac{1}{\beta_3} - \frac{I_{B5}}{I_{C1}}\right)$$

Trascurando i termini di ordine $1/\beta^3$:

$$\frac{I_{out}}{I_R} \cong k_S \left(1 - \frac{1}{\beta_4} + \frac{1}{\beta_3} - \frac{1}{\beta_3\beta_4} - \frac{I_{B5}}{I_{C1}}\right) = k_S \left(1 + \frac{\beta_3 - \beta_4}{\beta_3\beta_4} - \frac{1}{\beta_3\beta_4} - \frac{I_{B5}}{I_{C1}}\right)$$

Con un buon matching tra β_3 e β_4 , la differenza $\beta_3 - \beta_4$ si può ridurre a poche unità e pertanto l'errore relativo sul guadagno dello specchio:

$$e_r = \left(\frac{\beta_3 - \beta_4}{\beta_3\beta_4} - \frac{1}{\beta_3\beta_4} - \frac{I_{B5}}{I_{C1}}\right),$$

risulta composto da soli termini dell'ordine di $1/\beta^2$.

Specchio di corrente di Wilson.

Lo specchio di corrente di Wilson costituisce un'alternativa allo specchio cascode quando si vuole unire una elevata resistenza di uscita (pari a quella dello specchio cascode) alla precisione tipica degli specchi con amplificazione di corrente. È quindi equivalente allo specchio cascode con amplificazione di corrente. Esso presenta tuttavia il vantaggio di utilizzare solo 4 BJT e di richiedere una minor tensione d'ingresso ma anche lo svantaggio di lavorare correttamente soltanto con rapporti di corrente unitari. Lo schema di questo specchio è mostrato nella Figura 3.3.6.

Funzionamento. Anche in questo specchio il rapporto tra le correnti è governato dalla coppia Q1-Q2. Notiamo però che in questo caso è Q2 ad essere collegato a diodo.

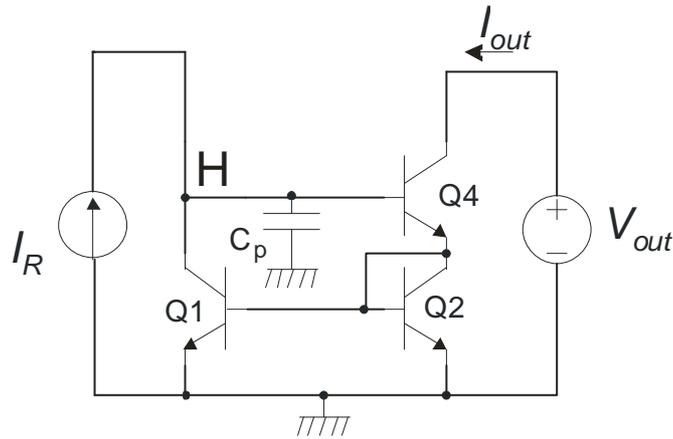


Figura 3.3.6

Questo specchio può essere scomposto in un blocco di azione (Q4) e un blocco di reazione costituito dallo specchio semplice Q2-Q1, connesso in modo da riportare la corrente di uscita verso l'ingresso. Il blocco di azione viene comandato dalla tensione sul nodo H (V_H) in modo tale da fare sì che la corrente assorbita da Q1 sia, a meno della corrente di base di Q4, uguale alla corrente in ingresso (I_R). Per capire come ciò avvenga, possiamo considerare il transitorio di accensione. Questo, illustrato nella Figura 3.3.7, prevede che la I_R passi da 0 al valore finale nell'istante t_1 .

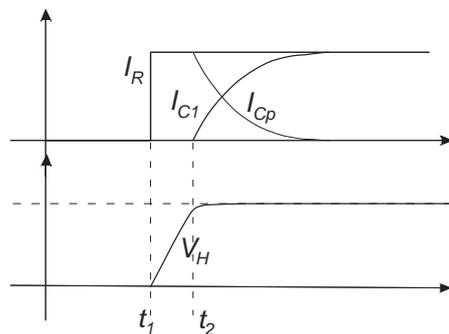


Figura 3.3.7

Transitorio di accensione dello specchio di Wilson.

Per quanto riguarda la V_{out} , si supponga che sia già al valore richiesto (superiore alla V_{min} , da determinare in seguito). Prima di t_1 , la corrente di base di Q4 è nulla, pertanto Q4 è interdetto e $I_{E4}=0$. Quindi nessuna corrente entra nello specchio semplice Q2-Q1, ovvero $I_{C1}=0$. Questa situazione si protrae anche negli istanti immediatamente successivi a t_1 , in quanto le capacità parassite manterranno i transistori in interdizione. Focalizziamoci sulla capacità parassita indicata in figura con C_p ed indichiamo con I_{Cp} la corrente che entra in essa. Semplificando, la C_p si può considerare rappresentativa di tutte le capacità parassite che hanno un terminale in H. Possiamo scrivere:

$$\frac{dV_H}{dt} = \frac{I_{Cp}}{C_p} \quad \text{con: } I_{Cp} = I_R - I_{C1} - I_{B4} \quad (3.3.7)$$

Negli istanti immediatamente successivi a t_1 , la I_{cp} coincide con la I_R , visto che $I_{B4}=I_{C1}=0$. Questa situazione rimane invariata fino all'istante t_2 , in cui la V_H raggiunge una tensione sufficiente ad accendere le giunzioni V_{BE2} e V_{BE4} . Questa tensione sarà di poco inferiore alla V_γ dei BJT interessati, considerando, in modo qualitativo, che la V_γ è la tensione in cui i transistori portano la corrente finale. Non appena Q4 si accende, la sua corrente aumenta esponenzialmente con la tensione V_H . La corrente di Q4 viene riflessa all'indietro dallo specchio Q2-Q1, per cui, trascurando le correnti di base:

$$I_{C1} \cong \frac{area_1}{area_2} I_{E4} \cong \frac{1}{k_S} I_{out} \quad \text{dove: } k_S \cong \frac{area_2}{area_1} \quad (3.3.8)$$

Considerando la (3.3.7) e trascurando la I_{B4} , otteniamo che la tensione V_H si stabilizzerà (derivata nulla) quando I_{C1} raggiungerà il valore di I_R . Quindi, considerando la (3.3.8) con la situazione di regime $I_{C1}=I_R$:

$$I_{out} \cong k_S I_R$$

Si può facilmente dimostrare che la reazione è negativa supponendo che la I_{out} subisca un aumento (per esempio per un disturbo). Questa variazione si riflette all'indietro (ovvero su I_{C1}), e pertanto non ci sarà più equilibrio al nodo H. Una corrente negativa scorrerà in C_p innescando una diminuzione della tensione V_H . Questa variazione ridurrà la corrente in Q4, opponendosi all'iniziale aumento. Ovviamente questo è vero in continua e a basse frequenze, dove i transistori si possono considerare "resistivi" (facendo eccezione per la sola C_p). Per garantire la stabilità di questo anello di reazione occorre prendere in considerazione tutta la risposta in frequenza del guadagno di anello ed applicare, se necessario, tecniche di compensazione. La tecnica più comune è aumentare la capacità C_p mediante l'inserimento di una capacità aggiuntiva tra H e massa. L'analisi dei problemi di stabilità esula da questa trattazione.

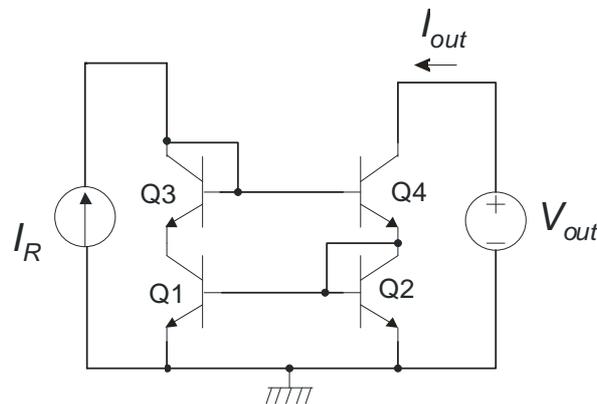


Figura 3.3.8

Specchio di corrente di Wilson a quattro transistori.

Un problema che si incontra con lo specchio di Wilson a tre transistori è il fatto che Q2 e Q1 lavorano con tensioni V_{CE} diverse. Infatti, $V_{CE2}=V_{BE2}$. $V_{CE1}=V_{BE2}+V_{BE4}$. Per ovviare a questo problema, che introduce un errore sistematico sul k_S , si ricorre allo specchio di Wilson a quattro transistori, mostrato in Figura 3.3.8. Il principio di funzionamento è perfettamente analogo, con la sola differenza che Q3 produce una traslazione di una V_{BE} rispetto a V_H , per cui

adesso: $V_{CE1}=V_H-V_{BE3}=V_{BE2}+V_{BE4}-V_{BE3}$. Con un dimensionamento analogo a quello degli specchi cascode, ovvero:

$$\frac{area_4}{area_3} = k_S$$

Si ottiene che $V_{BE3}=V_{BE4}$, per cui $V_{CE1}=V_{BE2}=V_{CE2}$, come richiesto.

Calcolo del contributo delle correnti di base all'errore sul valore del guadagno dello specchio (I_{out}/I_R) rispetto al valore nominale $area_2/area_1$.

Dal punto di vista dell'effetto delle correnti di base, le versioni a tre e quattro transistori sono completamente equivalenti, in quanto Q3 è un bipolo (diodo). La I_R risulta legata alla I_{C1} da:

$$I_R = I_{C1} + I_{B4}$$

La I_{out} , invece, risulta:

$$I_{out} = I_{C4} = I_{E4} - I_{B4} \quad \text{con} \quad I_{E4} = I_{C2} + I_{B2} + I_{B1}$$

Quindi, il rapporto I_{out}/I_R risulta:

$$\frac{I_{out}}{I_R} = \frac{I_{C2} + I_{B2} + I_{B1} - I_{B4}}{I_{C1} + I_{B4}}$$

Facendo le solite approssimazioni, valide solo ai fini dell'errore sulle correnti di base: $I_{C4} \cong I_{C2}$. Si ha:

$$\frac{I_{out}}{I_R} \cong \frac{I_{C2} + \frac{I_{C2}}{\beta_2} + \frac{I_{C1}}{\beta_1} - \frac{I_{C2}}{\beta_4}}{I_{C1} + \frac{I_{C2}}{\beta_4}} = \frac{I_{C2}}{I_{C1}} \frac{1 + \frac{1}{\beta_2} + \frac{1}{k_S \beta_1} - \frac{1}{\beta_4}}{1 + \frac{k_S}{\beta_4}}$$

Applicando la solita approssimazione di Taylor per portare il denominatore a numeratore otteniamo:

$$\frac{I_{out}}{I_R} \cong k_S \left[1 + \left(\frac{1}{\beta_2} - \frac{1}{\beta_4} \right) + \left(\frac{1}{k_S \beta_1} - \frac{k_S}{\beta_4} \right) + O\left(\frac{1}{\beta^2} \right) \right] \quad (3.3.9)$$

Dove $O(1/\beta^2)$ indica termini che vanno a zero almeno come $1/\beta^2$.

Il primo termine di errore nella parentesi quadra è dell'ordine di $1/\beta^2$, purché l'errore di matching tra β_2 e β_4 sia al più di qualche unità. Il secondo termine, invece potrà ridursi ad un'espressione simile a quella del primo, solo se $k_S=1$. Pertanto: il contributo delle correnti di base all'errore relativo sul guadagno dello specchio di Wilson è dell'ordine di $1/\beta^2$ solo per guadagni unitari.

Se si progetta lo specchio di Wilson con k_S diverso dall'unità, si perderanno i benefici in termini di precisione tipici di questa configurazione.

Pertanto, nello specchio di Wilson il rapporto tra I_{out} e I_R è effettivamente uguale al rapporto delle aree (a meno di un errore $1/\beta^2$) se e solo se i beta dei transistori si possono considerare uguali e il rapporto tra le aree è unitario. Più ci si discosta da queste condizioni più l'errore relativo cresce tendendo a diventare dell'ordine di $1/\beta$ come negli specchi normali (semplice o cascode) senza amplificazione di corrente di base.

Resistenza di uscita. Il calcolo della resistenza di uscita è piuttosto laborioso e, nel caso di rapporto unitario I_{out}/I_R , coincide con quella del cascode.

Dinamica di uscita. Coincide con quella del cascode (la condizione $V_{out}=V_{MIN}$ si ha quando Q4 arriva alla soglia della zona di saturazione).

3.4 Riepilogo delle caratteristiche degli specchi di corrente.

Specchi di corrente a transistori MOS

Topologia	Semplice	Cascode	Cascode a larga dinamica
R_{out}	r_d	$r_d(2+g_m r_d)$	$r_d(2+g_m r_d)$
$V_{Th}=R_{out}I_{out}$	λ^{-1}	$\lambda^{-1}(2+g_m r_d)$	$\lambda^{-1}(2+g_m r_d)$
I_{out}/I_R nominale	$\frac{\beta_2}{\beta_1} = \frac{W_2 L_1}{W_1 L_2}$	$\frac{\beta_2}{\beta_1} = \frac{W_2 L_1}{W_1 L_2}$	$\frac{\beta_2}{\beta_1} = \frac{W_2 L_1}{W_1 L_2}$
V_{MIN}	$V_{GS}-V_t$	$2(V_{GS}-V_t)+V_t$	$2(V_{GS}-V_t)$
V_{in} (appross.)	V_{GS}	$2V_{GS}$	$2V_{GS}$
Commenti			(1)

(1) Meno preciso dello specchio cascode classico.

Specchi di corrente a transistori Bipolari.

Topologia	Semplice	Cascode	Wilson	Semplice con amplificazione e della I_B	Cascode con amplificazione e della I_B	Widlar
R_{out}	$r_0 = \frac{1}{h_{oe}}$	$\cong \frac{h_{fe}}{2} r_0$	$\cong \frac{h_{fe}}{2} r_0$	$r_0 = \frac{1}{h_{oe}}$	$\cong \frac{h_{fe}}{2} r_0$	$\cong r_0(1 + g_m R_E)$
$V_{Th}=R_{out}I_{out}$	V_A	$\cong \frac{h_{fe}}{2} V_A$	$\cong \frac{h_{fe}}{2} V_A$	V_A	$\cong \frac{h_{fe}}{2} V_A$	$\cong V_A(1 + g_m R_E)$
I_{out}/I_R nominale	$\frac{A_{E2}}{A_{E1}}$	$\frac{A_{E2}}{A_{E1}}$	$\frac{A_{E2}}{A_{E1}}$	$\frac{A_{E2}}{A_{E1}} v$	$\frac{A_{E2}}{A_{E1}}$	$\exp\left(-\frac{I_{out} R_E}{V_T}\right)$
V_{MIN}	V_{CEsat}	$V_{CEsat} + V_\gamma$	$V_{CEsat} + V_\gamma$	V_{CEsat}	$V_{CEsat} + V_\gamma$	$V_{CEsat} + I_2 R_E$
effetto I_B	$\propto \frac{1}{\beta}$	$\propto \frac{1}{\beta}$	$\propto \frac{1}{\beta^2}$	$\propto \frac{1}{\beta^2}$	$\propto \frac{1}{\beta^2}$	$\propto \frac{1}{\beta}$
V_{in} (appross.)	V_γ	$2V_\gamma$	$2V_\gamma$	$2V_\gamma$	$3V_\gamma$	V_γ
Commenti			(2), (3)	(2)	(2), (4)	(5)

(2) Reazionato: può richiedere compensazione.

(3) Adatto solo a rapporti unitari.

(4) Richiede matching dei β per fornire la precisione indicata (errore dell'ordine di $1/\beta^2$).

(5) Il rapporto I_{out}/I_R non è costante ma diminuisce all'aumentare di I_R .

3.5 Effetto degli errori di matching sugli specchi di corrente.

Per semplicità di trattazione, in questo paragrafo si considereranno gli errori di matching per specchi di corrente aventi guadagno unitario. La figura 5 mostra uno specchio di corrente semplice a MOSFET ed una versione a BJT. I risultati ottenuti con gli specchi semplici si possono riportare direttamente al caso di specchi più complessi, quali gli specchi Cascode, in quanto il contributo agli errori di matching deriva principalmente dallo specchio semplice che essi contengono.

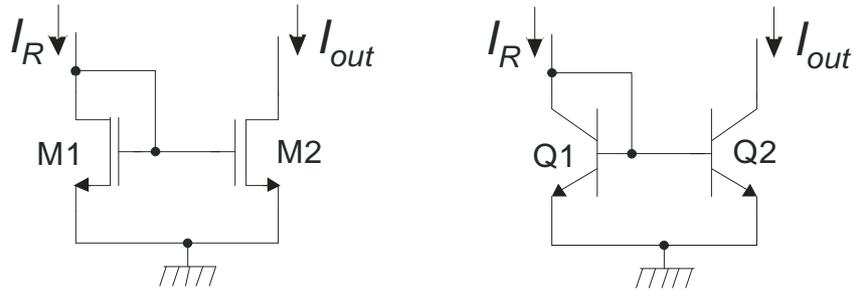


Figura 5

Specchio di corrente a MOSFET.

Siamo interessati all'errore relativo di corrente:

$$e_r = \frac{\Delta I}{I_R} \cong \frac{\Delta I}{\bar{I}}, \text{ dove: } \Delta I = I_R - I_{out}, \quad \bar{I} = (I_R + I_{out})/2$$

Considerando che M1 e M2 siano nominalmente identici e che si trovino nelle medesime condizioni di polarizzazione, possiamo considerare l'errore $I_R - I_{out}$ come l'errore di matching tra le I_D di M1 e M2.

Trascurando l'effetto della VDS, possiamo scrivere la corrente di un MOSFET in saturazione e forte inversione come:

$$I_D = \frac{\beta}{2} (V_{GS} - V_t)^2$$

Indicando con $B = (\beta/2)(V_{GS} - V_t)$, si può applicare l'espressione notevole dell'errore relativo, ottenendo:

$$\frac{\Delta I_D}{I_D} = \frac{\Delta \beta}{\beta} + 2 \frac{\Delta B}{B}$$

Siccome V_{GS} è la stessa per i due dispositivi (gate e source connessi assieme), allora, $\Delta B = -\Delta V_t$.

Pertanto:

$$\frac{\Delta I_D}{I_D} = \frac{\Delta \beta}{\beta} - 2 \frac{\Delta V_t}{(V_{GS} - V_t)} \quad (3.5.1)$$

Passando alle deviazioni standard:

$$\sigma_{\frac{\Delta I_D}{I_D}} = \sqrt{\sigma_{\frac{\Delta \beta}{\beta}}^2 + \left(\frac{2\sigma_{V_t}}{(V_{GS} - V_t)} \right)^2} \quad (3.5.2)$$

Esempio:

Consideriamo un caso reale, relativo ad un processo CMOS a 0.35 μm di lunghezza minima. Nel manuale di processo si trovano le seguenti informazioni:

$$\sigma_{\frac{\Delta \beta}{\beta}} = \frac{C_\beta}{\sqrt{WL}} \quad \text{con } C_\beta = 0.03 \mu\text{m}$$

$$\sigma_{V_t} = \frac{C_{V_t}}{\sqrt{WL}} \quad \text{con } C_{V_t} = 8.5 \text{ mV} \cdot \mu\text{m}$$

Ipotizziamo di realizzare uno specchio di corrente con M1 e M2 aventi dimensioni $L=W=1\text{mm}$ e

$V_{GS}-V_t$ pari a 100 mV. Otteniamo:

$$\sigma_{\frac{\Delta \beta}{\beta}} = 0.03, \quad \sigma_{V_t} = 8.5 \text{ mV} \quad \frac{2\sigma_{V_t}}{(V_{GS} - V_t)} = 0.17$$

Quindi:

$$\sigma_{\frac{\Delta I_D}{I_D}} = 0.173 \quad (= 17.3\%)$$

Otteniamo quindi un errore molto grande. Se prendiamo il dato a 3σ (ovvero quasi la totalità dei campioni), osserviamo che ci possiamo aspettare anche discrepanze del 50 % tra la corrente d'ingresso e quella d'uscita.

Osserviamo inoltre che il contributo degli errori sulla V_t è largamente prevalente. Il progettista ha la possibilità di mitigare questo errore aumentando la $V_{GS}-V_t$. Passando a $V_{GS}-V_t=500 \text{ mV}$, e mantenendo le stesse dimensioni per i MOSFET, avremmo:

$$\frac{2\sigma_{V_t}}{(V_{GS} - V_t)} = 0.034 \Rightarrow \sigma_{\frac{\Delta I_D}{I_D}} = 0.0453 \quad (4.53\%)$$

In questo caso i contributi degli errori sul beta e sulla V_t sarebbero paragonabili. Ovviamente, abbiamo visto che un aumento nella tensione di *overdrive* ($V_{GS}-V_t$) porta ad un peggioramento proporzionale nella V_{min} dello specchio. Aumentando entrambe le dimensioni dei MOS di un fattore 10 (quindi $L=W=10\mu\text{m}$), si può ulteriormente ridurre l'errore dello stesso fattore, quindi:

$$\left. \begin{array}{l} L = W = 10 \mu\text{m} \\ V_{GS} - V_t = 500 \text{ mV} \end{array} \right\} \Rightarrow \sigma_{\frac{\Delta \beta}{\beta}} = 0.003\sigma_{V_t} = 0.858 \text{ mV} \quad \frac{2\sigma_{V_t}}{(V_{GS} - V_t)} = 0.0034$$

Da cui: $\sigma_{\Delta I_D/I_D} = 0.0045$ (0.45 %)

Specchio di corrente a BJT

Ripetendo le stesse considerazioni, andremo a calcolare l'errore relativo sulle correnti di collettore. Sappiamo dai precedenti paragrafi che, anche se l'errore in questione fosse nullo, vi sarebbe comunque sovrapposto un errore sistematico dovuto alle correnti di base, il quale dipende dalla topologia prescelta.

Trascurando la dipendenza della corrente di collettore dalla tensione V_{CE} , possiamo fare riferimento all'espressione:

$$I_C = I_S e^{\frac{V_{BE}}{V_T}} \quad \text{con} \quad V_T = \frac{kT}{q}$$

Se consideriamo, inoltre che Q1 e Q2 hanno la stessa temperatura, allora essi hanno anche uguali fattori esponenziali (stessa V_{BE} e V_T). Pertanto:

$$\frac{\Delta I_C}{I_C} = \frac{\Delta I_S}{I_S} \quad (3.5.3)$$

La fonderia non fornisce di norma la deviazione standard di $\Delta I_S/I_S$, ma riferisce tutti gli errori di matching alla V_{BE} . In particolare, si fa riferimento a coppie di BJT nominalmente identici polarizzati in zona attiva diretta con delle V_{BE} tali da far sì che le correnti di collettore risultino uguali. Questo non è, ovviamente il caso dello specchio di corrente, dove le V_{BE} sono semmai forzate ad essere uguali e le differenze nelle I_S dovute agli errori di matching si traducono in differenze nelle I_C . Questo valore della deviazione standard ci permette però di risalire a quella della I_S . Se consideriamo la ΔV_{BE} di due BJT nominalmente identici (ma soggetti ad errori di matching), che si trovano alla medesima temperatura, troviamo

$$\Delta V_{BE} = V_T \Delta \left[\ln \left(\frac{I_C}{I_S} \right) \right] = V_T \left(\frac{\Delta I_C}{I_C} - \frac{\Delta I_S}{I_S} \right) \quad (3.5.4)$$

Il valore fornitoci dalla fonderia è relativo al caso di stesse correnti di collettore, ovvero $\Delta I_C=0$. Pertanto:

$$\Delta V_{BE}^* = V_T \left(-\frac{\Delta I_S}{I_S} \right) \Rightarrow \sigma_{\Delta V_{BE}^*} = V_T \sigma_{\frac{\Delta I_S}{I_S}} \quad (3.5.5)$$

Dalla 12, nota la $\sigma_{\Delta V_{BE}^*}$ fornita dalla fonderia, posso risalire alla $\sigma_{\frac{\Delta I_S}{I_S}}$.

Esempio.

Per un processo BCMOS la $\sigma_{V_{BE}^*}$ di un transistor elementare NPN sia pari a 0.3 mV. Considerando $V_T=26$ mV si ricava:

$$\sigma_{\frac{\Delta I_S}{I_S}} = 0.011$$

Pertanto, dalla (3.5.3) si ricava che la deviazione standard dell'errore di matching relativo per uno specchio di corrente realizzato con BJT standard risulta pari a 1.1 %. Questo errore si può ulteriormente utilizzando per Q1 e Q2 dispositivi ottenuti connettendo in parallelo più transistori elementari (ovvero aumentando l'area). L'effettiva riduzione nell'errore di matching è proporzionale alla radice del numero di transistori connessi in parallelo,

Capitolo 4. Amplificatori differenziali: concetti e strutture di base.

4.1 Definizioni generali riguardanti gli amplificatori differenziali.

Gli amplificatori differenziali si distinguono in base al numero di uscite nei due tipi seguenti:

- A singola uscita (single ended output)
- A doppia uscita (fully differential)

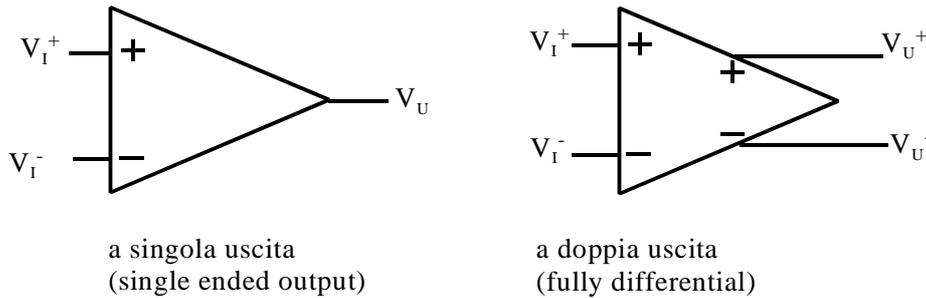


Figura 4.1.1

Tensioni di modo comune e di modo differenziale.

Richiamiamo le definizioni di tensioni di modo comune V_{ic} e differenziale V_{id} di ingresso.

$$V_{id} = V_I^+ - V_I^- \quad V_{ic} = \frac{V_I^+ + V_I^-}{2}$$

Per l'amplificatore a doppia uscita vengono definite anche una tensione di uscita di modo differenziale V_{ud} e di modo comune V_{uc}

$$V_{ud} = V_U^+ - V_U^- \quad V_{uc} = \frac{V_U^+ + V_U^-}{2}$$

In molti casi si omette il pedice "i" che indica le tensioni di ingresso che diventano semplicemente V_d e V_c .

Amplificazioni: Si definiscono 2 amplificazioni per l'amplificatore a singola uscita e 4 amplificazioni per il *fully differential*.

Nel caso di amplificatore a singola uscita si ha l'amplificazione differenziale (A_d) e di modo comune (A_c) che si definiscono come:

$$A_d = \left. \frac{V_U}{V_{id}} \right|_{V_{ic}=0} \quad A_c = \left. \frac{V_U}{V_{ic}} \right|_{V_{id}=0} \quad CMRR = \left| \frac{A_d}{A_c} \right|$$

Nel caso di amplificatore a doppia uscita possiamo definire le seguenti 4 amplificazioni e il CMRR:

$$A_{dd} = \frac{V_{ud}}{V_{id}} \quad A_{cc} = \frac{V_{uc}}{V_{ic}} \quad A_{cd} = \frac{V_{ud}}{V_{ic}} \quad A_{dc} = \frac{V_{uc}}{V_{id}} \quad CMRR = \left| \frac{A_{dd}}{A_{cd}} \right|$$

L'amplificazione A_{dd} fa le veci dell'amplificazione A_d per il single ended in quanto misura il rapporto tra i segnali utili in uscita e in ingresso (segnali differenziali). Le amplificazioni A_{cc} e A_{dc} rappresentano la risposta del segnale di modo comune di uscita rispetto al modo comune e differenziale in ingresso. Queste amplificazioni non sono particolarmente dannose in quanto si traducono in segnale di modo comune in uscita che di solito viene bloccato dagli stadi successivi o dall'utilizzatore. Si richiede comunque che il modo comune in uscita si mantenga sempre a livelli sufficientemente bassi da ridurre significativamente la dinamica dei due segnali di uscita. L'amplificazione A_{cd} , invece deve essere minimizzata in quanto rappresenta la risposta del segnale utile in uscita (differenziale) rispetto al segnale di disturbo in ingresso (modo comune).

Circuito equivalente di ingresso.

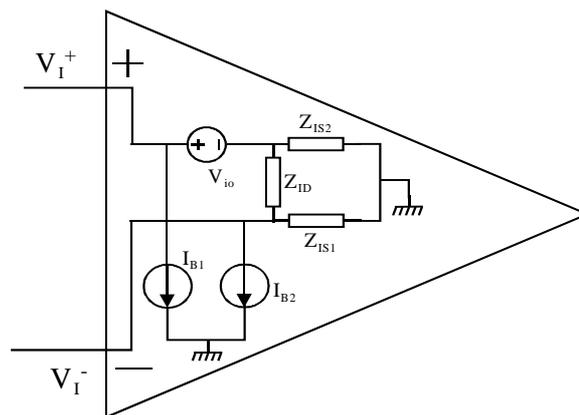


Figura 4.1.2

Si definiscono due impedenze di ingresso:

- Impedenza differenziale di ingresso (o impedenza di modo differenziale, Z_{ID} nella figura precedente)
- Impedenza di isolamento di ingresso (o impedenza di modo comune) pari a:

$$Z_{IS} = \frac{Z_{IS1} + Z_{IS2}}{2}.$$

In molti casi, in virtù della simmetria dello stadio di ingresso, si può assumere: $Z_{IS2} = Z_{IS1} = Z_{IS}$

Correnti di polarizzazione e tensione di offset.

La tensione di offset si definisce come la tensione di ingresso di modo differenziale in corrispondenza della quale la tensione di uscita è nulla.

La corrente di polarizzazione I_B si definisce come:
$$I_B = \frac{I_{B1} + I_{B2}}{2}.$$

La corrente di offset è la differenza $I_{io} = I_{B1} - I_{B2}$ che si ha quando in ingresso è applicata una tensione differenziale pari alla tensione di offset.

Importante: si noti che gli amplificatori vengono progettati in modo che le quantità V_{io} e I_{io} siano nulle. A causa della dispersione dei valori dei parametri elettrici dei dispositivi reali, la V_{io} e I_{io} non saranno nulle ma varieranno da circuito integrato a circuito integrato in modo casuale (errore casuale). Se il progetto è stato compiuto in modo corretto la media di tali quantità, calcolata su un grande numero di amplificatori sarà pari a zero. Lo scostamento rispetto al valor medio segue tipicamente una distribuzione gaussiana. Una misura della tipica tensione (o corrente) di offset che possiamo aspettarci è data dalla deviazione standard σ della distribuzione. Si sa che in un intervallo di $\pm 3\sigma$ attorno al valore nominale rientra il 99.7 % di tutti i valori effettivamente misurati.

Se neppure la media coincide con il valore di progetto (che, ripetiamo, nel caso di V_{io} e I_{io} deve essere nullo) allora è presente un errore sistematico.

Altri parametri importanti:

- Dinamica di ingresso di modo differenziale.
- Dinamica di ingresso di modo comune (CMR, Common Mode Range).
- Dinamica di uscita (output swing).
-

La dinamica di ingresso di modo differenziale è legata all'amplificazione e, spesso, al carico connesso sull'uscita. Un amplificatore con elevato guadagno avrà generalmente una piccola dinamica di ingresso differenziale, in quanto basta un piccolo valore di V_d per portare l'uscita al valore massimo o minimo ammissibile ("saturazione degli stadi di uscita"). Una bassa resistenza di carico sull'uscita, riducendo il guadagno, può ampliare la dinamica di ingresso.

Si possono definire grossolanamente due limiti per il segnale differenziale di ingresso:

- 1) Limite entro il quale il legame tra ingresso e uscita si può considerare lineare (intervallo di linearità).
- 2) Limite oltre il quale l'amplificatore "satura" e l'uscita non dipende più dal segnale di ingresso (vera e propria dinamica di ingresso).

La dinamica di ingresso di modo comune e la dinamica di uscita devono essere le più ampie possibile (per lo meno adeguate all'applicazione). Spesso questi due parametri non sono indipendenti (la dinamica di uscita è influenzata dal modo comune applicato in ingresso). Se la dinamica di ingresso e/o la dinamica di uscita si estendono dalla tensione di alimentazione negativa a quella positiva si parla di amplificatori "rail-to-rail".

Prodotto guadagno banda:

Se f_H è il limite superiore di banda il prodotto $A_d(0)f_H = \text{PGB}$ si definisce prodotto guadagno banda ($A_d(0)$ indica il valore di A_d a frequenza zero, ovvero in continua). Per amplificatori la cui risposta in frequenza risulta del tipo a polo dominante, il PGB coincide in pratica con la frequenza a cui il modulo del guadagno assume valore unitario. Per tale tipo di amplificatori, fissata una frequenza di lavoro f_w , una stima del guadagno massimo ottenibile è data da:

$$A_d(f_w) = \min \left[A_d(0), \frac{\text{PGB}}{f_w} \right]$$

4.2 Coppia di transistori MOSFET accoppiati di source (source coupled pair)

Questa configurazione, mostrata nella figura seguente, è la base di praticamente tutti gli amplificatori differenziali a MOSFET.

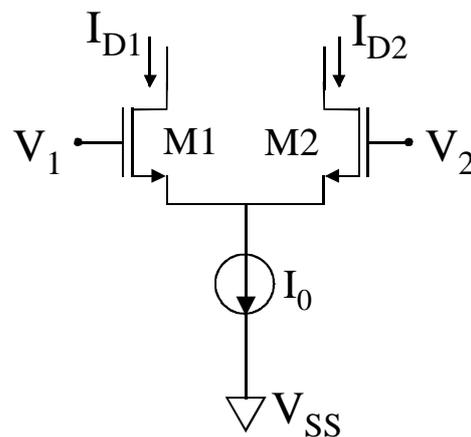


Figura 4.2.1

Il circuito ripartisce la corrente I_0 nelle due componenti I_{D1} e I_{D2} . Quando la tensione di modo differenziale $V_d = V_1 - V_2$ è nulla, I_0 si divide in parti uguali tra I_{D1} e I_{D2} . Per V_d che cresce progressivamente assumendo valori positivi ($V_1 > V_2$) la percentuale di I_0 che finisce in I_{D1} diventa sempre più grande a scapito di I_{D2} . Per V_d negative succede l'esatto opposto. Ci poniamo ora l'obiettivo di valutare quantitativamente la relazione tra le correnti I_{D1} e I_{D2} e la tensione V_d .

Calcolo delle correnti di drain in funzione della tensione differenziale di ingresso.

Ipotesi:

- I transistori lavorano in zona di saturazione.
- Si può trascurare l'effetto della V_{DS} sulle I_D . Perché ciò sia garantito è sufficiente che le V_{DS} siano uguali per $V_d = 0$ e che le variazioni causate dall'applicazione di V_d siano più piccole o al limite dello stesso ordine di grandezza delle rispettive variazioni delle V_{GS} .

- Le variazioni delle tensioni di source non producono variazioni della corrente di polarizzazione I_0 . Dato che tale corrente è prodotta da uno specchio di corrente, ciò consiste nel considerare infinita la resistenza di uscita dello specchio stesso.
- I due transistori sono identici.

Con queste ipotesi possiamo scrivere:

$$V_d = V_1 - V_2 = V_{GS1} - V_{GS2} = V_{t1} + \sqrt{\frac{2I_{D1}}{\beta_1}} - V_{t2} - \sqrt{\frac{2I_{D2}}{\beta_2}} \quad (4.2.1)$$

Siccome i due transistori sono identici, ovvero $V_{t1}=V_{t2}$ e $\beta_1=\beta_2=\beta$, otteniamo:
Pertanto otteniamo:

$$V_d = \sqrt{\frac{2I_{D1}}{\beta}} - \sqrt{\frac{2I_{D2}}{\beta}} = \sqrt{\frac{2}{\beta}} (\sqrt{I_{D1}} - \sqrt{I_{D2}})$$

La dipendenza delle correnti dalla tensione differenziale può essere ottenuta risolvendo questa equazione con il vincolo $I_{D1}+I_{D2}=I_0$. Si procede elevando al quadrato ambo i membri e si ottiene:

$$V_d^2 \frac{\beta}{2} = I_{D1} + I_{D2} - 2\sqrt{I_{D1}I_{D2}}$$

Elevando al quadrato abbiamo perso la seguente informazione che pertanto deve essere annotata per essere richiamata alla fine del calcolo:

Condizione 1: se $V_d > 0$ allora $I_{D1} > I_{D2}$.

A questo punto occorre osservare che $I_{D1}+I_{D2}=I_0$. Sostituendo si ottiene:

$$V_d^2 \frac{\beta}{2} - I_0 = -2\sqrt{I_{D1}I_{D2}}$$

Si procede elevando ancora al quadrato ma occorre ancora osservare che, essendo la radice quadrata positiva per definizione dovrà essere rispettata la condizione:

$$\underline{\text{Condizione 2}}: V_d^2 \frac{\beta}{2} - I_0 < 0 \Rightarrow -\sqrt{\frac{2I_0}{\beta}} < V_d < \sqrt{\frac{2I_0}{\beta}}$$

Questa condizione impone un limite inferiore e superiore alla tensione differenziale di ingresso affinché l'equazione abbia un risultato. Vedremo che a questo limite analitico corrisponde un ben preciso limite fisico.

Procedendo con l'elevazione al quadrato e considerando che $I_{D2}=I_0 - I_{D1}$ si ottiene l'equazione:

$$I_{D1}^2 - I_0 I_{D1} + \frac{1}{4} \left(V_d^2 \frac{\beta}{2} - I_0 \right)^2 = 0 \Rightarrow I_{D1} = \frac{I_0}{2} \pm V_d \frac{\beta}{4} \sqrt{\frac{4I_0}{\beta} - V_d^2}$$

L'ambiguità sul segno della radice si risolve subito richiamando la condizione che avevamo posto nell'effettuare la prima operazione di quadratura: siccome per V_d positivi la I_{D1} deve essere maggiore di I_{D2} , essa dovrà essere anche maggiore di $I_0/2$, altrimenti la somma $I_{D1}+I_{D2}$ risulterebbe inferiore a I_0 . Pertanto nell'espressione trovata dovrà risultare il segno "+". Otteniamo infine:

$$I_{D1} = \frac{I_0}{2} + V_d \frac{\beta}{4} \sqrt{\frac{4I_0}{\beta} - V_d^2} \quad I_{D2} = \frac{I_0}{2} - V_d \frac{\beta}{4} \sqrt{\frac{4I_0}{\beta} - V_d^2}$$

Mettendo in evidenza all'interno della radice il termine $\frac{2I_0}{\beta} = V_{MAX}^2$ si ottengono le formule:

$$\begin{aligned} I_{D1} &= \frac{I_0}{2} \left(1 + \frac{V_d}{V_{MAX}} \sqrt{2 - \frac{V_d^2}{V_{MAX}^2}} \right) \\ I_{D2} &= \frac{I_0}{2} \left(1 - \frac{V_d}{V_{MAX}} \sqrt{2 - \frac{V_d^2}{V_{MAX}^2}} \right) \end{aligned} \quad (4.2.2)$$

e, per quanto riguarda la differenza delle correnti $I_{D1}-I_{D2}$:

$$I_{D1} - I_{D2} = I_0 \frac{V_d}{V_{MAX}} \sqrt{2 - \frac{V_d^2}{V_{MAX}^2}}$$

Nella Figura 4.2.2 sono mostrate le correnti normalizzate a I_0 in funzione della tensione differenziale di ingresso normalizzata rispetto a $V_{max} = \sqrt{\frac{2I_0}{\beta}}$.

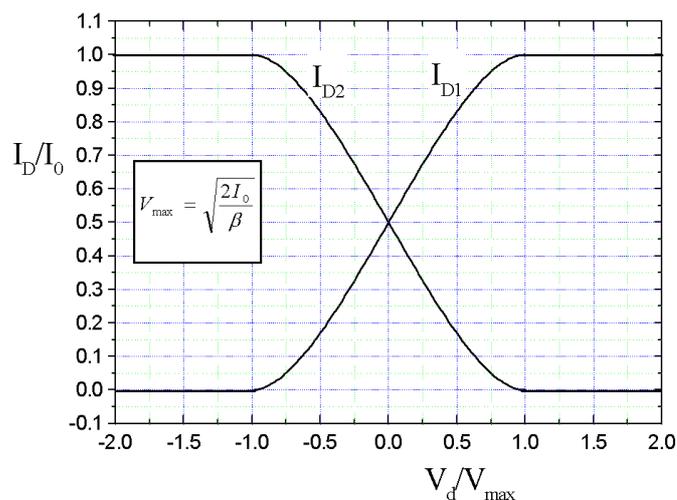


Figura 4.2.2

Queste curve sono state tracciate usando le espressioni ricavate per le I_D fermandoci però a

V_{max} , in quanto avevamo trovato che la risoluzione era valida solo per $-V_{max} < V_d < V_{max}$. Il tratto oltre V_{max} è stato ricavato osservando che per $V_d=V_{max}$ la I_{D1} è pari a I_0 e quindi conduce solo il transistor M1. In queste condizioni osserviamo che:

$$V_{GS1} = V_t + \sqrt{\frac{2I_0}{\beta}} \quad (\text{poiché } I_{D1} = I_0)$$

$$V_{GS2} = V_{GS1} - V_d = V_{GS1} - V_{max} = V_t$$

Aumentando ulteriormente la $V_d = V_{GS1} - V_{GS2}$ deve accadere una delle due cose (o entrambe): (1) aumenta V_{GS1} , (2) diminuisce V_{GS2} . Ovviamente V_{GS1} non può aumentare perché la I_{D1} non può superare la I_0 , allora diminuisce V_{GS2} che scende sotto la tensione di soglia e, continuando ad aumentare la V_d può diventare anche negativa. Pertanto all'aumentare della V_d , I_{D2} rimane nulla e deve risultare $I_{D1}=I_0$, come mostrato in figura. Ovviamente lo stesso discorso si può ripetere per l'estremo inferiore $V_d < -V_{max}$ dove è M2 a portare tutta la corrente I_0 mentre M1 rimane spento.

La dinamica di ingresso (ovvero V_{max}) può essere aumentata aumentando I_0 o diminuendo β (ovvero W/L).

Per piccoli valori di V_d (ovvero molto minori di V_{max}), il comportamento si può considerare approssimativamente lineare. Utilizzando per le equazioni (4.2.2) lo sviluppo di Taylor arrestato al primo ordine, si ottiene la nota espressione di piccolo segnale per le I_D :

$$I_{D1} \cong \frac{I_0}{2} + \frac{1}{2} g_m V_d \quad I_{D2} \cong \frac{I_0}{2} - \frac{1}{2} g_m V_d \quad \text{con } g_m = \sqrt{\beta I_0} = \sqrt{2\beta I_{DQ}}$$

dove con I_{DQ} è stato indicato il valore di riposo ($V_d=0$) di I_{D1} e I_{D2} , pari a $I_0/2$.

Nella figura seguente è mostrato l'effetto di un aumento di I_0 e di una diminuzione di β di un fattore 2. I grafici riportano in ordinate la differenza tra I_{D1} e I_{D2} che nella maggioranza dei casi pratici rappresenta il segnale utile.

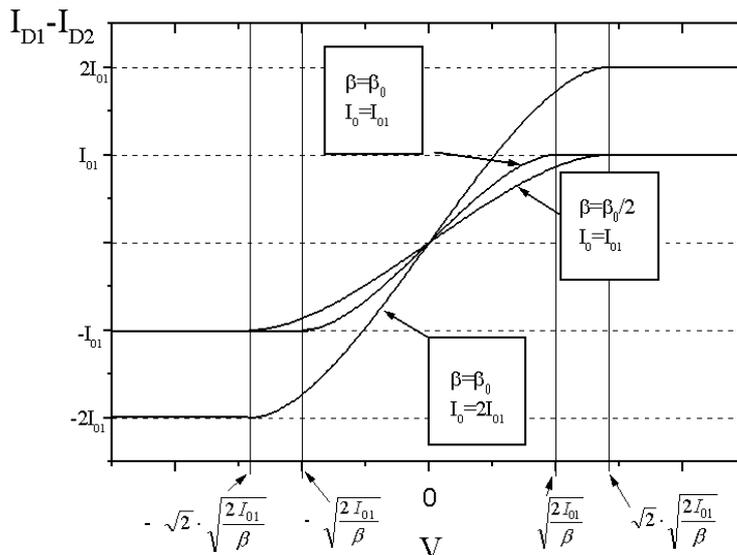


Figura 4.2.3

Si può osservare che:

- un aumento di I_0 aumenta la dinamica e contemporaneamente la pendenza nell'origine,

ovvero il guadagno di trasduzione tensione di ingresso - corrente di uscita.

- una diminuzione di β comporta un analogo aumento di dinamica ma anche una diminuzione del guadagno.
- La dinamica complessiva del segnale differenziale in ingresso è pari a $\pm V_{dmax}$. Si può verificare immediatamente che V_{dmax} è pari a $\sqrt{2}(V_{GS} - V_t)$, dove con V_{GS} si è indicato il valore assunto a riposo.

Per concludere calcoliamo la tensione di source V_S dei due MOS M1 e M2. Possiamo scrivere:

$$V_S = \frac{V_d}{2} - V_{GS1} = \frac{V_d}{2} - V_t - \sqrt{\frac{2I_{D1}}{\beta}}$$

$$V_S = -\frac{V_d}{2} - V_{GS2} = -\frac{V_d}{2} - V_t - \sqrt{\frac{2I_{D2}}{\beta}}$$

Sommando le due equazioni precedenti si ottiene:

$$2V_S = -2V_t - \sqrt{\frac{2I_{D1}}{\beta}} - \sqrt{\frac{2I_{D2}}{\beta}} \Rightarrow V_S = -V_t - \sqrt{\frac{1}{2\beta}}(\sqrt{I_{D1}} + \sqrt{I_{D2}})$$

Osserviamo che siccome $I_{D1} + I_{D2} = I_0 = \text{costante}$, la somma $\sqrt{I_{D1}} + \sqrt{I_{D2}}$ non potrà essere costante e quindi la tensione di source dei due MOSFET non è costante ma varierà in funzione del segnale differenziale applicato. Ciò evidenzia la differenza di comportamento tra piccoli segnali e grandi segnali: in condizioni di piccolo segnale la tensione di source rimane costante (ovvero a massa per le variazioni) in virtù delle proprietà di antisimmetria. Nel caso di ampi segnali i source non possono più considerarsi a massa ma variano con V_D . In particolare risulta:

$$\text{Per } V_d = 0 \Rightarrow V_S = -V_t - \sqrt{\frac{I_0}{\beta}}$$

$$\text{Per } V_d = \pm V_{MAX} \Rightarrow V_S = -V_t - \sqrt{\frac{I_0}{2\beta}}$$

4.3 Coppia di transistori bipolari connessi di emettitore.

Il funzionamento della coppia differenziale di transistori bipolari, mostrata nella figura seguente, è perfettamente analogo a quello della corrispondente configurazione a MOSFET.

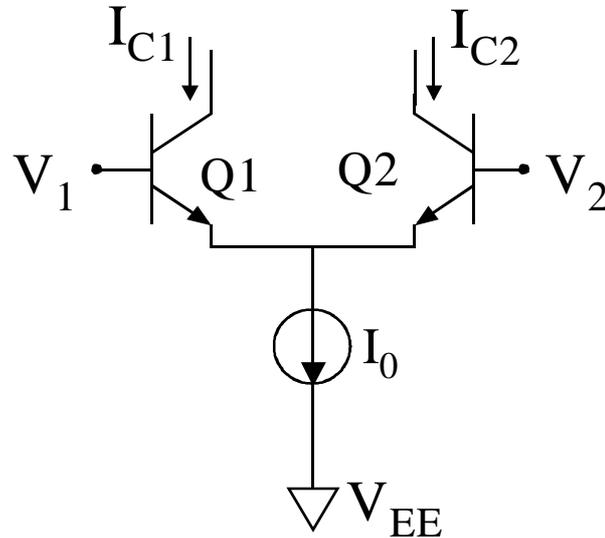


Figura 4.3.1

La tensione differenziale di ingresso V_d risulta pari a:

$$V_d = V_{BE1} - V_{BE2} = V_T \left(\ln \left(\frac{I_{C1}}{I_{S1}} \right) - \ln \left(\frac{I_{C2}}{I_{S2}} \right) \right) = V_T \ln \left(\frac{I_{C1}}{I_{C2}} \cdot \frac{I_{S2}}{I_{S1}} \right) \quad (4.3.1)$$

dove $V_T = kT/q$. Se i transistori sono uguali, $I_{S1} = I_{S2}$ e otteniamo:

$$V_d = V_T \ln \left(\frac{I_{C1}}{I_{C2}} \right) \Rightarrow I_{C2} = I_{C1} e^{-\frac{V_d}{V_T}}$$

Se si trascurano le correnti di base rispetto alle I_C si ha $I_{C1} + I_{C2} = I_0$. Da questa espressione si ricava facilmente:

$$I_{C1} = I_0 \frac{1}{1 + e^{-\frac{V_d}{V_T}}}; \quad I_{C2} = I_0 \frac{e^{-\frac{V_d}{V_T}}}{1 + e^{-\frac{V_d}{V_T}}}$$

La differenza tra le correnti di collettore risulta data da:

$$I_{C1} - I_{C2} = I_0 \frac{1 - e^{-\frac{V_d}{V_T}}}{1 + e^{-\frac{V_d}{V_T}}} = I_0 \frac{e^{\frac{V_d}{2V_T}} e^{\frac{V_d}{2V_T}} - e^{-\frac{V_d}{2V_T}} e^{-\frac{V_d}{2V_T}}}{e^{\frac{V_d}{2V_T}} e^{\frac{V_d}{2V_T}} + e^{-\frac{V_d}{2V_T}} e^{-\frac{V_d}{2V_T}}} = I_0 \tanh \left(\frac{V_d}{2V_T} \right)$$

L'andamento delle correnti di collettore e della loro differenza in funzione della V_d è mostrato nelle figure seguenti:

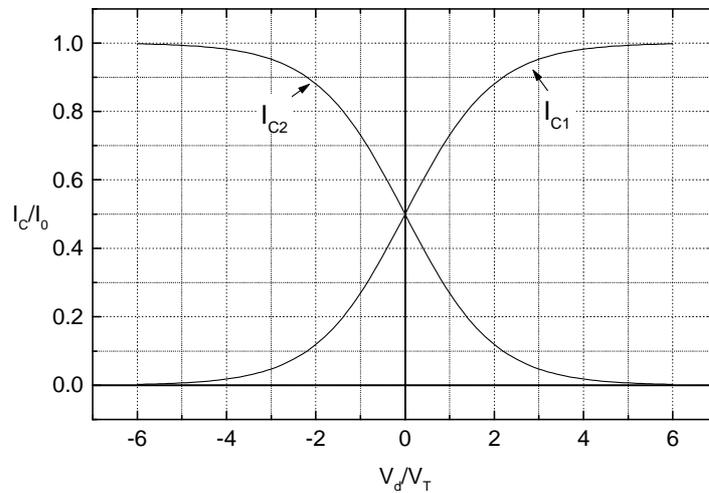


Figura 4.3.2

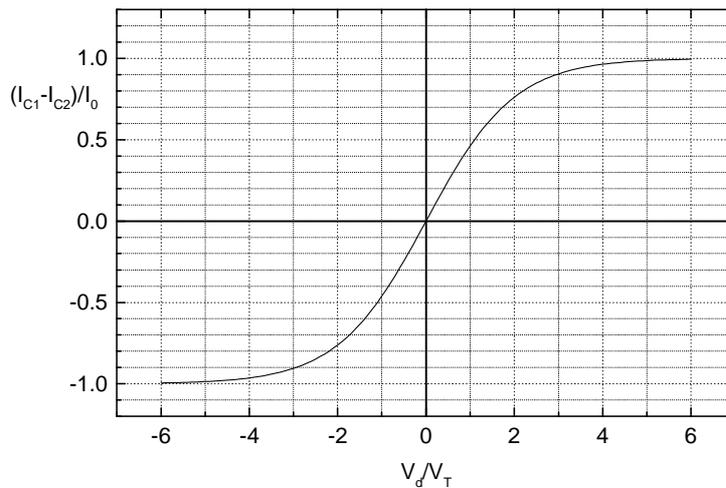


Figura 4.3.3

Possiamo osservare che,

- dal punto di vista matematico le correnti non raggiungono mai gli asintoti 0 e I_0 . In pratica, come si può osservare, per $V_d = \pm 4V_T$ le correnti I_{C1} e I_{C2} hanno praticamente raggiunto i due asintoti a meno di uno scarto il più delle volte trascurabile.
- La dinamica differenziale di ingresso è fissata solo dal valore di V_T , non è possibile agire su di essa attraverso il dimensionamento dei transistori. Ciò è un'importante differenza rispetto al circuito a MOSFET.

4.4 Amplificatori differenziali a carichi resistivi.

Amplificatore a MOSFET.

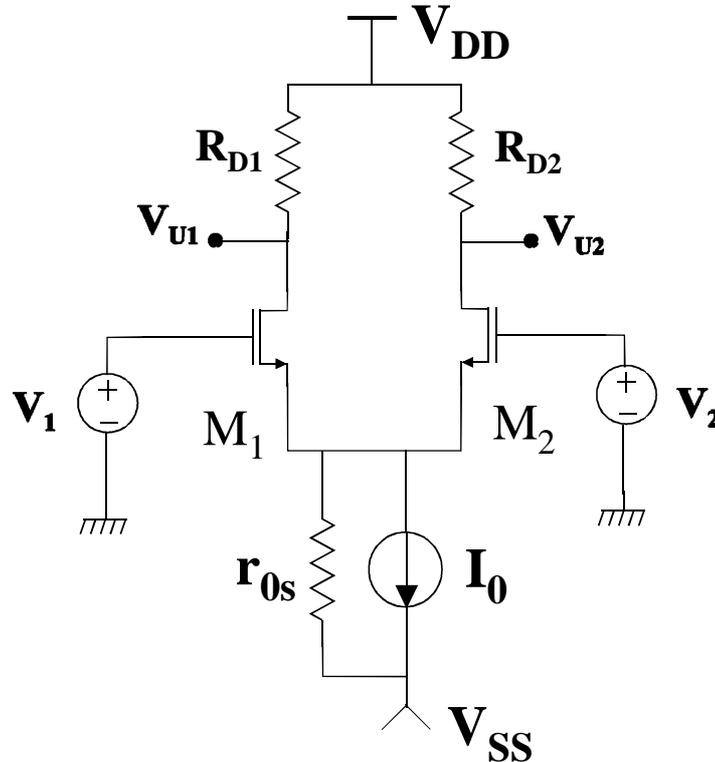


Figura 4.4.1

Il circuito della figura precedente può essere utilizzato sia come amplificatore fully differential sia come amplificatore differenziale con uscita unipolare (single ended). I due casi verranno affrontati insieme. Come analisi verranno calcolate le amplificazioni relative al funzionamento per piccolo segnale e si procederà poi al calcolo della tensione di offset. La resistenza r_{0s} rappresenta la resistenza (differenziale) dello specchio di corrente che polarizza l'amplificatore.

IPOTESI DI PARTENZA: in tutte le analisi che seguiranno si considererà che:

$$R_{D1}, R_{D2} \ll r_{d1}, r_{d2} \quad (4.4.1)$$

Questo comporta che si possa trascurare l'effetto delle V_{DS} sulle correnti di drain di M1 e M2.

Amplificazioni.

Tenendo presente l'ipotesi (4.4.1) si può subito scrivere:

$$v_{u1} = -i_{d1}R_{D1}; \quad v_{u2} = -i_{d2}R_{D2}; \quad v_{ud} = v_{u1} - v_{u2} = i_{d2}R_{D2} - i_{d1}R_{D1}$$

Per procedere all'analisi delle amplificazioni consideriamo che:

- Quando viene applicato un piccolo segnale di modo differenziale, il source si può considerare a massa alle variazioni. Questo a rigore è vero solo se il circuito è perfettamente simmetrico, ovvero $M1=M2$, $R_{D1}=R_{D2}=R_D$. In realtà se si può trascurare l'effetto delle v_{ds} sulla corrente di drain, ovvero se $R_{D1}, R_{D2} \ll r_{d1}, r_{d2}$, è solo sufficiente che siano uguali i due MOSFET, in quanto quello che accade sul drain non influenza la corrente di drain e quindi la tensione sul source. Quindi le variazioni delle correnti di drain saranno date da: $i_{d1} = g_m \frac{v_d}{2}$, $i_{d2} = -g_m \frac{v_d}{2}$.
- Quando viene applicato un segnale di modo comune v_c con segnale di modo differenziale v_d nullo, le V_{GS} di M1 e M2 sono uguali ($V_{GS1}=V_{GS2}$). Quindi se $M1=M2$ le correnti I_{D1} e I_{D2} rimarranno uguali, a prescindere dal segnale v_c . Senza commettere un grande errore si può considerare che anche le variazioni delle V_{GS} siano nulle. Infatti, se le V_{GS} variassero molto entrambe le correnti I_{D1} e I_{D2} subirebbero grandi variazioni e, dovendo valere sempre $I_{D1}=I_{D2}$ (solo perché $v_d=0$, altrimenti ciò non è più vero), anche la loro somma avrebbe grandi variazioni. Ma ciò non è possibile, perché la somma $I_{D1}+I_{D2}$ è fissata da uno specchio di corrente caratterizzato da un'elevata resistenza di uscita. Quindi essendo trascurabili le variazioni delle VGS si ha:

$$v_s = v_c - v_{gs} \cong v_c \Rightarrow i_s = i_{d1} + i_{d2} \cong \frac{v_c}{r_{0s}}$$

Amplificazioni: Caso single ended.

1) Amplificazione di modo differenziale. $A_d = \left. \frac{v_{u1}}{v_d} \right|_{v_c=0} = -\frac{g_m}{2} R_{D1}$

2) Amplificazione di modo comune: indicando come più sopra con i_s la somma delle variazioni di I_{D1} e I_{D2} , dovendo queste due ultime rimanere uguali, si avrà: $i_{d1}=i_{d2}=i_s/2$. Pertanto:

$$A_c = \left. \frac{v_{u1}}{v_c} \right|_{v_d=0} = -i_{d1} R_{D1} \frac{1}{v_c} = -\frac{i_s}{2} R_{D1} \frac{1}{v_c} = -\frac{R_{D1}}{2r_{0s}}$$

3) $CMRR = g_m r_{0s}$

Amplificazioni: Caso fully differential.

1) Amplificazione di modo differenziale: $A_{dd} = \left. \frac{v_{ud}}{v_d} \right|_{v_c=0} = -\frac{g_m}{2} (R_{D2} + R_{D1})$

se si considera in prima approssimazione $R_{D1}=R_{D2}$ allora $A_d = -g_m R_D$.

2) Amplificazione di modo comune. In questo caso, come indicato nell'introduzione sugli amplificatori differenziali, si considera come amplificazione di modo comune la A_{cd} , ovvero:

$$A_{cd} = \left. \frac{v_{ud}}{v_c} \right|_{v_d=0} = (i_{d2} R_{D2} - i_{d1} R_{D1}) \frac{1}{v_c} = \left(\frac{i_s}{2} R_{D2} - \frac{i_s}{2} R_{D1} \right) \frac{1}{v_c} = \frac{R_{D2} - R_{D1}}{2r_{0s}}$$

Quindi l'amplificazione di modo comune è tanto più bassa quanto più piccolo è l'errore di matching sulle resistenze. In realtà nell'analisi precedente si sono considerati uguali i due

transistori M1 e M2. Se essi sono diversi si ha anche differenza tra i_{d1} e i_{d2} , ovvero $i_{d1}/i_{d2}=g_{m1}/g_{m2}$ quindi il quadro sarebbe ulteriormente peggiorato. Nel calcolo del CMRR considereremo comunque per semplicità M1=M2.

$$3) CMRR = \frac{|A_{dd}|}{|A_{cd}|} = \frac{R_{D1} + R_{D2}}{R_{D2} - R_{D1}} \cdot g_m r_{0s}$$

Tensione di offset.

Per il calcolo della tensione di offset si prenderà in considerazione l'amplificatore fully differential. Infatti, come vedremo, in questo caso contribuiranno a generare la tensione di offset solo degli errori di matching. Si può dimostrare che nel caso single-ended la tensione di offset dipende soprattutto da errori sul valore nominale che, come è noto, sono molto più grandi di quelli di matching. Questo costituisce un ulteriore svantaggio di usare amplificatori a carichi resistivi quando serve un'uscita single-ended.

La tensione di offset V_{io} , è la tensione differenziale che occorre applicare per avere uscita V_{Ud} nulla, ovvero per avere:

$$I_{D1}R_{D1} = I_{D2}R_{D2} \quad (4.4.2)$$

Si ha quindi:

$$V_{io} = (V_{GS1} - V_{GS2})_{I_{D1}R_{D1}=R_{D2}I_{D2}}$$

La tensione di offset si può considerare come l'errore di matching che si ha tra le V_{GS} di M1 e M2 nel momento in cui è realizzata la (4.4.2). Se scriviamo la V_{GS} come la somma di due grandezze G_1 e G_2 definite da:

$$V_{GS} = G_A + G_B \quad \text{con: } G_A = V_t, \quad G_B = \sqrt{\frac{2I_D}{\beta}} = \sqrt{2} \cdot I_D^{\frac{1}{2}} \beta^{-\frac{1}{2}}$$

si ha:

$$V_{io} = \Delta V_{GS} = \Delta G_A + \Delta G_B$$

A questo punto $\Delta G_A = \Delta V_t$, mentre:

$$\Delta G_B = G_B \frac{\Delta G_B}{G_B} = G_B \left(\frac{1}{2} \frac{\Delta I_D}{I_D} - \frac{1}{2} \frac{\Delta \beta}{\beta} \right) = \frac{1}{2} \sqrt{\frac{2I_D}{\beta}} \left(\frac{\Delta I_D}{I_D} - \frac{\Delta \beta}{\beta} \right)$$

Per ottenere questa espressione è stata utilizzata l'equazione (2.3.5). A questo punto possiamo ricavare $\Delta I_D/I_D$ in funzione dell'errore di matching sulle resistenze, sfruttando la (4.4.2), in quanto è proprio la differenza tra le due resistenze a far sì che le due correnti di drain debbano essere diverse. Se definiamo la variabile $Z=I_D R_D$, osserviamo che la (4.4.2) significa che, per $V_d=V_{io}$, Z deve avere un errore di matching $\Delta Z=0$. Ovvero:

$$0 = \Delta Z = Z \frac{\Delta Z}{Z} = R_D I_D \left(\frac{\Delta I_D}{I_D} + \frac{\Delta R_D}{R_D} \right) \Rightarrow \frac{\Delta I_D}{I_D} = -\frac{\Delta R_D}{R_D}$$

L'espressione della tensione di offset risulta quindi:

$$V_{io} = \Delta V_t + \frac{V_{GS} - V_t}{2} \left(-\frac{\Delta R_D}{R_D} - \frac{\Delta \beta}{\beta} \right) \quad (4.4.3)$$

Considerazioni sulla tensione di offset in amplificatori a MOSFET a carichi resistivi.

Nell'equazione (4.4.3) compare l'errore di matching relativo sulle R_D . Per considerare come questo dipenda dagli errori relativi sulle dimensioni dei resistori e sulla resistenza di stato si faccia riferimento al capitolo 2. Per esprimere l'errore (assoluto, generalmente espresso in mV) sulla tensione di soglia e sul β , i manuali di processi dedicati alla progettazione riportano delle formula empiriche che li esprimono in funzione delle dimensioni dei transistori. Per quanto riguarda l'errore sulle tensioni di soglia si ha di solito:

$$\Delta V_t = C_{Vt} \frac{1}{\sqrt{WL}} \quad (4.4.4)$$

dove C_{Vt} è una costante empirica fornita nel manuale. Per quanto riguarda il β , possiamo osservare che valendo esso, $K_n W/L$, dove $K_n = \mu_n C_{OX}$, si ha:

$$\frac{\Delta \beta}{\beta} = \frac{\Delta K_n}{K_n} + \frac{\Delta W}{W} - \frac{\Delta L}{L} \quad (4.4.5)$$

Spesso anche per l'errore sul β viene fornita nei manuali di processo una legge empirica simile a quella delle tensioni di soglia. Questa legge, che tiene conto complessivamente degli errori su K_n e sulle dimensioni L e W è la seguente:

$$\frac{\Delta \beta}{\beta} = C_\beta \frac{1}{\sqrt{WL}} \quad (4.4.6)$$

dove C_β è una costante determinata empiricamente.

Le leggi (4.4.4) e (4.4.6) sono in genere valide per transistori notevolmente più grandi del transistor a geometria minima. Il manuale fissa solitamente un valore minimo di L sotto il quale le formule del matching non sono più applicabili e gli errori di matching divergono. Questo limite inferiore su L , solitamente tre-quattro volte la lunghezza minima di canale permessa dal processo, è dovuto agli effetti di canale corto che peggiorano la riproducibilità delle caratteristiche elettriche (specialmente della tensione di soglia). Attualmente, per esigenze di velocità di risposta più che di ingombro, vi è anche in campo analogico la tendenza a progettare circuiti che utilizzano MOSFET con lunghezze di canale prossime a quelle minime. Per tale ragione molti manuali di processo riportano anche espressioni empiriche degli errori di matching valide per dimensioni fino a quelle minime.

Per ridurre la tensione di offset devono essere rispettate le indicazioni fornite nel paragrafo 2.3. L'espressione (4.4.3) unitamente alle equazioni (4.4.4) e (4.4.6) conferma la necessità di utilizzare transistori di grandi dimensioni. La (4.4.3), suggerisce inoltre che, per minimizzare la componente proporzionale a $V_{GS} - V_t$, conviene dimensionare il circuito scegliendo il punto di lavoro in modo che tale quantità sia minima (senza scendere sotto i soliti 100 mV per i soliti problemi di debole inversione).

Amplificatore differenziale a BJT a carichi resistivi.

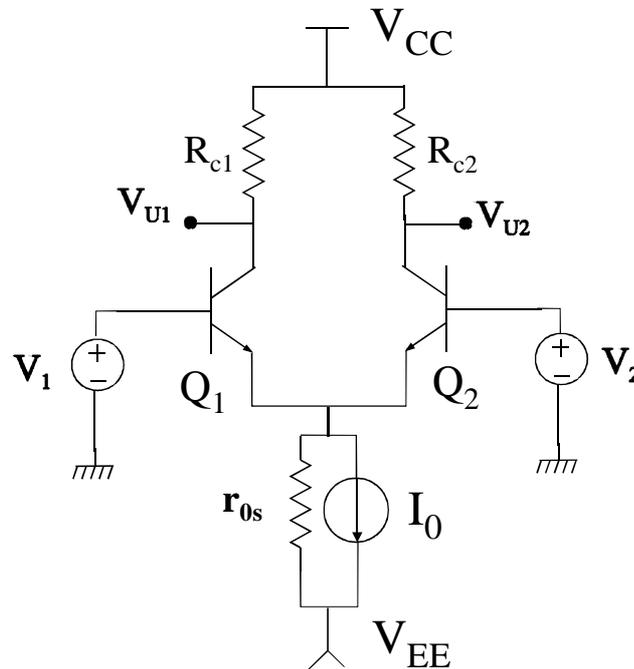


Figura 4.4.2

Per quanto riguarda le amplificazioni valgono esattamente le considerazioni e le formule presentate per l'amplificatore a carichi resistivi a MOSFET, qualora si sostituiscano le R_C in tutte le espressioni in cui compaiono le R_D e si tenga conto che il g_m dei BJT ha una dipendenza diversa dalla corrente di polarizzazione rispetto al g_m dei MOSFET.

Un vantaggio di questo tipo di amplificatore rispetto al corrispondente a MOSFET è dato dal fatto che il guadagno dipende linearmente dalla I_0 (perché così fa il g_m) e quindi l'amplificatore si presta ad essere usato come elemento base di moltiplicatori analogici.

Tensione di offset.

Ripetendo le considerazioni fatte per l'amplificatore a MOSFET, si arriva all'espressione:

$$V_{io} = (V_{BE1} - V_{BE2})_{I_{C1}R_{C1} = I_{C2}R_{C2}}$$

Quindi la tensione di offset è esprimibile come un errore di matching delle V_{BE} dei due transistor, nella condizione indicata di $I_{C1}R_{C1} = I_{C2}R_{C2}$.

Abbiamo quindi:

$$V_{io} = \Delta \left(V_T \ln \left(\frac{I_C}{I_S} \right) \right)$$

Trascurando le differenze di temperatura tra i due transistori si ottiene, applicando le espressioni riportate nel capitolo 2:

$$V_{io} = V_T \left(\frac{\Delta I_C}{I_C} - \frac{\Delta I_S}{I_S} \right)$$

Infine, considerando che come per l'amplificatore a MOSFET $\Delta I_C/I_C = -\Delta R_C/R_C$, si ottiene:

$$V_{io} = V_T \left(-\frac{\Delta R_C}{R_C} - \frac{\Delta I_S}{I_S} \right) \quad (4.4.7)$$

Per quanto riguarda il contributo dovuto all'errore di matching sulla R_C , vale quanto visto per l'amplificatore a MOSFET. Riguardo all'errore sulla corrente di saturazione, essendo essa esprimibile come $I_S = J_S A_E$, dove J_S dipende solo dai parametri fisici del processo, mentre A_E è l'area di emettitore, si ha:

$$\frac{\Delta I_S}{I_S} = \frac{\Delta A_E}{A_E} + \frac{\Delta J_S}{J_S}$$

Nel manuale di processo viene in genere fornita un'indicazione chiamata ΔV_{BE} (o, talvolta $\sigma_{V_{BE}}$). Questa è la deviazione standard della differenza di V_{BE} di due BJT identici e polarizzati con la stessa corrente. La ΔV_{BE} è data dalla (4.4.7) per $\Delta R_C/R_C = 0$. Quindi consente di trovare subito la $\Delta I_S/I_S$ in quanto:

$$\frac{\Delta I_S}{I_S} = \frac{\Delta V_{BE}}{V_T}$$

La ΔV_{BE} viene data per transistori che superino una certa dimensione (per esempio abbiano $area > 10$), per i quali il termine di errore geometrico (ovvero relativo ad A_E) nella I_S sia trascurabile rispetto al termine di errore su J_S .

Deriva della tensione di offset.

Derivando la (4.4.7) rispetto alla temperatura e considerando che gli errori relativi $\Delta I_S/I_S$ e $\Delta R_C/R_C$, non dipendono dalla temperatura (poiché sono rapporti in cui il denominatore e il numeratore variano ugualmente con la temperatura) si ottiene:

$$\frac{dV_{io}}{dT} = \frac{k}{q} \left(-\frac{\Delta R_C}{R_C} - \frac{\Delta I_S}{I_S} \right) = \frac{1}{T} \frac{kT}{q} \left(-\frac{\Delta R_C}{R_C} - \frac{\Delta I_S}{I_S} \right) = \frac{V_{io}}{T}$$

L'equazione precedente indica che se è nota la tensione di offset è nota anche la deriva termica della stessa. In particolare, se un amplificatore non presenta offset, esso, almeno in prima approssimazione non presenta deriva. In pratica l'offset può essere annullato aggiustando opportunamente i resistori R_{C1} e R_{C2} , per esempio mediante "laser trimming", ovvero introducendo nella (4.4.7) un termine $\Delta R_C/R_C = -\Delta I_S/I_S$. L'amplificatore a carichi resistivi presenta il vantaggio che se si annulla l'offset anche la deriva risulta annullata. Per questo motivo l'amplificatore fully differential a carichi resistivi e transistori bipolari è usato come primo stadio in amplificatori operazioni a bassissimo offset (per esempio l'amplificatore OP07 della Analog Devices). L'aggiustamento delle resistenze mediante inserimento di resistenze variabili esterne non è altrettanto efficace in quanto queste ultime presenterebbero un coefficiente di temperatura diverso da quello delle resistenze integrate e, inoltre, anche una temperatura diversa da quella del chip e pertanto $\Delta R_C/R_C$ non sarebbe più indipendente dalla temperatura.

Correnti di polarizzazione e corrente di offset.

A differenza dello schema a MOSFET, l'amplificatore a BJT necessita di correnti di polarizzazione I_{B1} e I_{B2} che devono essere fornite dalle sorgenti connesse in ingresso. Ciò comporta che se i generatori V_1 e V_2 non sono ideali come si è supposto finora ma hanno resistenze interne rispettivamente R_1 e R_2 , la tensione differenziale V_d che entra effettivamente in ingresso all'amplificatore non è pari a $V_1 - V_2$, come si vorrebbe, ma vale:

$$V_d = V_1 - V_2 + (I_{B2}R_2 - I_{B1}R_1)$$

Di solito $R_2 \neq R_1$, pertanto anche se fosse $I_{B1} = I_{B2}$ l'errore non sarebbe nullo. Nella maggioranza dei casi occorre cercare di tenere basso il valore delle I_B , che nello schema di Figura 4.4.2 è dato da:

$$I_B = \frac{I_C}{\beta}$$

I metodi per ottenere ciò sono i seguenti:

- Ridurre la I_C compatibilmente con la banda e la velocità di risposta.
- Usare configurazioni di Darlington con facendo però attenzione che dei due transistori della coppia, quello di ingresso viene ad avere una piccola I_C di riposo e quindi il suo β può diventare tanto basso da vanificare i vantaggi.
- Usare transistori "superbeta" per Q1 e Q1, se presenti nella tecnologia adottata.
- Adottare tecniche di cancellazione delle correnti di base.

In alcuni casi si riesce a garantire che $R_1 = R_2$. Allora l'errore introdotto dalle correnti di polarizzazione è proporzionale al termine $I_{B1} - I_{B2} = I_{io}$ = corrente di offset. La I_{io} risulta pari a:

$$I_{io} = \left(\frac{I_{C1}}{\beta_1} - \frac{I_{C2}}{\beta_2} \right)_{V_u=0}$$

Al solito si può notare come la corrente di offset sia data dall'errore di matching di una quantità $G = I_C/\beta$. Si ha pertanto:

$$I_{io} = \Delta G = G \frac{\Delta G}{G} = G \left(\frac{\Delta I_C}{I_C} - \frac{\Delta \beta}{\beta} \right) = \frac{I_C}{\beta} \left(\frac{\Delta I_C}{I_C} - \frac{\Delta \beta}{\beta} \right) = I_B \left(\frac{\Delta I_C}{I_C} - \frac{\Delta \beta}{\beta} \right)$$

Ricordando che per tensione di uscita nulla $\Delta I_C/I_C = -\Delta R_C/R_C$, si arriva all'espressione finale:

$$I_{io} = I_B \left(\frac{\Delta R_C}{R_C} - \frac{\Delta \beta}{\beta} \right)$$

4.5 Considerazioni generali sugli amplificatori a carichi resistivi.

Gli amplificatori a carichi resistivi trovano poche applicazioni in campo integrato. Questo è particolarmente vero per l'amplificatore con uscita single-ended, in quanto il suo CMRR è troppo ridotto e troppo dipendente dalla resistenza di uscita dello specchio di polarizzazione. In ogni caso le amplificazioni raggiungibili sono piccole in quanto le resistenze di carico non possono essere fatte troppo grandi, a causa della caduta di tensione statica ai loro capi. Infatti, se facciamo riferimento all'amplificazione del fully differential (ma lo stesso vale per il single ended, a parte un fattore 1/2):

$$A_{dd} = g_m R_D = g_m \frac{V_{RD}}{I_D}$$

dove con V_{RD} si è indicata la caduta di tensione sulle R_D (supposte uguali, in assenza di errori di matching). Facendo riferimento all'espressione di g_m/I_D indicata nel capitolo 2 si ha:

$$A_{dd} = 2 \frac{V_{RD}}{V_{GS} - V_t}$$

Ora, siccome per ovvi motivi V_{RD} non può essere più grande di $V_{DD} - V_{SS}$, ed anzi, per avere una dinamica di uscita accettabile possiamo considerare che V_{RD} sia dell'ordine di $(V_{DD} - V_{SS})/2$, il guadagno massimo risulta essere dell'ordine di:

$$\max(A_{dd}) \cong \frac{V_{DD} - V_{SS}}{V_{GS} - V_t}$$

Si comprende come, non potendo $V_{GS} - V_t$ scendere sotto il centinaio di mV per mantenere i MOSFET in forte inversione, il guadagno massimo ottenibile sia molto ridotto. In pratica, con il continuo abbassarsi delle tensioni di alimentazione esso si riduce a qualche decina.

Ripetendo gli stessi ragionamenti per un amplificatore differenziale a transistori bipolari, si ottiene il risultato:

$$\max(A_{dd}) = \frac{V_{CC} - V_{EE}}{2V_T}$$

Anche qui i guadagni che si possono ottenere, specie con tensioni di alimentazioni basse non sono molto elevati, tuttavia essendo $2V_T$ dell'ordine di 50 mV si ha almeno un fattore 2 di vantaggio rispetto al caso a MOSFET. Dal punto di vista dell'uso come amplificatore differenziale integrato (solo fully differential) vi è un ulteriore vantaggio dovuto alla possibilità di eliminare la deriva della tensione di offset semplicemente azzerando la tensione di offset stessa mediante un'operazione di trimming sulle resistenze di carico (si veda il paragrafo relativo).

Capitolo 5. Amplificatori differenziali: circuiti

5.1 Amplificatore differenziale semplice a MOSFET con uscita single-ended

Lo schema del tipo più semplice di amplificatore differenziale CMOS a carichi attivi e singola uscita è mostrato nella figura seguente:

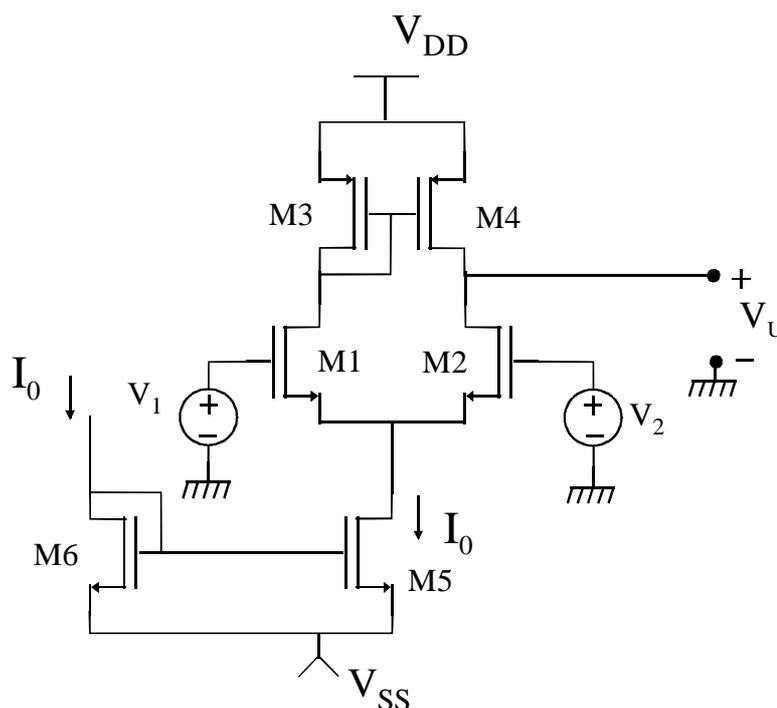


Figura 5.1.1

Punto di riposo. Il punto di riposo si ottiene con $V_1 = V_2 = 0$ (ovvero anche $V_d = V_c = 0$). Si dimostra semplicemente che se $V_d = 0$ (quindi anche con $V_c \neq 0$) $I_{D1} = I_{D2}$. Infatti $V_d = V_{GS1} - V_{GS2} = 0$ significa che $V_{GS1} = V_{GS2}$. Se fosse per assurdo $I_{D1} > I_{D2}$, allora dovrebbe essere $V_{DS1} > V_{DS2}$, ma anche $|V_{DS3}| > |V_{DS4}|$. Poiché risulta:

$$V_{DS1} + |V_{DS3}| = V_{DS2} + |V_{DS4}|$$

questa equazione non potrebbe essere verificata (entrambi gli addendi a primo membro sarebbero più grandi dei corrispondenti a secondo membro) e pertanto l'ipotesi è falsa. Ripetendo il ragionamento si conclude che è anche impossibile che sia $I_{D2} > I_{D1}$ e pertanto deve essere $I_{D1} = I_{D2} = I_0/2$. Per avere questa eguaglianza deve anche essere: $V_{DS1} = V_{DS2}$ e $V_{DS3} = V_{DS4}$. Si deduce immediatamente che la tensione a riposo (o in presenza di solo modo comune) sul nodo di uscita è uguale alla tensione nel nodo K. Si verifica immediatamente quindi che:

$$V_U = V_K = V_{DD} - |V_{GS3}|$$

Dato che i transistori M1 e M2 sono identici e hanno stesso punto di riposo (e così pure M3 e M4) definiremo:

$$g_{mn} = g_{m1} = g_{m2}; \quad g_{mp} = g_{m3} = g_{m4}; \quad r_{dn} = r_{d1} = r_{d2}; \quad r_{dp} = r_{d3} = r_{d4}$$

Principio di funzionamento. Questo circuito può essere rappresentato ricorrendo all'equivalente di Norton (valido per le variazioni) mostrato nella Figura 5.1.2(a).

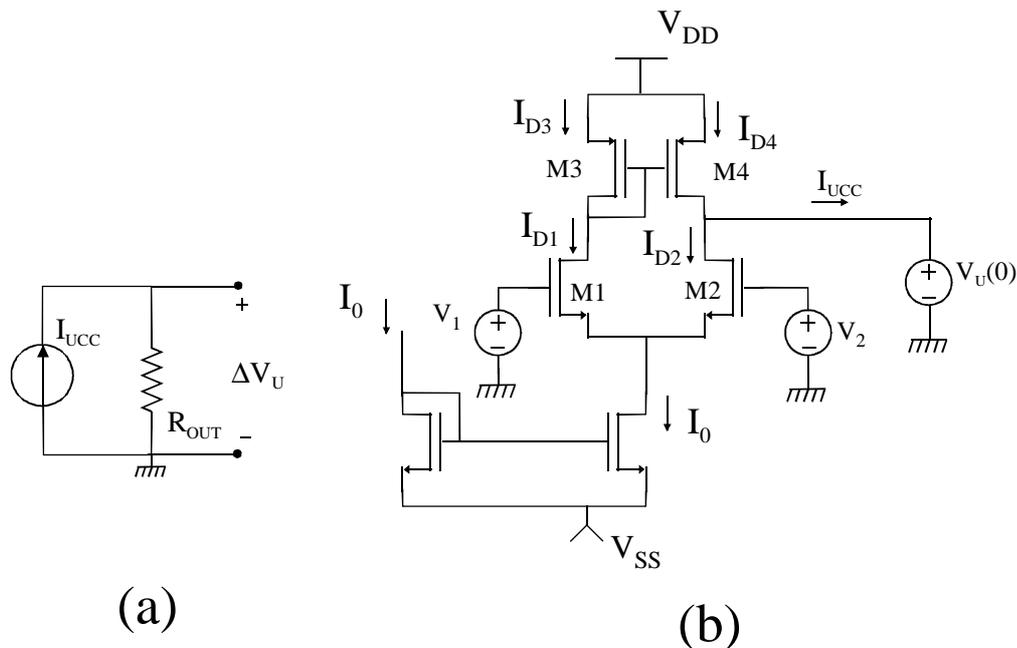


Figura 5.1.2

Occorre quindi determinare la dipendenza della corrente di cortocircuito dalle tensioni di ingresso (di modo comune e modo differenziale) e la resistenza di uscita. Si noti che il circuito di Figura 5.1.2(a) rappresenta le variazioni della tensione di uscita rispetto al punto di riposo. Per calcolare la corrente di corto circuito si può impiegare lo schema di Figura 5.1.2(b) dove il corto circuito alle variazioni viene realizzato fisicamente connettendo all'uscita un generatore di tensione ideale di tensione pari alla tensione assunta a riposo $V_U(0)$ dall'uscita (ovvero per $V_{id} = V_{ic} = 0$). Questo procedimento può essere usato vantaggiosamente nelle simulazioni. Osserviamo che:

$$I_{UCC} = I_{D4} - I_{D2} \quad (5.1.1)$$

dove I_{D4} e I_{D2} si intendono riferite ai versi indicati in Figura 5.1.2(b). Se si ipotizza che lo specchio M3-M4 è perfetto, ovvero che $I_{D3} = I_{D4}$, allora siccome è anche ovviamente $I_{D1} = I_{D3}$ si ha che l'equazione (5.1.1) diventa:

$$I_{UCC} = I_{D1} - I_{D2} \quad (5.1.2)$$

Come principio generale di funzionamento si può osservare in prima approssimazione che tensioni di modo differenziale producono variazioni delle I_D di segno opposto che quindi si sommano come intensità nella I_{UCC} mentre tensioni di modo comune provocano variazioni

uguali e concordi che dunque tendono ad elidersi a vicenda nella I_{UCC} .

Calcolo della resistenza di uscita. La resistenza di uscita si può calcolare inserendo (nel circuito alle variazioni) un generatore di prova V_P sull'uscita, con ingressi nulli, e valutando la corrente I_P fornita dal generatore V_P . In sostanza significa aggiungere un generatore in serie al generatore $V_U(0)$ della Figura 5.1.2 e calcolare la corrente erogata. Si ricorda che con il solo $V_U(0)$ la corrente erogata da quest'ultimo a riposo (ovvero proprio con i due ingressi nulli) sarebbe nulla. La configurazione considerata è quindi quella della Figura 5.1.3.

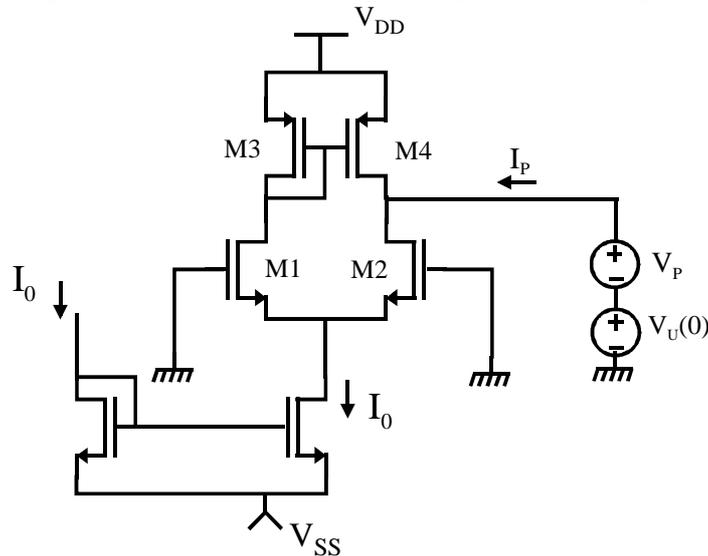


Figura 5.1.3

Per calcolare la I_P conviene per semplicità passare al circuito alle variazioni (mantenendo però i simboli dei transistori per poter meglio riconoscere configurazioni circuitali di cui si conoscono già le proprietà). Inoltre si può sdoppiare il generatore V_P in due generatori di uguale tensione, $V_{P1}=V_{P2}=V_P$, e dividere il nodo di uscita in due nodi di cui uno è connesso solo a V_{P1} , l'altro a V_{P2} . Si arriva quindi alla configurazione mostrata in Figura 5.1.4(a).

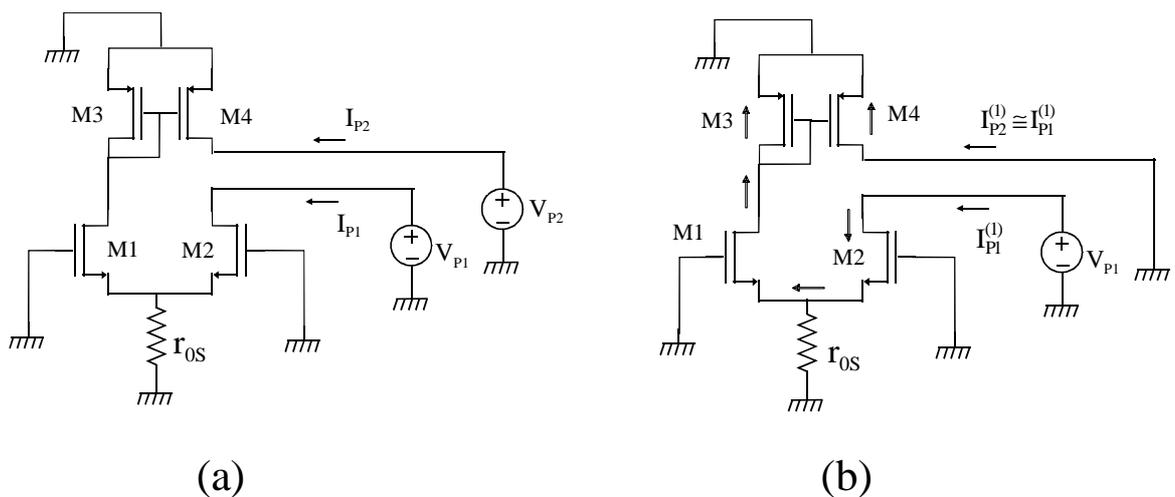


Figura 5.1.4

La corrente I_P cercata sarà la somma delle correnti I_{P1} e I_{P2} . Per calcolare I_{P1} e I_{P2} si può usare la sovrapposizione degli effetti e far agire V_{P1} e V_{P2} alternativamente. Si ha la situazione:

$$\text{agisce } V_{P1} \Rightarrow I_{P1} = I_{P1}^{(1)}; \quad I_{P2} = I_{P2}^{(1)}$$

$$\text{agisce } V_{P2} \Rightarrow I_{P1} = I_{P1}^{(2)}; \quad I_{P2} = I_{P2}^{(2)}$$

La corrente totale I_P sarà quindi data da:

$$I_P = I_{P1}^{(1)} + I_{P2}^{(1)} + I_{P1}^{(2)} + I_{P2}^{(2)} \quad (5.1.3)$$

Calcoliamo quindi tutte le componenti:

agisce V_{P2} : $V_{P2}=V_P$, $V_{P1}=0$; $I_{P1}^{(2)}$ è nulla perché V_{P2} entra sull'uscita dello specchio M3-M4 e quindi non ha una strada per arrivare sul generatore V_{P1} . $I_{P2}^{(2)}$ è invece pari a: V_P / r_{d4} .

agisce V_{P1} : $V_{P1}=V_P$, $V_{P2}=0$; La corrente $I_{P1}^{(1)}$ è data da: V_P / R_{VP1} , dove R_{VP1} è la resistenza vista dal generatore V_{P1} . Si osservi che entriamo sul drain di M2 il cui gate è a massa e il cui source vede una resistenza circa pari a $1/g_{m1}$ poichè dal source di M2 entriamo nel source di M1 (che ha il gate a massa e il drain quasi a massa, visto che dal drain di M1 andiamo verso massa con M3 connesso a diodo). La resistenza del generatore di corrente costituito da M5 si considera molto più grande di $1/g_{m1}$ e quindi si trascura nel parallelo. Pertanto:

$$R_{VP1} \cong r_{d2} \left(1 + g_{m2} \cdot \frac{1}{g_{m1}} \right)$$

Dato che $g_{m1}=g_{m2}$ si ha che $R_{VP1}=2r_{d2}$. Per valutare $I_{P2}^{(1)}$ si consideri il percorso compiuto dalla $I_{P1}^{(1)}$ dopo che essa è entrata nel drain di M2. Facendo riferimento alla Figura 5.1.4(b) si vede che $I_{P1}^{(1)}$ entra quasi completamente nel source di M1 verso il quale vede la minima resistenza, viene riflessa nello specchio M3-M4 (supponendo che esso specchi in modo unitario) e giunge sul generatore V_{P2} . Pertanto $I_{P2}^{(1)} \cong I_{P1}^{(1)} = V_P / 2r_{d2}$. Sommando i contributi si trova:

$$I_P \cong V_P \left(\frac{1}{r_{d4}} + 2 \frac{1}{2r_{d2}} \right) = V_P \left(\frac{1}{r_{d4}} + \frac{1}{r_{d2}} \right) \Rightarrow R_{out} = r_{d4} \parallel r_{d2} = r_{dp} \parallel r_{dn} \quad (5.1.4)$$

Sostituendo le espressioni delle r_d in termini di parametro λ e di I_D (uguale per M1, M2, M3, M4) si ottiene:

$$R_{out} = \frac{1}{I_D} \cdot \frac{1}{\lambda_p + \lambda_n}$$

Amplificazione di modo differenziale. Considerando l'effetto della sola tensione di modo differenziale, (si veda lo studio dell'amplificatore differenziale a carichi resistivi) risulta:

$$I_{D1} - I_{D2} = I_0 \frac{V_d}{V_{MAX}} \sqrt{2 - \frac{V_d^2}{V_{MAX}^2}} \quad (5.1.5)$$

$$\text{se } V_d \ll V_{dMAX} : I_{D1} - I_{D2} = g_{mn} V_d$$

dove g_{mn} indica il g_m dei transistori della coppia differenziale di ingresso M1e M2. Il guadagno di modo differenziale risulta quindi:

$$A_d = g_{mn} R_{out} = \frac{g_{mn}}{I_D} \cdot \frac{1}{\lambda_p + \lambda_n} \quad (5.1.6)$$

Sostituendo al rapporto g_{mn}/I_D il suo valore trovato nel capitolo 2, si giunge all'espressione del guadagno:

$$A_d = \frac{2}{(V_{GS} - V_t)} \frac{1}{(\lambda_n + \lambda_p)}$$

Dalle considerazioni fatte nel capitolo 2, sappiamo che la possibilità di ottenere amplificazioni qualsivoglia grandi facendo tendere $V_{GS}-V_t$ a zero è solo apparente, in quanto quando tale grandezza scende sotto il centinaio di mV, M1 e M2 entrano in debole inversione e il guadagno non aumenta più. Si potrebbe pensare allora di agire sulle lunghezze di canale, aumentandole in modo da ridurre i λ . Come vedremo ciò va a detrimento della risposta in frequenza, ovvero del prodotto guadagno-banda.

Per avere un'idea dell'entità di questa amplificazione si può far riferimento al caso in cui tutte le r_d e i g_m dell'amplificatore sono uguali. In tale caso l'amplificazione risulta:

$$A_d = \frac{1}{2} g_m r_d \quad (5.1.7)$$

I valori di amplificazione ottenibili sono quindi dell'ordine di $g_m r_d$, ovvero, a meno di ricorrere a transistori con lunghezze elevate, si possono normalmente ottenere guadagni dell'ordine del centinaio.

Amplificazione di modo comune. Per quanto visto nello studio del punto di riposo, applicando una tensione di modo comune V_c le correnti I_{D1} e I_{D2} rimangono uguali, pertanto le loro variazioni ΔI_{D1} e ΔI_{D2} sono identiche. Se lo specchio fosse ideale ($\Delta I_{D4}=\Delta I_{D1}$, $\Delta V_{DS3}=0$) allora si avrebbe una cancellazione perfetta e pertanto $I_{UCC}=0$ ovvero una $V_U=R_{out}I_{UCC}=0$ e quindi una $A_C=0$. In realtà lo specchio non è ideale: la tensione di ingresso allo specchio (V_{DS3}) varia al variare della I_{D1} mentre la tensione di uscita è stabilizzata dal corto circuito imposto per calcolare la I_{UCC} . Inoltre anche la V_{D2} è bloccata mentre la V_{D1} varia, pertanto con l'uscita in corto circuito possiamo aspettarci che neanche la relazione $\Delta I_{D2}=\Delta I_{D1}$ sia rigorosamente vera. Pertanto la cancellazione è imperfetta e si ha $V_C \neq 0$. Per calcolare l'amplificazione di modo comune si procede quindi rimuovendo il corto circuito in uscita e considerando che, come visto precedentemente, per $V_d=0$, $V_U=V_K$. Nota la variazione ΔI_{D1} (che sarà in queste condizioni, ovvero senza corto circuito, uguale a ΔI_{D2}) si ottiene la V_U come:

$$V_U|_{V_d=0} = V_K \cong \Delta I_{D1} \cdot \frac{1}{g_{m3}} \quad (5.1.8)$$

Questo perché dal nodo V_K si vede una resistenza verso massa che è sostanzialmente pari a quella di M3 connesso a diodo ($\cong 1/g_{m3}$).

Per calcolare le variazioni delle correnti I_{D1} e I_{D2} possiamo fare riferimento a quanto visto nel capitolo per l'amplificatore a carichi resistivi. Si era trovato che, in prima approssimazione le variazioni della V_S (tensione di source di M1 e M2) erano uguali alla tensione di modo comune, e quindi le variazioni delle I_D erano pari a $V_c/2r_{0s}$, dove r_{0s} è la resistenza di uscita dello specchio che fornisce la I_0 .

$$A_c = \left. \frac{V_U}{V_c} \right|_{V_d=0} \cong -\frac{1}{2g_{mp}r_{0s}} \quad (5.1.9)$$

Il **CMRR** risulta quindi:

$$CMRR = \left| \frac{A_d}{A_c} \right| = 2g_{mn}(r_{dp} \parallel r_{dn})g_{mp}r_{0s}$$

L'ordine di grandezza del CMRR è quello di $(g_m r_d)^2$, quindi dell'ordine di 10^4 (80 dB).

Cenni alla risposta in frequenza. Il limite superiore di banda di A_d , con ingressi prodotti da generatori ideali di tensione, può essere approssimato considerando che, quando i generatori di ingresso sono ideali, ovvero a resistenza interna nulla, l'unico nodo ad alta impedenza è l'uscita. Se C_o è la capacità tra uscita e massa (dovuta alla capacità di carico e alle capacità drain-body dei transistori M2 e M4) il limite superiore di banda f_H risulta:

$$f_H = \frac{1}{2\pi C_o R_{out}}$$

Il prodotto guadagno banda (PGB) risulta:

$$PGB = A_d f_H = g_{mn} R_{out} \frac{1}{2\pi C_o R_{out}} = \frac{g_{mn}}{2\pi C_o} \quad (5.1.10)$$

Per avere un'idea sugli effetti del dimensionamento del circuito sul PGB conviene esprimere il g_{mn} in funzione di $V_{GS}-V_t$ e considerare che la capacità C_o è, in prima approssimazione, data dalla seguente espressione:

$$C_o = C_L + k_C W$$

dove W è la larghezza di canale di M1 e M2, k_C è una costante di proporzionalità (dipendente da L e dalle regole di layout per i contatti) e C_L è la capacità di carico. Si consideri che con M2 e M4 in saturazione la C_{GD} è trascurabile rispetto alle capacità drain-body e che queste ultime sono proporzionali a W . Inoltre le W di M2 e M4 sono correlate, ovvero se si aumenta W_2 si aumenta di solito proporzionalmente la W_4 e viceversa. Si ottiene:

$$PGB = \frac{\mu_n C_{ox} \frac{W}{L} (V_{GS} - V_t)}{2\pi(C_L + k_C W)} = \frac{\mu_n C_{ox} \frac{1}{L} (V_{GS} - V_t)}{2\pi\left(\frac{C_L}{W} + k_C\right)}$$

Osserviamo che inizialmente aumentare W "paga", in quanto si riduce al denominatore il termine C_L/W . Quando questo valore diventa trascurabile rispetto a k_C , aumentare W significa solo aumentare gli ingombri senza benefici. Vediamo invece che ridurre L e aumentare $V_{GS}-V_t$ aumenta il PGB ma, per quanto visto precedentemente, peggiora l'amplificazione in continua (A_d). Occorre quindi ricercare una soluzione di compromesso tra guadagno in continua e PGB.

Se, come accade in pratica, i generatori di segnali non sono ideali ma hanno una resistenza interna R_t , vi è anche un polo di ingresso dovuto alla capacità di ingresso dello stadio. Particolarmente importante può essere la capacità sull'ingresso 2 (generatore V_2) in quanto il condensatore C_{GD2} risulta collegato tra ingresso e uscita e quindi viene riportato ingresso

moltiplicato per il fattore $(1+A_d)$ a causa dell'effetto Miller. Questo effetto è rilevante solo sull'ingresso 2 (V_{DI} è connesso a massa tramite M_3 connesso a diodo, ovvero tramite una bassa resistenza differenziale e quindi l'amplificazione tra V_I e V_{DI} è piccola). Ciò è causa di sbilanciamento e non consente l'uso di questo amplificatore in circuiti in cui è necessario un bilanciamento tra i due ingressi anche a frequenze elevate.

Comportamento con ampi segnali. Il circuito equivalente di Figura 5.1.2(a) può essere utilizzato anche per segnali ampi purché si consideri che R_{out} non è costante ma varia in funzione della V_u e pure il legame tra I_{UCC} e la tensione V_d non è lineare ma dato dall'equazione (5.1.5). Si può calcolare la caratteristica per punti, considerando che una variazione della V_U si può calcolare come segue:

$$\Delta V_U = \Delta V_d \cdot g_m(V_d) \cdot R_{out}(V_U)$$

dove è stato esplicitato che la resistenza di uscita è una funzione della V_U mentre il g_m equivalente (pendenza della curva corrispondente all'equazione (5.1.5)) è una funzione di V_d . Il circuito di Figura 5.1.2(a) rappresenta ovviamente le variazioni alle quali va aggiunta la $V_U(0)$. La curva tensione di uscita in funzione della tensione V_d è riportata in Figura 5.1.5.

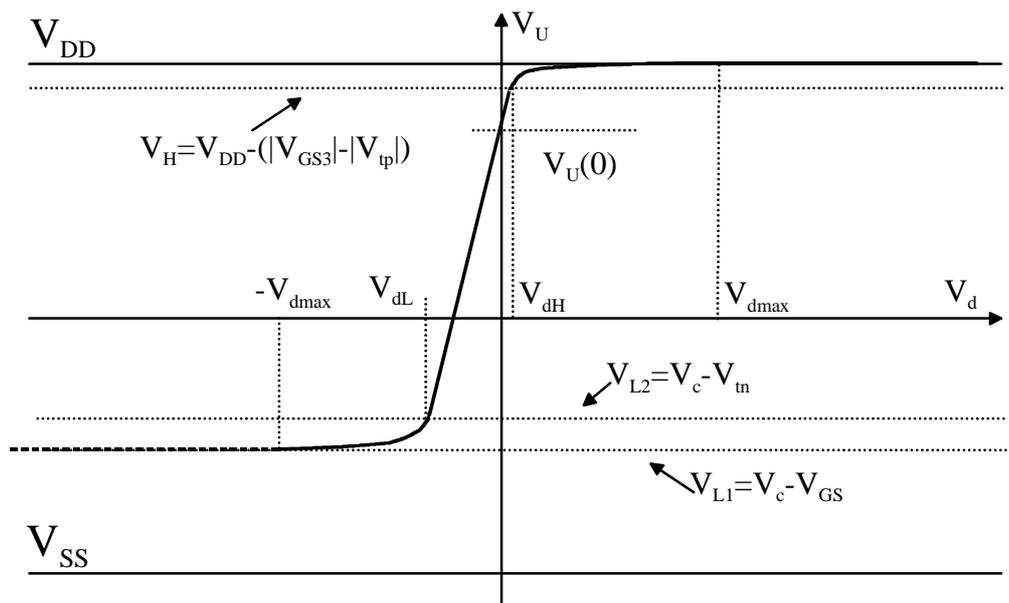


Figura 5.1.5

Si noti che la pendenza della curva è pari all'amplificazione per piccoli segnali. La curva rimane lineare nell'intervallo $V_{L2} < V_U < V_H$ nel quale entrambi i MOS M_2 e M_4 sono in saturazione. Per $V_U > V_H$ M_4 va in zona triodo e la sua r_d (ovvero la r_{dp}) si abbassa, riducendo l'amplificazione (Eq. (5.1.4) e (5.1.6)) e quindi, progressivamente, la pendenza. Lo stesso accade per $V_U < V_{L2}$, ma qui è M_2 ad andare in zona triodo e quindi la R_{out} si riduce a causa di r_{dn} . Imponendo la condizione di saturazione per M_4 e poi per M_2 si ottengono le relazioni:

$$V_H = V_{DD} - (|V_{GS4}| - |V_{tp}|) \quad (5.1.11)$$

$$V_{L2} = V_2 - V_m = V_c - \frac{V_d}{2} - V_m \cong V_c - V_m$$

L'espressione finale di V_{L2} è stata ottenuta considerando che la V_d non può eccedere la dinamica di ingresso $V_{d\max} = \sqrt{2I_0 / \beta_n}$ e quindi a trascurare $V_d/2$ si compie al più un errore di alcune decine di mV. All'interno dell'intervallo $V_{dL} < V_d < V_{dH}$ (ovvero per $V_{L2} < V_U < V_H$) la resistenza di uscita rimane approssimativamente costante. Inoltre, siccome generalmente l'intervallo $[V_{dL}, V_{dH}]$ è solo un piccolo sottoinsieme dell'intervallo $[-V_{d\max}, V_{d\max}]$, la relazione tra I_{UCC} e V_d rimane lineare in $[V_{dL}, V_{dH}]$. Per questi motivi il tratto di curva tra V_{dL} e V_{dH} è sostanzialmente lineare. Aumentando la V_d oltre V_{dH} si ottiene che, superata la $V_{d\max}$, M2 è completamente spento e M1, M3 conducono tutta la I_0 . Anche M4 è in conduzione, ($V_{gs4} < V_{tp}$) ma dovendo risultare e $I_{ds4} = I_{ds2} = 0$ dovrà anche essere $V_{ds4} = 0$. Per cui per $V_d > V_{d\max}$ si ha $V_U = V_{dd}$. La situazione per $V_U < V_{L2}$ è più complessa, in quanto M2 entra in zona triodo e l'analisi della coppia differenziale che ha portato all'equazione (5.1.5) non è più valida. In prima istanza la tensione tende ad arrivare ad un limite inferiore indicato in figura con V_{LI} , corrispondente all'annullamento della V_{DS2} (andamento a tratteggio in Figura 5.1.5). Pertanto $V_{LI} = V_c - V_{GS}$. Riducendo ulteriormente la V_d si può osservare un graduale aumento della tensione di uscita. In ogni caso l'amplificatore funziona correttamente (amplificazione elevata) solo per $V_{dL} < V_d < V_{dH}$.

Dinamica di uscita. Si ricava dall'analisi precedente ed è data da:

$$V_c - V_m < V_U < V_{DD} - (|V_{GS3}| - |V_{tp}|)$$

Osserviamo che la tensione di uscita può arrivare, in condizioni di funzionamento normali, quasi alla V_{DD} ma è limitata in basso dalla tensione di modo comune. Questo limita fortemente l'applicazione di questo schema qualora siano richieste ampie dinamiche di uscita.

Dinamica di modo comune (di ingresso). Si supponga che sia applicata solo una tensione di modo comune V_c e pertanto si ha $V_{GS1} = V_{GS2} = V_{GS}$. Il limite inferiore deriva dal fatto che diminuendo V_c , diminuisce praticamente della stessa entità la tensione di source di M1 e M2 e quindi la tensione di uscita dello specchio di corrente che fornisce la I_0 . Se indichiamo con V_{MIN} la tensione minima tra i terminali di uscita dello specchio affinché quest'ultimo funzioni correttamente si ha:

$$V_c - V_{GS} - V_{SS} > V_{MIN} \quad \Rightarrow \quad V_c > V_{SS} + V_{GS} + V_{MIN} \quad (5.1.12)$$

Il modo comune di ingresso non può quindi estendersi fino alla V_{SS} ma deve essere superiore a questa di una V_{GS} (quindi di una tensione dell'ordine del volt) sommata alla V_{MIN} . Nel caso dell'amplificatore di Figura 5.1.1 viene utilizzato uno specchio semplice e quindi la V_{MIN} è pari a $V_{GS5} - V_T$, ovvero al più qualche centinaio di mV. L'utilizzo di specchi più complessi come lo specchio cascode avrebbe benefici sul CMRR (r_{os} più elevata, si veda l'Eq. (5.1.9)) ma peggiorerebbe il limite inferiore del modo comune. Se si scende sotto il livello minimo indicato dalla (5.1.12) si riduce la I_0 fornita dallo specchio la quale si annulla totalmente (e quindi l'amplificatore cessa totalmente di funzionare) quando sullo specchio si ha una differenza di potenziale nulla, ovvero quando $V_c = V_{SS} + V_{GS}$. (con V_{GS} che vale V_m a causa dell'annullamento delle I_D). Per quanto riguarda il limite superiore si noti che essendo $I_{D1} = I_{D2} = I_0/2$ circa costante, il drain di M1 è vincolato al potenziale: $V_{D1} = V_{DD} - |V_{GS3}|$ mentre suo gate è pari proprio a V_c . Pertanto si arriverà ad un valore massimo di V_c , oltre il quale M1

va in zona triodo e l'amplificatore non funziona più correttamente, riducendosi il guadagno progressivamente a zero per la riduzione di g_{m1} . Il massimo valore di V_c sarà quindi dato da:

$$V_{D1} > V_{G1} - V_m \Rightarrow V_{DD} - |V_{GS3}| > V_c - V_m \Rightarrow V_c < V_{DD} - (|V_{GS3}| - V_m) \quad (5.1.13)$$

Si noti che il limite superiore è molto meno stringente in quanto V_c si può avvicinare alla V_{DD} a meno del termine $|V_{GS3}| - V_m = |V_{GS3}|/|V_{tp}| + (|V_{tp}| - V_m)$. Si noti che $|V_{GS3}|/|V_{tp}|$ può essere ridotto a qualche centinaio di mV, mentre $(|V_{tp}| - V_m)$ è di solito negativo, perchè la V_m è affetta da effetto body e quindi è molto più grande del valore per $V_{BS}=0$, che generalmente è prossimo a $|V_{tp}|$. Si comprende allora che $|V_{GS3}| - V_m$ può essere negativo e quindi la V_c può addirittura superare, anche se di poco, la V_{DD} . Non si dimentichi però che in queste condizioni (valori estremi di V_c) la dinamica di uscita risulta assai ridotta.

5.2 Amplificatori differenziali cascode a MOSFET

Abbiamo visto che l'amplificatore differenziale semplice ha un guadagno dato dal prodotto $g_m R_{out}$. È stato anche rilevato che un aumento del guadagno porta inevitabilmente una diminuzione del prodotto guadagno banda. Per aumentare il guadagno senza incorrere in questa limitazione si può cambiare la topologia circuitale in modo da aumentare la R_{out} senza modificare il g_m . Il metodo più semplice è ricorrere ad un amplificatore cascode, ovvero ad una cascata di un source comune e di un gate comune.

Amplificatore differenziale cascode.

La struttura di un amplificatore cascode differenziale è mostrata nella Figura 5.2.1

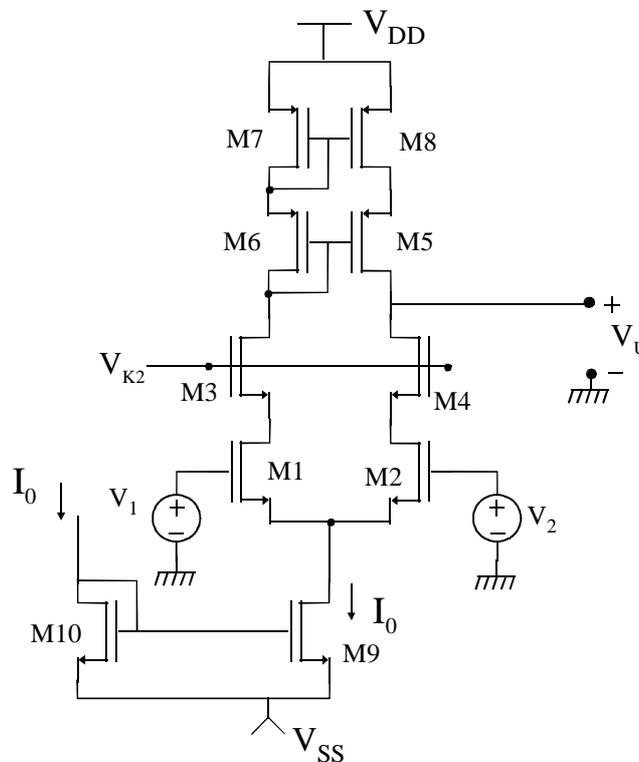


Figura 5.2.1

Lo stadio a source comune è costituito dai transistori M1 e M2 (si consideri che per piccoli segnali di modo differenziale essi hanno effettivamente il source a massa); lo stadio a gate comune (differenziale) è costituito dai transistori M3 e M4. Lo specchio cascode M5,M6,M7,M8 sostituisce lo specchio semplice dell'amplificatore di Figura 5.1.1. Ciò è necessario in quanto l'alta impedenza di uscita dello stadio amplificatore cascode non potrebbe essere sfruttata usando uno specchio di carico a bassa resistenza come lo specchio semplice. Si dimostra come nel caso precedente che la corrente di cortocircuito per piccoli segnali è data da: $g_{m1}V_d$. Con procedimento esattamente analogo a quello utilizzato per l'amplificatore di Figura 5.1.1 si ricava che la resistenza di uscita è data da:

$$R_{out} \cong (g_{mn}r_{dn})r_{dn} \parallel (g_{mp}r_{dp})r_{dp}$$

dove g_{mn} e r_{dn} sono relativi a M1,M2,M3,M4 (supposti uguali) e g_{mp} , r_{dp} sono relativi a M5,M6,M7,M8 (anche questi uguali). Si può notare che la resistenza di uscita è pari a al parallelo delle resistenze di uscita due specchi cascode (uno di tipo n e l'altro di tipo p). Come ordine di grandezza è pari a quella dell'amplificatore di Figura 5.1.1 moltiplicata per un termine $g_m r_d$, che tipicamente è dell'ordine di un centinaio. Pertanto questo stadio può raggiungere amplificazioni notevolmente superiori a quelle dello stadio di Figura 5.1.1. Una limitazione all'impiego di questo stadio è il fatto che anche qui la dinamica di uscita è legata al modo comune in ingresso: la tensione V_{K2} limita in alto il range del modo comune e in basso la dinamica di uscita. Infatti, affinché M2 e M1 rimangano in saturazione si dovrà avere:

$$V_{D1} = V_{K2} - V_{GS3} > V_{G1} - V_m = V_c - V_m \Rightarrow V_c < V_{K2} - (V_{GS3} - V_m)$$

Inoltre affinché M4 rimanga in saturazione si dovrà avere:

$$V_{D4} = V_U > V_{G4} - V_m \Rightarrow V_U > V_{K2} - V_m$$

Pertanto la V_{K2} agisce come "spartiacque" tra la dinamica dell'uscita e quella del modo comune. Lo stadio può essere quindi impiegato solo nei casi in cui non si voglia contemporaneamente un'ampia dinamica del modo comune in ingresso e dell'uscita. Un altro inconveniente di questo stadio è il fatto che in virtù del grande numero di transistori "impilati" esso non è adatto a lavorare con basse tensioni di alimentazione.

Amplificatore folded cascode.

L'amplificatore folded cascode (cascode ripiegato) supera i limiti visti per l'amplificatore cascode. La sua struttura è mostrata in Figura 5.2.2.

In questo amplificatore il source comune è sempre costituito da due n -MOS (M1 e M2) mentre il gate comune è a p -MOS (M6 e M7). Naturalmente esiste anche una versione duale con tutti i transistori di tipo opposto a quelli di figura Figura 5.2.2. V_{K2} e V_{K3} sono tensioni costanti di polarizzazione. M3 e M4 agiscono da generatori di corrente per la corretta polarizzazione del circuito.

Punto di riposo: $I_{D1}=I_{D2}=I_0/2$. $I_{D8}=I_{D9}=I_1-I_0/2$. Le correnti I_0 e I_1 sono imposte, I_{D8} ($=|I_{D6}|$) e I_{D9} ($=|I_{D7}|$)) derivano da esse e, per la corretta polarizzazione dello stadio, devono essere non nulle (e, naturalmente, del verso consentito dal tipo di MOSFET, ovvero dirette verso la V_{SS}).

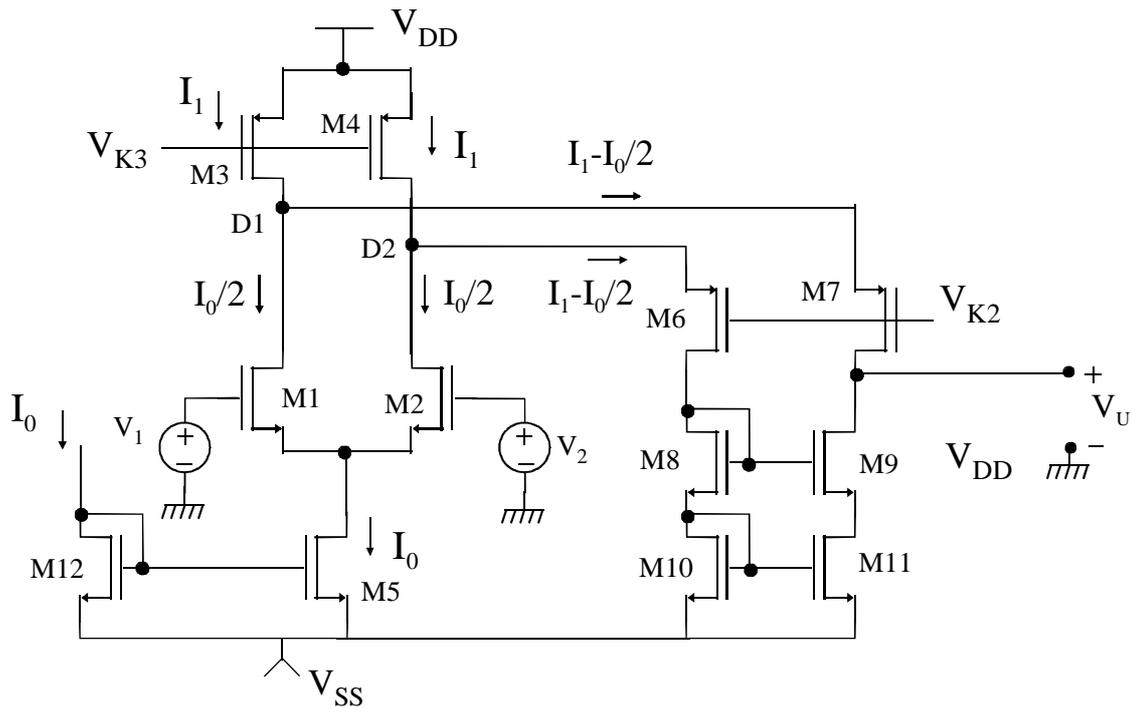


Figura 5.2.2

Affinché ciò accada dovrà essere:

$$I_1 - \frac{I_0}{2} > 0$$

Funzionamento del circuito. Si procede come nel caso dell'amplificatore differenziale semplice, calcolando prima una corrente di cortocircuito di uscita e poi calcolando la R_{out} . La corrente di cortocircuito risulta data da:

$$I_{UCC} = I_{D7} - I_{D9} = (I_1 - I_{D1}) - (I_1 - I_{D2}) = -(I_{D1} - I_{D2})$$

Si faccia attenzione al fatto che per il calcolo dell'espressione precedente si è considerato che in un transistor p -MOS una I_{DS} positiva è una corrente che scorre dal source verso il drain (verso naturale della corrente).

La corrente di cortocircuito ha quindi un'espressione analoga a quella dell'amplificatore differenziale semplice, essendo anche qui costituita dalla differenza tra le correnti di drain di due MOSFET accoppiati di emettitore e quindi è data dall'equazione (5.1.5) la quale è perfettamente valida in quanto le tensioni di drain di M1 e M2 sono bloccate dalla bassa impedenza degli stadi a gate comune M6-M7. Per piccole tensioni V_d anche qui abbiamo:

$$I_{UCC} = g_{m1} V_d$$

La resistenza di uscita si calcola come al solito introducendo un generatore di prova V_P in uscita e calcolando la I_P . Come per il differenziale semplice si sdoppia V_P in $V_{P1} = V_{P2} = V_P$ e si calcolano le quattro componenti della I_P :

$$I_P = I_{P1}^{(1)} + I_{P2}^{(1)} + I_{P1}^{(2)} + I_{P2}^{(2)}$$

La Figura 5.2.3 mostra il circuito utilizzato per il calcolo della I_P .

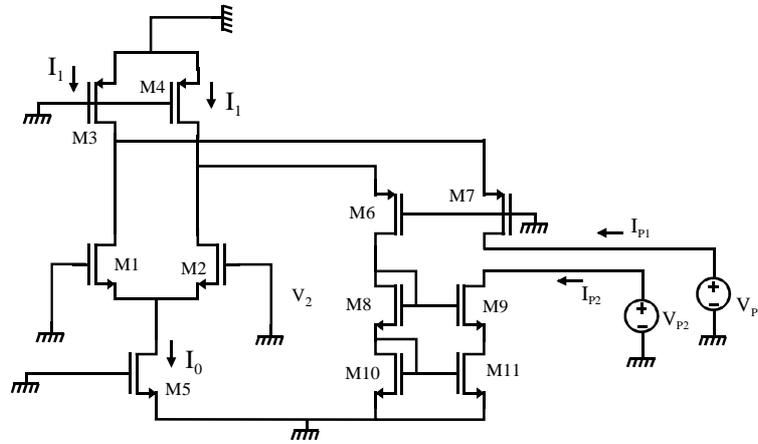


Figura 5.2.3

Quando agisce V_{P2} (e V_{P1} è cortocircuitato) si nota che V_{P2} entra sull'uscita di uno specchio e vede quindi una resistenza circa pari a $R_{VP2} = (g_{m9} r_{d11}) r_{d9}$. Pertanto la corrente erogata da V_{P2} sarà:

$$I_{P2}^{(2)} \cong \frac{V_P}{(g_{m9} r_{d11}) r_{d9}}$$

D'altro canto, la corrente erogata da V_{P2} si ferma sul lato di uscita dello specchio e V_{P2} non ha modo di indurre corrente nel ramo del generatore V_{P1} . Pertanto:

$$I_{P1}^{(2)} = 0$$

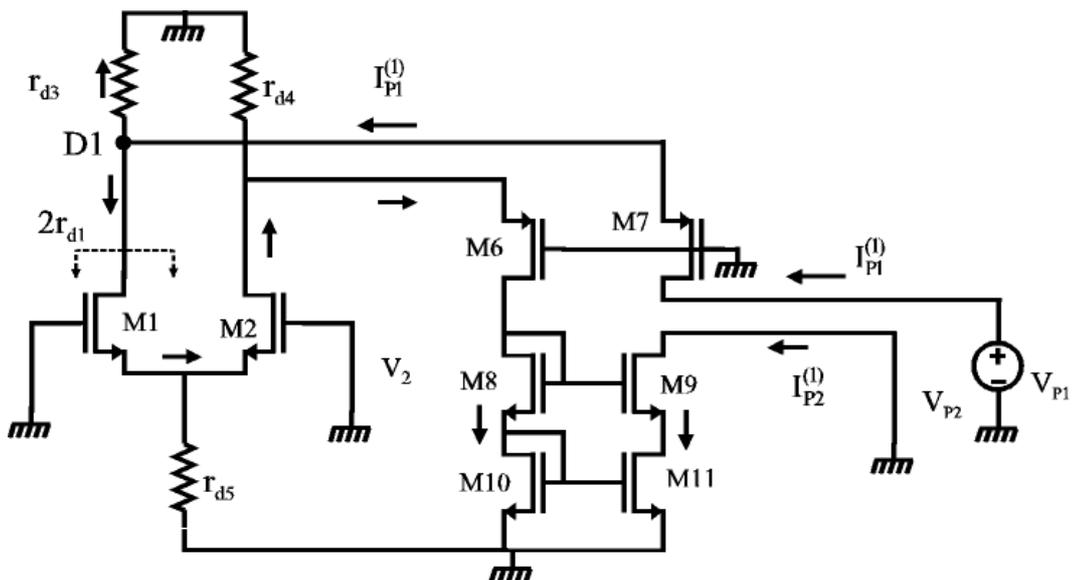


Figura 5.2.4

Quando agisce V_{P1} la situazione è un po' più complessa. Come si può osservare nella Figura 5.2.4, V_{P1} "entra" sul drain di M7 che ha sul source il parallelo di r_{d3} e di quello che si vede verso il drain di M1.

Ripetendo la considerazione fatta nel caso del differenziale semplice si osserva che dal drain di M1 si vede una resistenza circa pari a $2r_{d1}$ (si ricordi che il source di M1 è connesso verso massa attraverso una resistenza che sostanzialmente è pari a $1/g_{m2} = 1/g_{m1}$). Pertanto la resistenza vista da V_{P2} sarà data da:

$$R_{VP1} = (g_{m7} \cdot r_{d3} \parallel 2r_{d1})r_{d7} = g_{m7}r_{d7} \frac{2r_{d1}r_{d3}}{2r_{d1} + r_{d3}}$$

e sarà inoltre: $I_{P1}^{(1)} = V_P / R_{VP1}$. La corrente $I_{P1}^{(1)}$, arrivata al nodo D1 si divide in due componenti, di cui una scorre verso massa attraverso r_{d3} , mentre l'altra arriva praticamente inalterata sul generatore V_{P2} seguendo il percorso indicato nella figura precedente (incontra solo diramazioni dove una via è a impedenza molto più bassa e raccoglie quindi praticamente tutta la corrente). La componente che prosegue verso V_{P2} si può calcolare applicando la formula del partitore di corrente a $I_{P1}^{(1)}$ che giunge sul nodo D1. Pertanto abbiamo:

$$I_{P2}^{(1)} = I_{P1}^{(1)} \frac{r_{d3}}{2r_{d1} + r_{d3}}$$

Raccogliendo infine le componenti della corrente I_P si ottiene:

$$\begin{aligned} I_P &= I_{P1}^{(1)} + I_{P2}^{(1)} + I_{P2}^{(2)} = \frac{V_P}{g_{m9}r_{d11}r_{d9}} + I_{P1}^{(1)} \left(1 + \frac{r_{d3}}{2r_{d1} + r_{d3}} \right) = \\ &= \frac{V_P}{g_{m9}r_{d11}r_{d9}} + 2I_{P1}^{(1)} \left(\frac{r_{d1} + r_{d3}}{2r_{d1} + r_{d3}} \right) = \frac{V_P}{g_{m9}r_{d11}r_{d9}} + \frac{2V_P(2r_{d1} + r_{d3})}{2g_{m7}r_{d7}r_{d1}r_{d3}} \left(\frac{r_{d1} + r_{d3}}{2r_{d1} + r_{d3}} \right) \end{aligned}$$

Semplificando l'ultima espressione trovata si ottiene:

$$R_{out}^{-1} = \frac{I_P}{V_P} = \frac{1}{g_{m9}r_{d11}r_{d9}} + \frac{1}{g_{m7}r_{d7}r_{d3}} \left(\frac{r_{d1} + r_{d3}}{r_{d1}} \right)$$

La resistenza di uscita R_{out} è quindi pari al parallelo delle due resistenze:

$$R_n = (g_{m9}r_{d11})r_{d9}; \quad R_P = \left(g_{m7}r_{d3} \frac{r_{d1}}{r_{d1} + r_{d3}} \right) r_{d7}$$

La R_n è la resistenza di uscita dello specchio di carico, la R_P è pari alla resistenza di uno specchio costruito con i transistori M7 e M3, moltiplicata per il fattore $r_{d1}/(r_{d1}+r_{d3})$. Per valutare l'ordine di grandezza della R_{out} si può, per semplicità, supporre che tutte le r_d e i g_m dell'amplificatore siano uguali: si ottiene un valore di resistenza di uscita:

$$R_{out} \approx \frac{1}{3} (g_m r_d) \cdot r_d$$

L'amplificazione che si può ottenere è quindi dell'ordine di:

$$A_d = g_m R_{out} \approx \frac{1}{3} (g_m r_d)^2$$

Confrontando questa amplificazione con quella dell'equazione (5.1.7) si evince che il folded cascode ha un guadagno superiore a due amplificatori semplici collegati in cascata.

Amplificazione di modo comune. Si dimostra semplicemente che, nonostante l'asimmetria dovuta al collegamento a diodo di M8-M10, per tensioni differenziale nulle (ovvero in presenza di sola V_c), si ha $I_{D8} = I_{D10}$. Con argomentazioni analoghe a quelle utilizzate per il punto di riposo dell'amplificatore differenziale semplice se ne deduce che $V_{DS8} = V_{DS9}$ e $V_{DS10} = V_{DS11}$. Pertanto la tensione sul drain di M9 (ovvero la V_U) deve essere uguale alla tensione sul drain di M8 (ricordiamo, solo per $V_d = 0$). Quindi basta calcolare le variazioni della V_{D8} e si ha il segnale in uscita (variazioni). Come per l'amplificatore differenziale semplice, anche qui un modo comune provoca variazioni di corrente in M1 e M2 date da:

$$\Delta I_D \Big|_{V_d=0} = \frac{V_c}{2r_{os}}$$

dove r_{os} è la resistenza di uscita dello specchio (M5 in Figura 5.2.2) che fornisce la I_0 . La V_{D8} è chiaramente data dal prodotto di questa variazione di corrente per la resistenza che connette a massa il drain di M8. Questa è pari alla serie delle resistenze differenziali di M8 e M10 connessi a diodo ed è quindi pari a $2/g_m$. Si ha quindi infine:

$$A_c = \frac{V_U}{V_c} \Big|_{V_d=0} = \frac{V_{D8}}{V_c} \Big|_{V_d=0} = -\frac{1}{g_m r_{os}}$$

L'amplificazione di modo comune è quindi doppia rispetto a quella dell'amplificatore differenziale semplice ma A_d è molto più grande e quindi anche il **CMRR** è molto più elevato. Dinamica di ingresso (modo comune) e di uscita. Affrontiamo prima la dinamica di uscita: in basso è limitata dalla dinamica dello specchio cascode di carico e potrà arrivare quindi fino a $V_{SS} + 2(V_{GS} - V_{Tn}) + V_{Tn}$. Dunque:

$$V_U > V_{SS} + 2(V_{GS} - V_{Tn}) + V_{Tn} \quad (5.2.1)$$

In alto il limite sarà dato dalla saturazione di M7:

$$|V_{DS7}| > |V_{GS7}| - |V_{tp}| \Rightarrow V_{S7} - V_{D7} > V_{S7} - V_{G7} - |V_{tp}|$$

Si noti che è stato possibile sostituire l'operatore di modulo con un cambio di segno dell'argomento perché già si sapeva che gli argomenti sono negativi. Semplificando e sostituendo $V_{D7} = V_U$ e $V_{G7} = V_{K2}$ si ottiene:

$$V_U < V_{K2} + |V_{tp}| \quad (5.2.2)$$

Per quanto riguarda il modo comune in ingresso, è chiaro che in basso sussiste la stessa

limitazione dell'amplificatore differenziale semplice, per identità di struttura. Ovvero:

$$V_c > V_{SS} + V_{GS} + V_{MIN} \quad (5.2.3)$$

dove V_{MIN} è la minima tensione sui terminali di uscita dello specchio che fornisce la I_0 (nel nostro caso $V_{MIN}=(V_{GS}-V_{Tn})$). In alto, invece il limite è dato dall'entrata in zona triodo di M1 e M2. Fissiamoci su M1. Deve risultare:

$$V_{D1} = V_{K2} + |V_{GS7}| > V_{G1} - V_m = V_c - V_m$$

Pertanto si avrà:

$$V_c < V_{K2} + |V_{GS7}| + V_m$$

Osserviamo subito che, a differenza di quanto accadeva per il cascode (non folded), fissare V_{K2} ad un valore più alto possibile ha qui effetti benefici sia sulla dinamica di uscita sia su quella di modo comune in quanto estende verso l'alto il limite superiore di entrambe. Il massimo valore che può essere assegnato alla V_{K2} è quello che posizione M3 e M4 al limite della zona di saturazione. Andare oltre li porterebbe in zona triodo. Pertanto:

$$V_{K2MAX} = V_{DD} - (|V_{GS3}| - |V_{tp}|) - |V_{GS7}|$$

Con questo valore di V_{K2} si ottengono i seguenti limiti superiori per la dinamica di uscita e il modo comune di ingresso:

$$V_U < V_{DD} - 2(|V_{GSP}| - |V_{tp}|) \quad (5.2.4)$$

$$V_c < V_{DD} - (|V_{GSP}| - |V_{tp}|) + V_m \quad (5.2.5)$$

Le due espressioni precedenti sono state ottenute considerando che le $V_{GS3} = V_{GS7} = V_{GSP}$ (almeno come ordine di grandezza). Importante: nel folded cascode la V_U può raggiungere la V_{DD} a meno di due ($V_{GS}-V_t$), ovvero di qualche centinaio di mV e contemporaneamente la dinamica di modo comune può addirittura superare la V_{DD} !. In basso le condizioni sono notevolmente peggiori in quanto le equazioni (5.2.1) e (5.2.3) indicano che la distanza dalla V_{SS} contiene un termine V_t e quindi non può scendere sotto un valore che tipicamente è dell'ordine di 1 V.

5.3 Amplificatori a MOSFET a larga dinamica (OTA e OTA rail-to-rail)

L'amplificatore differenziale semplice mostrato in Figura 5.1.1 ha una limitazione importante che consiste nel fatto che il modo comune limita in basso la dinamica di uscita. La versione duale di questo amplificatore, con transistori p -MOS di ingresso, mostrata nella figura seguente, ha limitazione opposta, ovvero il modo comune limita in alto la dinamica di uscita. Pertanto, se è importante che la dinamica di uscita arrivi fino alla V_{SS} (ma non alla V_{DD}), in alternativa alla versione a n -MOS, di ingresso si può usare quella a p -MOS.

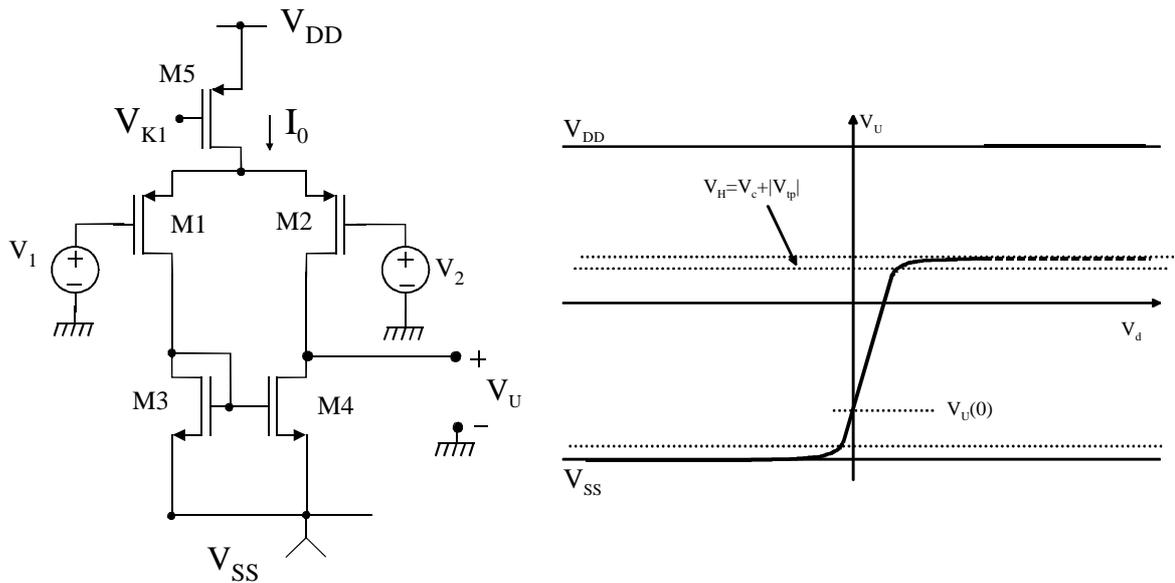


Figura 5.3.1

L'amplificatore a p -MOS non risolve però il problema di avere una dinamica di uscita che, indipendentemente dal modo comune in ingresso, si estende dalla V_{DD} alla V_{SS} . Per fare questo bisogna che il terminale di uscita e uno dei due di ingresso non si trovino sullo stesso transistor. Un modo per ottenere ciò potrebbe essere il folded cascode visto nel paragrafo 5.2, che però presenta qualche limitazione sulla dinamica di uscita, introdotta dallo specchio cascode. Inoltre il folded cascode presenta maggiore complessità circuitale, e, soprattutto, tensioni di alimentazione più elevate. Una soluzione più semplice che mantiene tutte le caratteristiche dell'amplificatore differenziale semplice (A_d , A_c , R_{out} ...) ed è molto adatta ad essere utilizzata con bassi valori delle tensioni di alimentazione è mostrata nello schema della Figura 5.3.2. Il funzionamento di questo schema è il seguente

- La corrente I_{D7} è pari alla corrente I_{D2} , a causa dello specchio M4-M7.
- La corrente I_{D6} è pari a I_{D1} , a causa dello specchio M3-M6.
- La corrente I_{D9} è pari a I_{D6} , a causa dello specchio M8-M9; pertanto $I_{D9} = I_{D1}$.

A questo punto basta notare che, come al solito, chiudendo in cortocircuito l'uscita si ottiene la corrente:

$$I_{UCC} = I_{D7} - I_{D9} = I_{D2} - I_{D1}$$

Quindi, a parte un'inversione di segno, la corrente di cortocircuito è uguale a quella dello stadio differenziale semplice (che, ricordiamo era pari a $I_{D1} - I_{D2}$). In questo schema il terminale non invertente è quindi quello contrassegnato con V_2 e quello invertente con V_1 . Per cui $V_d = V_2 - V_1$. La resistenza di uscita di questo circuito si calcola semplicemente osservando che dall'uscita vedo solo il parallelo delle resistenze di drain di M7 e M9. Per piccoli segnali $i_{ucc} = g_{m1} v_d$. Quindi:

$$A_d = g_{m1} R_{out} = g_{m1} (r_{d7} \parallel r_{d9})$$

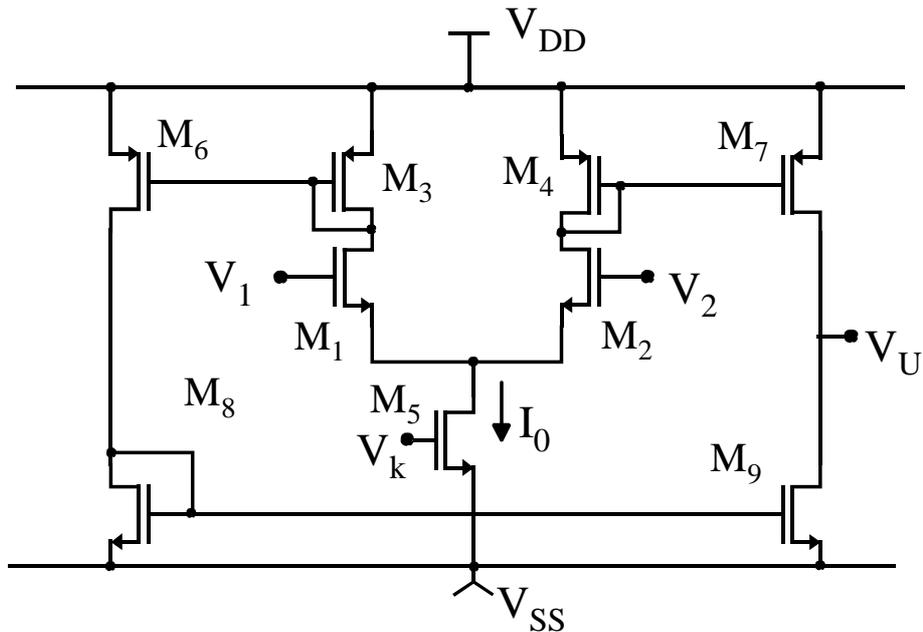


Figura 5.3.2

Si verifica facilmente che anche questo circuito si comporta bene per quanto riguarda l'amplificazione di modo comune in quanto le variazioni di i_{d1} e i_{d2} in presenza di solo modo comune sono uguali e, a parte l'errore introdotto dagli specchi di corrente, si sottraggono in uscita.

La dinamica di uscita è praticamente estesa dalla V_{SS} alla V_{DD} . Più precisamente l'amplificazione (ovvero la pendenza della caratteristica) rimane praticamente costante per:

$$V_{SS} + (V_{GS9} - V_m) < V_U < V_{DD} - (|V_{GS7}| - |V_{tp}|)$$

Quindi ci possiamo avvicinare alla V_{DD} e alla V_{SS} a meno di una tensione di saturazione ($V_{GS} - V_t$) ovvero di una quantità che può essere ridotta fino a un centinaio di mV. Oltre questi estremi la resistenza di uscita diminuisce rapidamente e con essa pure la pendenza della caratteristica. Aumentando la V_d fino al valore V_{dmax} della coppia differenziale M1-M2 si raggiungono comunque i rail di alimentazione. L'amplificatore di Figura 5.3.2 realizza la forma più semplice di OTA, ovvero Operational Transconductance Amplifier. Per quanto visto esso è *rail-to-rail* in uscita.

Per quanto riguarda la dinamica del modo comune in ingresso, l'amplificatore di Figura 5.3.2 non presenta alcun vantaggio rispetto all'amplificatore differenziale semplice di Figura 5.1.1. Infatti esso smette di funzionare quando V_c scende sotto $V_{SS} + V_{GS1} + (V_{GS5} - V_m)$. In alto, abbiamo visto che V_c può invece raggiungere la V_{DD} .

La versione duale, a transistori *p*-MOS in ingresso, permette invece di arrivare con la V_c fino alla V_{SS} , ma presenta limitazioni in alto. La figura seguente riassume schematicamente i limiti della dinamica del modo comune per i due tipi di OTA.

La Figura 5.3.3 suggerisce un modo per ottenere un amplificatore il cui range di ingresso copra tutto l'intervallo VSS-VDD. L'idea è quella di connettere "in parallelo" due OTA duali, in modo che almeno uno dei due sia sempre in funzione. Per fare ciò occorre che i due amplificatori abbiano lo stesso segnale di ingresso e producano correnti di corto circuito che

si sommano nel nodo di uscita. La Figura 5.3.4 mostra lo schema più comune di amplificatore differenziale rail-to-rail in ingresso e uscita. In pratica i transistori M11-M15 realizzano l'amplificatore a p -MOS, mentre M21-M25 quello a n -MOS.

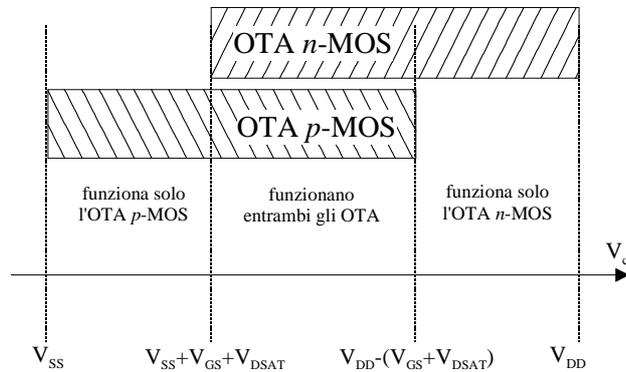


Figura 5.3.3

Si può verificare rapidamente che $I_{D24} = I_{D12} + I_{D22}$ e $I_{D13} = I_{D11} + I_{D21}$. I_{D24} e I_{D13} si specchiano nei due transistori dello stadio di uscita.

Per cui si verifica facilmente che $I_{UCC} = (I_{D12} - I_{D11}) + (I_{D22} - I_{D21})$. Osservando che le differenze nelle due parentesi sono proprio le correnti di cortocircuito dei due stadi differenziali p ed n , si verifica che il circuito funziona come anticipato. L'unico inconveniente è che il g_m complessivo dello stadio (pari a i_{ucc}/v_d) varia a seconda della zona di in cui l'amplificatore si trova a lavorare. I transistori MA1 e MA2 servono a sommare le correnti dello stadio, mentre MB1-MB3 forniscono le correnti di polarizzazione.

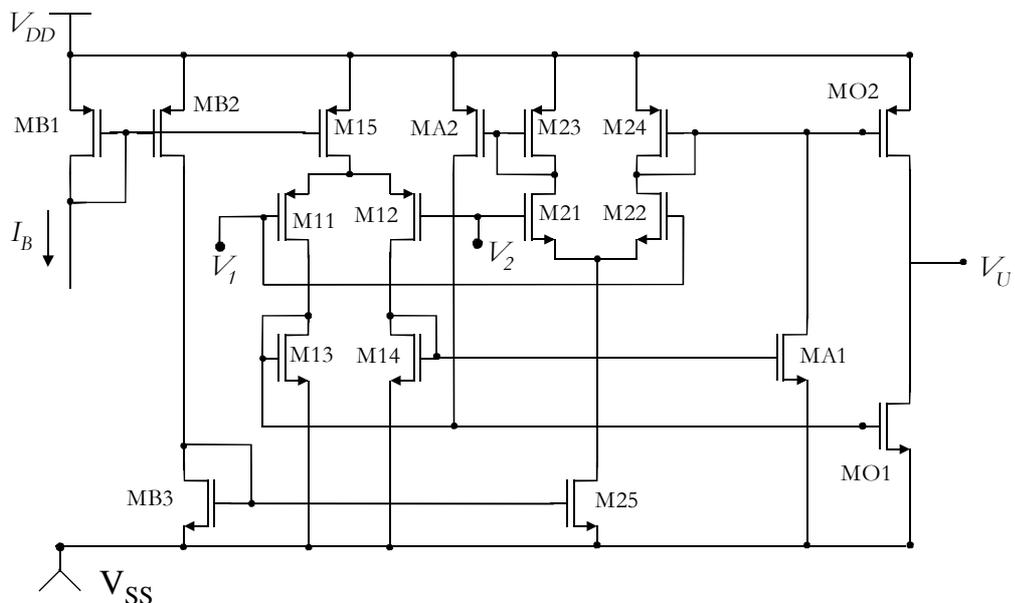


Figura 5.3.4

5.4 Amplificatore differenziale a BJT, carichi attivi e uscita single ended.

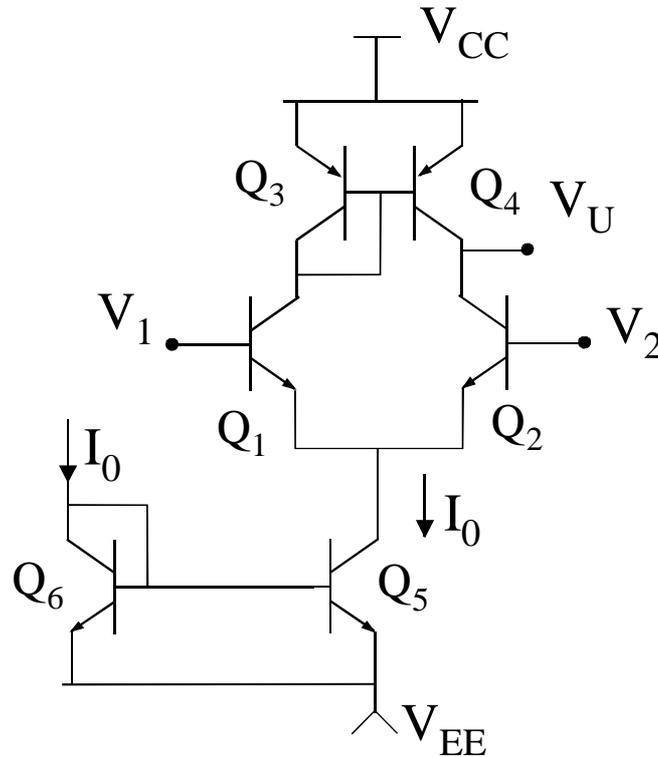


Figura 5.4.1

l'amplificatore mostrato in Figura 5.4.1 è analogo alla versione a MOSFET mostrato in Figura 5.1.1. L'analisi del circuito può essere svolta in modo simile a quella effettuata per il circuito a MOS. In particolare si ritrova l'espressione del guadagno di modo differenziale:

$$A_d = g_{m1} (r_{0n} \parallel r_{0p})$$

dove r_{0n} e r_{0p} sono le resistenze differenziali di collettore (pari a $1/h_{oe}$) rispettivamente dei transistori NPN (M1 e M2) e di quelli PNP (M3 e M4). Esplicitando il g_m e le r_0 si ottiene:

$$A_d = \frac{V_{Aeq}}{V_T} \quad \text{con} \quad V_{Aeq} = \left(\frac{1}{V_{An}} + \frac{1}{V_{Ap}} \right)^{-1}$$

dove la V_{Aeq} è una tensione di Early equivalente che combina la tensione di Early degli NPN (V_{An}) e quella dei PNP (V_{Ap}). Il guadagno non può quindi essere influenzato in alcun modo, in quanto è dato dal rapporto di due tensioni che dipendono dal processo (V_A) e dalla temperatura ($V_T = kT/q$). Essendo le tensioni di Early dell'ordine delle decine di volt (fino al centinaio) e V_T a temperatura ambiente circa 26 mV, si ottengono guadagni che raggiungono facilmente il migliaio.

Per quanto riguarda il modo comune la situazione è più complessa che per il circuito a MOSFET in quanto, a causa della corrente di base nello specchio Q3-Q4, anche se $V_d = 0$ non si ha $I_{C1} = I_{C2}$. Si può fare comunque una stima approssimata del guadagno di modo comune

facendo riferimento alla figura seguente che rappresenta la situazione nell'amplificatore, con uscita in cortocircuito, quando viene applicata una tensione di modo comune.

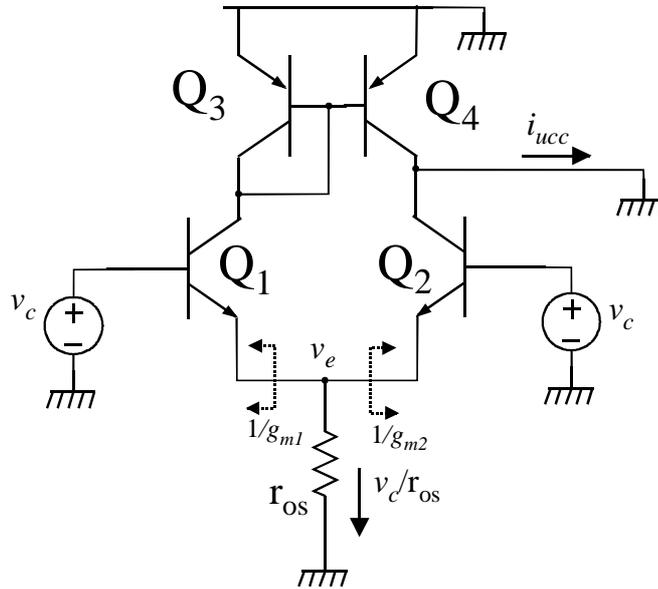


Figura 5.4.2

Innanzitutto osserviamo che le correnti a riposo I_{C1} e I_{C2} non sono uguali ma sono legate dal rapporto:

$$\frac{I_{C2}}{I_{C1}} = k_s$$

dove k_s è il rapporto fissato dallo specchio Q3-Q4. L'applicazione di un modo comune in ingresso provoca una variazione della tensione di emittore, v_e , che, con ottima approssimazione è proprio pari a v_c . (Dato che $V_E = V_C - V_{BE}$ e V_{BE} è praticamente costante). La variazione di corrente nello specchio di polarizzazione (Q5 in Figura 5.4.1, rappresentato dalla resistenza differenziale r_{os}), pari a v_c/r_{os} , entra nel nodo di emittore e si divide nelle due variazioni i_{c1} e i_{c2} , ciascuna proporzionale al g_m del corrispondente transistor. Ciò è vero perché risulta (trascurando la dipendenza dalla V_{CE}):

$$i_{c1} = g_{m1} v_{be1}; \quad i_{c2} = g_{m2} v_{be2}$$

(le v_{be} indicate sono le variazioni) e, essendo $v_d = 0$, $v_{be1} = v_{be2}$. Quindi:

$$v_d = 0 \quad \Rightarrow \quad \frac{i_{c2}}{i_{c1}} = \frac{g_{m2}}{g_{m1}} = \frac{I_{C2}}{I_{C1}} = k_s$$

La i_{ucc} risulta data da:

$$i_{ucc} = i_{c4} - i_{c2} = k'_s i_{c1} - i_{c2} = k'_s i_{c1} - k_s i_{c1} = i_{c1} (k'_s - k_s)$$

dove abbiamo indicato con k'_s il rapporto dello specchio Q3-Q4 alle variazioni (che differisce

di poco da quello statico). A questo punto basta calcolare la variazione i_{c1} in funzione di v_c , considerando che k_s non differisce molto dall'unità:

$$i_{c2} = \frac{g_{m2}}{g_{m1}} i_{c1} \Rightarrow i_{c1} + k_s i_{c1} = \frac{v_c}{r_{os}} \Rightarrow i_{c1} = \frac{v_c}{(1+k_s)r_{os}} \cong \frac{v_c}{2r_{os}}$$

Sostituendo la i_{c1} nell'espressione della i_{ucc} e considerando che $v_u = i_{ucc}R_{out}$ si ha infine:

$$A_c = \frac{R_{out}(k'_s - k_s)}{2r_{os}}$$

Considerando che R_{out} e r_{os} sono dello stesso ordine di grandezza e che la differenza $k'_s - k_s$ è pari a qualche centesimo, l'amplificazione di modo comune risulta solitamente dell'ordine di 10^{-2} .

Il CMRR vale quindi:

$$CMRR = \frac{2g_{m1}r_{os}}{(k'_s - k_s)}$$

Per quanto riguarda la dinamica di uscita, ripetendo le considerazioni fatte per l'amplificatore differenziale a MOSFET si arriva alla rappresentazione grafica di Figura 5.4.3

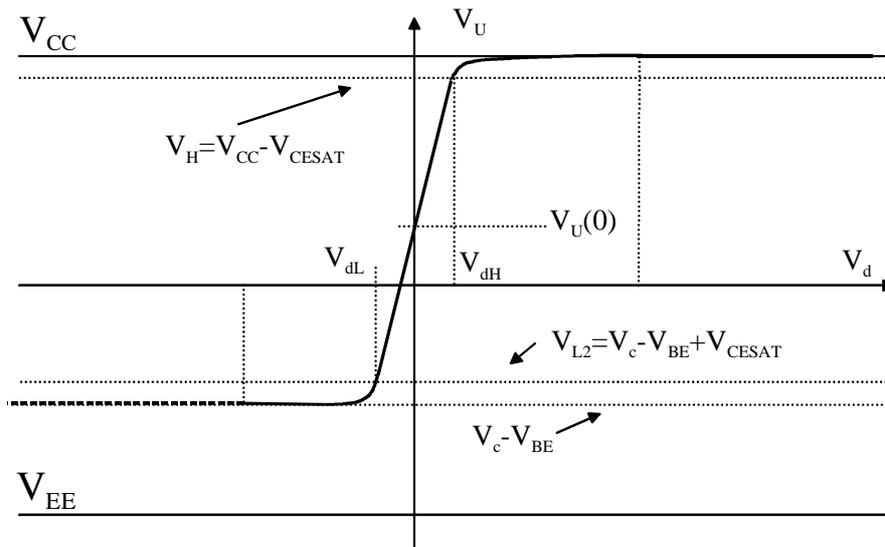


Figura 5.4.3

Anche qui la dinamica di uscita è limitata in basso dal valore del modo comune.

Per quanto riguarda il range del modo comune in ingresso, con considerazioni analoghe a quelle fatte per l'amplificatore a MOSFET, si trova che:

$$V_{EE} + V_{BE} + V_{CESAT} < V_c < V_{DD} - V_{CESAT}$$

Capitolo 6. Circuiti analogici di largo impiego

6.1 Riferimenti di tensione.

Per la realizzazione di sorgenti di tensione costante sia nei confronti della temperatura sia delle tensioni di alimentazioni si possono sfruttare due tipi differenti di approccio:

- Circuiti utilizzanti diodi zener.
- Circuiti a Band Gap.

I circuiti che utilizzano diodi zener presentano alcune controindicazioni, tra cui le più importanti sono:

1. Il diodo zener non è disponibile nelle tecnologie standard. Come sostituto si può usare la giunzione base emettitore dei transistori bipolari, che ha una tensione di breakdown tipicamente di 7-10 V. Questa tensione non è tuttavia riproducibile. L'inserimento di diodi zener affidabili e con tensione voluta comporta una complicazione del processo.
2. I diodi zener hanno una V_{BV} che dipende dalla temperatura (cresce con T per $V_{BV} > 6$ V, decresce per $V_{BV} < 6$ V)
3. Diodi zener per tensioni inferiori a 3 V non sono disponibili e presenterebbero comunque una forte dipendenza dalla temperatura. Siccome la tensione di alimentazione deve essere maggiore di V_{BV} , gli stabilizzatori a zener non sono adatti a lavorare con basse tensioni di alimentazione.
4. Le tensioni ottenute mediante diodi zener sono molto rumorose. Per eliminare le componenti a bassa frequenza del rumore degli zener sarebbero necessari condensatori di elevata capacità, non sono integrabili su silicio.

Per questi motivi sono stati sviluppati riferimenti di tensione che fanno uso solo di componenti standard (transistori). Questi circuiti, denominati “riferimenti di tensione a band gap” sono basati su stringenti requisiti di matching e possono essere realizzati solo in forma integrata.

Lo schema di principio di un circuito a band gap è mostrato nella Figura 6.1.1:

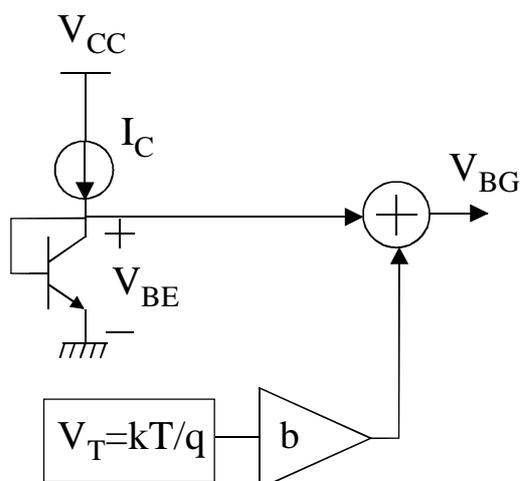


Figura 6.1.1

In pratica il transistoro bipolare montato a diodo viene alimentato da un generatore di corrente I_C . La sua V_{BE} vale circa 0.6-0.7, in dipendenza dalla corrente e dall'area del transistoro. Questa tensione non può essere usata come un riferimento stabile in quanto essa dipende fortemente dalla temperatura applicata con un coefficiente di temperatura negativo, tipicamente variabile da -3×10^{-3} V/K a -2×10^{-3} V/K. Per annullare questa variazione si può sommare una tensione proporzionale a V_T , dove V_T è kT/q . In questo modo, facendo riferimento alla figura, la derivata della tensione di uscita (V_{BG}) è pari a:

$$\frac{dV_{BD}}{dT} = \frac{dV_{BE}}{dT} + b \frac{k}{q}; \quad \text{con: } \frac{k}{q} = 8.61 \times 10^{-5} \text{ V/K} \quad (6.1.1)$$

Per esempio, se la derivata della V_{BE} è -2.6 mV/K, per ottenere una V_{BG} indipendente dalla temperatura occorre fissare un'amplificazione $b=30.2$. Sfortunatamente la derivata della V_{BE} cambia con la temperatura, pertanto l'annullamento perfetto della dV_{BG}/dT si avrà solo per una specifica temperatura. Comunque si constata che la dipendenza della V_{BG} dalla temperatura risulta molto ridotta in un ampio intervallo di temperature (si possono ottenere variazioni complessive di pochi mV su tutto l'intervallo $-40 \div +150$ °C).

Prima di passare ad alcuni esempi circuitali che realizzano questa funzione, approfondiamo il comportamento della V_{BG} .

Innanzitutto calcoliamo la V_{BE} :

$$V_{BE} = V_T \ln \left(\frac{I_C}{I_S} \right)$$

La I_C non è costante, ma viene fatta variare con la temperatura assoluta T secondo la legge:

$$I_C = GT^\alpha$$

dove G e α sono costanti. Questo tipo di andamento non è essenziale per il funzionamento del band-gap e, come verrà mostrato in seguito, non introduce sostanziali differenze rispetto al caso in cui I_C sia una costante. Occorre inoltre osservare che è difficile ottenere correnti costanti, mentre si possono facilmente ottenere correnti proporzionali alla temperatura ($\alpha=1$).

La corrente di saturazione I_S può essere espressa da:

$$I_S = \frac{qA_E n_i^2 D_n}{Q_B} = F \cdot n_i^2 D_n$$

dove F è una costante che congloba i termini non dipendenti dalla temperatura, Q_B è il drogaggio totale della base, D_n la costante di diffusione degli elettroni, A_E , l'area di emettitore e n_i la concentrazione di portatori nel silicio intrinseco.

Esplicitando D_n e n_i^2 come:

$$n_i^2 \propto T^3 e^{-\frac{E_g}{kT}}; \quad D_n = \mu_n \frac{kT}{q}$$

dove E_g è l'energia del gap del silicio (circa 1.1 eV) e μ_n la mobilità degli elettroni. Infine, esprimendo la dipendenza dalla temperatura della mobilità come: $\mu_n \propto T^{-n}$ si ottiene:

$$I_S = BT^\gamma e^{-\frac{V_{G0}}{V_T}}$$

dove B è una costante indipendente dalla temperatura, $\gamma=4-n$ e $V_{GO} = Eg/q$ paria circa 1.12 V. Sostituendo le espressioni della I_S e della I_C nell'espressione della V_{BE} si ottiene:

$$V_{BE} = V_T \ln \left(\frac{GT^\alpha}{BT^\gamma \exp\left(\frac{-V_{GO}}{V_T}\right)} \right) = V_{GO} + V_T [\ln(G \cdot E) - (\gamma - \alpha) \ln(T)] \quad (6.1.2)$$

dove con E abbiamo indicato B^{-1} .

La tensione di uscita del regolatore band-gap risulta allora:

$$V_{BG} = V_{GO} + V_T [\ln(G \cdot E) + b - (\gamma - \alpha) \ln(T)] \quad (6.1.3)$$

Il coefficiente b viene calcolato in modo tale che la derivata della V_{BG} risulta nulla per un valore di temperatura stabilito in fase di progetto. Detto T_0 questo valore si ha:

$$\frac{dV_{BG}}{dT} = \frac{k}{q} [\ln(G \cdot E) + b - (\gamma - \alpha) - (\gamma - \alpha) \ln(T_0)] = 0 \quad (6.1.4)$$

L'equazione precedente consentirebbe di stabilire il valore di b qualora fossero noi i parametri G , E , γ , α ed il valore prescelto per T_0 . In realtà, in fase di progetto si procede più semplicemente utilizzando la formula (6.1.1), nella quale si usa il valore di dV_{BE}/dT stimato dalle simulazioni.

L'equazione (6.1.4) è comunque utile per ricavare un'espressione semplice della V_{BG} che ne permetta di valutare l'andamento. Dall'equazione (6.1.4) si ricava infatti:

$$\ln(G \cdot E) + b = (\gamma - \alpha) + (\gamma - \alpha) \ln(T_0)$$

Sostituendo questa espressione nella (6.1.3) si ottiene:

$$V_{BG} = V_{GO} + V_T (\gamma - \alpha) \left(1 + \ln\left(\frac{T_0}{T}\right) \right) \quad (6.1.5)$$

Questa espressione descrive il comportamento della V_{BG} in funzione della temperatura qualora siano noti la temperatura T_0 a cui la derivata si annulla e il valore del parametro $\gamma-\alpha$. A titolo di esempio, nella Figura 6.1.2 è mostrato l'andamento della tensione di band-gap fornita dall'equazione (6.1.5) per due diversi valori di T_0 , pari a 27 °C e 50 °C e per $\gamma-\alpha = 2.5$. Si osservi che le curve presentano una derivata nulla solo per la temperatura ottimale di funzionamento (T_0). Inoltre la curva a T_0 maggiore presenta anche un valore di tensione di uscita più alto, come si può verificare dall'analisi dell'equazione (6.1.5). I grafici della Figura 6.1.2 mostrano anche che la tensione di uscita (V_{BG}) subisce una variazione molto piccola (inferiore allo 0.1 %) su tutto l'intervallo di temperature (0-70 °C). In sede di progetto e dimensionamento di un riferimento di tensione a band-gap, la temperatura T_0 può essere fissata scegliendo opportunamente il termine b (si veda la Figura 6.1.1).

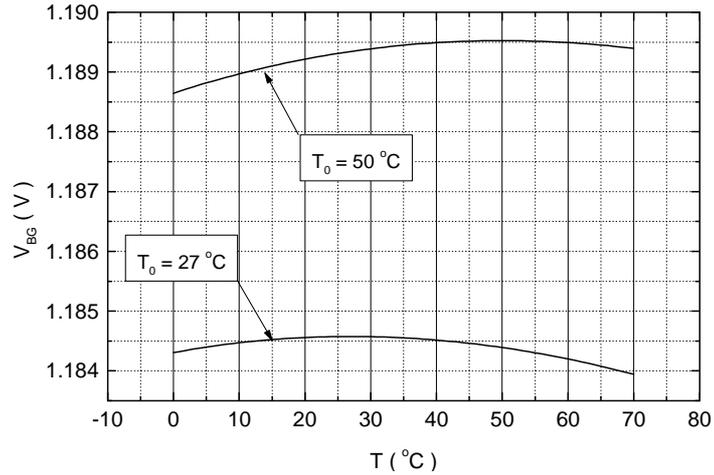


Figura 6.1.2

Esiste una grande varietà di circuiti che forniscono una tensione di riferimento stabile sfruttando il principio del band-gap. In ogni caso è sempre necessario disporre di transistori bipolari o, al limite, diodi a giunzione. In tecnologia CMOS si possono realizzare riferimenti a band-gap utilizzando i BJT parassiti di substrato. In un processo CMOS *n*-well i transistori di substrato sono di tipo pnp e presentano il collettore connesso con il substrato stesso e quindi alla V_{SS} . La figura seguente mostra un esempio di riferimento di tensione a band-gap in tecnologia CMOS *n*-well.

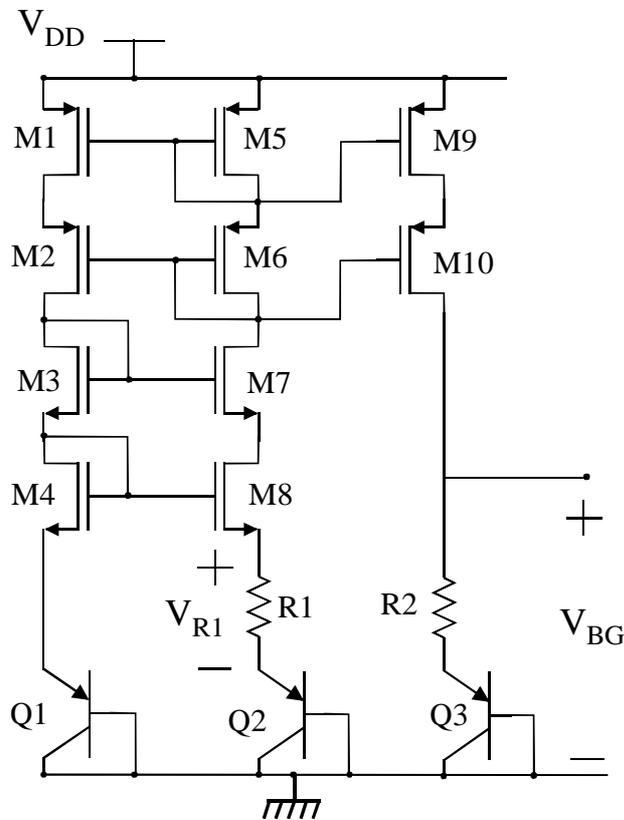


Figura 6.1.3

Il circuito è composto dai tre rami: (1): transistori M1-M2-M3-M4-Q1; (2) transistori M5-M6-M7-M8-Q2 e resistore R1; (3) M9-M10 Q3 e resistore R2. I p -MOS sono tutti uguali e realizzano quindi uno specchio cascode che impone che la corrente sia uguale nei tre rami. Inoltre anche tutti gli n -MOS sono identici. Ciò, assieme all'eguaglianza delle correnti nei rami, impone $V_{GS8} = V_{GS4}$. Questa relazione è verificata con elevata precisione in quanto M3 e M7 fanno sì che M4 e M8 abbiano anche stessa V_{DS} . Quindi si ricava facilmente $V_{S4} = V_{S8}$. (si noti che $V_{G4} = V_{G8}$ poichè i gate di M8 e M4 sono connessi). Ma:

$$V_{S4} = V_{BE1} = V_{S8} = V_{BE2} + V_{R1}$$

dove le V_{BE} vanno intese come valore assoluto. Si ricava quindi la tensione sulla resistenza R1:

$$V_{R1} = V_{BE1} - V_{BE2} = V_T \ln\left(\frac{I_{C1} I_{S2}}{I_{C2} I_{S1}}\right) = V_T \ln\left(\frac{I_{S2}}{I_{S1}}\right)$$

Il transistorore Q2 viene fatto appositamente di area più grande di quella di Q1. In particolare:

$$\frac{A_{E2}}{A_{E1}} = \frac{I_{S2}}{I_{S1}} = n > 1 \quad \Rightarrow \quad V_{R1} = V_T \ln(n)$$

La corrente in R1 (e quindi in tutti e tre i rami) sarà data da:

$$I = \frac{V_{R1}}{R_1} = \frac{V_T}{R_1} \ln(n)$$

Se si trascura la dipendenza di R_1 dalla temperatura la corrente è proporzionale alla temperatura assoluta. Questa corrente scorre anche nel terzo ramo e quindi in Q3 e R2. La tensione di uscita sarà quindi:

$$V_{BG} = V_{BE3} + IR_2 = V_{BE3} + V_T \frac{R_2}{R_1} \ln(n)$$

La tensione di uscita è quindi data dalla somma di una V_{BE} e un termine proporzionale alla V_T . Riconosciamo quindi lo schema tipico del band-gap. Essendo la corrente che scorre in Q3 proporzionale a T , sarà $\alpha=1$. Inoltre si può osservare che il termine b di Figura 6.1.1 è pari a:

$$b = \frac{R_2}{R_1} \ln(n)$$

Agendo sul rapporto R_2/R_1 si può quindi ottenere il valore di T_0 voluto, ovvero l'annullamento della derivata alla temperatura richiesta.

Infine occorre osservare che il circuito mostrato, come la maggioranza dei circuiti band-gap, presenta anche un altro punto di lavoro possibile, ovvero quello caratterizzato da correnti nulle in tutti e tre i rami. Si osservi che nel circuito è presente un anello di reazione positiva, costituito dai MOS M1 ... M8, dai BJT Q1,Q2 e dal resistore R1. Si può facilmente dimostrare che in corrispondenza del punto di riposo a correnti nulle il guadagno di anello è >1 , mentre è <1 nel punto di lavoro corretto. Solo quest'ultimo risulterebbe quindi stabile. In pratica anche il punto a correnti nulle è stabile perché in tali condizioni i MOS entrano sotto-soglia e il loro guadagno risulta degenerato. Affinché il circuito si porti nel punto di lavoro utile occorre quindi aggiungere un semplice circuito di accensione che inietti una piccola corrente nel momento in cui il sistema viene alimentato e si disconnetta da solo una volta che il riferimento, grazie alla reazione positiva, si porta nel punto di lavoro corretto.

6.2 Traslatori di tensione.

I traslatori di tensione hanno la funzione di sommare una tensione costante al segnale. Il fatto che all'interno di un circuito integrato i vari stadi di elaborazione del segnale siano accoppiati in continua rende spesso necessaria l'operazione di traslazione di tensione.

Nel campo dei circuiti **CMOS**, il più semplice traslatore di tensione è costituito dal source follower. Con riferimento alla figura seguente, in (a) è mostrato un traslatore di tensione basato su un transistor *n*-MOS. Il traslatore è polarizzato con una corrente continua I_0 . Questo circuito è in grado di traslare in basso di una quantità pari a V_{GSn} . Se si richiede una traslazione verso l'alto è necessario ricorrere ad un source follower che impiega un *p*-MOS, come mostrato in (b). In questo caso si trasla verso l'alto di una quantità pari a $|V_{GSp}|$.

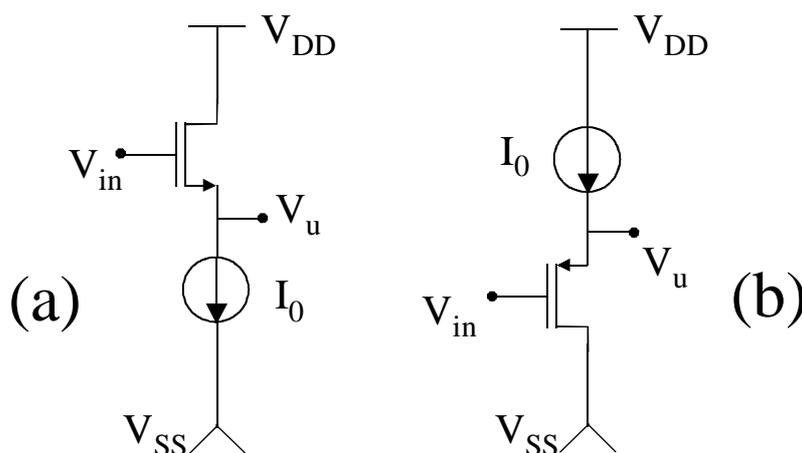


Figura 6.2.1

La traslazione può essere variata entro ampi limiti agendo sulla corrente I_0 e sulle dimensioni (W/L) dei transistori. In particolare si ha:

$$\text{Traslatore } n : V_u = V_{in} - \left(V_{in} + \sqrt{\frac{2I_0}{\beta_n}} \right)$$

$$\text{Traslatore } p : V_u = V_{in} + \left(|V_{tp}| + \sqrt{\frac{2I_0}{\beta_p}} \right)$$

Circuiti con uguali rapporti I_0/β si comportano quindi nello stesso modo per quanto riguarda la traslazione. Questo grado di libertà viene sfruttato per fissare la resistenza di uscita del traslatore, pari a:

$$R_{out} = \frac{1}{g_m} = \frac{1}{\sqrt{2\beta I_0}}$$

La resistenza di uscita è spesso un parametro critico in quanto assieme alla capacità di carico fissa il limite superiore di banda del traslatore. Pertanto, a parità di traslazione e capacità di carico, grandi valori di I_0 e β comportano bande più estese, ma, di contro, elevata potenza dissipata e ingombro. La scelta cade quindi sulle minime dimensioni che consentono di soddisfare il limite di banda richiesto.

Come si può osservare non è possibile traslare mai di una tensione inferiore alla tensione di soglia. Nel caso si richiedano traslazioni inferiori a questo limite, si può usare una successione di un traslatore *n* e di uno *p* opportunamente dimensionati.

Osservazione importante: nelle espressioni della tensione di traslazione compare la tensione di soglia: quest'ultima, a causa dell'effetto body, dipende dal valore della tensione tra source e body, quindi dal valore della tensione di uscita. In altre parole la tensione di cui si trasla il segnale dipende dal valore del segnale stesso. È facile dimostrare che questo si traduce in una distorsione non lineare sul segnale. Questa distorsione potrebbe risultare inaccettabile se l'ampiezza del segnale è elevata. In un processo CMOS *n*-well è possibile connettere il body di un transistor *p*-MOS al suo source, inserendolo in una propria *n*-well. In questo modo si riduce notevolmente la distorsione del traslatore di tipo *p*. Questa operazione non è possibile per i transistori *n*-MOS e quindi il traslatore di tipo *n* presenta un comportamento corretto solo per piccoli segnali. In un processo CMOS *p*-well la situazione è opposta.

Per quanto riguarda i circuiti bipolari, si possono usare traslatori di livello ad emitter-follower, mostrati nella Figura 6.2.2.

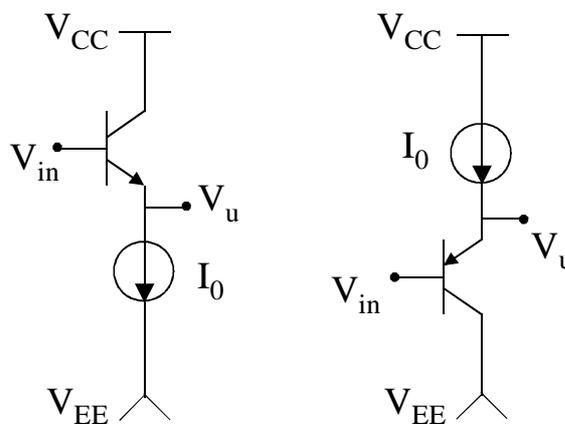


Figura 6.2.2

Per traslare verso il basso si utilizza il circuito con BJT npn, per traslare verso l'alto il pnp. Tuttavia è facile comprendere che con i circuiti mostrati si può traslare solo di una quantità pari a V_{BE} che cambia poco con la corrente e le dimensioni dei transistori. Per traslare di quantità multiple di V_{BE} si possono impilare alcuni diodi tra l'emettitore dell'emitter follower e l'uscita.

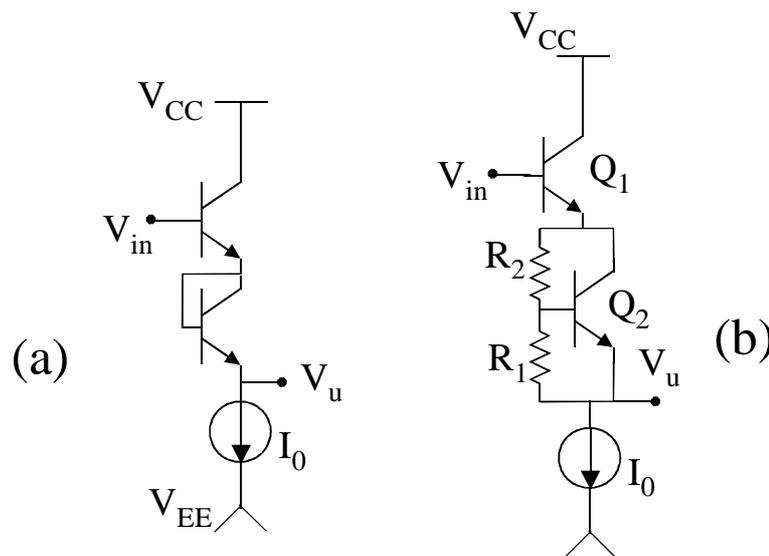


Figura 6.2.3

Nella Figura 6.2.3 (a) è stato aggiunto un transistor collegato a diodo per ottenere una traslazione verso il basso pari a $2 V_{BE}$. Per traslare di quantità che non sono multipli interi della V_{BE} si può usare un circuito “moltiplicatore di V_{BE} ”, costituito in Figura 6.2.3 (b) da Q2 e dalle resistenze R1 e R2. In questo caso si trasla verso il basso di una tensione pari a:

$$V_{BE1} + V_{BE2} \frac{R_1 + R_2}{R_1}$$

L'espressione precedente fissa il rapporto tra R_1 e R_2 . Per quanto riguarda il loro valore occorre rispettare la condizione:

$$I_{B2} \cong \frac{I_0}{\beta_2} \ll \frac{V_{BE}}{R_1} \ll I_0$$

che garantisce che la I_0 coincida praticamente con la I_{C2} e, contemporaneamente, che nel partitore R1-R2 si possa trascurare la corrente di base I_{B2} .

6.3 Stadi di uscita.

Lo scopo degli stadi di uscita è quello di trasferire il segnale sul carico (resistivo o capacitivo). Il principale requisito di uno stadio di uscita è quello di erogare il livello di corrente necessario a pilotare il carico stesso. Se il carico è resistivo, questa caratteristica consente di avere sul carico stesso il livello di tensione richiesto. Ovviamente più bassa è la resistenza del carico, più elevate sono le correnti che lo stadio stesso deve fornire. In presenza di carico capacitivo un'elevata corrente di uscita si traduce in elevata velocità e quindi, possibilità di mantenere i livelli di tensione di uscita richiesti anche a frequenze relativamente alte.

Stadi di uscita a source follower e emitter-follower.

Uno stadio di uscita molto semplice che in molti casi può essere utilizzato è il source follower. Questo circuito, già visto come traslatore di tensione è mostrato nella figura seguente dove è evidenziato il collegamento del carico tra l'uscita e il comune (ground). Il circuito si suppone alimentato da due tensioni V_{DD} e V_{SS} di valore opposto.

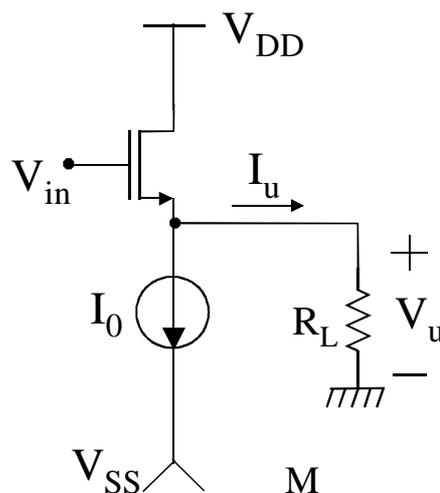


Figura 6.3.1

Questo circuito presenta il vantaggio di avere un'amplificazione praticamente unitaria, quindi può essere usato per trasferire il segnale sul carico senza modifiche sostanziali (a parte

l'inevitabile traslazione), ovvero si comporta da buffer. Inoltre esso ha una bassa resistenza di uscita che fa sì che l'amplificazione (unitaria) non diminuisca sostanzialmente anche per resistenze di carico relativamente basse. Presenta però i seguenti svantaggi importanti:

- Impossibilità di fornire tensioni di uscita superiori a $V_{DD} - V_m$ (con V_m aggravata dall'effetto body). Ciò si può facilmente comprendere considerando che sul gate non può essere presente una tensione superiore a V_{DD} .
- Questo circuito può assorbire una corrente (ovvero erogare una corrente negativa) che al massimo raggiunge I_0 . Ovviamente entrambe le correnti erogabili, negativa e positiva, devono essere paragonabili in quanto il segnale sul carico deve poter assumere valori negativi e positivi senza sostanziali asimmetrie. In sostanza la massima corrente di uscita è limitata dalla corrente assorbita a riposo. Questo fa sì che il circuito in questione possa essere usato solo quando le correnti erogabili sono modeste.

Come vedremo in seguito il primo svantaggio è tipico dei circuiti basati su source follower e potrà essere risolto solo cambiando configurazione circuitale. Per quanto riguarda il secondo punto osserviamo invece che è tipico degli stadi in classe A, ovvero quegli stadi in cui il dispositivo attivo deve condurre per tutto il periodo del segnale di ingresso (supposto sinusoidale). Infatti, nello schema di Figura 6.3.1, se per una frazione del periodo il MOSFET non conduce abbiamo come conseguenza una distorsione del segnale, in quanto l'uscita rimane costante a $-R_L I_0$.

Per ovviare a questa limitazione si può usare sempre un source follower ma in classe AB. Nel ben noto schema di Figura 6.3.2 (a) è mostrato un circuito in cui si fa uso di una topologia push-pull con transistori di uscita M1 e M2 complementari.

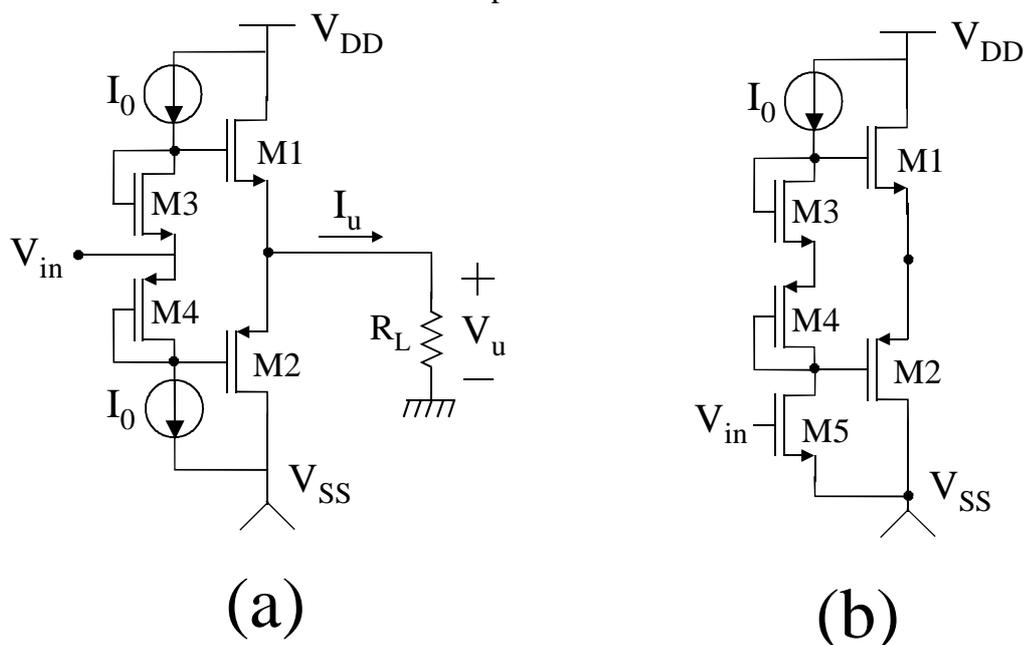


Figura 6.3.2

Si verifica immediatamente che: $V_{GS1} + |V_{GS2}| = V_{GS3} + |V_{GS4}|$, ovvero:

$$V_{t1} + \sqrt{\frac{2I_{D1}}{\beta_1}} + |V_{t2}| + \sqrt{\frac{2I_{D2}}{\beta_2}} = V_{t3} + \sqrt{\frac{2I_0}{\beta_3}} + |V_{t4}| + \sqrt{\frac{2I_0}{\beta_4}} \quad (6.3.1)$$

Trascurando in prima approssimazione l'effetto body, si ha: $V_{t1} = V_{t3}$ e $V_{t2} = V_{t4}$. Inoltre un dimensionamento tipico prevede:

$$\beta_1 = \beta_2 \equiv \beta_U \quad \text{e} \quad \beta_3 = \beta_4 \equiv \beta_I$$

Questo fa sì che i transistori n -MOS siano equivalenti (complementari) dei transistori p -MOS. Con il dimensionamento l'equazione (6.3.1) diventa:

$$\sqrt{I_{D1}} + \sqrt{I_{D2}} = 2\sqrt{I_0} \sqrt{\frac{\beta_U}{\beta_I}} \quad (6.3.2)$$

Dall'equazione precedente si possono ricavare due informazioni importanti. La prima ci consente di fissare la corrente a riposo I_{DQ} nei MOS di uscita (M1 e M2) a partire da I_0 . A riposo, infatti, $I_U = 0$, ovvero $I_{D1} = I_{D2} = I_{DQ}$ e quindi si trova facilmente:

$$I_{DQ} = I_0 \frac{\beta_U}{\beta_I}$$

L'equazione (6.3.2) e la (6.3.1) da cui essa deriva sono valide solo se entrambi M1 e M2 conducono. Questo è vero solo per piccoli segnali. Quando si applica, per esempio, un segnale V_{in} positivo, la corrente di uscita non è più nulla e diventa positiva. Deve quindi crescere I_{D1} a scapito di I_{D2} , ovvero V_{GS1} a scapito di V_{GS2} , in quanto la somma di V_{GS1} e di V_{GS2} (modulo) deve rimanere costante. Per un certo valore di corrente erogata in uscita, la V_{GS2} sarà pari alla tensione di soglia di M2. Se, in seguito ad un aumento del segnale di ingresso, la corrente di uscita aumenta ulteriormente M2 si interdice rimarrà acceso solo M1. Per tensioni di ingresso (e quindi correnti di uscita) negative accadrà esattamente il contrario con M1 che si interdice e rimane solo M2 a lavorare. Siccome per un intervallo di correnti erogate comprendente piccoli valori negativi e positivi entrambi i transistori conducono, ciascun transistorore lavora per un po' di più di un semiperiodo del segnale di ingresso e quindi la classe è di tipo AB. Importante: in questo amplificatore la massima corrente di uscita, sia positiva sia negativa, data da:

$$I_U > 0: \quad |I_{UMAX}| = \frac{\beta_1}{2} (V_{DD} - V_m)^2$$

$$I_U < 0: \quad |I_{UMAX}| = \frac{\beta_2}{2} (|V_{SS}| - |V_p|)^2$$

non dipende dalla corrente (e quindi dal consumo) a riposo. La corrente di riposo I_{DQ} fissa invece la resistenza di uscita per piccoli segnali, pari a $1/(g_{m1} + g_{m2})$.

Il source follower complementare viene spesso incorporato in uno stadio di amplificazione come mostrato in Figura 6.3.2 (b), in cui M5, insieme allo specchio di carico che fornisce la I_0 , costituisce un amplificatore a source comune.

La versione a transistori bipolari dello schema di figura Figura 6.3.2 (a) è mostrato nella Figura 6.3.3.

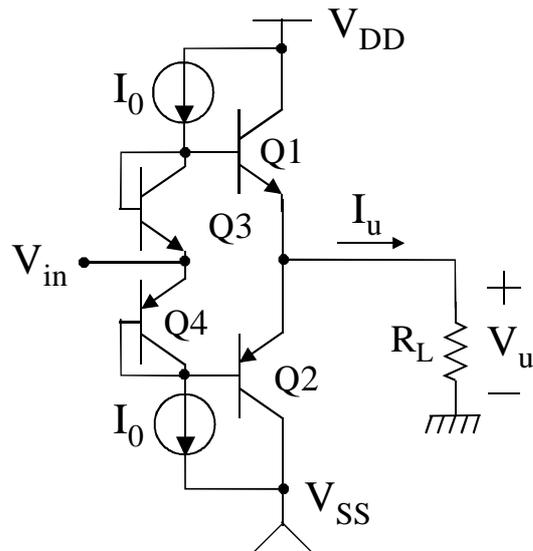


Figura 6.3.3

Senza dilungarci su questo schema, il cui funzionamento è simile a quello dell'equivalente a MOSFET, osserviamo che anche qui si ricerca il comportamento complementare dimensionando Q1 e Q2 in modo che abbiano lo stesso valore della corrente di saturazione. Per fare ciò si può lavorare sulle aree dei BJT, tenendo presente che questo parametro spesso si può solo variare di quantità intere. Il rapporto tra le aree dei transistori di uscita Q1 e Q2 e quelle dei corrispondenti transistori di ingresso Q3 e Q4, fissa la corrente di riposo in Q1 e Q2. Un aspetto che va però sottolineato riguarda il transistor pnp Q2: nei tipici processi bipolari ottimizzati per il BJT npn, il pnp disponibile è di tipo laterale. Questo transistor mal si presta ad essere impiegato in stadi di uscita, a causa della sua limitata capacità di portare grosse correnti. Per sopperire a questa limitazione si impiega, quando è disponibile, il transistor pnp di substrato la cui caratteristica di avere il collettore connesso al substrato (cioè alla V_{SS}) non è qui, ovviamente, una limitazione. Altre possibilità di superare la limitazione del transistor pnp prevedono l'impiego di transistori pnp equivalenti, formati connettendo un npn con un pnp laterale (configurazioni a simmetria "quasi complementare") oppure circuiti con soli npn connessi in configurazione "totem pole". La trattazione degli schemi che impiegano queste configurazioni esula dagli scopi del presente trattato.

Limiti degli stadi di uscita a source follower e loro superamento mediante schemi a source comune.

Come evidenziato precedentemente, gli stadi basati su source follower (nel caso a bipolari, emitter follower) presentano la limitazione che la tensione di uscita non può mai eguagliare i rail di alimentazione. Infatti, se prendiamo in esame il source-follower di Figura 6.3.2, la massima e la minima tensione di uscita saranno date da:

$$V_{UMAX} = V_{DD} - \left(V_m + \sqrt{\frac{2I_U}{\beta_1}} \right); \quad V_{UMIN} = V_{SS} + \left(|V_{tp}| + \sqrt{\frac{2|I_U|}{\beta_2}} \right) \quad (6.3.3)$$

Quindi, non solo perdiamo sicuramente una tensione di soglia sia rispetto al rail negativo, sia rispetto al rail positivo, ma abbiamo anche un'ulteriore perdita che dipende fortemente dalla corrente che dobbiamo erogare. Il corrispondente stadio a bipolari (Figura 6.3.3) si comporta un po' meglio, in quanto la caduta di tensione si limita ad una V_{BE} (spesso anche questa non è

comunque accettabile) e, soprattutto, questo valore aumenta poco con la corrente di uscita, in virtù del legame logaritmico tra V_{BE} e corrente di collettore.

La continua riduzione della tensione di alimentazione dettata da esigenze di riduzione della potenza assorbita comporta che assai spesso la limitazione indicata nelle espressioni (6.3.3) rende i circuiti a source follower completamente inutilizzabili. Si pensi che molti circuiti analogici devono ormai funzionare con una singola tensione di alimentazione dell'ordine di 3 V. Perdere quasi due volt a causa delle tensioni di soglia (senza considerare la perdita ulteriore dovuta alla corrente) significa ridurre la dinamica del segnale di uscita a poco più di 1 V.

Pertanto ogni volta che si richiedono prestazioni *rail-to-rail* in uscita è necessario fare ricorso a schemi che non facciano uso di source follower.

Lo schema che si impiega è quindi il source comune. Il più semplice stadio a source comune per circuiti integrati è mostrato nella figura seguente:

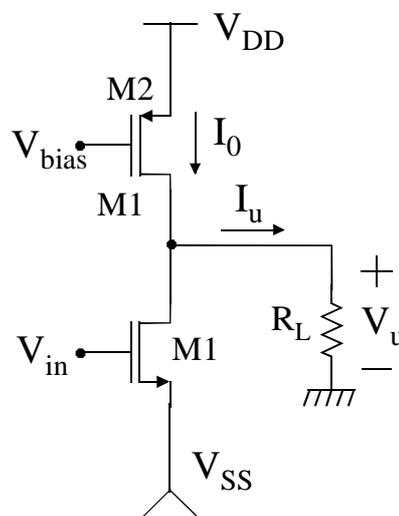


Figura 6.3.4

Il transistor M1 costituisce il vero e proprio source comune, mentre M2 ha funzione di generatore di corrente di polarizzazione e di carico dinamico. Questo stadio fornisce un'amplificazione che, a vuoto è pari a $-g_{m1}(r_{d1}||r_{d2})$. Quindi, quando il source comune è usato come stadio di uscita, oltre ad avere lo scopo di erogare la corrente richiesta sul carico, fornisce anche un'amplificazione (tra l'altro negativa, quindi lo stadio è invertente) di cui occorre tenere conto. Questo stadio non può quindi essere usato come buffer ma deve essere inserito nella catena di amplificazione.

Si verifica facilmente che a vuoto (ovvero con R_L infinita) la tensione di uscita può raggiungere la V_{DD} , quando M1 è interdetto, e la V_{SS} , (a meno di una piccola caduta di tensione su M1), quando M1 conduce con una V_{GS} elevata.

Lo stadio mostrato in Figura 6.3.4 non è però adatto ad erogare elevate correnti, ovvero a pilotare bassi carichi resistivi. Infatti esso è uno stadio in classe A e quindi può erogare al massimo una corrente pari alla corrente di polarizzazione I_0 .

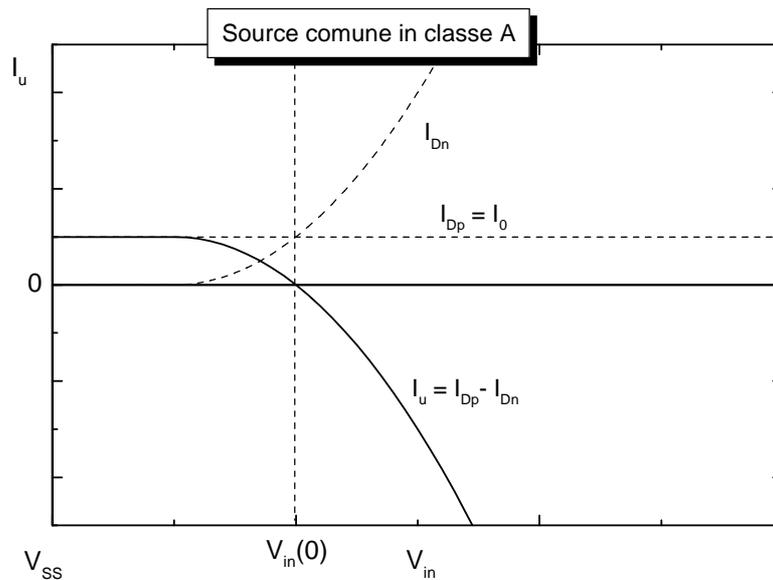


Figura 6.3.5

La Figura 6.3.5 mostra l'andamento delle correnti di M1 e M2 e di uscita in funzione di V_{in} , in condizione di carico resistivo approssimabile ad un corto circuito. Si osserva che la corrente di uscita può assumere valori negativi di elevata intensità, limitati solo dalla massima V_{in} e dalle dimensioni di M1. Per quanto riguarda le correnti positive, esse possono essere erogate solo da M2, che lavorando a V_{GS} costante (pari a $V_{DD} - V_{bias}$), può fornire solo una corrente massima pari a I_0 , che, appunto, è anche la corrente a riposo. Abbiamo già visto nel caso di source follower, che questa condizione comporta che per progettare uno stadio capace di erogare elevate correnti (negative e positive), è necessario prevedere elevati consumi e dissipazioni di potenza a riposo. Quando le correnti da erogare superano le centinaia di mA, questa saturazione diventa inaccettabile e occorre ricorrere a stadi in classe AB.

Uno stadio source comune in classe AB molto usato è mostrato schematicamente in Figura 6.3.6. La differenza tra questo stadio e quello in classe A sta nel fatto che il segnale, previa opportuna traslazione di una tensione V_B , viene anche assegnato al transistor di carico (M2). In questo modo la corrente I_0 non è più costante ma può aumentare al variare del segnale di ingresso, proprio come la corrente in M1. In questo stadio entrambi, M1 e M2 hanno funzione di amplificazione con configurazione source comune. Il funzionamento dello stadio è illustrato schematicamente con la Figura 6.3.7 nella quale viene raffigurato un possibile andamento nel tempo dei segnali assegnati ai gate di M1 (V_{GN}) e M2 (V_{GP}). La distanza di V_{GN} dal livello $V_{SS} + V_m$, è la $V_{GS} - V_t$ di M1. La distanza di V_{GP} da $V_{DD} - |V_{tp}|$ è, in modulo, la $V_{GS} - V_t$ di M2.

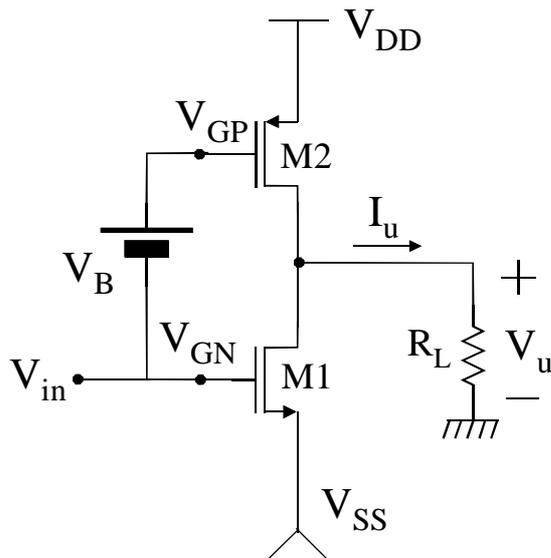


Figura 6.3.6

In condizioni di riposo abbiamo $V_{GN} = V_{in} = V_{GN}(0)$ e $V_{GP} = V_{GP}(0)$. In queste condizioni le $V_{GS} - V_t$ di entrambi i transistori sono tali da mantenerli in conduzione. In particolare, essendo questa la condizione di riposo, si richiederà che $I_U = I_{D1} - I_{D2} = 0$, ovvero $I_{D1} = I_{D2}$. Nel primo tratto abbiamo raffigurato un aumento del segnale di ingresso. Si nota che la $V_{GS} - V_t$ del transistore M1 aumenta, mentre quella di M2, in modulo, diminuisce. Quindi M1 conduce sempre di più e M2 sempre di meno. Pertanto la corrente di uscita assume valori negativi via via sempre più crescenti. Raggiunto l'istante t_1 , un ulteriore aumento del segnale V_{in} causa lo spegnimento di M2. La corrente continua tende però ad aumentare di intensità grazie all'aumento della $V_{GS} - V_t$ di M1. Viceversa, a partire da t_2 abbiamo raffigurato un tratto in cui il segnale di ingresso diminuisce. All'istante t_3 sarà quindi M1 a spegnersi e M2 a condurre sempre di più. Supponendo di applicare una sinusoide di ampiezza sufficiente, vi sarà quindi un intervallo del periodo in cui entrambi i transistori sono accesi, e altri due intervalli, inferiori ad un semiperiodo, in cui alternativamente lavora solo M1 o M2. Il funzionamento in classe AB è quindi dimostrato.

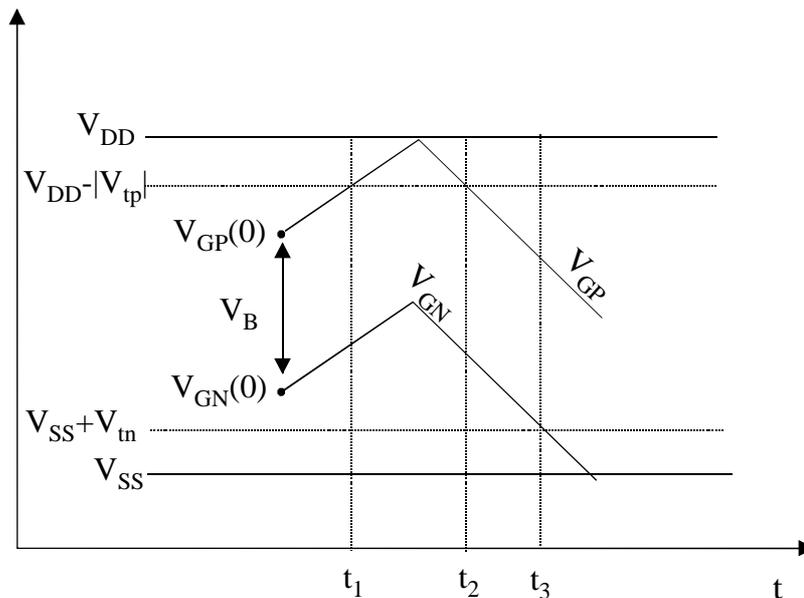


Figura 6.3.7

L'andamento delle correnti in funzione del segnale di ingresso è rappresentato in Figura 6.3.8. Si osservi come la massima corrente erogabile, sia negativa sia positiva, sia molto più grande del valore I_{DQ} a riposo.

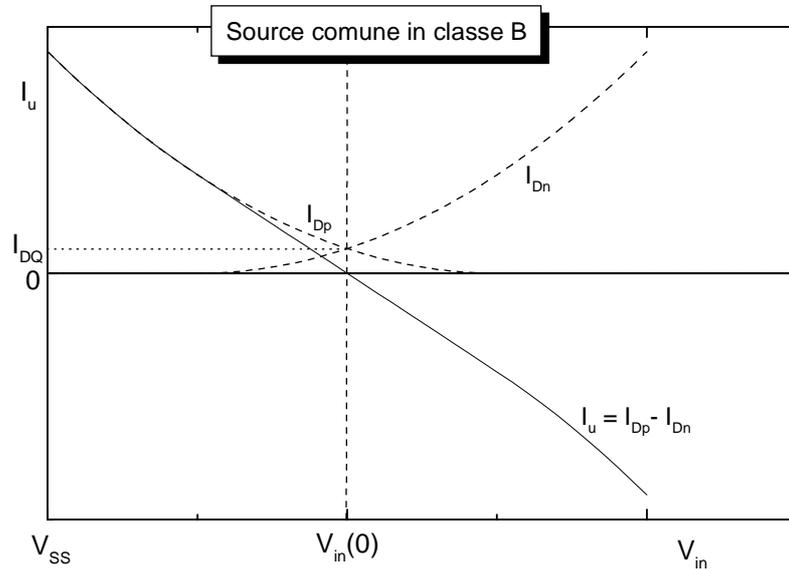


Figura 6.3.8

La traslazione V_B tra V_{GN} e V_{GP} è necessaria per fissare la corrente I_{DQ} . Per trovare un legame tra I_{DQ} e V_B , consideriamo innanzitutto il legame tra V_{GS1} , V_{GS2} e V_B . Dall'analisi del circuito si ha subito:

$$V_{GS1} + V_B + |V_{GS2}| = V_{DD} - V_{SS}$$

Quindi:

$$\sqrt{\frac{2I_{D1}}{\beta_1}} + V_m + V_B + \sqrt{\frac{2I_{D2}}{\beta_2}} + |V_{tp}| = V_{DD} - V_{SS}$$

ovvero:

$$\sqrt{\frac{2I_{D1}}{\beta_1}} + \sqrt{\frac{2I_{D2}}{\beta_2}} = V_{DD} - V_{SS} - (V_m + V_B + |V_{tp}|)$$

Considerando per semplicità M1 e M2 complementari, ovvero $\beta_1 = \beta_2 = \beta$ e tenendo conto che a riposo $I_{D1} = I_{D2} = I_{DQ}$ si ottiene con facili passaggi:

$$I_{DQ} = \frac{\beta}{8} [V_{DD} - V_{SS} - (V_m + V_B + |V_{tp}|)]^2$$

Variando la V_B si può ridurre o aumentare la I_{DQ} . In questo modo si può regolare l'amplificazione per piccoli segnali. Questa, in presenza di un carico R_L risulta:

$$A_V = -(g_{m1} + g_{m2}) \cdot (r_{d1} \| r_{d2} \| R_L)$$

Se, come spesso accade, R_L è molto più piccola di r_{d1} e r_{d2} , A_V si riduce a $-(g_{m1}+g_{m2})R_L$. Affinché il guadagno si mantenga sufficientemente elevato fino al minimo valore di R_L per cui l'amplificatore è stato progettato, i g_m di M1 e M2 a riposo devono essere sufficientemente elevati. Questo viene garantito da un corretto valore della corrente I_{DQ} . Naturalmente I_{DQ} non può essere troppo elevata per non incorrere in elevati consumi a riposo. Uno schema di stadio di uscita in cui viene impiegato un source follower per effettuare la traslazione di V_B è mostrato in Figura 6.3.9.

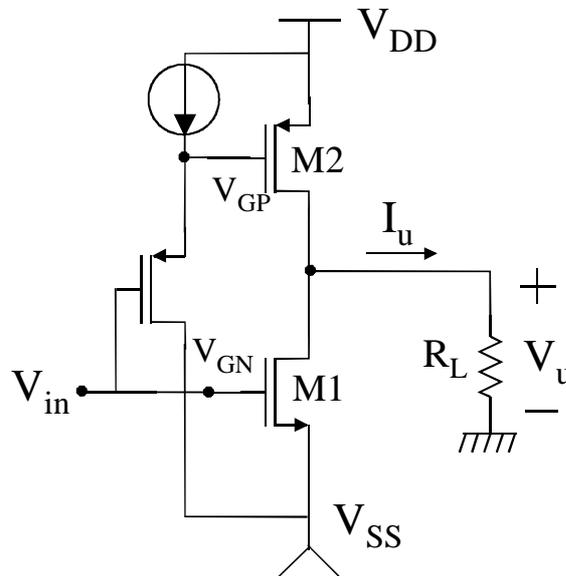


Figura 6.3.9

Stadi di guadagno.

Gli stadi di uscita a source comune in classe A (Figura 6.3.4) e in classe AB (schema di principio di Figura 6.3.6) possono essere usati come stadi di guadagno ad ingresso e uscita unipolari (single-ended). Se lo stadio di guadagno non deve fungere anche da stadio di uscita, si preferisce utilizzare, per la sua semplicità, lo stadio in classe A.

6.4 Cenni agli amplificatori operazionali CMOS.

L'architettura più comune di un amplificatore operazionale CMOS è quella a due stadi. Con questa denominazione si intende un amplificatore con due stadi di amplificazione. Possono poi essere presenti stadi buffer, traslatori, circuiti di polarizzazione. L'amplificatore a due stadi ha quindi il seguente schema equivalente:

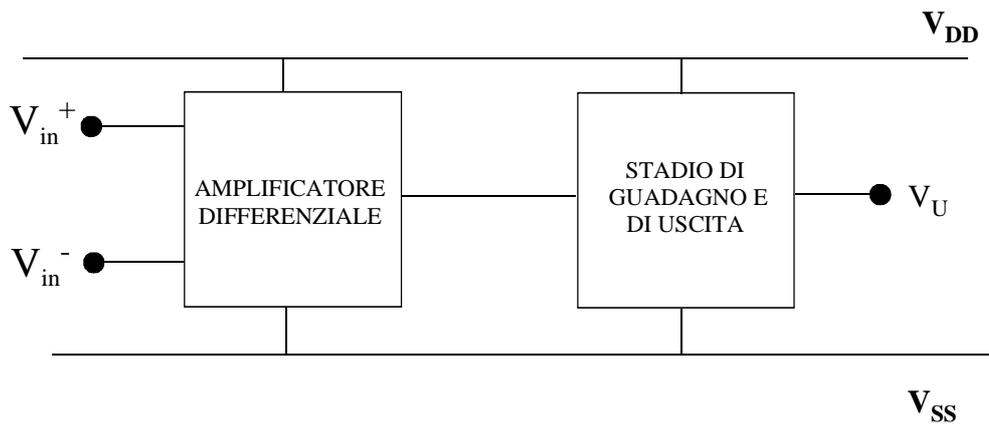


Figura 6.4.1

Lo scopo del primo stadio è quello di operare la conversione da differenziale a unipolare e, inoltre, fornire una parte del guadagno complessivo. Lo stadio di ingresso determina alcune proprietà importanti dell'amplificatore operazionale, quali il CMRR, il rumore equivalente di ingresso, la tensione di offset, la dinamica del modo comune di ingresso (CMR=Common Mode Range). Lo stadio di guadagno fornisce un'ulteriore amplificazione richiesta per raggiungere l'amplificazione complessiva richiesta. Negli operazionali CMOS lo stadio di guadagno ha spesso anche funzione di stadio di uscita e deve pertanto essere dimensionato in modo tale da:

- fornire la massima corrente richiesta, dipendente dalle applicazioni per le quali l'amplificatore è stato progettato.
- fare sì che per tutto l'intervallo di resistenze di carico per le quali l'amplificatore è stato pensato, l'amplificazione si mantenga abbastanza elevata da soddisfare le specifiche di progetto. Si ricorda che in un amplificatore operazionale un'amplificazione troppo bassa si traduce in una tensione di ingresso che non può più considerarsi trascurabile come previsto dal metodo del corto circuito virtuale e quindi che introduce un errore non trascurabile nelle funzioni sintetizzate usando l'amplificatore stesso. Per esempio, un buffer sintetizzato con un amplificatore operazionale con guadagno insufficiente presenta un errore sistematico tra ingresso e uscita.

Lo stadio di guadagno/uscita è responsabile della distorsione operata sul segnale dall'amplificatore perché in esso non è più valida l'approssimazione di piccolo segnale.

Quando la dipendenza del guadagno dalla resistenza di carico fornita da uno schema quale quello di Figura 6.4.1 non è più accettabile, occorre inserire in uscita uno stadio buffer. Per questo stadio si può usare un source follower push pull, con lo svantaggio di perdere la possibilità di avere un'uscita rail-to-rail. Come alternativa si può connettere in uscita un altro operazionale come quello di Figura 6.4.1 ma chiuso a buffer. In questo caso la reazione aiuta a mantenere la resistenza di uscita bassa.

Prima di mostrare uno schema completo di un semplice amplificatore operazionale, occorre richiamare alcuni concetti riguardo alla compensazione degli amplificatori operazionali. Lo scopo della compensazione in frequenza è quello di far sì che l'amplificatore possa essere chiuso in reazione senza che esso diventi instabile e si inneschino delle oscillazioni. La struttura di Figura 6.4.1 può essere schematizzata secondo lo schema seguente:

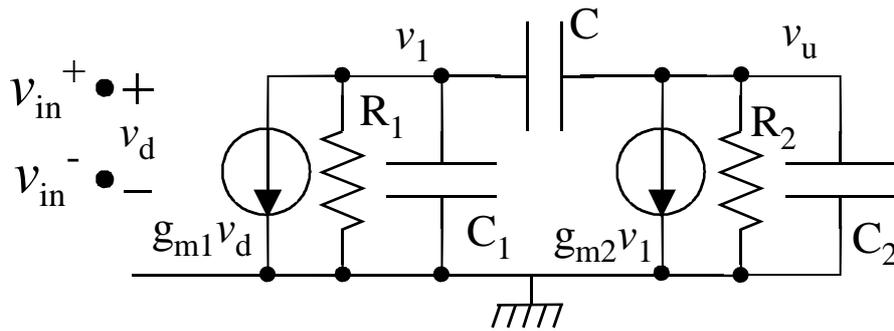


Figura 6.4.2

In questo schema sono stati utilizzati equivalenti di Norton per il primo e per il secondo stadio. Le resistenze R_1 e R_2 sono rispettivamente le resistenze di uscita dell'amplificatore differenziale e dello stadio di guadagno/uscita. La capacità C_1 è la somma della capacità di uscita del primo stadio e la capacità di ingresso del secondo. La capacità C_2 include la capacità di uscita del secondo stadio e la capacità di carico C_L . Il condensatore C viene introdotto opportunamente per attuare la compensazione.

Senza il condensatore di compensazione lo schema della figura precedente avrebbe sicuramente due poli a frequenza comparabile, di pulsazione ω_{1A} e ω_{2A} . Questo perché vi sono nel circuito due punti ad alta impedenza, ovvero l'uscita del primo stadio e l'uscita del secondo. Ciò porta quasi sicuramente ad una risposta in frequenza ad anello aperto del tipo di quella mostrata nella Figura 6.4.3.

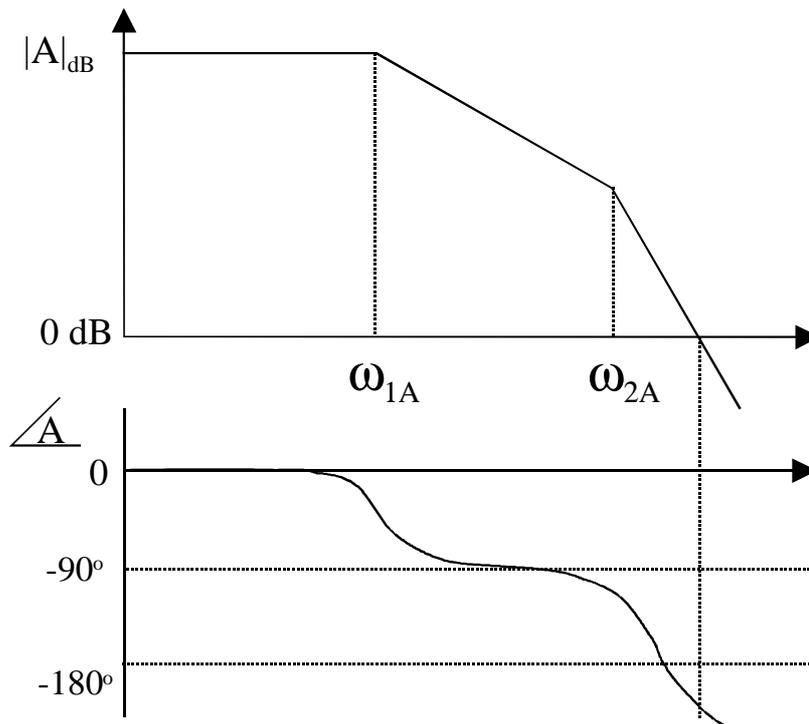


Figura 6.4.3

Siccome ciascun polo introduce un ritardo di fase di 90° , avremmo a frequenza sufficientemente superiore ad entrambi, una rotazione di -180° . Siccome i due poli sono a frequenze comparabili e il guadagno in continua relativamente elevato, si può comprendere come quando il modulo raggiunge il valore unitario (0 dB) la rotazione di fase dovuta a ω_{1A} e

ω_{2A} , sia ormai completata, ovvero valga praticamente -180° . In queste condizioni il margine di fase è nullo. Basta quindi la presenza di un'ulteriore piccola rotazione di fase dovuta a poli ad alta frequenza che, come mostrato in figura, la fase in corrispondenza del punto $|A|=0$ dB sia minore di -180° . In questa situazione non è possibile usare l'amplificatore chiuso in reazione con valori del β (guadagno della catena di reazione) prossimi all'unità, quindi non sarebbe possibile usare l'operazionale chiuso come buffer (ma neanche usarlo come integratore etc..).

L'aggiunta del condensatore di compensazione (C in Figura 6.4.3) modifica la risposta in frequenza provocando il cosiddetto "pole splitting", ovvero uno dei poli viene spostato a bassa frequenza e l'altro a frequenza più elevata. Se C è sufficientemente grande si raggiunge la situazione mostrata in Figura 6.4.4.

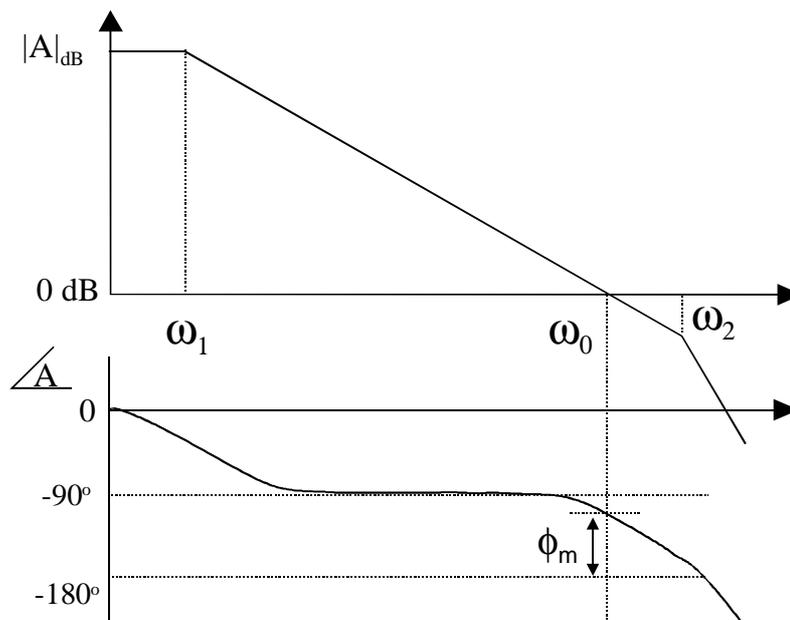


Figura 6.4.4

In questo caso il secondo polo, ω_2 , si trova oltre la pulsazione alla quale il diagramma di $|A|$ incrocia l'asse 0 dB. Questa pulsazione che, a parte un fattore 2π , è pari a prodotto guadagno banda (PGB), è stata indicata con ω_0 nella figura. In questo modo solo il polo a pulsazione ω_1 , avrà completato la rotazione di fase di 90° e pertanto la fase per $\omega = \omega_0$ sarà sufficientemente lontana dal valore -180° . La misura di questa distanza è data dal margine di fase, indicato in figura con ϕ_m .

Si dimostra che per C sufficientemente elevato (è sufficiente che sia molto più grande di C_1 oppure di C_2), si hanno le seguenti relazioni:

$$\omega_1 = \frac{1}{R_1 C g_{m2} R_2}; \quad \omega_2 = \frac{g_{m2}}{C_1 + C_2} \quad \omega_0 = \frac{g_{m1}}{C}$$

Affinché la risposta in frequenza sia del tipo di quella della Figura 6.4.4 (ovvero a polo dominante), si richiede che il secondo polo (ω_2) sia più grande di ω_0 . In genere si introduce un fattore di sicurezza σ , di solito pari a 3, in modo tale che sia:

$$\omega_2 = \sigma \omega_0$$

Sostituendo le espressioni di ω_2 e ω_0 , si ha una relazione che consente di calcolare il valore del condensatore C :

$$C = \sigma \frac{g_{m1}}{g_{m2}} (C_1 + C_2)$$

Se abbiamo già progettato il nostro amplificatore (quindi sono noti g_{m1} e g_{m2} , C_1 e C_2) allora l'espressione precedente ci fornisce univocamente il valore di C . Si noti che di solito $C_1 + C_2$ è quasi solo composto dalla capacità di carico C_L . C si dimensiona quindi in modo tale che l'amplificatore rimanga stabile per la massima capacità di carico prevista. Si dimostra facilmente che se la capacità di carico è inferiore o nulla la condizione di stabilità è verificata con un margine di fase addirittura superiore.

Una regola pratica di dimensionamento dell'amplificatore è quella di fare sì che il rapporto g_{m1}/g_{m2} sia pari a $1/\sigma$, ovvero a circa $1/3$. In queste condizioni si ha che $C = C_1 + C_2$ ovvero circa uguale al massimo valore di C_L .

Resta da fare una considerazione molto importante: l'analisi del circuito di Figura 6.4.2 mostra che nella risposta in frequenza è presente anche uno zero. Esso si trova a pulsazione:

$$\omega_z = \frac{g_{m2}}{C}$$

Questo zero va a posizionarsi in prossimità di ω_2 . Purtroppo questo zero è a parte reale positiva, quindi non solo non annulla il polo ω_2 , ma produce un ulteriore ritardo di fase (e non un benefico anticipo, come farebbe uno zero a parte reale negativa). Ciò può degenerare il margine di fase in modo inaccettabile. Per ovviare a questo problema si introduce in serie al condensatore di compensazione una resistenza R , come mostrato nel seguente schema modificato:

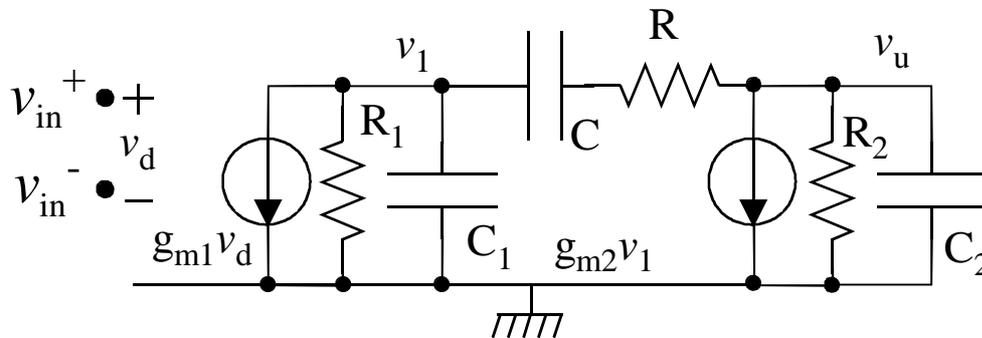


Figura 6.4.5

Si dimostra facilmente che lo zero assume in queste condizioni valore:

$$s_z = \frac{1}{C \left(\frac{1}{g_{m2}} - R \right)}$$

La situazione ottimale sarebbe quella di scegliere $R = 1/g_{m2}$ in modo da spostare lo zero all'infinito, ovvero eliminarlo. Di solito si sceglie R un po' più grande di $1/g_{m2}$ in modo da cambiare il segno allo zero e far sì che esso produca un anticipo di fase invece che un ritardo. Inoltre il fatto che R sia prossima a $1/g_{m2}$ sposterà comunque a frequenze più alte lo zero. La resistenza R , che spesso viene realizzata con MOSFET tenuti in zona lineare, si chiama resistore di annullamento dello zero (zero-nulling resistor).

Uno schema completo di un semplice amplificatore operazionale è mostrato in Figura 6.4.6.

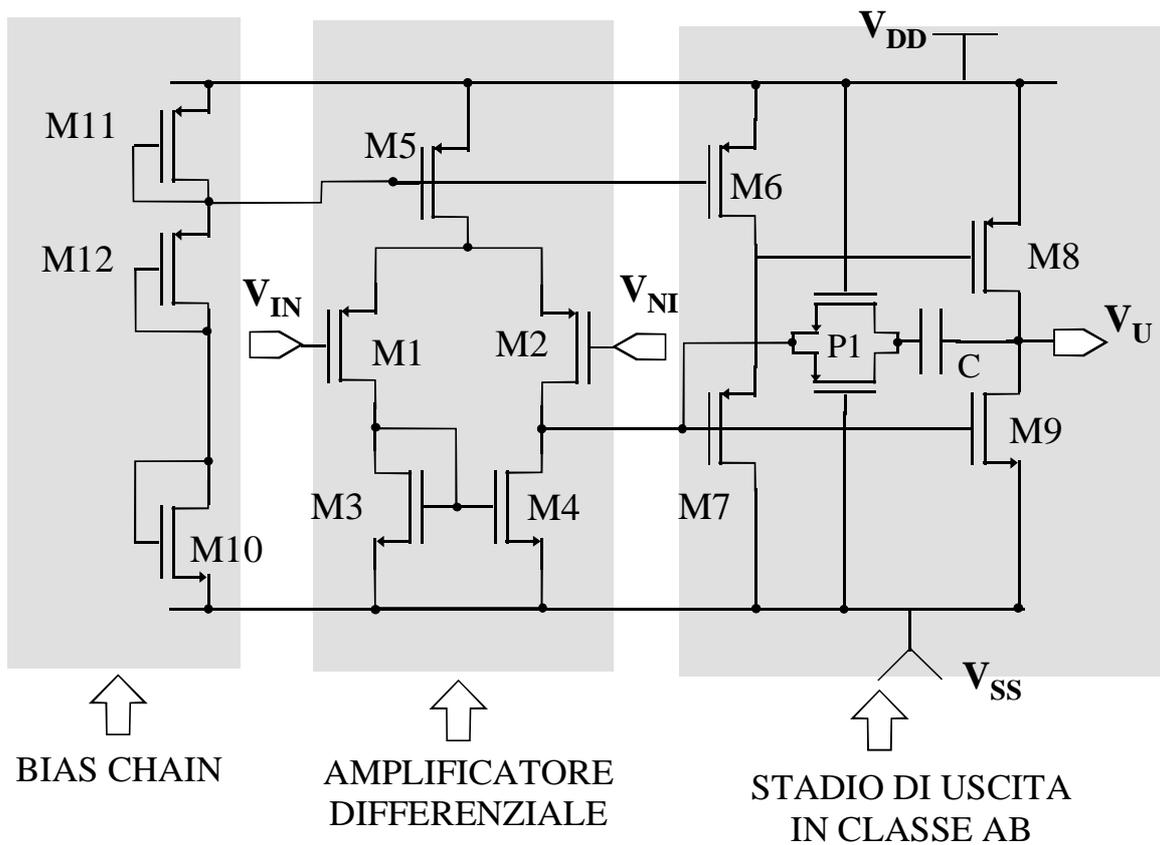


Figura 6.4.6

In questo schema possono essere distinti facilmente i seguenti blocchi funzionali:

1. Amplificatore differenziale di ingresso a transistori di ingresso di tipo p , costituito da M1-5.
2. Stadio di guadagno/uscita common source in classe AB, costituito dai transistori di uscita M8, M9 e dal traslatore di tensione M7, polarizzato dal generatore di corrente M6.
3. Condensatore di compensazione C con in serie una pass-gate (P1) in funzione di resistore di zero nulling. Si preferisce una pass-gate a un singolo n -MOS o p -MOS perché la pass-gate presenta una resistenza più costante al variare della tensione a cui è sottoposta.
4. Una bias chain di polarizzazione costituita dai transistori M11, M12, M10. La bias chain ha funzione di partitore di tensione per fornire la tensione di polarizzazione dei generatori di corrente M5 e M6.

Appendice A. Breve guida al simulatore elettrico SPICE

Lo scopo dei simulatori elettrici è l'analisi del comportamento di reti elettriche rappresentate da insiemi di nodi ai quali sono connessi i terminali dei dispositivi. I simulatori elettrici devono consentire la previsione dell'andamento di tutte le grandezze elettriche (tensioni e correnti) del circuito nel massimo dettaglio possibile. I componenti vengono rappresentati da equazioni che legano, nel modo più possibile aderente alla realtà, le varie grandezze elettriche (correnti e tensioni) presenti ai terminali dei componenti stessi. Il simulatore elettrico si distingue quindi dai simulatori logici, che non trattano tensioni e correnti ma solo gli stati logici che esse codificano e dai simulatori misti, che rappresentano i componenti in modo estremamente semplificato. I simulatori elettrici costituiscono uno strumento indispensabile per il progetto di circuiti analogici integrati e per l'analisi delle prestazioni di ogni nuova cella digitale che viene sviluppata.

Capostipite dei simulatori elettrici è il programma SPICE (*Simulation Program with Integrated Circuit Emphasis*), sviluppato nella sua prima versione all'inizio degli anni '70 presso l'università di Berkeley.

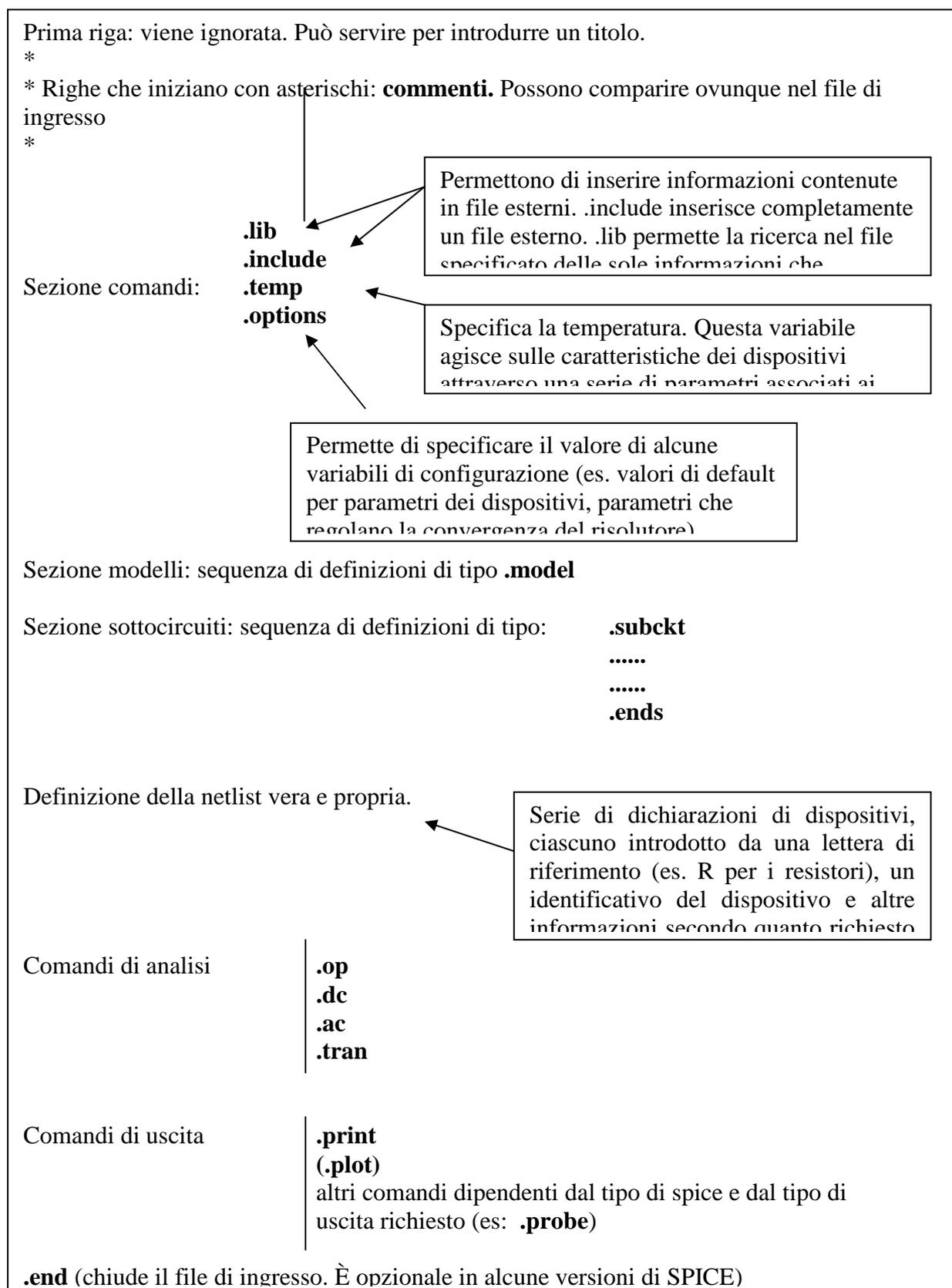
SPICE utilizza diverse strategie di risoluzione della rete a seconda del tipo di analisi che deve effettuare. Nel caso di calcolo del punto di lavoro, operazione preliminare a tutte le altre analisi, vengono risolte le equazioni della rete mediante il metodo iterativo di Newton Raphson, opportunamente modificato per migliorare la convergenza con le tipiche equazioni non lineari introdotte dai componenti a semiconduttore. Nel caso di analisi in transitorio vengono integrate le equazioni differenziali che regolano l'evoluzione delle varie variabili di stato (cariche sulle capacità e dei flussi magnetici nelle induttanze). L'analisi nel dominio della frequenza viene invece effettuata sul circuito linearizzato attorno al punto di riposo e prevede quindi la soluzione di sistemi lineari a coefficienti complessi.

La rete elettrica, gli stimoli applicati ad essa, il tipo di analisi da effettuare, i risultati da presentare in uscita ed altri parametri della simulazione (opzioni) vengono specificati in un file di ingresso testuale (ASCII), il cui formato verrà descritto nel seguito di questa appendice. Grazie a comandi opportuni inseriti nel file di input, si può comandare l'inclusione di altri file o indicare al simulatore quali file di libreria devono essere scanditi per la ricerca delle eventuali informazioni mancanti (tipicamente modelli di dispositivi o descrizione di sottocircuiti). Il simulatore produce sempre un file di uscita (tipicamente di estensione .out), sempre in formato ASCII, contenente i risultati della simulazione oppure, nell'eventualità di errori, la descrizione di questi ultimi.

Si fa osservare che la maggior parte dei simulatori elettrici più diffusi (SPICE2 e SPICE3 di Berkeley, HSPICE, ELDO, Spectre, PSPICE), utilizzano formati praticamente coincidenti per i file testuali di ingresso e uscita. Come risultato della simulazione si possono generalmente produrre file binari contenenti le funzioni di uscita in un formato adeguato per la rappresentazione grafica che viene effettuata da programmi distinti dal simulatore stesso. I formati di questi file dipendono esclusivamente dal tipo di visualizzatore supportato dal simulatore.

In questa guida riassuntiva viene descritta la sintassi dei comandi e delle dichiarazioni più comuni che possono essere introdotti nel file di input. Nella pagina seguente vengono mostrati gli elementi principali che costituiscono la struttura del file di input.

Struttura semplificata di un file di ingresso in formato SPICE



Convenzioni

In questa breve guida a SPICE verrà seguita la convenzione di indicare tra parentesi angolari “< >” i parametri variabili; e tra parentesi quadre [] gli argomenti opzionali. La sintassi di SPICE non fa distinzione tra caratteri maiuscoli e minuscoli. Accanto al formato generico, supportato universalmente da tutti i simulatori della famiglia di SPICE, verranno indicate quelle variazioni che caratterizzano due simulatori di interesse in ambito didattico: PSPICE incluso nell’ambiente di progettazione Orcad per schede a circuito stampato, e T-SPICE, che fa parte dell’ambiente di progettazione di circuiti integrati della Tanner.

Fattori moltiplicativi.

La maggior parte dei parametri indicati nel file di input sono espressi in unità del sistema internazionale. Vi sono alcune eccezioni, quali il parametro mobilità (UO) della “model” del MOSFET e i drogaggi, legati alla consuetudine di utilizzare unità differenti. I valori dei parametri possono essere indicati nei più comuni formati usati per i numeri, con eventuale aggiunta dell’esponente. Esempi: 1.26, 0.001, 1.7E-6. Per migliorare la leggibilità si possono usare i seguenti moltiplicatori:

F	10 ⁻¹⁵	(femto)	K	10 ³	(kilo)
P	10 ⁻¹²	(pico)	MEG	10 ⁶	(mega)
N	10 ⁻⁹	(nano)	G	10 ⁹	(giga)
U	10 ⁻⁶	(micro)	T	10 ¹²	(tera)
M	10 ⁻³	(milli)			

Componenti passivi.

Resistori

Sintassi:

```
R<nome> <N+> <N-> <valore>
```

Esempio:

```
R10 1 3 10k
```

```
***** Espansione PSPICE *****
```

```
R<nome> <N+> <N-> [nome_model] <valore>
```

La model per un resistore assume la forma:

```
.model <nome_model> RES (R=<moltiplicatore> TC1=<tc1> TC2=<tc2>
```

La resistenza risulta:

$$\text{Resistenza} = R \cdot \langle \text{valore} \rangle \cdot \left(1 + TC1(T - T_{nom}) + TC2(T - T_{nom})^2 \right).$$

```
***** Espansione T-Spice *****
```

```
R<nome> <N+> <N-> <valore> [TC=A, [B]]
```

con A che fa le veci di TC1 e B di TC2 in PSPICE.

Induttanze

Sintassi:

L<nome> <N+> <N-> <valore> [IC = <corrente iniziale>]

Nota: l'opzione IC che fissa la condizione iniziale (corrente) per il transitorio non è valida in T-SPICE.

Condensatori

Sintassi:

C<nome> <N+> <N-> <valore>

Esempio:

Cload 23 12 100n

***** Espansione PSPICE *****

C<nome> <N+> <N-> [<nome_model>] <valore> [IC=<Tensione iniziale>]

Tipo della model: CAP

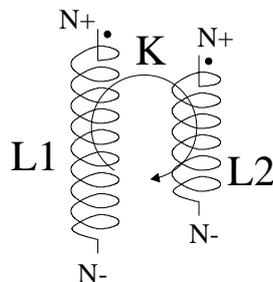
Parametri della model: C, VC1, VC2, TC1,TC2.

Espressione della capacità:

$$Capacità = C \cdot valore \cdot (1 + VC1 \cdot V + VC2 \cdot V^2) \cdot (1 + TC1(T - T_{nom}) + TC2(T - T_{nom})^2)$$

V è la tensione ai capi della capacità presa con polarità dal nodo N+ al nodo N-. Il parametro IC fissa il valore iniziale della tensione per le simulazioni in transitorio.

Mutue Induttanze



Date due induttanze già definite con l'apposita sintassi, si può introdurre un accoppiamento tra le due nel seguente modo:

K<nome> <induttanza 1> <induttanza 2> <valore di k>

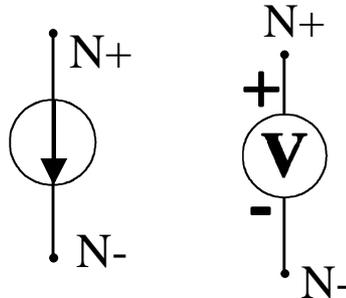
La figura chiarisce la relazione tra l'ordine con cui sono introdotti i nodi nella dichiarazione delle due induttanze e i "pallini" convenzionali che indicano la polarità delle tensioni indotte. Si riporta la relazione tra la mutua induttanza e il coefficiente K:

$$M = K \cdot \sqrt{L_1 L_2}$$

K deve essere compreso tra 0 e 1.

Generatori indipendenti

Esistono due tipi di generatori indipendenti: generatori di corrente e generatori di tensione. La figura mostra la convenzione sulle polarità dei generatori.



Generatori indipendenti di tensione

Sintassi:

V<nome> <N+> <N-> DC <valore in d.c.> [AC <ampiezza> <fase in gradi>]
[<funzione transitorio>]

Generatori indipendenti di corrente

Sintassi:

I <nome> <N+> <N-> DC <valore in d.c.> [AC <ampiezza> <fase in gradi>]
[<funzione transitorio>]

Significato dei campi che caratterizzano i generatori indipendenti:

DC: Fissa il valore del generatore per il calcolo del punto di riposo.

AC: Fissa i valori di ampiezza e fase da usare nell'analisi nel dominio della frequenza. Si tenga presente che in tale analisi tutti i generatori che hanno una specifica AC vengono accesi contemporaneamente (con i valori di ampiezza e fase specificati). La frequenza dei generatori viene fissata uguale per tutti e variata su un intervallo prestabilito dai parametri dell'analisi AC (si veda il paragrafo sul tipo di analisi). I generatori che non possiedono un campo AC sono spenti durante la suddetta analisi. Essi contribuiscono comunque a fissare il punto di lavoro attorno al quale il circuito viene linearizzato prima di effettuare l'analisi nel dominio della frequenza (analisi AC).

Funzione transitorio: specifica la forma d'onda che produce il generatore durante l'analisi in transitorio. I generatori che non possiedono questo campo rimangono fissati al valore specificato dal campo DC. La sintassi delle funzioni usate più comunemente è descritta nella pagina seguente.

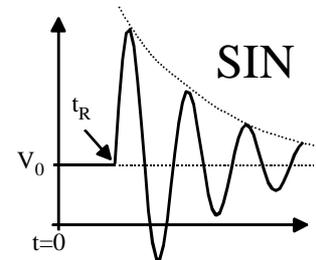
Sintassi di alcune funzioni che definiscono il comportamento dei generatori durante l'analisi in transitorio.

1) Sinusoidale con eventuale smorzamento.

Sintassi (varia molto a seconda del tipo di implementazione del simulatore):

SIN(<V₀> <A> <f> <t_R> <d> <φ>)

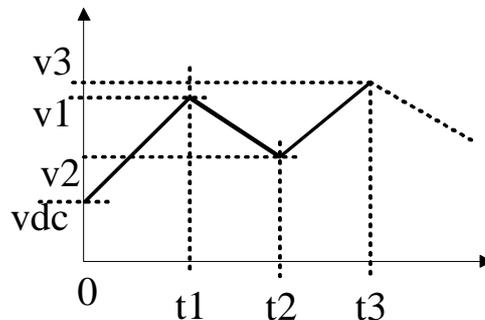
$$f(t) = \begin{cases} V_0 + A \cdot e^{-d \cdot (t-t_R)} \sin(2\pi f(t-t_R) + \varphi / 360) & \text{per } t > t_R \\ V_0 & \text{per } t < t_R \end{cases}$$



2) Lineare a tratti.

Sintassi:

PWL(<t1> <v1> <t2> <v2> <t3> <v3>)

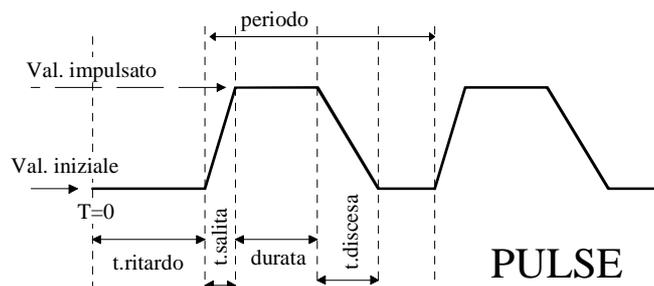


se t1 ≠ 0 per t=0 vale il valore in DC (come in figura).

3) Sequenza di impulsi.

Sintassi:

PULSE(<val. iniziale> <val. impulsato> <t. ritardo> <t. salita> <t. discesa> <durata> <periodo>)



Nota: in T-SPICE non è ammesso specificare un valore DC e contemporaneamente una funzione transitorio. Nel caso sia specificata la funzione transitorio il valore in continua è dato dal valore della funzione transitorio per t=0.

ESEMPI:

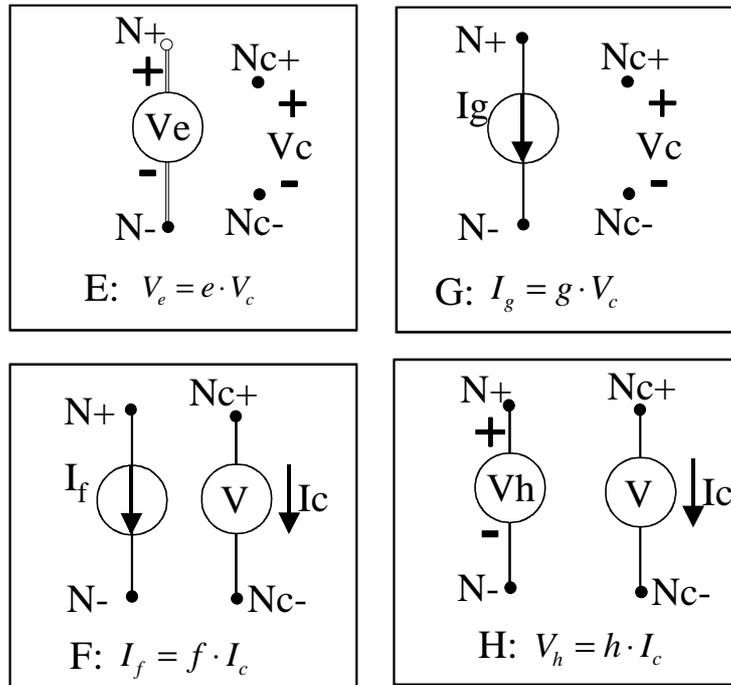
Vdd 11 0 DC 5

Vin 23 12 DC 0 AC 1 180

I0 4 0 DC 1m PWL(0 1m 1u 1m 1.1u 0 2u 0)

Generatori dipendenti

La figura seguente illustra i quattro possibili generatori dipendenti in SPICE. Si osservi che per i generatori comandati da una corrente (H e F), la corrente di comando deve scorrere in un generatore indipendente di tensione. Per far riferimento ad una corrente che scorre in un ramo in cui non sono presenti generatori indipendenti di tensione, si può introdurre in serie al ramo stesso un generatore opportuno di tensione nulla che non perturba il circuito e, nello stesso tempo, funziona da “amperometro”.



Di tensione controllato in tensione:

E<nome> <N+> <N-> <Nc+> <Nc-> <e>

Di corrente controllato in tensione:

G<nome> <N+> <N-> <Nc+> <Nc-> <g>

Di corrente controllato in corrente

F<nome> <N+> <N-> <sorgente> <f>

Di tensione controllato in corrente

H<nome> <N+> <N-> <sorgente> <h>

Nota: con “sorgente” si è indicato il generatore di tensione usato per poter riferire la corrente di comando.

Esempi di generatori dipendenti

```
E2 1 2 3 44 -1
Gms 2 33 21 3 1m
Fhfe1 1 2 Vre 300
H23 11 2 Vin 1e3
```

Componenti a semiconduttore

Nella dichiarazione di un qualsiasi dispositivo a semiconduttore si deve fare sempre riferimento alla model, ovvero a quella parte di file di input (o di libreria) dove vengono specificati parametri del dispositivo. Per la sintassi delle model si rimanda all'appendice 2.

Diodi

D<nome> <N+> <N-> <nome model> [<area>]

Il nodo N+ è l'anodo, mentre n- è il catodo.

Esempio di dichiarazione di un diodo.

```
D22 1 8 mio_diodo 2
```

Nota: il parametro area agisce come un parametro adimensionale di scala su alcuni parametri specificati nella model del diodo. Per esempio la corrente di saturazione IS viene moltiplicata per l'area, le resistenze serie vengono divise per l'area.

Transistori Bipolari

Q<nome> <collettore> <base> <emettitore> [<substrato>] <nome model>
<area>

Esempio di dichiarazione di BJT (si è omessa l'area che per default vale 1)

```
Q11 2 5 6 78 miobjt
```

Nota: collettore, base, etc. sono i nodi a cui sono collegati i corrispondenti terminali dei transistori.

Per quanto riguarda l'area vale quanto detto per il diodo.

Transistori MOS.

M<nome> <drain> <gate> <source> <substrato> <nome model> [L=<l>]
[W=<w>] [...altri parametri ..]

Esempio di definizione di MOSFET:

```
M2 2 5 11 3 mio_mos L=1u W=2u AD=2p AS=8.5p
```

La lista di parametri è costituita da definizioni di tipo: Parola chiave = <valore> come mostrato per i due parametri L e W. I parametri specificabili sono i seguenti:

Parametro	Significato	Default PSPICE	Default T-SPICE
L	Lunghezza del canale	defl	defl
W	Larghezza del canale	defw	defw
AD	Area di drain	defad	LW oppure defad
PD	Perimetro di drain	0	2(L+W)
AS	Area di source	defas	LW oppure defas
PS	Perimetro di source	0	2(L+W)
NRD	Numero di quadri della diffusione di drain	1	defnrd
NRS	Numero di quadri della diffusione di source	1	defnrs

Transistori JFET.

J<nome> <drain> <gate> <source> <nome_model> [<area>]

Esempio di dichiarazione di JFET:

J75 12 34 11 mio_jfet

Sottocircuiti

I sottocircuiti corrispondono ai moduli (o celle) dello schematic editor. L'uso dei sottocircuiti consente di velocizzare la scrittura della netlist quando una stessa parte di un circuito complessa è ripetuta molte volte (per esempio un operazionale, una porta logica, un flip-flop etc. etc.). Un altro vantaggio inerente l'uso di sottocircuiti è la maggior leggibilità della netlist, con conseguente maggior facilità diagnostica.

Occorre rammentare che SPICE, prima di risolvere la netlist, la espande sostituendo a ciascun sottocircuito il suo contenuto. Se sono presenti più livelli di gerarchia questo processo arriva fino a generare una netlist in cui compaiono solo componenti (celle primitive). Questa operazione viene indicata come "flattening".

Definizione:

```
.SUBCKT <nome_sottocircuito> <lista ordinata di nodi formali>
.....
elementi della rete costituente il sottocircuito
.....
.ENDS
```

Istanza:

```
X<nome> <lista ordinata di nodi > <nome_sottocircuito>
```

Esempio:

Definizione del sottocircuito:

```
.SUBCKT PASSA_BASSO 1 2
R1 1 2 100K
C1 2 0
.ENDS
```

Istanze del sottocircuito (2 passa basso in cascata)

```
X3 44 55 PASSA_BASSO
X4 55 66 PASSA_BASSO
```

Note sull'uso dei sottocircuiti.

- All'interno della rete possono essere introdotti nodi ausiliari non visibili all'esterno.
- Il nodo "0" (nodo di riferimento, denominato anche "gnd" nei programmi che accettano nodi letterali) è sempre un nodo globale, ovvero è visibile ed è lo stesso all'interno di tutti i sottocircuiti. Non occorre quindi passarlo nei nodi formali e non può essere usato come nodo locale (ovvero visibile solo nel sottocircuito).
- In alcune versioni del simulatore è possibile introdurre altri nodi globali (per esempio i *rail* di alimentazione), in modo da migliorare la leggibilità della *netlist*. Si noti l'analogia con quanto accade nello schematic editor.
- Ovviamente un sottocircuito può non contenere connessioni interne a nodi globali.
- La corrispondenza tra nodi formali (nella definizione del sottocircuito) e nodi del circuito utilizzatore (nell'istanza) è stabilita esclusivamente dall'ordine.
- Un sottocircuito può contenere istanze di altri sottocircuiti (struttura gerarchica).
- I le grandezze elettriche relative a componenti e nodi interni ai sottocircuiti possono essere sempre rintracciati, indicandoli con la seguente sintassi:
<Nome istanza>.<grandezza elettrica>. Esempio: X27.V(6) (tensione del nodo 6 interno all'istanza di sottocircuito denominata X27).

Comandi di Spice

I comandi di SPICE si distinguono dalle istanze dei dispositivi che costituiscono la *netlist* attraverso il punto con il quale iniziano le relative parole chiave. Sono già stati introdotti i comandi relativi alla definizione di un sottocircuito all'introduzione della *model* dei dispositivi. In questa sezione verrà completata la descrizione dei comandi più comuni.

Scelta del tipo di analisi

Determinazione del punto di riposo.

Sintassi:

.OP

il comando .OP ha una funzione che dipende dal tipo di Spice che si sta utilizzando:

-) in T-SPICE comanda effettivamente il calcolo del punto di riposo che viene effettuato considerando come note le variabili (tensioni e correnti) fissate con il comando .IC. I valori specificati con .NODESET vengono utilizzati per fornire l'inizio del procedimento iterativo. Per riportare i risultati nel file .out occorre specificare nel comando ".print DC" le variabili di cui si desidera conoscere il valore. Sempre in T-Spice, per avere in uscita le informazioni sui parametri di piccolo segnale dei dispositivi attivi occorre aggiungere una riga di tipo:

.ACMODEL { * }

-) in PSPICE il punto di riposo viene calcolato sempre e .OP fa sì che vengano riportate nel file .out informazioni dettagliate sul punto di riposo (parametri dei circuiti linearizzati) che altrimenti sarebbero limitate alle sole tensioni nodali.

Nel calcolo del punto di riposo i generatori indipendenti vengono posti al loro valore in continua. Anche qui c'è una differenza tra T-SPICE e PSPICE: in T-SPICE nel caso un generatore abbia specificata una forma d'onda per l'analisi in transitorio viene usato come valore in continua il valore assunto dalla forma d'onda per $t=0$. Nel caso di PSPICE è invece possibile specificare un valore in continua (parola chiave DC) indipendente dalla funzione usata per il transitorio.

Scansione in continua (“sweep d.c.”).

Sintassi:

```
.DC V<nome1> <Vi1> <Vf1> <Vstep1> [V<nome2> <Vi2> <Vf2> <Vstep2>]
```

Il generatore V<nome1> viene incrementato da Vi1 a Vf1 con passo pari a Vstep1. Per ogni valore del generatore V<nome1> viene effettuato il calcolo del punto di riposo (analisi in continua). Se è presente anche la parte opzionale il ciclo viene ripetuto incrementando il generatore V<nome2> da Vi2 a Vf2 con passo Vstep2. In pratica alla fine di ogni ciclo del generatore 1 viene incrementato di un passo il generatore 2. Tutti gli altri generatori indipendenti sono mantenuti fissi al loro valore in continua.

Analisi “AC” (analisi nel dominio della in frequenza).

Sintassi:

```
.AC <suddivisione > <numero_di_punti> <freq. iniz.> <freq. finale>
```

Con l'analisi AC viene linearizzato il circuito attorno al punto di riposo (stabilito dal valore DC dei generatori) e su di esso vengono fatti agire contemporaneamente tutti i generatori indipendenti che hanno specificati i dati per l'analisi AC. Gli altri generatori indipendenti sono cortocircuitati (V) oppure aperti (I) nel circuito linearizzato. Per ogni valore di frequenza richiesto viene calcolato il regime sinusoidale e i risultati sono fasori delle tensioni o correnti. Ovviamente anche il circuito linearizzato è composto di impedenze complesse che vengono ricalcolate per ogni valore di frequenza a cui si effettua l'analisi. Se ad agire è un solo generatore, di ampiezza AC unitaria e fase nulla, le tensioni (o correnti) fornite dall'analisi AC coincidono, in modulo e fase, con la risposta in frequenza rispetto al dato generatore.

Il parametro **suddivisione** può assumere i tre valori: **LIN**, **DEC** e **OCT** e fissa la suddivisione dell'asse delle frequenze in lineare, a decadi o a ottave.

Il numero dei punti è il numero totale di punti sull'asse delle frequenze nel caso **LIN** oppure è il numero di punti per decade o per ottava nei casi **DEC** o **OCT**.

Per esempio la dichiarazione:

```
.AC DEC 9 1 1k
```

Comanda l'esecuzione del calcolo del regime sinusoidale (sul circuito linearizzato!) alle frequenze: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40600, 700, 800, 900, 1000.

Funzione di trasferimento in continua.

Sintassi:

```
.TF <variabili dipendendenti> <sorgente indipendente>
```

.TF consente il calcolo della funzione di trasferimento di piccolo segnale (si opera sul circuito linearizzato), in continua, tra un generatore indipendente specificato nel parametro “sorgente indipendente” e una serie di variabili (correnti o tensioni). In sostanza l'analisi è

equivalente a valutare il limite della funzione di trasferimento per la frequenza che tende a zero. Per la sintassi da utilizzare per specificare le variabili dipendenti si veda più avanti il comando “PRINT”.

PSPICE consente una sola variabile dipendente per comando .TF e non necessita di comando ulteriore (es. .PRINT) per produrre il risultato sul file di output.

T-SPICE consente più variabili ma necessita che si specifichi un comando “.PRINT DC <variabili dipendenti>”.

Analisi in transitorio.

Viene effettuata un’analisi di transitorio facendo agire i generatori indipendenti per i quali è stata specificata una forma d’onda. Gli altri generatori indipendenti vengono mantenuti al valore in continua. Il circuito su cui viene calcolato il transitorio non viene linearizzato ma vengono mantenute tutte le non linearità dei componenti.

La sintassi e il significato dei parametri è piuttosto differente tra PSPICE e TSPICE e conviene quindi mostrarle separatamente.

Sintassi PSPICE (comune alla maggioranza delle implementazioni di SPICE)

```
.TRAN <step di stampa> <tempo finale> [<inizio stampa>  
<step interno massimo>] [UIC]
```

L’analisi inizia sempre dall’istante $t=0$ e finisce per il tempo specificato con <tempo finale>. Si può evitare di riportare nei risultati quello che accade prima di un certo istante specificando il parametro opzionale “inizio stampa”.

È importante chiarire il significato del parametro “step interno massimo”. SPICE varia il passo temporale di risoluzione delle equazioni differenziali che regolano il transitorio in modo tale da mantenere la precisione richiesta. In particolare, in presenza di grandezze elettriche che tendono a variare molto più velocemente del passo di risoluzione in vigore, SPICE riduce il passo stesso. Viceversa, se tutte le grandezze elettriche variano lentamente, SPICE “allunga il passo”, in quanto gli è possibile prevedere correttamente il nuovo stato del circuito senza dover effettuare step intermedi. Spesso SPICE fissa un passo di risoluzione molto più grande dell’intervallo di stampa. In questo caso tutti i punti intermedi richiesti per la stampa non sono calcolati esattamente da SPICE ma interpolati tra due risultati esatti. L’effetto si manifesta in una eccessiva angolosità delle curve. Specificando il parametro step interno massimo si fissa un valore massimo per il passo di risoluzione delle equazioni. Fissando tale parametro ad un valore pari o inferiore allo step di stampa si può far sì che i punti forniti in uscita siano tutti caratterizzati da una precisione pari a quella stabilita per la risoluzione del circuito. Ovviamente in questo modo si impedisce a SPICE di “allungare il passo”, e ne risulta aumentato il tempo totale di calcolo.

Nota! se si vuole specificare lo <step interno massimo> occorre anche specificare <inizio stampa>.

La parola chiave [UIC] forza il programma a utilizzare le condizioni iniziali specificate con il comando .IC o specificate direttamente nella definizione dei componenti (IC=...). Senza UIC le condizioni iniziali sono quelle valutate con l’analisi del punto di riposo.

Sintassi T-SPICE

```
.TRAN [/<mode>] < step interno massimo > <tempo finale>  
[START=<inizio intervallo di stampa>]
```

In questo caso viene fissato come parametro obbligatorio il passo di risoluzione delle equazioni (“step interno massimo”) e l’intervallo di tempo simulato (“tempo finale”). Anche qui il transitorio inizia sempre per $t=0$ ma i risultati possono essere forniti in uscita a partire da un <inizio intervallo di stampa>.

Lo step di stampa viene invece fissato pari allo “step interno massimo” e non si hanno quindi i problemi di precisione citati per PSPICE (e tipici anche di molte altre versioni di SPICE, compreso SPICE 2G). L’intervallo di stampa può comunque essere reso differente da “step interno massimo” specificando il parametro “prtdel” in una dichiarazione .OPTIONS.

Il parametro “mode” accetta tre valori (di cui il primo è di default).

dcoppt: viene calcolato inizialmente il punto di riposo usando le condizioni iniziali se specificate nel comando “.IC” e i suggerimenti per la convergenza fissati il comando “.NODESET”.

powerup: viene aiutata la convergenza facendo partire tutti i generatori da zero e facendoli arrivare al valore corretto (per $t=0$) gradualmente. Una volta trovato il punto di riposo per $t=0$ si procede alla simulazione.

preview: serve per controllare se le forme d’onda di ingresso sono corrette: non viene effettuata la simulazione ma vengono riportate in uscita le forme d’onda dei generatori indipendenti specificati nel comando “.PRINT TRAN”

Altri comandi di Spice:

Comandi di universali di output.

```
.PRINT <analisi> <variabili>
```

Il comando .PRINT consente di riportare in modalità di file testuale i risultati di una delle analisi eseguite attraverso gli appositi comandi (per esempio .TRAN). I risultati vengono inseriti nel file di output testuale (.out). Il parametro <analisi> stabilisce che tipo di analisi si vuole riportare come risultato nel file di output.

<analisi> può assumere i valori DC AC TRAN e NOISE che si riferiscono alle rispettive analisi circuitali.

Le variabili sono dichiarate con una sintassi del tipo:

V(<nodo>) per le tensioni.

Per le correnti, si può avere riferimento alle sole correnti nei generatori di tensione e quindi occorre usare la sintassi:

I(V<nome_gen>), dove nome_gen è il nome di un generatore indipendente di tensione.

***** Espansione TSPICE *****

TSPICE consente anche di dichiarare altre variabili tra cui la tensione tra due nodi con la sintassi:

V(<N+>,<N->)

e la corrente entrante in un dispositivo <disp> attraverso il nodo <N> con la sintassi:

I(<disp>,<N>).

Comandi di uscita per post-processor grafici:

TSPICE: è sufficiente il comando .PRINT: il post-processore grafico chiamato WINWED32 legge i file testuali di tipo “.out”.

PSPICE: si può produrre un file di dati (.dat) per il post-processore WPROBE attraverso il comando:

```
. PROBE
```

che fa sì che tutte le tensioni nodali e tutte le correnti in dispositivi attivi e generatori di tensione siano salvate nel file “.dat” Molte implementazioni di spice supportano il comando PROBE. Per restringere l’insieme di variabili salvate nel file si può far seguire al comando PROBE una lista di variabili, come per il comando PRINT.

Utilizzo di informazioni presenti su file esterni:

Consultazioni di file di libreria.

```
.LIB <file di libreria>
```

Questo comando fa sì che venga consultato il file specificato per ricercare definizioni di sottocircuiti o “model” di dispositivi non presenti nel file di input ma utilizzati per la descrizione della rete. I file di libreria sono file di testo che possono contenere solo definizioni di sottocircuiti o “model”. Il programma crea in memoria un file di input espanso nel quale inserisce solo quelle sezioni di librerie che gli servono per la rete in esame (model di dispositivi o definizioni di sottocircuiti citati nel file di input principale).

Inclusione completa di file esterni.

```
.INCLUDE < file di include> (sintassi di TSPICE)
```

```
.INC < file di include> (sintassi di PSPICE)
```

Questi comandi fanno sì che il file specificato venga inserito integralmente nel file di input nel punto esatto in cui tale comando compare. Il file di include può contenere qualsiasi comando o definizione di SPICE. Il file di input viene processato come un unico blocco dopo che sono stati inseriti tutti i file di include.

Specificazione della temperatura.

Sintassi:

```
.TEMP <valore di temperatura>
```

fissa il valore della temperatura alla quale viene effettuata la simulazione. Questo parametro agisce su tutte quelle equazioni costitutive dei dispositivi per le quale è stata inserita una

dipendenza dalla temperatura. Se non è presente il comando “.TEMP” la temperatura è pari all’ “option” TNOM che per default vale 27 °C.

Assegnazione dei parametri globali (options).

Sintassi:

```
.OPTIONS <nome par.>=<valore par.> <nome par.>=<valore par.> .....
```

Esempio:

```
.OPTIONS TNOM=30 DEFW=10u DEFL=2u
```

Il comando .OPTIONS consente di fissare il valore di alcuni parametri che possono influenzare lo svolgimento della simulazione. Il nome e il numero dei parametri cambia a seconda della versione del simulatore. Per ottenere la visualizzazione di tutte le opzioni correnti (nel file .out) si può usare la seguente sintassi:

```
.OPTIONS OPT (sintassi PSPICE)
```

```
.OPTIONS verbose=2 (sintassi T-SPICE)
```

Lista dei parametri di uso più frequente

PSPICE		T-SPICE		Significato del parametro
nome	default	nome	default	
TNOM	27 °C	TNOM	27 °C	temperatura nominale
DEFW	100 μm	DEFW	0	W di default per i MOSFET
DEFL	100 μm	DEFL	0	L di default per i MOSFET
DEFAD	0	DEFAD	0	AD di default per i MOSFET
DEFAS	0	DEFAS	0	AS di default per i MOSFET
GMIN	10 ⁻¹² Ω ⁻¹			Conduttanza messa in parallelo a tutte le
ITL1	40	numnd	100	Max Numero Iterazioni per il calcolo del punto
ITL4	10	numnt	50	Max. Numero di iterazioni per passo per il
ITL5	5000			Max numero di Iterazioni totali per il
LIMPTS	201			Numero massimo di punti in una tabella di
ABSTOL	1 pA	ABSTOL	5 nA	Errore assoluto sulle correnti
RELTOL	10 ⁻¹³	RELTOL	10 ⁻⁷	Errore relativo su correnti e tensioni (solo sulle tensioni in T-SPICE)
VNTOL	1 μV			Errore assoluto sulle tensioni
CHGTOL		CHARGETOL		Errore assoluto sulle cariche nel transitorio.
		RELCHARGETO	10 ⁻³	

Assegnazione di un punto iniziale per aiutare la convergenza in SPICE:

Questo comando consente di suggerire al simulatore il valore della tensione di alcuni nodi. Questi valori sono utilizzati da SPICE come punto di partenza per il metodo iterativo usato per il calcolo del punto di lavoro. Si rammenta che spesso il calcolo del punto di lavoro è il

compito più difficile che deve essere affrontato dal simulatore in quanto le equazioni in gioco sono fortemente non lineari, e, in mancanza di suggerimenti, il simulatore non possiede un valido punto di partenza da cui cominciare le iterazioni. Va precisato anche che tutte le altre analisi si basano sempre sul calcolo di un punto di lavoro come fase preliminare. Ovviamente il risultato del calcolo del punto di lavoro potrà essere completamente diverso da quello suggerito dall'utente mediante il comando NODESET.

Il comando NODESET può risultare utile quando non si riesce a raggiungere la convergenza e la tensione di alcuni nodi sia di facile determinazione da parte dell'utente.

Inoltre, nel caso di circuiti che presentino più stati stabili (per esempio flip-flop), l'uso del comando NODESET può favorire la selezione di una delle varie soluzioni.

Sintassi:

```
.NODESET V(<nodo1>)=<valore1> V(<nodo2>)=<valore2>
```

Assegnazione di condizioni iniziali per il transitorio

L'assegnazione di condizioni iniziali viene utilizzata per fissare la tensione sui condensatori e la corrente nelle induttanze al tempo $t=0$, nell'analisi in transitorio (.TRAN). Non si devono confondere le condizioni iniziali con i valori assegnati con il comando NODESET. Solo nel caso di condizioni iniziali si ha un effettivo significato fisico delle grandezze assegnate e la soluzione (analisi nel dominio del tempo) ne risulta fortemente influenzata. In tutte le analisi diverse dal transitorio le condizioni iniziali sono ignorate (non hanno significato).

L'uso delle condizioni iniziali può essere spesso evitato in quanto, nella maggior parte dei casi di interesse pratico, si sceglie una soluzione a regime, o comunque si inizia ad applicare il segnale (per esempio una sequenza di bit) dopo un intervallo di tempo iniziale introdotto ad arte per far raggiungere al circuito una condizione di regime.

Un caso in cui le condizioni iniziali sono di estrema importanza è quello degli oscillatori. Infatti, non essendo presente nel circuito simulato il rumore che garantisce l'innesco delle oscillazioni in un circuito reale, è possibile spostare dall'equilibrio l'oscillatore assegnando condizioni iniziali diverse da quelle determinate dal punto di lavoro su alcuni elementi reattivi e innescare così l'oscillazione.

Sintassi PSPICE:

```
.IC V(<nodo1>)=<valore1>....
```

Sintassi T-SPICE:

```
.IC V(<nodo1>)=<valore1> . V(<nodo2>,<nodo3>)=<valore>  
I(induttore1)=<valore>....
```

Si noti che PSPICE ha una sintassi apparentemente meno potente ma consente di assegnare le condizioni iniziali anche direttamente sui alcuni componenti (condensatori e induttori). Si veda in proposito la sintassi di dichiarazione dei singoli componenti.

Appendice B. Cenni ai modelli usati da SPICE per gli elementi a semiconduttore.

Tutti gli elementi a semiconduttore presenti nella netlist devono fare riferimento ad una “model”, ovvero ad una sezione del file di input (o di file di libreria), dove vengono specificati i valori dei parametri che definiscono le caratteristiche del componente stesso. Le model hanno il seguente formato:

```
.MODEL <nome_model> <tipo_model> <lista di parametri>
```

Il “nome_model” è un nome convenzionale assegnato dall’utente o, più comunemente, da chi ha compilato la libreria.

Il “tipo_model” specifica di che tipo di componente si tratta (diodo, transistor bipolare etc. etc.). Questa stringa può assumere solo alcuni valori predefiniti stabiliti dal tipo di implementazione di SPICE. I dispositivi di interesse microelettronico presenti in tutte le versioni di SPICE e di simulatori affini sono:

D	diodo
NPN	bjt npn
PNP	bjt pnp
NMOS	n-mosfet
PMOS	p-mosfet
NJF	n-jfet
PJF	p-jfet

La lista di parametri ha un formato del tipo: <nome_parametro>=<valore>. Per esempio: VTO=1.23 (vissa il valore della tensione di soglia di un mosfet). Una distinzione importante riguarda la model dei transistori MOS. In questo caso il primo parametro della lista è del tipo:

```
LEVEL=<numero>
```

Il LEVEL fissa il tipo di modello usato per descrivere il comportamento del mosfet. Per esempio, esistono tre livelli indicati con LEVEL=1, LEVEL=2 e LEVEL=3, supportati da tutte le implementazioni di SPICE e programmi affini. Questi tre livelli rappresentano il comportamento del MOSFET con ordine di complessità crescente. Il LEVEL=1, al quale per semplicità si farà riferimento in questa trattazione, aveva importanza solo i tempi in cui la ridotta potenza di calcolo degli elaboratori poteva sconsigliare l’uso di modelli più complicati.

Attualmente i simulatori supportano un numero elevato di modelli per il mosfet, introdotti per rappresentare quei fenomeni connessi con le dimensioni sub-micrometriche dei moderni dispositivi.

In questa trattazione si farà riferimento ad alcuni parametri validi per i tre livelli base (1,2,3), spesso mantenuti anche per i livelli superiori. In ogni caso, i manuali riportano sempre per ciascun modello supportato, una lista dettagliata dei parametri con il loro significato e l’unità di misura.

Nei LEVEL superiori al primo l'espressione della I_{DS} rimane una funzione delle tre tensioni V_{GS} , V_{DS} e V_{BS} ma risulta più complessa e, in particolare, risulta meno discontinuo il passaggio dalla saturazione alla zona triodo e, per modelli più recenti, dalla zona di forte inversione alla zona di sottosoglia. I parametri introdotti nella tabella precedente risultano ancora tutti validi tranne il parametro LAMBDA che non ha più significato nel LEVEL=3.

Nel LEVEL=3 l'espressione della corrente in saturazione è data da:

$$I_{DS} = I_{DSAT} \frac{1}{1 - \frac{\Delta L}{L_1}} \quad (A2.3)$$

dove I_{DSAT} è la corrente al limite della saturazione, L_1 è la lunghezza efficace del canale prima dello strozzamento mentre ΔL è l'accorciamento del canale (valore positivo) dovuto allo strozzamento. ΔL è quindi una funzione crescente di V_{DS} .

Dall'espressione (3) può essere calcolato un valore di λ considerando che, dalla (A2.1) in saturazione:

$$\lambda = \frac{1}{I_{DSAT}} \frac{\partial I_{DS}}{\partial V_{DS}}$$

e pertanto, derivando la (A2.3), se si può approssimare $1 - \Delta L/L_1$ a circa 1, si ottiene:

$$\lambda = \frac{1}{L_1} \frac{\partial \Delta L}{\partial V_{DS}} \quad (A2.4)$$

In genere $\frac{\partial \Delta L}{\partial V_{DS}}$ si può considerare che dipenda solo da $V_{DS} - V_{DSAT}$ e pertanto il valore di

λ diminuisce al crescere della lunghezza di canale. Ciò corrisponde alla realtà sperimentale ed è ben rappresentato da simulazioni effettuate sia con LEVEL=2 sia con LEVEL=3.

I diodi source-substrato (D_{BS}) e drain-substrato (D_{BD}) sono caratterizzati da una corrente di saturazione $I_S = JS \times AS$ per D_{BS} e $I_S = JS \times AD$ per il diodo D_{BD} . Il parametro JS è fornito nelle model, mentre AS e AD sono specificati nella dichiarazione del dispositivo nella netlist.

CAPACITÀ NEL MODELLO DEL MOS

Con riferimento al modello per grandi segnali abbiamo sei capacità, tutte variabili con la tensione a cui sono sottoposte. Pertanto si parlerà di capacità differenziali.

Capacità di gate: C_{GS} C_{GD} C_{GB} . Queste capacità possono essere scritte come:

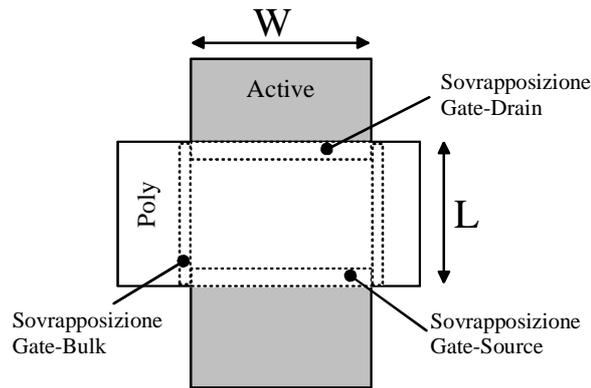
$$C_{GS} = C_{GS}^{(c)} + C_{GS}^{(s)}; \quad C_{GD} = C_{GD}^{(c)} + C_{GD}^{(s)}; \quad C_{GB} = C_{GB}^{(c)} + C_{GB}^{(s)}$$

dove l'apice (c) indica "capacità di canale" e l'apice (s) indica la capacità dovuta alla sovrapposizione del gate con il rispettivo elettrodo.

Le capacità di sovrapposizione si ottengono dai parametri delle model CGSO, CGDO e CGBO. In particolare:

$$C_{GS}^{(s)} = CGSO \times W; \quad C_{GD}^{(s)} = CGDO \times W; \quad C_{GB}^{(s)} = CGBO \times L$$

La figura seguente mostra il layout di un MOSFET evidenziando l'origine delle capacità di sovrapposizione.



Per quanto riguarda le capacità di canale, esse variano fortemente con la zona di funzionamento del MOS. Nella formulazione più semplice data dal LEVEL 1 si ha:

	$V_{GS} < V_T$	Zona Triodo	Saturazione
$C_{GS}^{(c)}$	0	$1/2 LWC_{OX}$	$2/3 LWC_{OX}$
$C_{GD}^{(c)}$	0	$1/2 LWC_{OX}$	0
$C_{GB}^{(c)}$	LWC_{OX}	0	0

Capacità di giunzione: C_{BS} e C_{BD}

Queste capacità sono date ciascuna da due termini corrispondenti rispettivamente al “fondo” e alle “pareti” delle diffusioni di drain e di source.

$$C_{BS} = \frac{CJ \times AS}{\left(1 - \frac{V_{BS}}{PB}\right)^{MJ}} + \frac{CJSW \times PS}{\left(1 - \frac{V_{BS}}{PB}\right)^{MJSW}} ; \quad C_{BD} = \frac{CJ \times AS}{\left(1 - \frac{V_{BD}}{PB}\right)^{MJ}} + \frac{CJSW \times PS}{\left(1 - \frac{V_{BD}}{PB}\right)^{MJSW}}$$

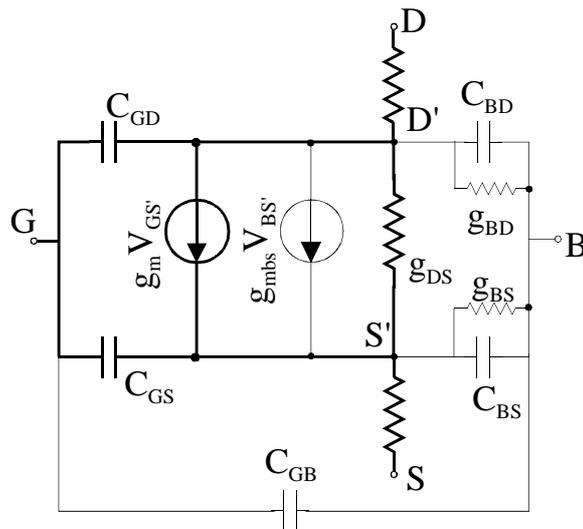
I parametri CJ, CGSW, MJ, MJSW e PB sono indicati nella model, mentre le aree AS e AD e i perimetri PS e PD vengono indicati nella dichiarazione del dispositivo nella netlist.

Resistenze serie: R_D e R_S

Le resistenze serie vengono calcolate attraverso il parametro RSH (resistenza di strato) fornito nella model.

Si ha : $R_D = NRD \times RSH$ e $R_S = NRS \times RSH$. NRD e NRS vengono indicati nella dichiarazione del MOSFET.

Modello a piccolo segnale dei MOSFET

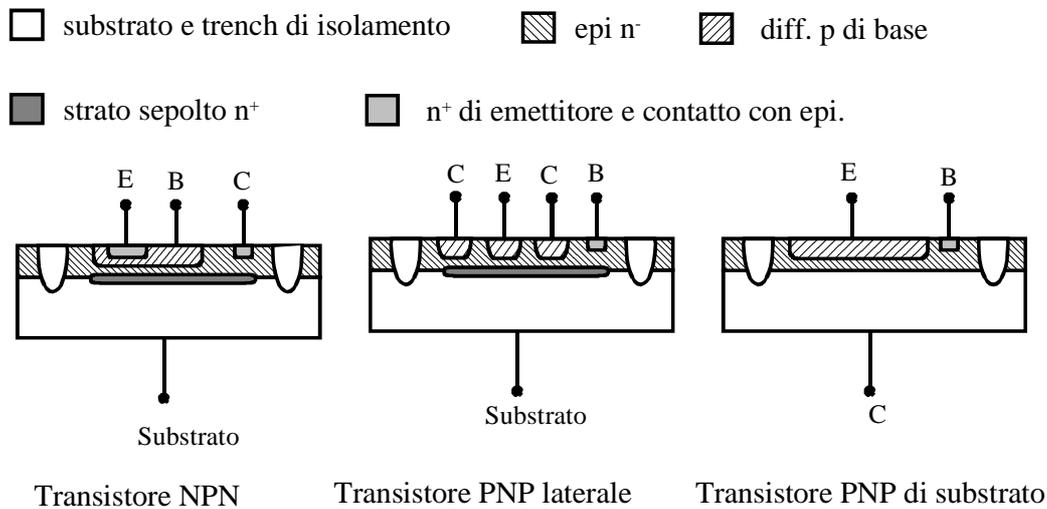


Riassunto di alcuni parametri della MODEL del MOSFET

LEVEL	Tipo di modello selezionato. (modelli universali di Spice: 1,2,3)
VTO	Tensione di soglia per $V_{BS}=0$ e in assenza di effetti di canale corto o stretto.
UO	Mobilità
LD	Riduzione della lunghezza del canale su ciascun lato
WD	Riduzione della larghezza del canale su ciascun lato.
LAMBDA	Pendenza delle caratteristiche in saturazione (non valido per il level 3)
GAMMA	Coefficiente dell'effetto body
PHI	Potenziale di superficie
RSH	Resistenza di strato delle diffusioni di Drain e Source
TOX	Spessore dell'ossido
JS	Densità di corrente di saturazione delle giunzioni tra diffusioni e substrato
CJ	Capacità di giunzione a potenziale nullo (da moltiplicare per l'area della giunz.)
CJSW	Capacità di giunzione a potenziale nullo (da moltiplicare per il perimetro)
MJ	Coefficiente di grading per le giunzioni (relativo al termine di area)
MJSW	Coefficiente di grading per le giunzioni (relativo al termine di perimetro)
PB	Potenziale di contatto delle giunzioni tra diffusioni e substrato
CGSO	Capacità di sovrapposizione gate-source (da moltiplicare per W)
CGDO	Capacità di sovrapposizione gate-drain (da moltiplicare per W)
CGBO	Capacità di sovrapposizione gate-bulk (da moltiplicare per L)

Transistori bipolari (BJT)

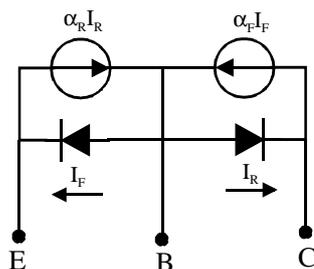
Nella figura seguente vengono indicate schematicamente le tipiche sezioni dei dispositivi bipolari per circuiti integrati. Lo scopo della figura è far comprendere l'origine di alcuni elementi, quali, per esempio, le capacità verso il substrato che non sono presenti in un componente discreto (provvisto dei soli tre terminali base emettitore e collettore). Gli schemi fanno riferimento a un processo bipolare planare epitassiale o in generale a processi in cui è presente uno strato epitassiale di tipo n , quali alcuni tipi di BiCMOS e di BCD. I componenti sono sempre racchiusi in sezioni di strato epitassiale, denominate "sacche" (*epi pocket*) isolate in basso dalla giunzione in inversa con il substrato (di tipo p) e lateralmente da diffusioni di tipo p o da scavi (trench) di isolamento. Le sezioni sono comunque egualmente rappresentative di quei processi BiCMOS nei quali la funzione delle "sacche" di epi è svolta da n -well.



La dichiarazione del transistor bipolare (vedi appendice 1) prevede nella sintassi un nodo di substrato opzionale. Con riferimento alla figura precedente, i transistori NPN (verticale) e PNP laterale necessitano di un terminale di substrato poiché non si può prescindere dalle capacità che il collettore (nel NPN) e la base (nel PNP) hanno verso il substrato. Il transistor PNP di substrato non necessita di un quarto terminale in quanto il substrato coincide fisicamente con il collettore. Come accennato precedentemente i transistori discreti non necessitano di terminale di substrato.

Modello statico del transistor bipolare.

Per il modello statico SPICE utilizza il modello di Ebers-Moll nella versione a trasporto. Questo è matematicamente equivalente al modello a iniezione della figura seguente.



$$IS = \alpha_F I_{ES} = \alpha_R I_{CS}$$

$$BF = \frac{\alpha_F}{1 - \alpha_F} \quad BR = \frac{\alpha_R}{1 - \alpha_R}$$

Sono riportate anche le equazioni che legano i parametri BF, BR e IS di Spice (tipici del modello a trasporto) ai parametri del modello di Ebers-Moll a iniezione. Ricordiamo che le correnti I_F e I_R sono date da:

$$I_F = I_{ES} \left(e^{\frac{V_{BE}}{\eta_F V_T}} - 1 \right); \quad I_R = I_{CS} \left(e^{\frac{V_{BC}}{\eta_R V_T}} - 1 \right);$$

Inoltre sono specificati anche due coefficienti di idealità per i due diodi, ovvero η_F (NF nella MODEL) per il diodo B-E e η_R (NR nella MODEL) per il diodo B-C.

Il modello di Ebers-Moll così riportato non tiene conto dei seguenti fenomeni:

1. Resistenze in serie al collettore (R_{cc}), all'emettitore (R_{ee}) e alla base (R_{bb}).
2. Dipendenza di I_C in zona attiva da V_{BC} (ovvero in pratica da V_{CE}) dovuta all'effetto Early.
3. Variazione del guadagno di corrente β_{\square} in funzione della corrente di collettore.
4. Effetti di alta iniezione.

Le modifiche dell'Ebers- Moll introdotte per tenere conto dei problemi indicati consistono nel modello di Gummel-Poon statico.

Per quanto riguarda il correttivo per l'effetto dovuto alle resistenze serie abbiamo tre parametri, indicati con RC, RE, RB nella model, che appunto costituiscono tre resistenze inserite in serie rispettivamente a collettore, emettitore e base.

Per quanto riguarda l'effetto Early, vengono introdotti due parametri, VAF e VAR, corrispondenti alle tensioni di Early rispettivamente in zona attiva diretta e zona attiva inversa. A titolo di esempio riportiamo l'equazione della corrente di collettore utilizzata da SPICE per la zona attiva diretta (trascurando le correnti di saturazione della giunzione B-C):

$$I_C = IS \cdot e^{\frac{V_{BE}}{NF \cdot V_T}} \left(1 + \frac{V_{CB}}{VAF} \right)$$

Per quanto riguarda la variazione del β (e quindi di β_F e β_R) richiamiamo il fatto che il guadagno di corrente assume il valore massimo (corrispondente in SPICE a BF e BR) solo per un intervallo di correnti di collettore detta regione delle medie correnti. Per correnti inferiori (regione delle piccole correnti) il guadagno decresce significativamente (scendendo anche sotto l'unità) a causa delle correnti di ricombinazione nella zona di carica spaziale della giunzione polarizzata in diretta. Queste correnti, infatti, contribuiscono alla corrente di base ma non sono utili per la corrente di collettore. In SPICE questo fenomeno è modellato aggiungendo due diodi "di perdita" in parallelo ai diodi della I_F e I_R dell'Ebers Moll. Questi diodi hanno una corrente di saturazione inversa e un coefficiente di idealità (di solito intorno a 2) che costituiscono quattro parametri aggiuntivi di SPICE detti parametri di "emissione".

Per correnti superiori alla zona delle medie correnti si ha la regione delle alte correnti dove il guadagno di corrente scende ancora, questa volta per effetti di alta iniezione. Esula da questa trattazione una descrizione dettagliata dei parametri SPICE corrispondenti a questo fenomeno.

Modello dinamico del transistor bipolare.

Il modello dinamico per il transistor bipolare utilizzato da SPICE è il modello a controllo di carica. La carica risulta immagazzinata nelle due giunzioni dei diodi del modello dell'Ebers Moll e risulta data, per ciascuna delle due giunzioni dalla somma:

$$Q = Q_D + Q_J$$

dove Q_D è la capacità di diffusione dovuta all'accumulo di minoritari nella base e nel corrispondente elettrodo (collettore o emettitore, a seconda di quale dei due diodi si considera), mentre la Q_J è la capacità di giunzione associata alle cariche fisse nella zona di svuotamento.

La carica Q_J viene espressa mediante le capacità non lineari di giunzione. Come per il MOSFET la capacità di giunzione viene modellata con l'espressione:

$$C_J = \frac{C_{j0}}{\left(1 - \frac{V_D}{V_0}\right)^m}$$

dove C_{j0} è la capacità per tensione nulla (sul diodo), V_0 è il potenziale di contatto, m è il coefficiente di grading e V_D la tensione sul diodo. Nella model si trovano i seguenti parametri:

CJE, VJE, MJE, corrispondenti ai parametri C_{j0} , V_0 , e m per il diodo BE;

CJC, VJC, MJC, corrispondenti ai parametri C_{j0} , V_0 , e m per il diodo BC;

Vi sono inoltre i parametri CJS, VJS, MJS, corrispondenti ai parametri C_{j0} , V_0 , e m per la giunzione tra il collettore e il substrato (transistore verticale NPN) o tra la base e il substrato (transistore PNP laterale). T-Spice ha il parametro SUBS che assume valore -1 per struttura laterale e 1 per struttura verticale.

La carica di diffusione Q_D viene invece associata alla corrente mediante due parametri, detti tempi di transito: TF per la giunzione BE, TR per la giunzione BC.

Le cariche Q_{DE} (giunzione BE) e Q_{DC} (giunzione BC) risultano:

$$Q_{DE} = TF \cdot I_S \left(e^{\frac{V_{BE}}{V_T}} - 1 \right)$$
$$Q_{DC} = TR \cdot I_S \left(e^{\frac{V_{BC}}{V_T}} - 1 \right)$$

In zona attiva diretta, dove l'unità può essere trascurata rispetto all'esponenziale si ha:

$$C_{DE} = \frac{\partial Q_{DE}}{\partial V_{BE}} = TF \cdot \frac{I_S}{V_T} e^{\frac{V_{BE}}{V_T}} = TF \cdot \frac{I_C}{V_T} = TF \cdot g_m$$

La C_{DE} così calcolata è la capacità di diffusione del modello linearizzato. Un'espressione analoga si può ricavare per la C_{DC} , la capacità di diffusione della giunzione collettore-base

che in zona attiva diretta è trascurabile. Si osservi che in zona di saturazione entrambe le capacità di diffusione sono significative e preponderanti rispetto alle capacità di giunzione. Si può derivare una semplice approssimazione della frequenza di transizione f_T :

A rigore, in zona attiva diretta:

$$f_T = \frac{g_m}{2\pi(C_{jE} + C_{jC} + C_{DE})}$$

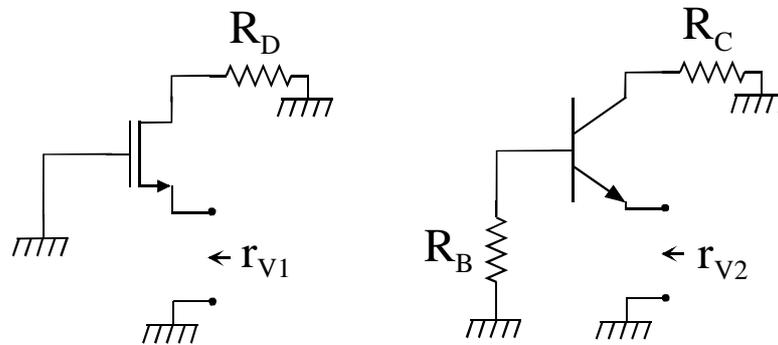
Poichè quando la giunzione BE è in diretta si può ritenere valida l'approssimazione $C_{DE} \gg C_{jE} + C_{jC}$, pertanto:

$$f_T \cong \frac{g_m}{2\pi C_{DE}} \cong \frac{1}{2\pi TF}$$

Questa espressione è utile per ottenere una stima approssimata della frequenza di transizione a partire dal parametro TF della MODEL.

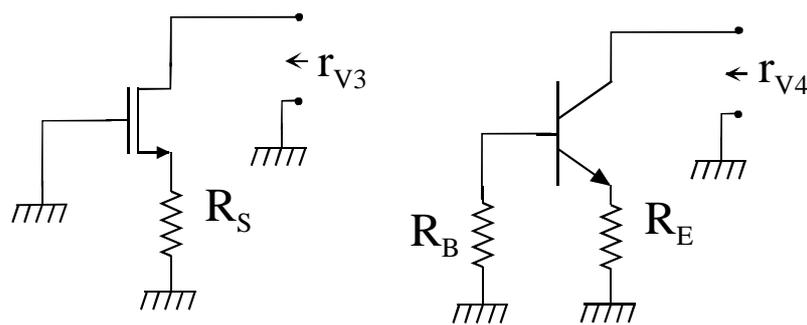
Appendice C. Casi salienti di calcolo della resistenza vista.

In questa sezione verranno esaminate le resistenze viste nei quattro casi indicati nella figura seguente. Questi risultati sono utili per l'analisi dei circuiti che vengono esaminati nel corso del presente trattato.



(a)

(b)



(c)

(d)

Caso (a)

Risultato esatto:

$$r_{v1} = \frac{R_D + r_d}{1 + g_m r_d}$$

Risultato approssimato.

$$\left. \begin{array}{l} \text{ipotesi 1: } g_m r_d \gg 1 \\ \text{ipotesi 2: } R_D \ll r_d \end{array} \right\} \Rightarrow r_{v1} \cong \frac{1}{g_m}$$

Caso (b)

Risultato esatto:

$$r_{v2} = \frac{R_C + r_o}{1 + g_{meq} r_o} \parallel (h_{ie} + R_B) \quad \text{dove: } g_{meq} = \frac{h_{ie}}{R_B + h_{ie}} g_m \quad \text{e } r_o = \frac{1}{h_{oe}}$$

Risultato approssimato:

$$\left. \begin{array}{l} \text{ipotesi 1: } g_{meq} r_o \gg 1 \\ \text{ipotesi 2: } R_C \ll r_o \end{array} \right\} \Rightarrow r_{v2} \cong \frac{1}{g_{meq}} \parallel (h_{ie} + R_B) = \frac{h_{ie} + R_B}{h_{fe} + 1}$$

Caso (c).

Risultato esatto:

$$r_{v3} = R_S + r_d (1 + g_m R_S)$$

Spesso $R_S \ll r_d$ per cui si arriva all'espressione:

$$r_{v3} \cong r_d (1 + g_m R_S)$$

Caso (d).

Risultato esatto.

Lavorando sul circuito equivalente alle variazioni si può ottenere un circuito perfettamente analogo a quello del caso (c), nel quale però compaiono un g_m equivalente e la resistenza R_S è il parallelo di R_E con $h_{ie} + R_B$. Si ottiene:

$$r_{v4} = R_{Eq} + r_o (1 + g_{meq} R_{Eq}) = R_{Eq} + r_o \left(1 + h_{fe} \frac{R_E}{R_E + h_{ie} + R_B} \right)$$

dove:

$$R_{Eq} = (R_B + h_{ie}) \parallel R_E \quad \text{e} \quad g_{meq} = \frac{h_{ie}}{h_{ie} + R_B} g_m = \frac{h_{fe}}{h_{ie} + R_B}$$

Risultati approssimati:

Si possono distinguere due casi opposti:

$$\text{caso 1: } R_E \ll (h_{ie} + R_B) \Rightarrow r_{v4} \cong r_o (1 + g_{meq} R_E)$$

$$\text{caso 2: } R_E \gg (h_{ie} + R_B) \Rightarrow r_{v4} \cong r_o (h_{fe} + 1)$$