Appunti di Comunicazioni Elettriche

Capitolo 10 - Parte I

Dispositivi e circuiti

INTRODUZIONE ALLA RICEZIONE DI UN SEGNALE PASSA-BANDA

Il problema che ci poniamo in questa sede è quello di ricevere un segnale passa-banda, il cui spettro cioè non comprenda la continua e componenti armoniche ad essa prossime.

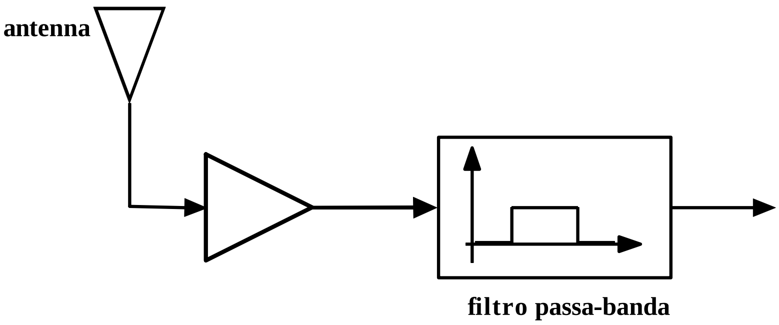
Dobbiamo necessariamente fare i conti con due aspetti:

- in primo luogo, con il rumore termico sovrapposto al segnale da ricevere: tale rumore è sia quello captato dall'antenna ricevente sia quello generato all'interno degli stessi dispositivi usati nel ricevitore);
- in secondo luogo, con tutti i segnali che "accompagnano" il segnale di nostro interesse e che sono separati da esso in quanto allocati in bande di frequenza diverse.

Sono altresì evidenti due considerazioni:

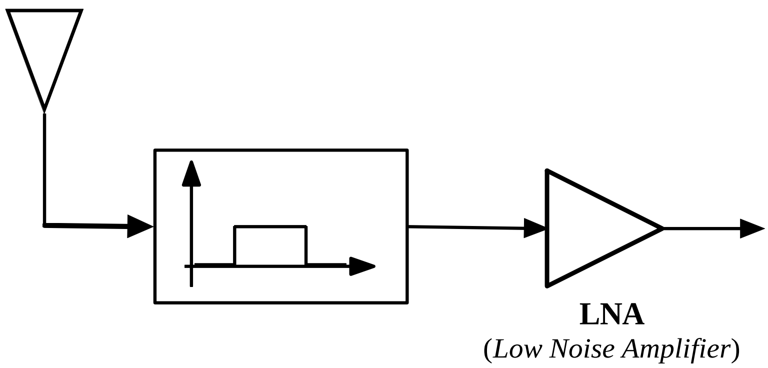
- in primo luogo, il segnale che giunge al nostro apparato ricevente è un segnale generalmente di livello molto basso, per cui è sempre necessario amplificarlo prima di poterlo elaborare;
- in secondo luogo, dovendo isolare il segnale utile da tutti gli altri segnali captati, incluso il rumore nelle bande adiacenti, dobbiamo necessariamente effettuare un filtraggio.

Deduciamo quindi che, in ricezione, le operazioni da compiere comunque sono amplificazione e filtraggio. Si tratta di capire in quale ordine effettuare queste operazioni. La prima possibilità è quella di effettuare prima l’amplificazione e poi il filtraggio:



In questo caso, quello che facciamo è amplificare tutti i segnali captati (quindi il segnale utile, gli eventuali altri segnali ed il rumore) e poi inviare il tutto ad un filtro che dovrebbe far passare solo tutto quanto è contenuto nella banda di nostro interesse. Il problema è, però, che l’amplificazione non sempre è lineare: infatti, dato che la dinamica di un qualsiasi amplificatore è sempre limitata, se amplifichiamo tutto quello che riceviamo in ingresso, è probabile che l’amplificatore venga portato a funzionare anche in regione non-lineare, il che porta ad una distorsione del segnale e quindi rende inutile la successiva operazione di filtraggio. Tanto per chiarirci le idee, supponiamo che la nostra antenna abbia captato solo due segnali (e, per ipotesi, nessun rumore): una sinusoide a frequenza f_0, che possiamo identificare come quella di nostro interesse, e una sinusoide a frequenza f_k molto minore di f_0; se le due sinusoidi sono entrambe di ampiezza molto piccola, è possibile che l’amplificatore, nell’amplificare la loro somma, rimanga in zona lineare di funzionamento, per cui effettivamente il successivo filtraggio consente la separazione; al contrario, se la sinusoide a frequenza f_k è di ampiezza notevole, al contrario dell’altra, essa porta l’amplificatore in saturazione, per cui la sua uscita è una forma d’onda fortemente distorta rispetto al caso precedente, dalla quale sarà quindi impossibile tirare fuori il segnale di interesse a frequenza f_0.

L’altra possibilità è quella di effettuare prima il filtraggio nella banda di interesse, al fine di eliminare sia il rumore sia eventuali altri segnali allocati all’esterno di tale banda, e poi l’amplificazione[1] che tiri nuovamente su il segnale di interesse:

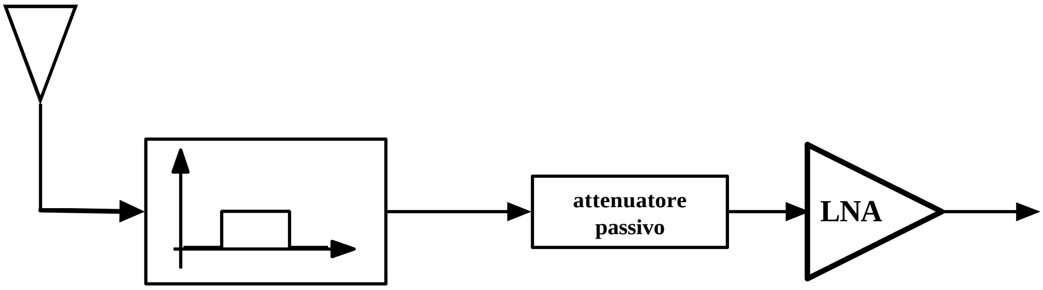


Questa soluzione, pur essendo sicuramente preferibile rispetto alla precedente, non è però esente da problemi. Il problema fondamentale deriva dal filtro passa-banda, che non potrà mai avere, nella

[1] Quest’ultima andrà ovviamente eseguita mediante un amplificatore che generi internamente il minor rumore possibile, al fine di sporcare il meno possibile il segnale in ingresso: dovrà essere un **LNA** (Low Noise Amplifier), cioè un amplificatore con un bassissimo fattore di rumore.

realtà, quelle caratteristiche ideali che noi desideriamo: in particolare, ogni filtro reale ha sempre, oltre ad un comportamento reattivo, un comportamento dissipativo, per cui esso introduce inevitabilmente una attenuazione anche alle componenti di segnale interne alla sua banda passante. Questa attenuazione prende il nome di **perdita di inserzione** del filtro.

Possiamo dunque schematizzare un qualsiasi filtro reale, mediante la cascata di due componenti: un *filtro ideale*, che ha attenuazione nulla (cioè modulo unitario della funzione di trasferimento) nella banda passante ed attenuazione infinita (cioè modulo nullo della funzione di trasferimento) al di fuori, ed un *attenuatore passivo*:



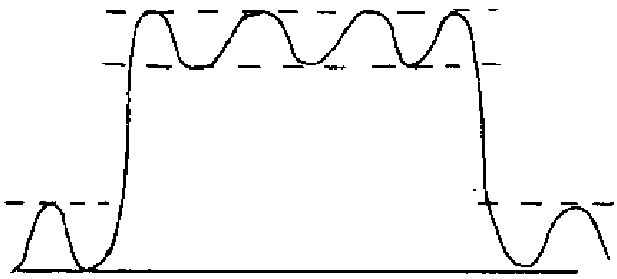
Il segnale all'uscita del filtro è allo stesso livello (basso) con cui è stato captato. L'attenuatore passivo interviene allora ad abbassare ulteriormente il livello del segnale, col pericolo che esso sprofondi rispetto al rumore. Quindi, si peggiorano comunque le prestazioni del sistema in termini di rapporto segnale-rumore, senza considerare il fatto che l'attenuatore passivo è esso stesso sede di rumore e quindi contribuisce ad un ulteriore abbassamento del rapporto S/N.

Siamo dunque giunti alla conclusione per cui *le funzioni fondamentali di un ricevitore passa-banda sono, nell'ordine, il filtraggio e l'amplificazione.*

CARATTERISTICHE DEL FILTRO PASSA-BANDA

Un filtro reale, oltre a presentare una perdita di inserzione (cioè una attenuazione non nulla anche nella banda passante), non ha mai una funzione di trasferimento con modulo perfettamente costante nella banda passante e al di fuori di essa:

Al contrario, esso presenta sempre delle ondulazioni sia in banda sia fuori banda, come indicato nella figura seguente:



Allora, per specificare le caratteristiche del filtro che serve, generalmente si fissa l'*escursione massima delle ondulazioni nella banda passante*: in particolare, si specifica una fascia di valori permessi per la suddetta escursione.

L'altra specifica fondamentale è evidentemente la *banda passante* del filtro, intesa come l'intervallo di frequenza al cui interno l'attenuazione non va oltre i 3dB.

Si specifica anche l'intervallo di frequenza all'interno del quale l'attenuazione supera un minimo prefissato: questo è l'intervallo di frequenza nel quale il filtro taglia, in modo sufficientemente efficace, i segnali che riceve in ingresso. Il legame tra la banda passante e quest'ultimo intervallo è evidentemente rappresentato dalle *bande di transizione* (o *fianchi*) del filtro, ossia gli intervalli di frequenza in cui il modulo della funzione di trasferimento passa dal valore minimo al valore massimo e viceversa. E' intuitivo comprendere che quanto più ripidi si vogliono i fianchi del filtro, tanto maggiore è la difficoltà di implementazione, data la necessità di introdurre poli molto ravvicinati tra loro.

Sempre dal punto di vista della difficoltà realizzativa, si verifica che *la realizzazione di un filtro stretto (cioè di banda passante stretta) è molto più agevole in bassa frequenza che non in alta frequenza:* infatti, ...(?).... Questo per dire che, nel progettare un filtro, non è tanto importante la banda passante in se, quanto la cosiddetta banda relativa, definita come rapporto tra la banda passante e la frequenza centrale attorno alla quale essa è allocata:

$$B_{relativa} = \frac{B_{passante}}{f_{centrale}}$$

Quanto maggiore è la banda relativa, tanto più facile è la realizzazione pratica del filtro.

Ad esempio, volendo un filtro con banda passante di 1MHz, sarà più facile realizzarlo se tale banda è centrata su una frequenza centrale di 10 MHz che non su una di 1 GHz: nel primo caso, la banda relativa è 1/10, mentre nel secondo è 1/1000.

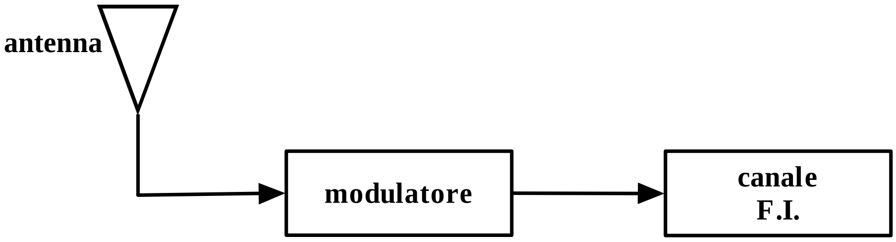
RICEVITORE A CONVERSIONE DI FREQUENZA (O A SUPERETERODINA)

Il fatto che il filtro (da porre a valle dell'antenna) presenti difficoltà realizzative diverse a seconda che la sua banda passante sia allocata in bassa o in altra frequenza, comporta una conseguenza molto importante. Consideriamo ad esempio un **ricevitore radio**: esso deve essere in grado di ricevere tutti quanti i *canali radiofonici* presenti e non solo uno di essi; ciò significa che esso dovrà essere **sintonizzabile**, nel senso che dovrà essere in grado di isolare, trattare e fornire all'utente un segnale presente in una qualsiasi banda di frequenza. Allora, volendo porre un filtro immediatamente a valle dell'antenna, dovremo fare in modo che anch'esso sia sintonizzabile: dovremmo poter spostare la sua banda passante a nostro piacimento, in modo da isolare il segnale di interesse dovunque sia sistemato nello spettro di frequenza. In base a quanto detto, però, spostare la banda passante in alta frequenza potrebbe non essere possibile, il che ci impedirebbe di realizzare un ricevitore sintonizzabile. Dobbiamo allora trovare un altro sistema.

Autore: Sandro Petrizzelli

L'opposto di un *ricevitore sintonizzabile* è un **ricevitore a frequenza fissa**, che cioè è ottimizzato per funzionare ad una frequenza centrale prefissata, situata al centro della banda passante. In effetti, quello che si fa nella pratica è proprio progettare dei ricevitori a frequenza centrale standardizzata: è allora ovvio che, per ricevere un qualunque segnale non allocato attorno a tale frequenza centrale, è necessaria una preventiva traslazione in frequenza, in modo che il suddetto segnale venga spostato appunto attorno alla frequenza centrale del ricevitore.

Lo schema a blocchi da adottare è dunque il seguente:



Il segnale di interesse, una volta captato (insieme al rumore e a tutti gli altri segnali presenti), viene spostato (mediante una semplice operazione di modulazione) a cavallo della frequenza centrale del ricevitore a frequenza fissa, che prende il nome di **canale a frequenza intermedia**. Quest'ultimo non è altro che un blocco di amplificazione, filtraggio e demodulazione, il cui funzionamento sarà descritto tra un attimo.

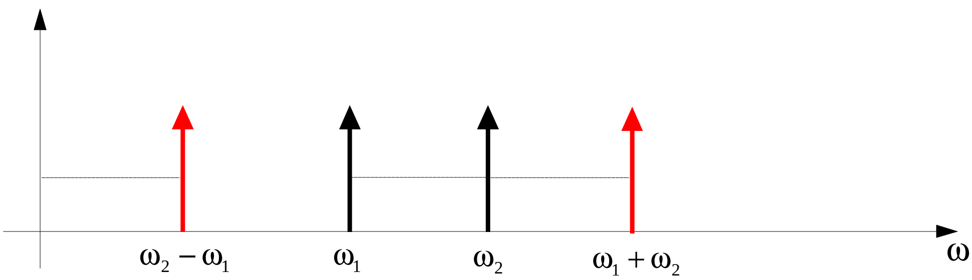
Questo è lo schema di un **ricevitore a conversione di frequenza** (detto anche *ricevitore a supereterodina*)[2].

Conversione di frequenza

Vediamo allora di scendere nei dettagli del funzionamento del modulatore. Abbiamo detto che dobbiamo shiftare in frequenza il segnale, in modo che finisca a cavallo della frequenza del **canale F.I.** (canale a frequenza intermedia). Possiamo allora sfruttare quanto sappiamo a proposito della modulazione di ampiezza: avendo due sinusoidi, di frequenza ω_1 ed ω_2 (con $\omega_2>\omega_1$) il loro prodotto è

$$\cos(\omega_1 t)\cos(\omega_2 t)=\frac{1}{2}\cos((\omega_1+\omega_2)t)+\frac{1}{2}\cos((\omega_2-\omega_1)t)$$

Il prodotto da cioè origine ad altre due sinusoidi, uno a frequenza somma ed uno a frequenza differenza. In frequenza, quindi, abbiamo due impulsi, alle frequenza $\omega_1+\omega_2$ e $\omega_1-\omega_2$:



[2] Ricordiamo che si parla di **eterodina** quando il risultato del prodotto (battimento) di un segnale con l'oscillazione locale è un segnale passa-banda, come nel caso considerato. Se, invece, il risultato del battimento è un segnale passa-basso, allora si parla di **omodina**.

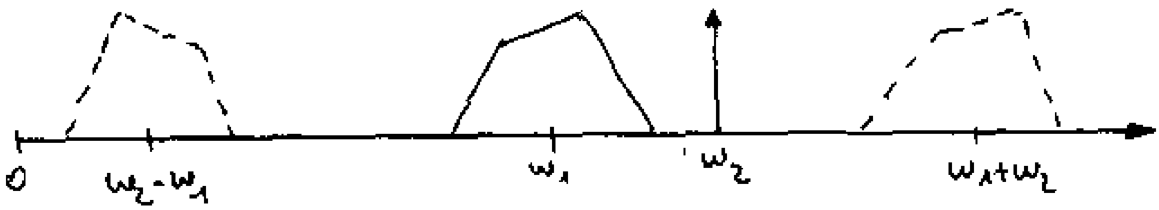
Se, al posto di usare la sinusoide $\cos(\omega_1 t)$, usiamo un generico segnale s(t), il discorso non cambia: infatti, dato che il moltiplicatore è un dispositivo lineare[3] e dato che un generico segnale s(t) è sempre interpretabile, secondo Fourier, come somma di sinusoidi, allora anche il prodotto di s(t) per $\cos(\omega_2 t)$ dà origine a due repliche traslate (e scalate di un fattore 2) dello spettro del segnale: lo spettro del segnale dato dal prodotto è infatti

$$S_t(\omega) = \frac{1}{2}S(\omega - \omega_2) + \frac{1}{2}S(\omega + \omega_2)$$

Gli spettri traslati sono identici allo spettro di partenza dato che ogni componente spettrale viene traslata della stessa quantità.

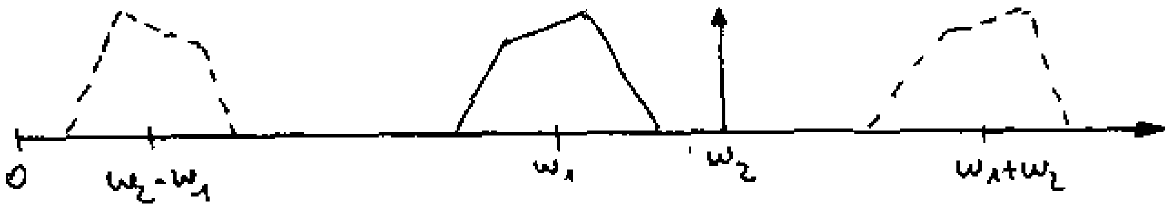
Con questo procedimento, riusciamo dunque a spostare lo spettro del segnale dove ci fa comodo, semplicemente variando il valore dell'oscillazione locale ω_2.

In realtà, c'è però da fare una importante osservazione circa la posizione reciproca dei vari spettri. Consideriamo il caso in cui lo spettro di s(t), che supponiamo centrato su una frequenza centrale ω_1, si trovi a frequenza più alta di ω_2. In questo caso, la situazione è indicata nella figura seguente:



Come si nota, lo spettro viene semplicemente spostato in alta e in bassa frequenza rispetto alla posizione originale, senza ribaltamenti.

Al contrario, se supponiamo che ω_2 sia superiore alla massima frequenza di S(f), la situazione è indicata nella figura seguente:



Si osserva che lo spettro S(f) non viene solo spostato in alta e in bassa frequenza, ma subisce anche un ribaltamento nella replica a bassa frequenza. Il motivo, senza addentrarci nei dettagli matematici, è il seguente:

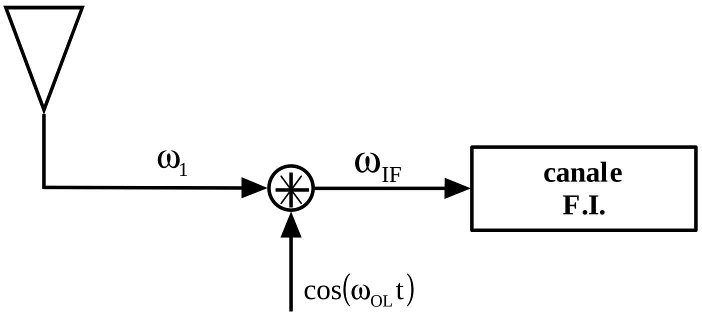
- quando $\omega_1 >> \omega_2$, le componenti di S(f) a maggior frequenza sono quelle più lontane all'oscillazione locale, per cui, dopo la moltiplicazione, subiscono lo spostamento maggiore da $\omega=0$;
- al contrario, quando $\omega_2 >> \omega_1$, le componenti di S(f) a maggior frequenza sono quelle più vicine all'oscillazione locale, per cui, dopo la moltiplicazione, subiscono lo spostamento minore da $\omega=0$.

[3] Il *moltiplicatore* (o **mixer** o *convertitore di frequenza*) è lineare ma non è tempo-invariante, per cui non ha senso parlare, per esso, di funzione di trasferimento.

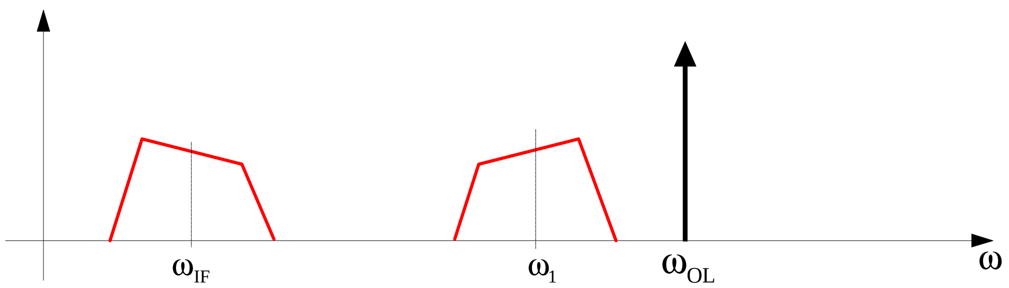
Quindi, *nell'effettuare il battimento, il termine a frequenza somma viene replicato così com'è, mentre quello a frequenza differenza può essere ribaltato o meno a seconda che l'oscillazione locale sia, rispettivamente, maggiore o minore della frequenza centrale del segnale in arrivo.*

Di questo bisognerà tener conto a monte del moltiplicatore, qualora si decida di utilizzare la replica a bassa frequenza[4].

In definitiva, quindi, utilizzando l'effetto del battimento tra una sinusoide ed un segnale passa-banda generico, possiamo realizzare un **ricevitore a conversione**:



Il segnale ricevuto, con banda centrata su ω_1, batte con l'oscillazione locale ω_{OL}, dando origine alle due repliche descritte prima, le quali vanno in ingresso al **canale F.I.**: quest'ultimo è un blocco di amplificazione, filtraggio e demodulazione, che, come detto, funziona ad una frequenza fissa, che indicheremo con ω_{IF} (**frequenza intermedia**). Se scegliamo di utilizzare la replica dello spettro S(f) a bassa frequenza, è evidente che, fissata ω_{IF} e nota ω_1, bisognerà impostare ω_{OL} in modo che risulti $\boxed{\omega_{IF} = \omega_{OL} - \omega_1}$:



Il problema della banda immagine

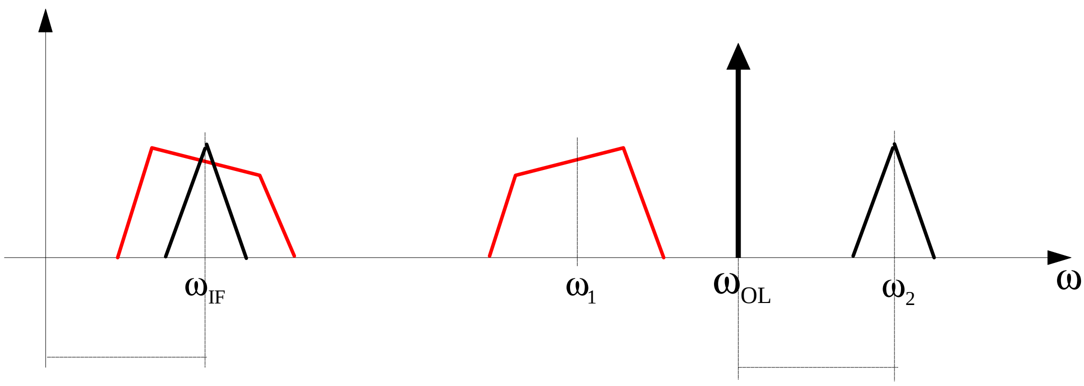
Lo schema descritto finora non è completo, in quanto c'è un altro problema da considerare: abbiamo detto che, tramite la moltiplicazione con ω_{OL}, il segnale S(f) di nostro interesse viene portato a cavallo di ω_{IF}; tuttavia, il segnale di interesse non è stato captato da solo dall'antenna, per cui ci sono anche altre componenti spettrali (oltre l'immancabile rumore) che vengono moltiplicate con ω_{OL}; è allora possibile che anche alcune di queste ulteriori componenti vengano spostate a

[4] Intuitivamente, è più logico utilizzare proprio la replica a bassa frequenza, visto che l'elaborazione di un segnale è sempre più agevole in bassa frequenza che non in altra frequenza. Tuttavia, come avremo modo di vedere, ci sono anche casi in cui si usa la replica ad alta frequenza. Si consideri, tra l'altro, che, allo stato attuale della tecnologia, circuiti funzionanti a diverse decine di MHz sono comunque ampiamente sperimentati ed efficienti.

cavallo ω_{IF}. E' anche facile verificare quali componenti spettrali possono darci "fastidio": indichiamo con ω_2 la frequenza di una generica componente spettrale che, essendo stata captata, viene moltiplicata per l'oscillazione locale a frequenza ω_{OL}; la moltiplicazione dà origine ai due termini $\omega_2+\omega_{OL}$ (somma) e $\omega_2-\omega_{OL}$ (differenza); imponendo allora che questi due termini siano entrambi uguali a ω_{IF}, otteniamo

$$\omega_X + \omega_{OL} = \omega_{IF} \longrightarrow \omega_X = \omega_{IF} + \omega_{OL}$$
$$\omega_X - \omega_{OL} = \omega_{IF} \longrightarrow \omega_X = \omega_{OL} + \omega_{IF} = \omega_2$$

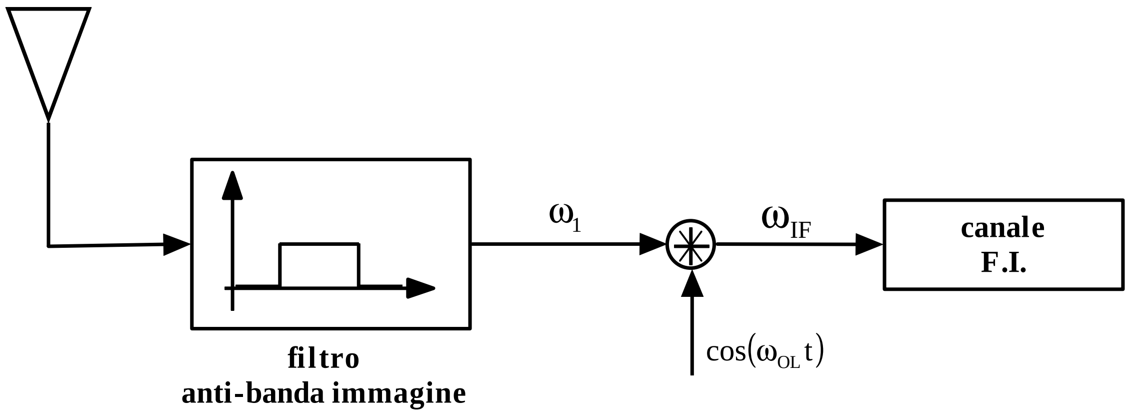
Dalla seconda uguaglianza otteniamo una pulsazione $\omega_2 = \omega_{IF} + \omega_{OL}$. Quindi, dopo la moltiplicazione con ω_{OL}, a cavallo della frequenza intermedia ω_{IF} troviamo sia il segnale utile, centrato su ω_1, sia un eventuale segnale centrato su ω_2, come indicato nella figura seguente:



Quindi, se l'antenna ha captato qualcosa a cavallo di ω_2, questo "qualcosa" si sposterà anch'esso a cavallo di ω_{IF}, ossia si andrà a sovrapporre al segnale utile, risultando così indistinguibile da questo.

La banda di frequenza che si trova a cavallo di ω_2 (cioè, nel dettaglio, la banda che si trova a distanza ω_{IF} dall'oscillazione locale ω_{OL} e in posizione simmetrica, rispetto sempre a ω_{OL}, al segnale utile) prende il nome di **banda immagine** del segnale considerato[5].

E' chiaro che <u>la banda immagine può essere un notevole problema</u>: se ci sono componenti spettrali contenute nella banda immagine, esse finiscono col sommarsi al segnale utile (a cavallo di ω_{IF}) impedendone l'estrazione. Non possiamo allora far altro che porci al riparo da questo problema, predisponendo un filtro che elimini qualsiasi cosa sia contenuta nella banda immagine:

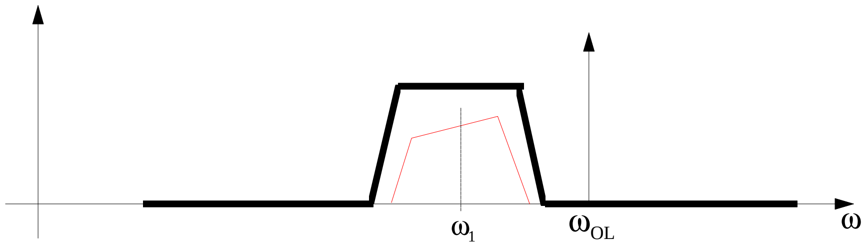


[5] Segnaliamo che, per definizione, l'ampiezza della banda immagine è pari all'ampiezza della banda del segnale di interesse.

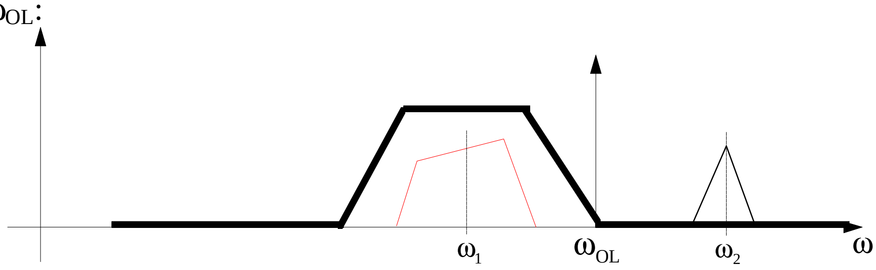
Questo **filtro anti-banda immagine** completa dunque lo schema di un **ricevitore a conversione**.

Sembrerebbe però di essere giunti ad una contraddizione: è evidente, infatti, che, dovendo variare di volta in volta il valore di ω_{OL} (in modo da portare a cavallo di ω_{IF} il segnale desiderato), dovremo anche variare la banda passante del filtro anti-banda immagine, ossia dovremmo adottare un filtro sintonizzabile. Ma noi siamo giunti a questo schema del ricevitore proprio partendo dal presupposto di non poter realizzare in pratica un buon filtro sintonizzabile. Come si spiega? Si spiega nel modo seguente:

- se ipotizziamo di usare un filtro passa-banda immediatamente a valle dall'antenna, in modo da lasciar passare solo il segnale di interesse, abbiamo la necessità di realizzare un filtro sintonizzabile con fianchi molto ripidi, in modo che il filtraggio sia sempre ottimale:



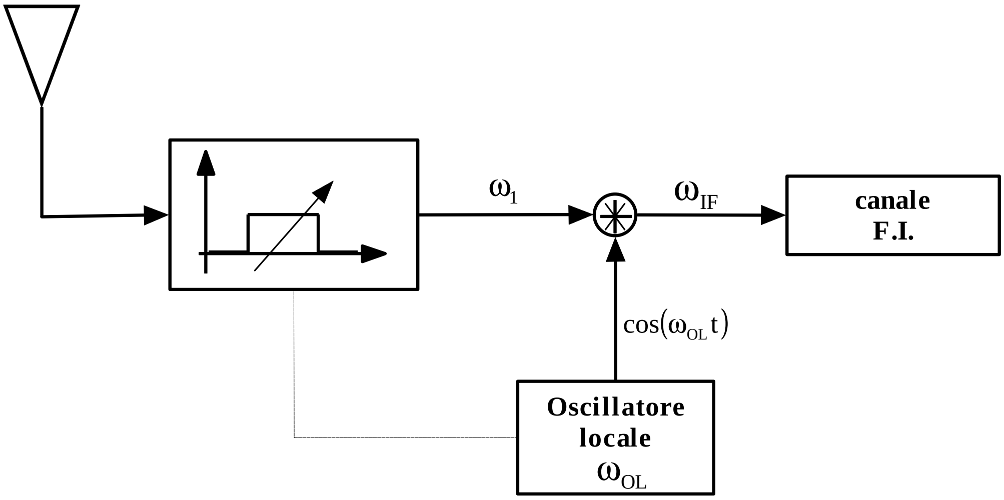
- al contrario, <u>il filtro anti-banda immagine non deve avere fianchi particolarmente ripidi</u>, in quanto non deve effettuare un filtraggio particolarmente selettivo: esso, infatti, deve lasciar passare il segnale, centrato in ω_1, e deve attenuare solo nella banda immagine, ossia a partire almeno da ω_{OL}:



La distanza tra la frequenza centrale (ω_1) del segnale utile e quella (ω_2) della banda immagine è $2\omega_{IF}$: facendo allora ω_{IF} abbastanza grande, la banda passante del filtro, centrata su ω_1, potrà avere un andamento piuttosto lento ai bordi, il che rende facile la realizzazione pratica del filtro: si può persino utilizzare un semplice *circuito risonante RLC* (del quale avremo modo di parlare anche in seguito), che è facilmente sintonizzabile. <u>Un filtro molto più selettivo andrà invece posto a valle del moltiplicatore</u>, nel blocco denominato canale F.I., ma si tratterà di un filtro a frequenza fissa (pari a ω_{IF}), per cui si potrà realizzare anch'esso con più facilità.

Quello trovato è dunque lo schema della cosiddetta **radio a supereterodina**: come visto, essa è costituita da un filtro passa-banda (anti-banda immagine) sintonizzabile, posto immediatamente a valle dell'antenna, da un convertitore di frequenza e dal canale a frequenza intermedia.

Una cosa interessante è che `il filtro passa-banda e il convertitore sono pilotati dallo stesso oscillatore,` proprio perché gli spostamenti dell'oscillazione locale e della banda passante vanno di comune accordo:



Si parla di **supereterodina** proprio perché il ricevitore riesce a sintonizzare di comune accordo, tramite l’oscillatore locale (OL), il filtro in ingresso e l’oscillatore locale.

Osserviamo che il filtro anti-banda immagine, pur essendo irrinunciabile, ha comunque delle controindicazioni, legate ancora una volta alla inevitabile perdita di inserzione e quindi alla inevitabile degradazione del rapporto S/N. D’altra parte, non ne possiamo proprio fare a meno, in quanto, anche nella ipotesi che la banda immagine non contenga alcun segnale, conterrà sicuramente del rumore: senza il filtro, all’uscita del moltiplicatore avremmo una doppia potenza di rumore, nella banda utile (centrata su ω_{IF}), rispetto a quella che otteniamo con il filtro.

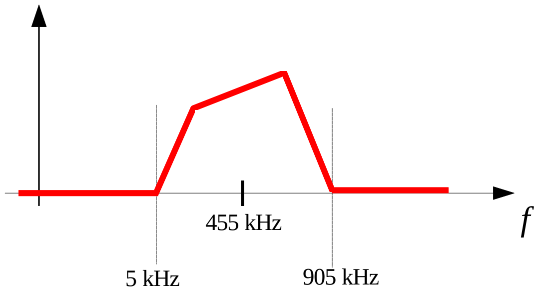
Esempio numerico: trasmissione radiofonica AM e trasmissione TV

Per scendere in dettagli ancora maggiori a proposito del problema della banda immagine, facciamo un esempio numerico. Partiamo da due ipotesi:

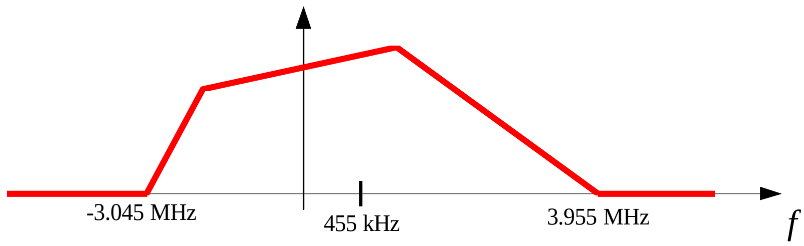
- in primo luogo, supponiamo di aver progettato un canale a frequenza intermedia che lavora a ω_{IF}=455 kHz[6];
- in secondo luogo, supponiamo che il segnale da ricevere sia un segnale AM in doppia banda laterale, che quindi occupa una banda, centrata sulla frequenza ω_1 della portante, doppia rispetto alla banda del segnale modulante: se B è la banda di quest’ultimo, 2B sarà la banda del segnale da ricevere.

Supponiamo che il segnale modulante sia un segnale audio: *per la diffusione AM dei segnali audio, generalmente si assegnano canali larghi 9 kHz,* il che significa che la banda del segnale da ricevere sarà 2B=9kHz. Per questa banda, la frequenza ω_{IF}=455kHz ci va bene, in quanto, dopo la moltiplicazione, avremo 9 kHz centrati su 455kHz:

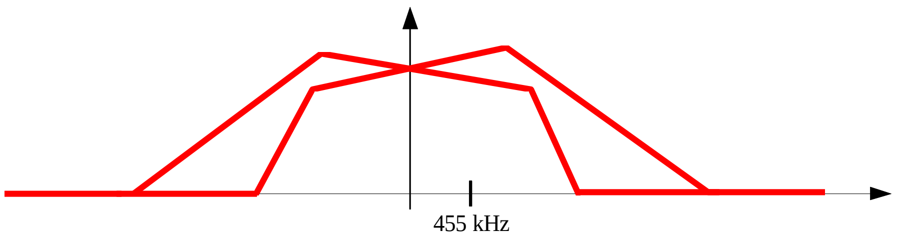
[6] Ricordiamo che i ricevitori a frequenza fissa lavorano a frequenze centrali standardizzate, che sono 455 kHz, 10.7 MHz, 35 MHz e così via.



Non andrebbe invece bene per un segnale TV, che ha una banda di circa 7 MHz (considerando 1 MHz per la banda vestigiale e l'ulteriore banda per il segnale audio): in questo caso, avremmo 7MHz centrati su 455 kHz, per cui avremmo ben 3.045 MHz nel campo delle frequenze negative:



Considerando che per le frequenze negative vale la situazione esattamente simmetrica, avremmo alla fine la sovrapposizione spettrale di due segnali a cavallo dello 0, per cui otterremmo alla fine un segnale passa-basso con caratteristiche profondamente diverse dal segnale ricevuto:



Questo discorso serve dunque a spiegare che la scelta della frequenza intermedia ω_{IF} va fatta sulla base di almeno due considerazioni:

- la prima, in base a quanto detto in precedenza, è che <u>al crescere di ω_{IF}, cresce anche la distanza della banda immagine dalla banda del segnale utile</u>, per cui il filtro anti-banda immagine può essere fatto più blando;
- la seconda riguarda appunto la banda del segnale utile da traslare a cavallo di ω_{IF}: <u>quanto maggiore è la banda del segnale utile, tanto maggiore dovrà essere ω_{IF}</u>. A questo proposito, esistono dei <u>valori standardizzati</u>, che, per esempio, nel caso del segnale televisivo (che richiede almeno 3.5MHz a destra e a sinistra di ω_{IF}), prevedono ω_{IF}=10.7 MHz.

Ottimizzare il valore di ω_{IF} non è comunque una operazione facile. Fin quando il ricevitore che si intende realizzare lavora a frequenza fissa[7] (cioè non è sintonizzabile), è possibile trovare una soluzione pratica abbastanza efficiente. Nei ricevitori sintonizzabili, invece, che devono coprire (o

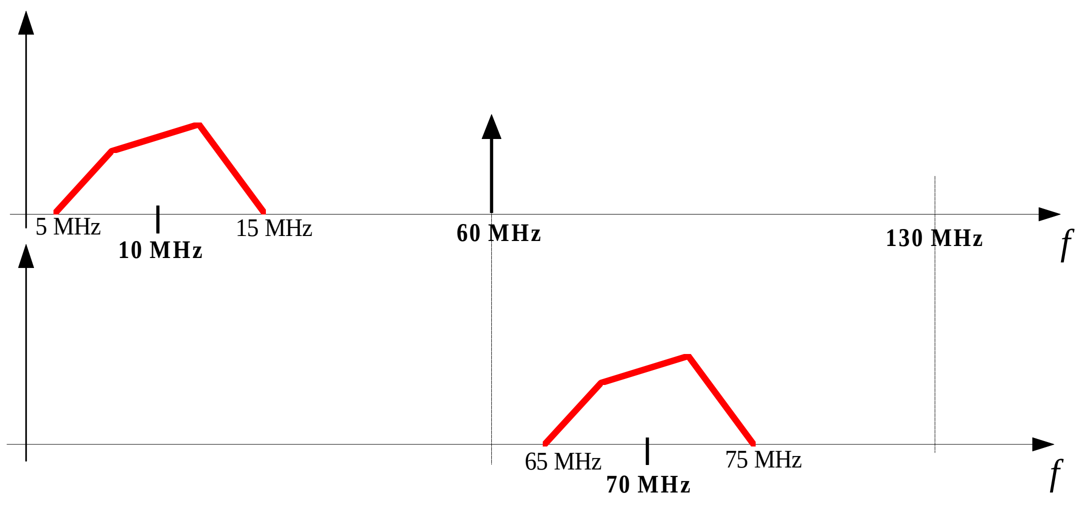
[7] Tutt'al più, si può fare in modo che ω_{IF} vari in un intervallo percentualmente piccolo

spazzolare) un intervallo di frequenza percentualmente grande[8], non è affatto facile determinare un valore di ω_{IF} fisso.

Per risolvere questi problemi, si sono adottate, negli anni, due soluzioni:

- inizialmente (e parliamo di circa di 20 anni fa), i ricevitori effettuavano più conversioni di frequenza, portando il segnale verso le basse frequenze in più passi successivi anziché in un unico passo. Questo modo di procedere consentiva di controllare meglio la reiezione della banda indesiderata;
- attualmente, invece, si usa un’altra strategia: anziché usare il termine differenza (ω_{OL}-ω_1) che viene fuori dal battimento del segnale con l’oscillazione locale, si usa il termine somma (ω_{IF}=ω_{OL}+ω_1), considerando che la gran parte dei dispositivi (filtri, amplificatori, ...) sono ormai tranquillamente in grado di operare a frequenze abbastanza elevate. Il vantaggio di questa procedura è che la distanza della banda immagine dalla banda centrata su ω_1 è ancora maggiore di prima: infatti, la banda immagine è ancora centrata su $\omega_2 = \omega_{IF} + \omega_{OL}$, ma ω_{IF} è cresciuta notevolmente rispetto a prima, proprio perché è la somma di ω_{OL} e ω_1 e non più la differenza. Di conseguenza, il filtraggio che richiediamo al filtro anti-banda immagine potrà essere ancora più blando, per cui il filtro stesso avrà caratteristiche ancora migliori in termini di sintonizzabilità.

Possiamo anche fare un esempio numerico di quanto appena detto: supponiamo di avere scelto una frequenza intermedia ω_{IF}=70 MHz e supponiamo che, in trasmissione, si sia usato un segnale modulante con banda di 5 MHz. Se questo segnale è andato a modulare (in DSB-SC) una portante a ω_1=10 MHz, il segnale che ci interessa ricevere è centrato su ω_1 e occupa una banda di 10 MHz disposta simmetricamente su ω_1. L’oscillazione locale da predisporre per portare il segnale a cavallo di ωIF sarà in questo caso $\omega_{OL} = \omega_{IF} - \omega_1 = 60\ \text{MHz}$. La banda immagine si trova invece centrata su $\omega_2 = \omega_{IF} + \omega_{OL} = 130\,\text{MHz}$, a ben 120 MHz da ω_1:



In questa situazione, basterà che il filtro anti-banda immagine elimini componenti spettrali a partire, per esempio, da 90 MHz in poi.

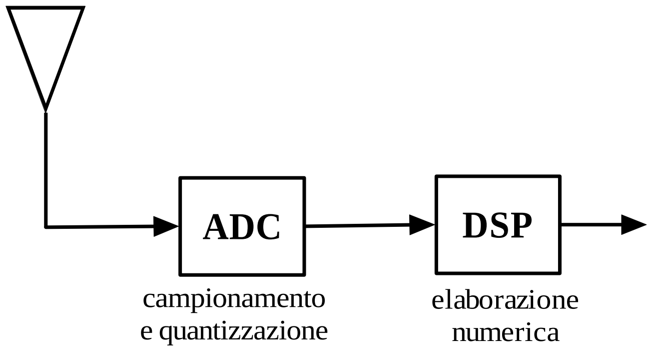
[8] Ad esempio, i **ricevitori per onde corte** (che ricordiamo essere onde con lunghezze d’onda che vanno da 600 m a 10 m) devono coprire le frequenze da 500 kHz a 30 MHz.

Il ricevitore potrà poi funzionare alla frequenza fissa di 70 MHz e potrà così essere ottimizzato. Eventualmente, si potrà scendere a frequenza più bassa in un secondo momento, magari in più passaggi, in modo da effettuare una demodulazione più agevole del segnale.

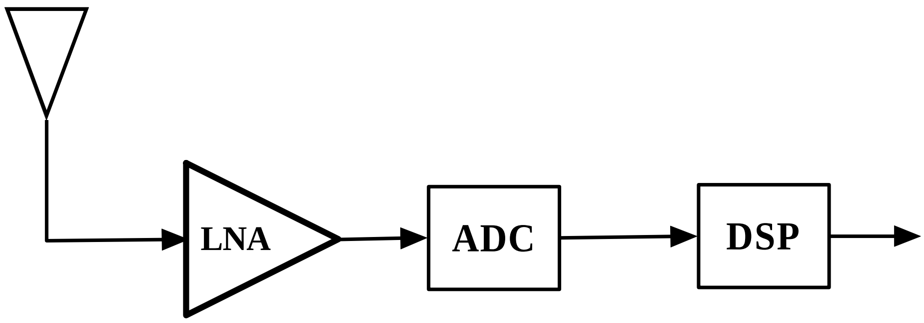
RICEVITORE NUMERICO

Il ricevitore a supereterodina descritto nei paragrafi precedenti è un tipico ricevitore analogico, che cioè effettua elaborazioni su segnali analogici. Ci chiediamo se si possa trasformare un simile ricevitore in un ricevitore numerico, che cioè compia le sue elaborazioni non più su segnali analogici, ma sui corrispondenti segnali numerici.

In generale, sarebbe auspicabile effettuare la digitalizzazione quanto prima possibile, in quanto è noto che l'elaborazione digitale risulta sempre più precisa e accurata dell'analoga elaborazione analogica. In base a questo criterio, specialmente nel caso di radio a onde medie o a onde corte, si potrebbe allora pensare di effettuare la conversione in numerico direttamente sul segnale captato dall'antenna:



Il segnale captato viene convertito immediatamente in numerico (mediante un **ADC**, ossia un *Analog-Digital Converter*), dopo di che ogni successiva elaborazione è di tipo digitale (ed avviene mediante un **DSP**, ossia un *Digital Signal Processor*). In realtà, però, essendo il segnale captato dall'antenna troppo basso per essere elaborato, è comunque necessaria una preventiva amplificazione:



E' dunque necessario affrontare un primo problema di amplificazione a larga banda. Anche risolvendo questo problema, bisogna poi considerare che l' ADC ha comunque una dinamica limitata, per cui non possiamo decisamente pensare di inviargli in ingresso, tutti amplificati, il segnale utile, il rumore e tutti gli altri segnali captati. Oltre a questo, c'è anche il fatto che l' ADC è anche intrinsecamente non lineare (pensiamo al processo di quantizzazione).

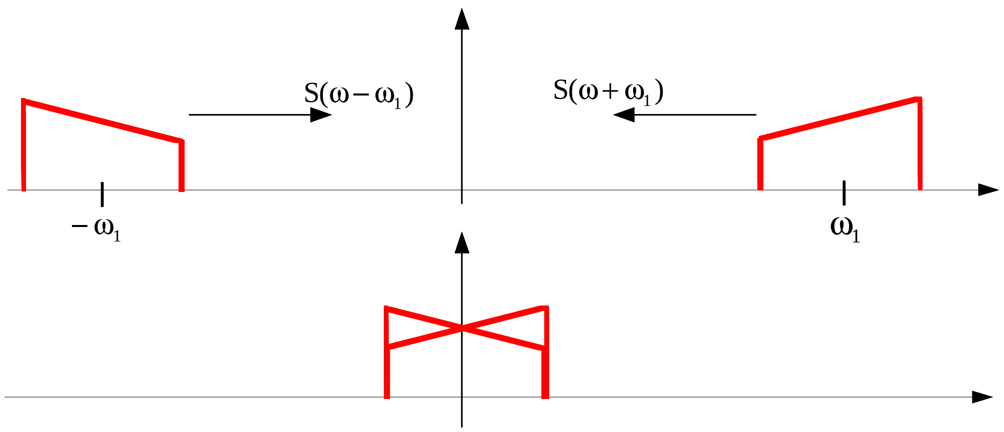
Per queste ragioni, uno schema con quello appena disegnato non è praticabile. Possiamo allora usare il concetto di **ricevitore omodina**, *nel quale cioè, partendo da un segnale* *passa-banda quale è quello che ci interessa, giungiamo ad un segnale passa-basso sul quale compiano una serie di operazioni.*

In primo luogo, abbiamo osservato prima, nel ricevitore a (super)eterodina, è opportuno rendere abbastanza distanti l’oscillazione locale ω_{OL} e la frequenza centrale ω_1 del segnale di interesse, per almeno due motivi: per evitare che le repliche del segnale finiscano troppo vicino allo zero e per evitare problemi con eventuali termini in banda immagine (in quanto $\omega_{IF}=\omega_1+\omega_{OL}$, per cui, maggiore è la distanza tra ω_1 e ω_{OL}, più grande risulta ω_{IF} e quindi più distante è la banda immagine).

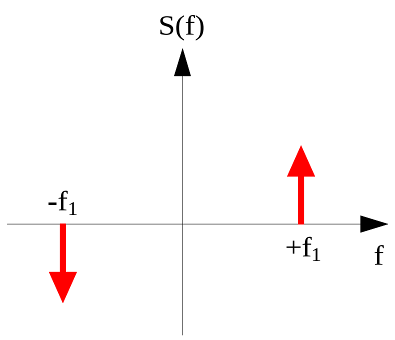
Al contrario, vediamo che succede se poniamo proprio $\omega_{OL}=\omega_1$: se il segnale utile fosse una sinusoide a frequenza ω_1, il prodotto con una sinusoide di uguale frequenza genererebbe una componente a frequenza 0 (cioè la continua) e una componente a frequenza $2\omega_1$; se, invece, il segnale utile è solo centrato su ω_1, sappiamo, dalla demodulazione coerente, che il battimento con una oscillazione locale di frequenza ω_1 sposterebbe lo spettro in banda base (oltre a spostarlo a cavallo di $2\omega 1$, ma questa replica non ci interessa e la cancelliamo con un filtraggio):

$$s(t)\cos(\omega_1 t) \xleftrightarrow{\text{Fourier}} S(\omega) * \frac{1}{2}\left(\delta(\omega-\omega_1)+\delta(\omega+\omega_1)\right) = \frac{1}{2}S(\omega-\omega_1)+\frac{1}{2}S(\omega+\omega_1)$$

Il problema è però nel fatto che si sovrapporrebbero la replica $S(\omega+\omega_1)$ a frequenze positive e quella $S(\omega-\omega_1)$ a frequenze negative, dando luogo ad un segnale inutilizzabile:

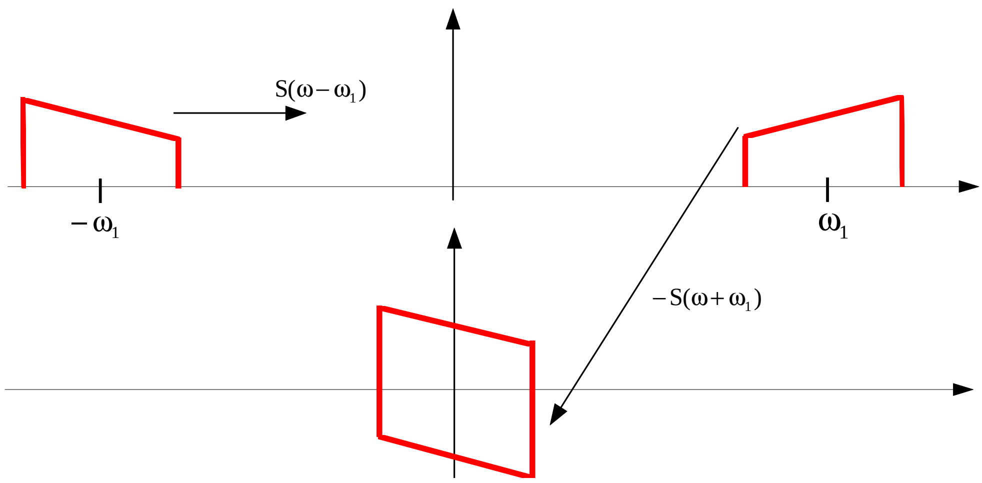


Possiamo provare ad effettuare la traslazione in banda base con una oscillazione sinusoidale $A\sin(\omega_1 t)$, che ha notoriamente positivo l’impulso a frequenza ω_1 e negativo quello a frequenza $-\omega_1$:



In questo caso, otteniamo quanto segue:

$$s(t)\sin(\omega_1 t) \xleftrightarrow{\text{Fourier}} S(\omega) * \frac{1}{2}\left(\delta(\omega-\omega_1)-\delta(\omega+\omega_1)\right) = \frac{1}{2}S(\omega-\omega_1)-\frac{1}{2}S(\omega+\omega_1)$$

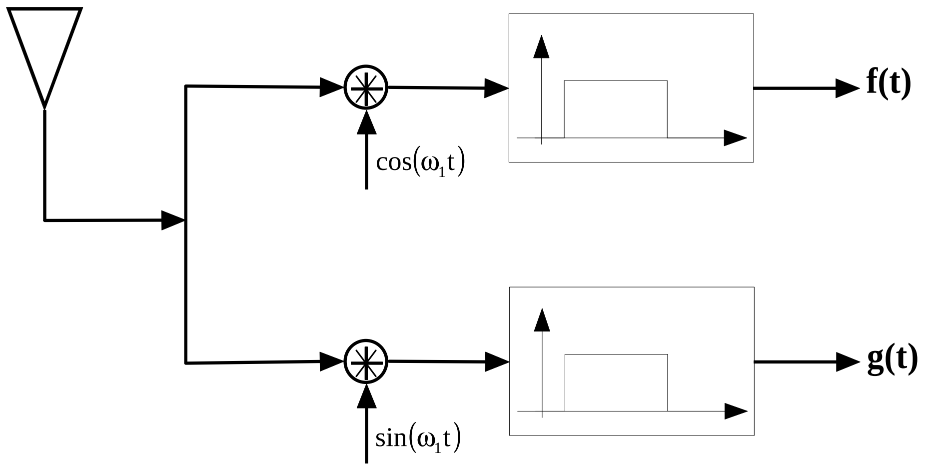


In questo caso si ottiene la differenza tra le due repliche, che rappresenta ancora una volta un segnale inutilizzabile, in quanto non è possibile separare l'una o l'altra replica.

A questo punto, viene spontaneo di effettuare entrambe le conversioni: infatti, disponendo della somma e della differenza delle due repliche, possiamo evidentemente ricostruire queste ultime. Indicando ad esempio con $S^+(f)$ la replica a frequenza positiva e con $S^-(f)$ quella a frequenza positiva, abbiamo evidentemente quanto segue:

$$\begin{cases} \cos(\omega_1 t) \longrightarrow F(f) = S^+(f) + S^-(f) \\ \sin(\omega_1 t) \longrightarrow G(f) = S^+(f) - S^-(f) \end{cases} \longrightarrow \begin{cases} F(f) + G(f) = 2S^+(f) \\ F(f) - G(f) = 2S^-(f) \end{cases}$$

Possiamo dunque adottare uno schema del tipo seguente:



(ovviamente, le due oscillazioni proverranno da uno stesso oscillatore: il Coseno sarà il segnale prodotto dall'oscillatore, mentre il Seno di otterrà dal Coseno per un semplice sfasamento di $\pi/2$).

Il segnale captato s(t) viene trasmesso ad entrambi i rami dell'apparato, nei quali viene moltiplicato per l'oscillazione Seno e per l'oscillazione Coseno; dopo il filtraggio[9], che ci serve ad eliminare le repliche in alta frequenza ($2\omega_1$), disponiamo di due funzioni f(t) e g(t) che, come è noto,

[9] La banda dei due filtri passa-basso sarà uguale a metà della banda del segnale captato, dato l'effetto del moltiplicatore di *ripiegare* (rispetto alla frequenza di oscillazione) lo spettro ricevuto e di traslarlo in banda base.

possono essere accomunate in una rappresentazione matematica: esse infatti forniscono una descrizione del segnale s(t) di partenza in termini di *segnale complesso* del tipo

$$\boxed{s(t) = f(t) + jg(t)}$$

Una volta che siano disponibili queste funzioni, possiamo pensare di effettuare qualsiasi operazione. Ad esempio, possiamo pensare di effettuare due tipi di demodulazione:

- per effettuare una <u>demodulazione di ampiezza</u> (in particolare, una demodulazione ad inviluppo, cioè non coerente), ci basta eseguire la formula

$$\sqrt{f^2(t) + g^2(t)}$$

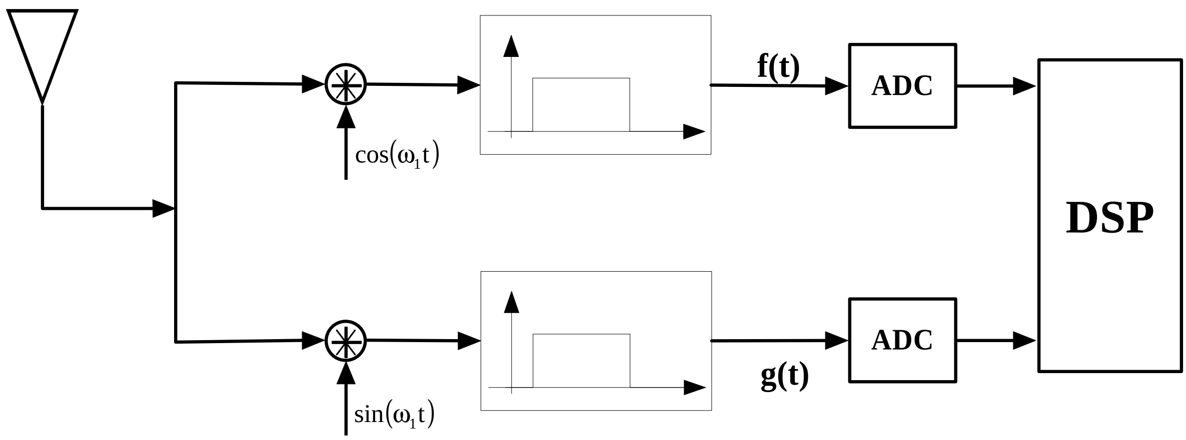
  (valida a patto che la fase della portante non cambi segno);

- per effettuare, invece, una <u>demodulazione di fase</u>, ci basta calcolare

$$\mathrm{arctg}\frac{g(t)}{f(t)}$$

Quindi, con questo sistema, agendo sulle funzioni f(t) e g(t) possiamo realizzare qualunque operazione in maniera del tutto equivalente a quello che potremmo fare sul segnale originario captato. Questo appunto perché la rappresentazione mediante f e g, che fornisce due segnali passa-basso, è del tutto equivalente, in termini di informazione convogliata, al segnale passa-banda di partenza.

Ritornando allora al discorso del ricevitore numerico, possiamo effettuare la conversione analogico→digitale di f(t) e g(t), in modo poi da usare un DSP che compia le elaborazione volute partendo appunto dalla forma numerica di f(t) e g(t):



<u>*Questa struttura ha anche l’enorme vantaggio di aver eliminato completamente la banda immagine:*</u> infatti, ricordando che la banda immagine è centrata su $\omega_2 = \omega_{IF} + \omega_{OL}$ ed inoltre che $\omega_{IF} = \omega_{OL} - \omega_1$, è evidente che, se $\omega_{OL}=\omega_1$, risulta $\omega_{IF}=0$ e quindi $\omega_2=\omega_{OL}=\omega_1$. Risulta cioè che la banda immagine è sovrapposta alla banda utile, per cui <u>non ha più senso parlare di banda immagine</u>.

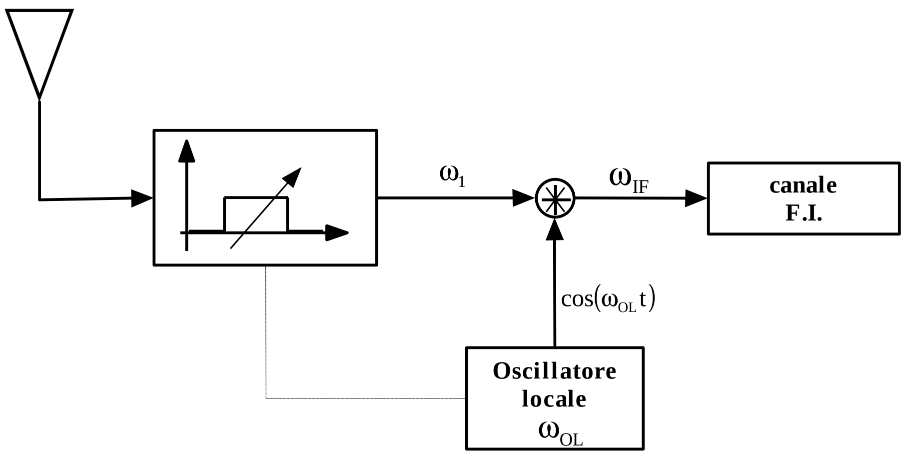
Questi vantaggi si pagano con l’evidente svantaggio di dover realizzare due canali di conversione in contemporanea. Non solo, ma ci sono anche difficoltà realizzative derivanti dal fatto che i moltiplicatori realizzati nella pratica forniscono in uscita dei termini in bassa frequenza indesiderati. Questi termini, dovuti essenzialmente alla non-idealità dei moltiplicatori, interferiscono ovviamente con le operazioni successive, pregiudicandole. Questo è il motivo per cui *le prestazioni dei ricevitore omodina realizzati con tecniche numeriche NON sono, attualmente, paragonabili né superiori a quelle dei ricevitori ad eterodina.*

In realtà, si possono pensare una serie di miglioramenti dello schema visto prima.

Amplificazione selettiva

CONCETTI GENERALI SULL'IMPIEGO DI RISONANTI RLC

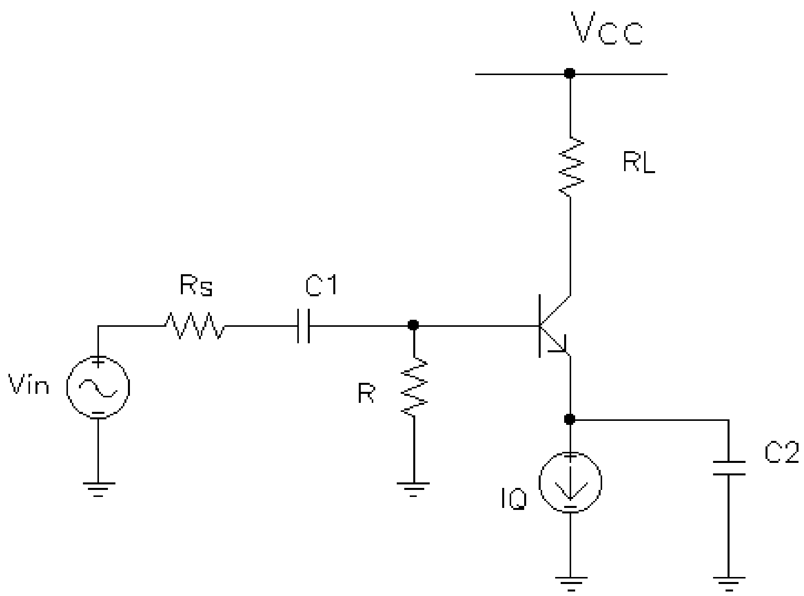
Riprendiamo lo schema di un ricevitore radio a supereterodina:



Abbiamo osservato che le funzioni principali, in questo schema, sono la **conversione di frequenza** (cioè la moltiplicazione del segnale utile captato, centrato su ω_1, con l'oscillazione locale ω_{OL}) e poi il filtraggio (o meglio l'**amplificazione selettiva**). In questi paragrafi ci occupiamo in particolare dell'amplificazione selettiva.

Il dispositivo che deve effettuare l'amplificazione selettiva dovrà dunque essere un amplificatore, con un certo guadagno all'interno della banda passante (intesa, in quanto caso, come la banda, centrata sulla frequenza intermedia ω_{IF} , pari a quella occupata dal segnale di interesse[10]) e con attenuazione elevata all'esterno di tale banda. In generale, un tipico amplificatore a transistor è uno stadio ad emettitore comune del tipo rappresentato nella figura seguente:

[10] Ricordiamo che, al contrario del filtro passa-banda posto immediatamente a valle dell'antenna, il quale deve solo eliminare la banda immagine, il filtro inserito nel canale a frequenza intermedia deve operare un filtraggio estremamente selettivo).



Nel disegnare questo circuito, abbiamo tenuto conto di una serie di considerazioni:

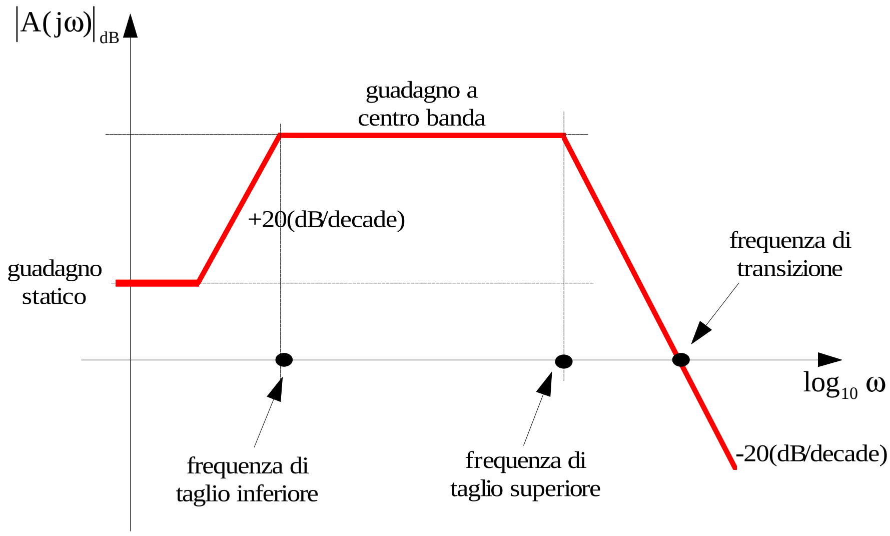
- in primo luogo, la presenza della capacità C_1 serve a disaccoppiare la sorgente dal circuito: infatti, questa capacità blocca la continua, ossia fa in modo che la sorgente non assorba la corrente di polarizzazione del transistor; d’altra parte, se non ci fosse la resistenza R, in assenza del generatore di segnale (schematizzato dalla tensione Vin in serie ad Rs), si avrebbe corrente di base I_B=0, il che comporterebbe chiaramente lo spegnimento del BJT;
- il secondo luogo, per la polarizzazione è stato considerato un generatore di corrente (ideale, cioè con resistenza di Norton infinita) che impone una prefissata corrente I_Q di emettitore (e quindi anche di collettore, visto che $I_B=I_C/\beta$ se il transistor funziona in zona lineare); in regime di piccolo segnale, però, il generatore di corrente scompare, per cui l’emettitore del transistor risulterebbe *floating* (flottante) se non avessimo messo la capacità C_2 che lo porta a massa per il segnale: ciò significa supporre che la capacità C_2 sia sufficientemente alta da comportarsi come un cortocircuito alle frequenze di lavoro che noi consideriamo.

Il fatto che l’emettitore sia a massa per il segnale fa sì che la tensione v_π, che pilota il transistor in regime appunto di piccolo segnale, sia univocamente determinata dalla tensione di segnale $v_{in}(t)$: quest’ultima, quindi, è in grado, grazie alla C2, si controllare la tensione $v_{BE}(t)$ applicata alla giunzione base-emettitore.

Il carico del transistor è rappresentato dalla resistenza R_L e sappiamo perciò che, in assenza di ulteriori stadi in cascata, il guadagno di tensione di piccolo segnale è

$$A_V = \frac{v_C}{v_{in}} \cong g_m R_L$$

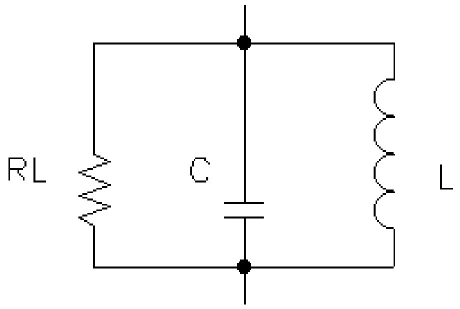
E’ bene però precisare che questo è il guadagno a centro banda, ossia per un intervallo di frequenze tale che sia gli effetti capacitivi dovuti a C_1 e C_2 sia gli effetti capacitivi dovuti alle capacità intrinseche del dispositivo siano trascurabili:



La frequenza di taglio inferiore ω_L è determinata dalle capacità di disaccoppiamento (C1) e da quelle di by-pass (C2), mentre la frequenza di taglio superiore ω_H è determinata dalle capacità intrinseche C_π (dell'ordine del centinaio di pF) e C_μ (dell'ordine di pochi pF).

Noi supponiamo di lavorare a frequenze intermedie tra ω_L ed ω_H. In questo intervallo di frequenze, il nostro interesse è quello di ottenere non un guadagno costante con la frequenza, ma un guadagno variabile (sagomato) e questo lo possiamo ottenere solamente se sostituiamo il carico reattivo R_L con una impedenza $Z_L(\omega) = R_L + jX_L(\omega)$ opportunamente dimensionata.

Come scegliamo Z_L? Dato che vogliamo un amplificatore selettivo, che cioè presenti un guadagno massimo in corrispondenza di una certa frequenza ω_0 e poi decrescente man mano che ci allontaniamo da essa, ci viene subito da pensare ad un circuito RLC risonante, quale ad esempio un risonante parallelo come quello della figura seguente:



Ricordiamo allora il funzionamento di un simile circuito.

Supponiamo di alimentare il collegamento mediante una corrente sinusoidale del tipo $I(t) = I_M \cos(\omega t)$. A regime, la risposta del circuito è la generazione di una tensione sinusoidale, isofrequenziale con l'ingresso, del tipo $V(t) = V_M \cos(\omega t + \varphi)$, dove φ è il cosiddetto *angolo di sfasamento*, ossia l'argomento dell'impedenza totale del circuito considerato:

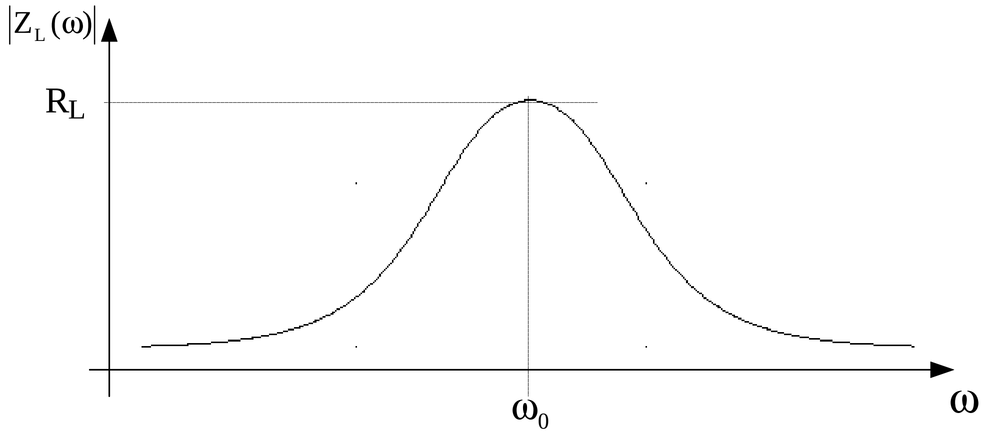
$$Y_L(\omega) = \frac{1}{Z_L(\omega)} = \frac{1}{R_L} + j\left(\omega C - \frac{1}{\omega L}\right) \longrightarrow Z_L(\omega) = \frac{1}{\frac{1}{R_L} + j\left(\omega C - \frac{1}{\omega L}\right)}$$

Quindi, alimentando il circuito con la corrente I(t), esso risponde con una tensione V(t), isofrequenziale con I(t), il cui modulo e la cui fase sono i seguenti:

$$I(\omega) = V|Y_L(\omega)| = V\sqrt{\frac{1}{R_L^2} + \left(\omega C - \frac{1}{\omega L}\right)^2}$$

$$\varphi(\omega) = \text{arctg}\left(R_L\left(\omega C - \frac{1}{\omega L}\right)\right)$$

Il modulo dell'impedenza $Z_L(\omega)$ rappresenta dunque il guadagno V/I del circuito ed ha un andamento, in funzione di ω, del tipo seguente:



(la scala usata in ascisse, per ω, è logaritmica)

In corrispondenza della pulsazione di risonanza $\boxed{\omega_0 = \frac{1}{\sqrt{LC}}}$, il guadagno assume il proprio valore massimo, pari al valore della resistenza R_L: infatti, si ha che

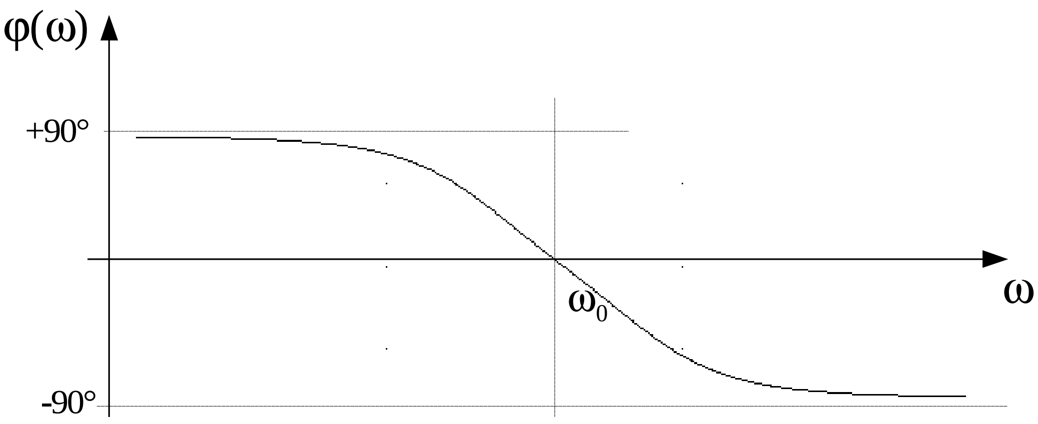
$$I(\omega_0) = V\sqrt{\frac{1}{R_L^2} + \left(\omega_0 C - \frac{1}{\omega_0 L}\right)^2} = \frac{1}{R_L} \longrightarrow \frac{V}{I} = R_L$$

$$\varphi(\omega_0) = \text{arctg}\left(R_L\left(\omega_0 C - \frac{1}{\omega_0 L}\right)\right) = \text{arctg}\,0 = 0°$$

Il valore di RL determina, quindi, l'apertura della "campana" rispetto al valore di ω_0.

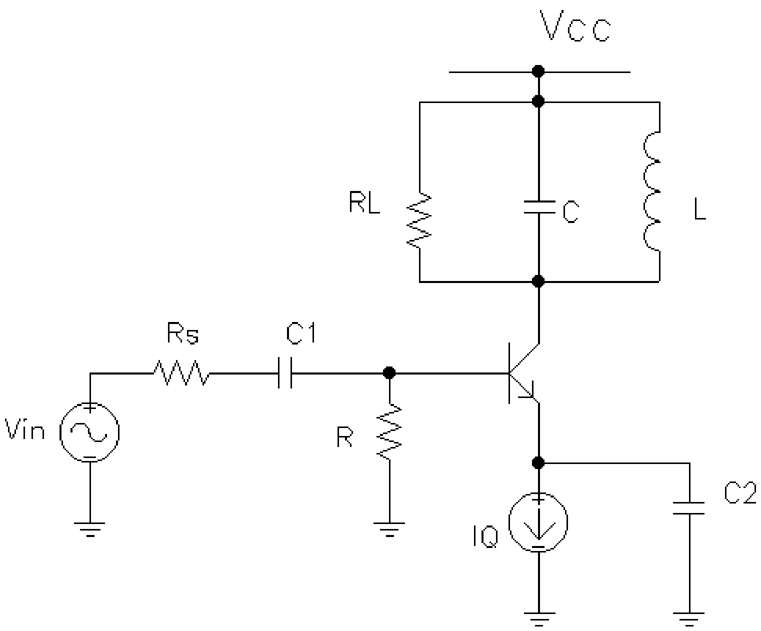
Come indicato, la fase del guadagno vale 0 in ω_0, a conferma quindi del fatto che il risonante, alla frequenza ω_0, fornisce una tensione perfettamente sincrona, in frequenza ed in fase, con la corrente in ingresso. In pratica, in condizioni di risonanza, la capacità e l'induttanza si compensano a vicenda.

Per pulsazioni diverse da ω_0, il guadagno diminuisce e la fase della tensione V(t) non è più 0: per $\omega<\omega_0$ prevale la suscettanza induttiva $1/\omega L$ rispetto alla suscettanza capacitiva ωC, per cui V(t) presenta un anticipo di fase ($\varphi>0$), mentre per $\omega>\omega_0$ prevale la suscettanza capacitiva e la V(t) presenta un ritardo di fase ($\varphi<0$). L'andamento della fase del guadagno è diagrammato nella figura seguente:



(la scala usata in ascisse, per ω, è logaritmica)

Torniamo al modulo dell'impedenza Z_L: dato che esso presenta un massimo alla frequenza di risonanza, è adatto ai nostri scopi, per cui possiamo pensare di utilizzare il seguente circuito:



In questo circuito, il guadagno di tensione per piccolo segnale è adesso pari (in modulo) a

$$|A_V(\omega)| \cong g_m |Z_L(\omega)| = g_m \frac{1}{\sqrt{\frac{1}{R_L^2} + \left(\omega C - \frac{1}{\omega L}\right)^2}}$$

Esso varia con ω allo stesso modo con cui Z_L varia con ω, per cui l'amplificatore effettua un filtraggio con le caratteristiche da noi desiderate. Dimensionando i valori di L e di C, fissiamo la frequenza centrale ω_0; fissando poi il valore di R fissiamo in pratica l'estensione della banda passante attorno ad ω_0. A questo proposito possiamo anche fare qualche calcolo. In particolare, definendo **banda passante** dell'amplificatore/filtro *l'intervallo di frequenza entro cui il guadagno dell'amplificatore si mantiene non più di 3dB al di sotto del valore massimo (pari a $g_m R_L$), possiamo calcolare gli estremi di tale banda passante.*

Fissato dunque il valore $g_m R_L$ massimo del guadagno (che si ha appunto in corrispondenza di ω_0), dobbiamo valutare per quali pulsazioni $|A_V(\omega)|$ diventa $\sqrt{2}/2$ volte più piccolo:

$$g_m \frac{1}{\sqrt{\frac{1}{R_L^2}+\left(\omega C-\frac{1}{\omega L}\right)^2}}=\frac{g_m R_L}{\sqrt{2}} \longrightarrow \frac{1}{\frac{1}{R_L^2}+\left(\omega C-\frac{1}{\omega L}\right)^2}=\frac{R_L^2}{2} \longrightarrow 2=1+R_L^2\left(\omega C-\frac{1}{\omega L}\right)^2$$

Dobbiamo dunque risolvere l'equazione di 2° grado

$$2=1+R_L^2\left(\omega C-\frac{1}{\omega L}\right)^2$$

Prende il nome di "**fattore di merito**" del risonante parallelo[11] il rapporto tra la corrente nel condensatore (o dell'induttore, in base a quanto appena visto) e la corrente in ingresso al circuito, quando il circuito è in condizioni di risonanza:

$$\boxed{Q=\frac{I_L}{I}=\frac{I_C}{I}=R\sqrt{\frac{C}{L}}}$$

L'equazione da risolvere diventa allora

$$1=Q^2\left(\omega LC-\frac{1}{\omega LC}\right)^2$$

Ricordando che $\omega_0=\frac{1}{\sqrt{LC}}$, possiamo porre l'equazione nella forma

$$1=Q^2\left(\frac{\omega}{\omega_0}-\frac{\omega_0}{\omega}\right)^2$$

Facendo la radice di entrambi i membri e andando poi a risolvere rispetto ad ω, si trovano le seguenti 2 soluzioni:

$$\omega_1=-\frac{\omega_0}{2Q}+\sqrt{\frac{\omega_0^2}{4Q_2}+\omega_0^2} \qquad \omega_2=+\frac{\omega_0}{2Q}+\sqrt{\frac{\omega_0^2}{4Q_2}+\omega_0^2}$$

Questi sono dunque gli estremi della banda passante, la cui ampiezza è allora

$$\boxed{B_{passante}=\omega_2-\omega_1=\frac{\omega_0}{Q}}$$

Abbiamo cioè trovato che la banda passante è il rapporto tra la pulsazione di risonanza del carico risonante ed il fattore di merito del risonante stesso.

[11] E' importante specificare che il fattore di merito (o anche fattore di qualità) è riferito al risonante parallelo, in quanto, se si considera il circuito duale (RLC serie con gli stessi valori degli elementi), si trova un fattore di merito serie pari al reciproco del fattore di merito parallelo.

Considerando allora che il rapporto tra passante passante e frequenza centrale è stato in precedenza definito come **banda relativa**, concludiamo che

$$Q = \frac{\omega_0}{B_{passante}} = \frac{1}{B_{relativa}}$$

Il fattore di merito è il reciproco della banda relativa.

Di conseguenza, dovendo dimensionare un carico risonante per un amplificatore ad emettitore comune come quello che stiamo considerando, diventa fondamentale il fattore di merito Q, che ci determina la banda passante dell'amplificatore stesso.

Intuitivamente, il fattore di merito viene dunque a rappresentare la selettività dell'amplificatore, in quanto quantifica la differenza che si ha nel modulo del guadagno a seguito di un piccolo cambiamento di ω rispetto ad ω_0.

Richiami sul fattore di merito

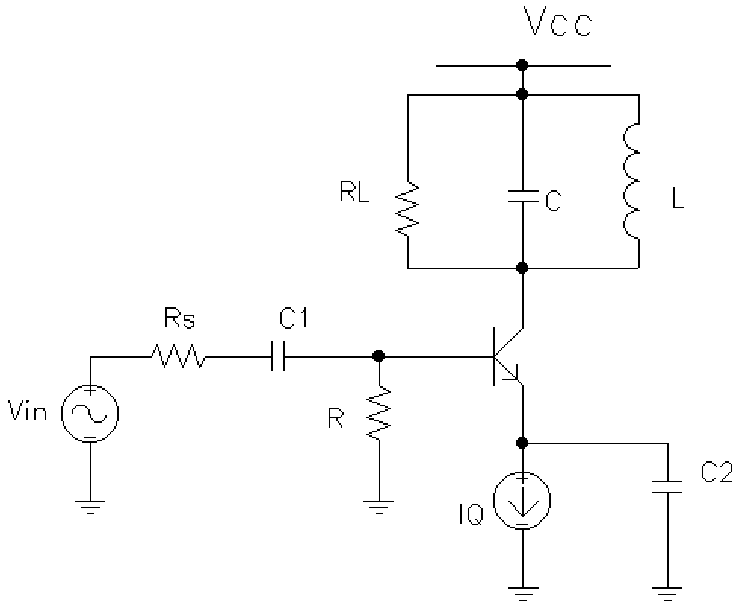
Il fattore di merito di un circuito RLC ha anche un'altra definizione, di tipo energetico: in regime sinusoidale, esso corrisponde infatti al rapporto tra la potenza reattiva e la potenza dissipata dal circuito. Esso si intende sempre calcolato alla frequenza di risonanza ed ha le seguenti espressioni:

RLC risonante parallelo $\rightarrow Q=R/\omega_0 L$
RLC risonante serie $\rightarrow Q=\omega_0 L/R$

Una importante osservazione è che, a rigore, la frequenza ω_0 non si trova esattamente al centro tra ω_1 ed ω_2. Tuttavia, dato che fa molto comodo una simmetria rispetto a ω_0, noi considereremo sempre che essa sussista.

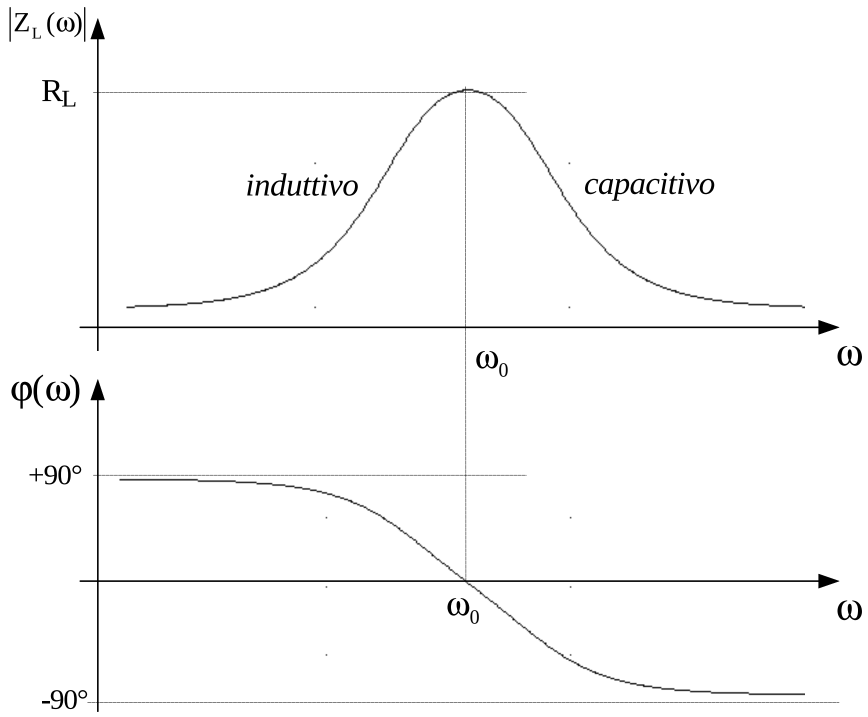
PROBLEMI DI INSTABILITÀ E DI ALLINEABILITÀ

Torniamo adesso nuovamente all'amplificatore ad emettitore comune il cui carico è costituito da un risonante RLC:



Ci sono una serie di problematiche da considerare.

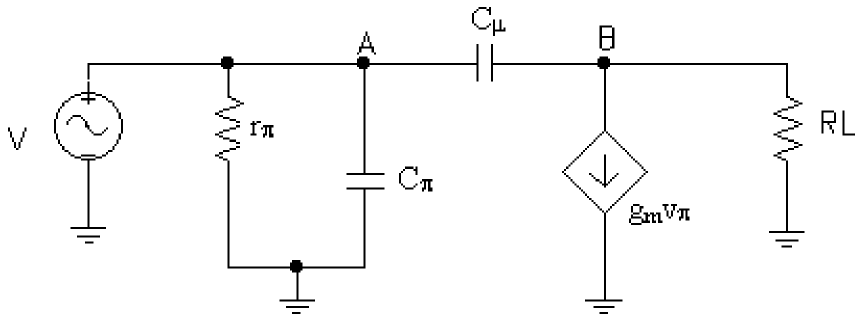
Abbiamo detto che il modulo dell'impedenza del risonante, cioè quindi il modulo del rapporto tra la tensione ai capi del risonante e la corrente che lo alimenta, ha il seguente andamento:



Il comportamento dell'impedenza varia dunque profondamente al variare della frequenza, passando da induttivo a capacitivo se ci spostiamo da sinistra a destra di ω_0. Dato che il circuito, in generale, riceve in ingresso un segnale avente più componenti spettrali oltre ω_0, esso si comporta in modo diverso per tali componenti:

- per l'unica componente a frequenza ω_0, il risonante si comporta come un resistore di valore R_L;
- per le componenti a frequenza $\omega<\omega_0$, il risonante si comporta come un induttore variabile con ω;
- per le componenti a frequenza $\omega>\omega_0$, il risonante si comporta come un condensatore variabile con ω.

Vediamo allora quali sono le conseguenze di questo fatto. In particolare, ci mettiamo in una ipotesi diversa dai casi precedenti: mentre, fino ad ora, abbiamo sempre supposto che le frequenze di lavoro non fossero tali da eccitare gli effetti capacitivi intrinseci al transistor, supponiamo di essere invece in condizioni di lavoro tali da dover necessariamente considerare anche C_π e C_μ. Questo significa che dobbiamo includere queste capacità nel modello di piccolo segnale del nostro circuito. Andiamo allora a disegnare tale modello, supponendo per il momento che il carico sia puramente resistivo:



In questo circuito abbiamo trascurato sia la resistenza intrinseca di base r_b sia la resistenza di uscita r_O. Abbiamo inoltre supposto che il transistor sia pilotato da un circuito esterno in modo che sia V la tensione applicata alla giunzione base-emettitore.

Ci interessa calcolare la corrente che il circuito esterno vede assorbire dal transistor (in modo da ricavare l'impedenza di ingresso V/i del circuito): si tratta evidentemente della somma di 3 correnti, che sono quella assorbita da r_π, quella assorbita da C_π e quella assorbita da C_μ:

$$i = i_{r\pi} + i_{C\pi} + i_{C\mu} = \frac{v_\pi}{r_\pi} + \frac{v_\pi}{1/j\omega C_\pi} + i_{C\mu} = \frac{V}{r_\pi} + j\omega C_\pi V + i_{C\mu}$$

Per calcolare la $i_{C\mu}$, osserviamo che $V=v_{C\mu}+v_L$, per cui

$$i_{C\mu} = \frac{v_{C\mu}}{1/j\omega C_\mu} = j\omega C_\mu v_{C\mu} = j\omega C_\mu (V - v_L) = j\omega C_\mu \left(V - R_L\left(i_{C\mu} - g_m V\right)\right)$$

Se, nella corrente $\left(i_{C\mu} - g_m V\right)$ che dà luogo alla tensione v_L, trascuriamo il contributo proveniente dalla C_μ stessa, possiamo approssimare

$$i_{C\mu} \cong j\omega C_\mu (V + g_m R_L V) = j\omega C_\mu (1 + g_m R_L) V$$

Questa formula esprime il cosiddetto effetto Miller, in base al quale il circuito esterno vede la capacità C_π in parallelo ad una capacità $C_\mu(1+g_m R_L)$, dove $g_m R_L$ è il modulo del guadagno del transistor a centro banda ($\omega=\omega_0$).

Tornando all'espressione di i, otteniamo

$$\boxed{i = \frac{V}{r_\pi} + j\omega C_\pi V + j\omega C_\mu (1 + g_m R_L) V}$$

Questa espressione ci dice che il circuito esterno, a fronte di una tensione applicata V, vede assorbire una corrente formata da 3 contributi: un contributo V/r_π, dovuto solo alla resistenza r_π e quindi in fase con V, una corrente $j\omega C_\pi V$, dovuta solo dalla C_π e in quadratura di fase con V, ed infine una corrente $j\omega C_\mu (1 + g_m R_L) V$, sempre in quadratura con V, dovuta alla capacità C_μ e all'effetto Miller cui essa è soggetta.

Questo discorso vale dunque se il carico è resistivo, il che accade, nel caso di aver usato il risonante RLC parallelo, solo alla pulsazione ω_0. Se la frequenza è invece inferiore ad ω_0, allora il risonante si comporta come una induttanza: il discorso non cambia, salvo a sostituire RL con una impedenza induttiva, cioè del tipo $j\omega L$, per cui la corrente assorbita è

$$i = \frac{V}{r_\pi} + j\omega C_\pi V + j\omega C_\mu (1 + j\omega g_m L) V = \frac{V}{r_\pi} + j\omega (C_\pi + C_\mu) V - \omega^2 g_m L C_\mu V$$

I primi due contributi non cambiano, in quanto non dipendono dal carico, mentre cambia il contributo di C_μ, che si scinde in un contributo $j\omega C_\mu V$ in quadratura e in un contributo $-\omega^2 g_m L C_\mu V$. La presenza del segno negativo indica che questa corrente non è assorbita dal transistor, ma torna indietro verso il circuito. E' come il contributo di corrente dovuto ad una resistenza negativa $-1/\omega^2 g_m L C_\mu$.

Se il carico, anziché induttivo, risulta capacitivo (il che avviene per frequenze $\omega>\omega_0$), allora dobbiamo considerare una reattanza $1/j\omega C$, per cui

$$i = \frac{V}{r_\pi} + j\omega C_\pi V + j\omega C_\mu \left(1 + \frac{g_m}{j\omega C}\right) V = \frac{V}{r_\pi} + j\omega\left(C_\pi + C_\mu\right)V + g_m \frac{C_\mu}{C} V$$

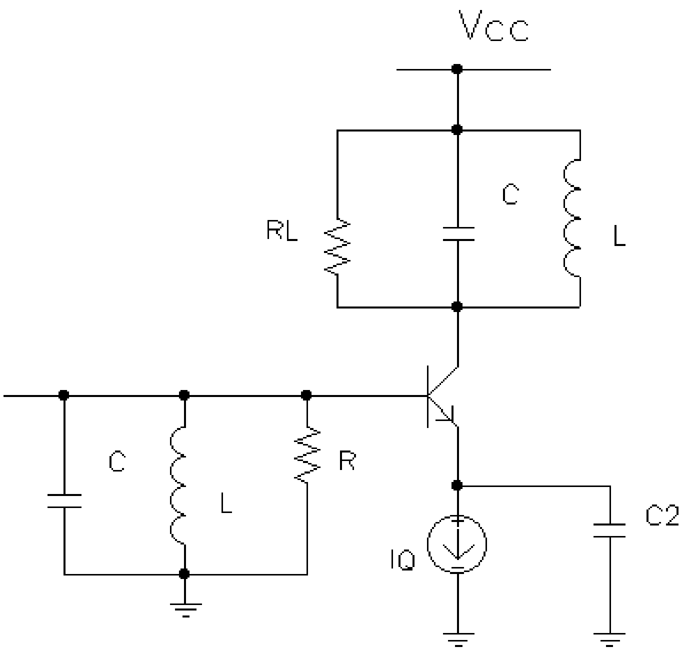
In questo caso, abbiamo una corrente positiva (cioè assorbita dal circuito), in fase con V, di valore $g_m \frac{C_\mu}{C} V$.

La capacità C_μ applica dunque una reazione che, a seconda dei valori reattivi assunti dal carico, può creare dei problemi notevoli: in particolare, il problema principale è che, nel caso di carico induttivo, la resistenza negativa può bilanciare ed anche superare il valore di resistenza positiva effettivamente presente nel circuito (la r_π e le eventuali resistenze di polarizzazione poste sulla base del transistor); se questo avviene, ***se cioè il circuito presenta una parte reale dell'impedenza di ingresso (V/i) di segno negativo, il circuito stesso non rappresenta più un amplificatore, ma un oscillatore***[12].

Se poi l'entità della resistenza negativa riportata in ingresso è particolarmente eccessiva, possiamo trovarci addirittura di fronte a problemi di instabilità.

C'è anche un altro problema che si fa sentire prima ancora di una eventuale oscillazione o di una eventuale instabilità: è il problema della *allineabilità* tra gli stadi di un amplificatore.

Un amplificatore, in generale, non è formato da un solo stadio come l'emettitore comune descritto fino ad ora, ma da una cascata di più stadi. Nel caso di un ricevitore sintonizzabile, quindi, il generico stadio amplificatore avrà sia come carico sia come ingresso un risonante:



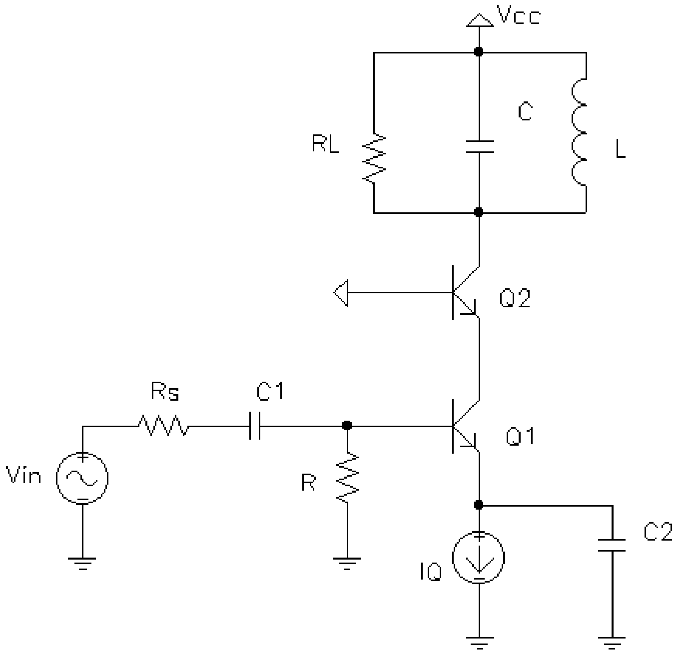
La presenza del risonante in ingresso è dovuta o al fatto che l'ingresso è accordato o, quanto meno, al fatto che è accordato il carico dello stadio precedente.

Per una struttura di questo tipo, ci sono grossi problemi, legati al fatto che, una volta montato il circuito, questo va tarato, ossia si deve poter operare delle regolazioni tali da compensare piccoli scostamenti rispetto ai valori teorici previsti. Queste regolazioni sono possibili solo se le capacità e

[12] In generale, se un circuito presenta una impedenza di ingresso con parte resistiva di segno negativo, esso non rappresenta più un utilizzatore, ma un generatore.

le induttanze sono variabili e quindi regolabili dall'operatore. Allora, quando si agisce, per esempio, sulla capacità C sul collettore, l'azione di retroazione della C_μ è tale da far variare anche l'impedenza vista dal risonante in ingresso, il che significa che varia anche la sintonia di quest'ultimo. Quindi, *prima ancora di ottenere problemi di instabilità, si presentano problemi di **allineabilità** degli stadi, dovuti agli effetti di reazione intrinseci ai vari dispositivi.*

A questo punto, è ovvio che dobbiamo limitare al minimo gli effetti della C_μ: in particolare, dato che l'effetto della C_μ è tanto maggiore quanto maggiore è l'effetto Miller, dobbiamo ridurre l'effetto Miller. Dagli studi di Elettronica applicata sappiamo che l'effetto Miller si riduce notevolmente con la **configurazione cascode**, nella quale, in pratica, si interpone, tra il transistor ad emettitore comune ed il carico, uno stadio a base comune, come nella figura seguente:



Con questa configurazione, l'emettitore comune vede un carico pari a $1/g_{m2}$, il che fa si che il suo guadagno di tensione diventi praticamente unitario: in questo modo, l'effetto Miller sulla $C_{\mu 1}$ è estremamente ridotto. Per quanto riguarda Q2, invece, non c'è effetto Miller, in quanto, essendo la base a massa per il segnale[13], la $C_{\mu 2}$ va semplicemente in parallelo al carico.

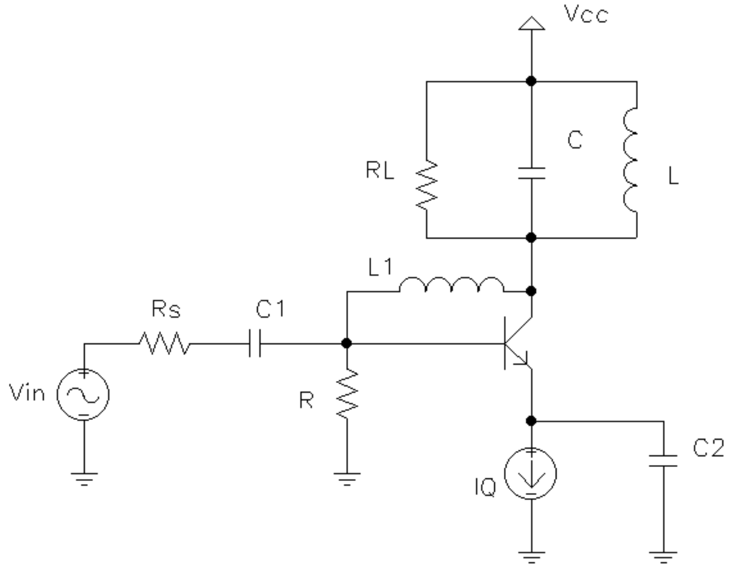
Lo stadio a base comune preleva, in pratica, la corrente di collettore di Q1 e la riporta sul carico (ha cioè un guadagno di corrente praticamente unitario tra collettore ed emettitore): in questo modo, *ottieniamo un guadagno di tensione complessivo pari a quello di uno stadio ad emettitore comune, ma conserviamo una banda passante pari a quella di uno stadio a base comune, che è notoriamente molto estesa.*

La configurazione cascode ci dà quindi tutte le caratteristiche amplificative e selettive di uno stadio CE con carico risonante, ma ci dà, in più, un bassissimo effetto di retroazione tra uscita ed ingresso, avendo neutralizzato l'effetto di $C_{\mu 1}$.

La configurazione cascode comporta, comunque, una complicazione del circuito, per cui possiamo pensare a qualche altra soluzione, che utilizzi ancora un solo transistor.

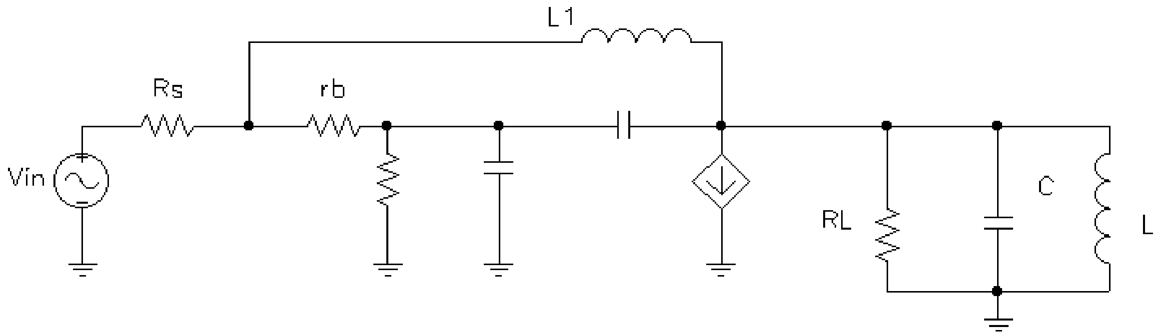
Un modo intuitivo di procedere sarebbe quello di porre in parallelo alla C_μ una induttanza che ne compensi gli effetti:

[13] Dal punto di vista della polarizzazione, la base di Q2 non è a massa, ma ad una opportuna tensione necessaria a garantire la polarizzazione diretta della giunzione base-emettitore di Q2 stesso



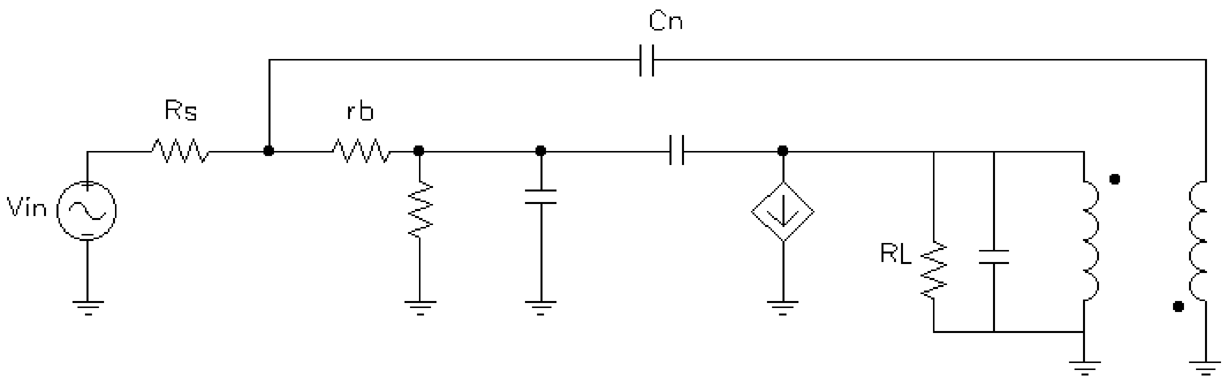
Si crea cioè un risonante LC parallelo, con Q molto alto, che si comporta o come una resistenza di valore molto grande, in corrispondenza della risonanza, o come una reattanza (capacitiva o induttiva a seconda della frequenza) di valore molto piccolo:

Questo modo di procedere non dà però i frutti sperati a causa della inevitabile presenza della resistenza intrinseca di base r_b. Infatti, per porre una induttanza L1 in parallelo a C_μ, dovremmo aver accesso alla base intrinseca ed al collettore intrinseco, in quanto la C_μ è situata tra questi due nodi; mentre la resistenza intrinseca di collettore è trascurabile, per cui abbiamo accesso al collettore intrinseco, la r_b non lo è, per cui andremmo a porre l'induttanza sulla base estrinseca, come nella figura seguente:



Oltre a questa considerazione, l'idea di usare una induttanza L_1 si scontra anche col fatto che la C_μ viene compensata solo alla frequenza di risonanza, mentre, alle altre frequenze, predomina una delle due reattanze.

Diventa invece molto più vantaggioso usare, al posto di L_1, una ulteriore capacità (detta di neutralizzazione) connessa, mediante un trasformatore, nel modo seguente:



Si opera in modo da garantire due condizioni:

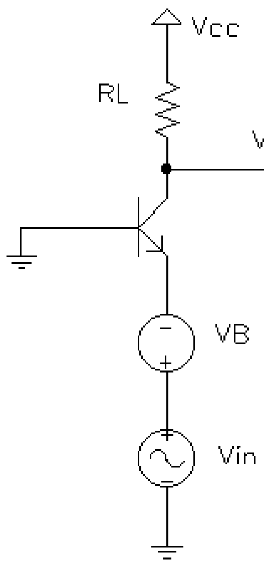
- in primo luogo, si fa in modo che la corrente nella capacità C_n sia di segno opposto rispetto a quella assorbita dal circuito, il che si ottiene invertendo il segno dei due avvolgimenti del trasformatore;
- in secondo luogo, combinando opportunamente il valore di Cn e il rapporto spire N del trasformatore, si fa in modo che, alla frequenza di lavoro, la corrente fornita da C_n sia assolutamente uguale a quella assorbita da C_μ: in questo modo, la C_n compensa la corrente assorbita da C_μ, eliminando i problemi di cui abbiamo parlato fino ad ora.

Il grosso vantaggio di questo meccanismo è nel fatto che, al variare della frequenza, aumentano, nello stesso identico modo (visto che sono entrambe delle capacità), il modulo della corrente che arriva da C_n e il modulo della corrente assorbita da C_μ. Il bilanciamento viene cioè effettuato con due termini che, al variare della frequenza, variano nella stessa direzione.

A conclusione del discorso sulla **amplificazione selettiva**, osserviamo quanto segue: *l'uso di circuiti risonanti negli amplificatori garantisce una buona selettività, ma non è proprio indolore, visto che bisogna ricorrere o a particolari configurazioni circuitali (nel senso di usare più dispositivi attivi opportunamente combinati, come nella configurazione cascode) o a circuiti un po' più complicati che consentano di effettuare la neutralizzazione di C_m.*

AMPLIFICATORE IN CLASSE C

Nei precedenti paragrafi abbiamo considerato un amplificatore selettivo realizzato mediante uno stadio ad emettitore comune con carico risonante sul collettore. Adesso consideriamo un altro circuito e, in particolare, uno stadio a base comune (inseguitore di corrente) anch'esso eventualmente con carico risonante sul collettore:



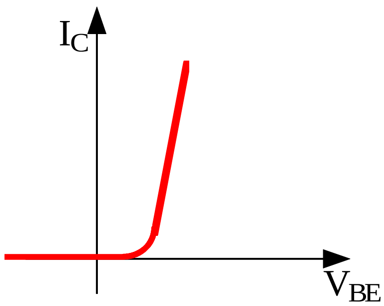
In questo circuito, la polarizzazione è garantita, oltre ovviamente che dall'alimentazione V_{CC}, dalla batteria V_B posta sull'emettitore: infatti, essendo la base a massa, il compito di tale batteria è di mantenere l'emettitore ad una tensione di circa -0.7V, in modo da garantire la polarizzazione diretta della giunzione base-emettitore[14] (e quindi il funzionamento del transistor in zona lineare) anche in assenza di segnale. Il segnale che lo stadio deve elaborare è la tensione $v_{in}(t)$ posta in serie alla batteria.

Per l'analisi del comportamento sotto segnale di questo circuito, facciamo delle considerazioni preliminari. In particolare, fino ad ora abbiamo supposto che il transistor funzioni, sotto segnale, sempre in modo lineare, il che accade, come sappiamo, quando le variazioni istantanee della tensione, ai capi della giunzione V_{BE}, rispetto al valore nel punto di lavoro sono dell'ordine della tensione termica kT/q e non superiori. L'ipotesi di comportamento lineare è utile in quanto possiamo sostituire il circuito con il suo equivalente lineare (per piccoli segnali), la cui analisi può essere condotta semplicemente con le leggi di Kirchoff applicate a elementi lineari (resistori, induttori, condensatori, generatori pilotati). Al contrario, quando il circuito è pilotato da un segnale passa-banda, è molto probabile che esso non funzioni sempre in zona lineare, data la presenza di più armoniche sovrapposte. Vogliamo allora capire cosa succede al circuito in esame quando siamo costretti ad abbandonare l'ipotesi di perfetta linearità dell'amplificatore.

Una volta stabilito non valido il modello lineare del transistor, dobbiamo considerare un opportuno modello non lineare e, nel caso del BJT, quello più comodo (ma anche più fedele) da usare è il modello di Ebers-Moll, ossia un modello puramente esponenziale. In base a tale modello, quando il transistor funziona in zona attiva diretta, la corrette di collettore è esprimile tramite una dipendenza esponenziale dalla tensione tra base ed emettitore:

$$\boxed{i_C(t) = I_{CS} e^{\frac{v_{BE}(t)}{V_T}}}$$

Questa relazione lega la corrente di collettore alla tensione base-emettitore secondo un relazione grafica del tipo seguente:



Se consideriamo un segnale applicato di piccola ampiezza (non oltre i 10 mV ai capi della giunzione B-E), possiamo sicuramente linearizzare la caratteristica nell'intorno del punto di lavoro e lavorare con il modello lineare del transistor. Supponiamo allora che l'ingresso del circuito sia puramente sinusoidale del tipo

$$v_{in}(t) = V_1 \cos(\omega_0 t)$$

Se V_1 è dell'ordine di 10 mV (o comunque inferiore ai 26mV della tensione termica[15]), l'argomento $v_{BE}(t)/V_T$ dell'esponenziale è piccolo, per cui l'esponenziale stesso può essere sviluppato in serie di Taylor: dallo sviluppo, trascurando i termini di ordine superiore al primo, si ottiene un legame lineare tra corrente di emettitore e tensione applicata (i passaggi analitici saranno fatti più avanti, visto che adesso ci interessa solo l'aspetto qualitativo della cosa).

[14] Se la batteria V_B (=0.7V) è posta con le polarità come nella figura (ossia col morsetto a potenziale maggiore connesso a massa), è evidente che la tensione continua applicata alla giunzione B-E è proprio uguale a V_B, per cui il punto di lavoro è (V_B,I_Q), dove I_Q è la corrente di collettore (approssimativamente pari a quella di emettitore) dovuta appunto alla polarizzazione.

[15] La tensione termica vale 26mV solo a temperatura ambiente

Man mano che la tensione applicata aumenta di ampiezza, interessiamo una zona della caratteristica I_C-V_{BE} che non più essere approssimata da un andamento lineare. Subentrano allora due considerazioni da fare:

- in primo luogo, bisogna considerare che la caratteristica esponenziale aumenta, rispetto al punto di lavoro, molto più rapidamente di quanto non decresca: ciò significa che, per scostamenti elevati dal punto di lavoro, la corrente di collettore aumenta, in corrispondenza dei picchi positivi del segnale, molto più di quanto diminuisce in corrispondenza dei picchi negativi;
- in secondo luogo, le escursioni dal punto di lavoro possono anche essere tali da portare il transistor in zona di saturazione (in corrispondenza dei picchi positivi del segnale) oppure in interdizione (in corrispondenza dei picchi negativi): ciò significa che, mentre prima, nel funzionamento lineare, il transistor conduceva per l'intero periodo della sinusoide in ingresso, adesso esso conduce per un intervallo di tempo minore del periodo, in quanto il transistor va in saturazione o in interdizione in corrispondenza dei picchi positivi e negativi).

In base a queste considerazioni, quindi, per un segnale (sinusoidale) applicato di ampiezza elevata, la forma d'onda della corrente di collettore è fortemente deformata rispetto alla sinusoide che invece si avrebbe se il funzionamento fosse sempre lineare.

A parità di segnale applicato, chi regola il periodo di conduzione è la tensione di polarizzazione V_B. Vediamo di chiarire il concetto mediante una simulazione con P-Spice.

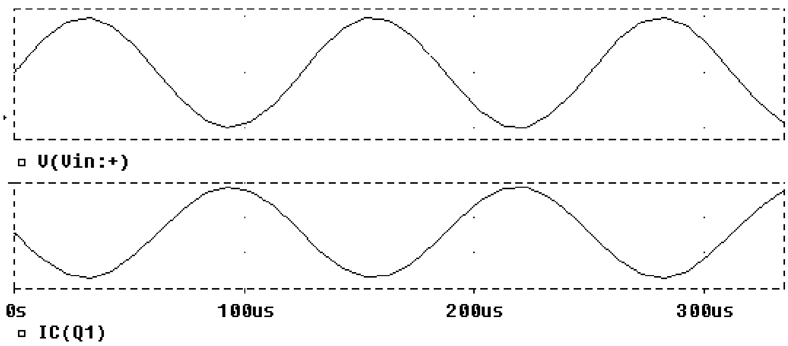
In primo luogo, abbiamo chiarito che il valore di I_C nel punto di lavoro è fissato dal valore di V_B: ad esempio, prendendo V_B=0.7V, la corrente di collettore vale

$$I_C^Q = I_{CS} e^{\frac{V_{BE}^Q}{V_T}} = I_{CS} e^{\frac{V_B}{V_T}} = 1.53\text{mA}$$

(dove abbiamo considerato I_{CS}=3fA). La simulazione P-Spice indica una corrente di collettore di 1.73mA.

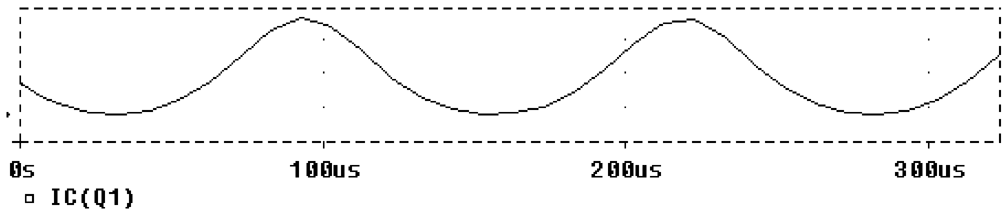
Per garantire che il punto di lavoro sia in zona attiva diretta, dobbiamo imporre che la V_{CE} sia al di sopra del valore di saturazione di circa 0.2V. Questo lo otteniamo, una volta nota l'alimentazione V_{CC}, impostando opportunamente il valore della resistenza R_L sul collettore: ad esempio, con una alimentazione da 5V e con R_L=1kΩ, si ottiene $V_{CE}^Q \cong 4\text{V}$.

Fissata la polarizzazione, possiamo applicare il nostro segnale. Supponendo di adottare una tensione sinusoidale a frequenza f_0=8kHz e di ampiezza 1 mV, la simulazione fornisce il seguente andamento temporale della I_C:



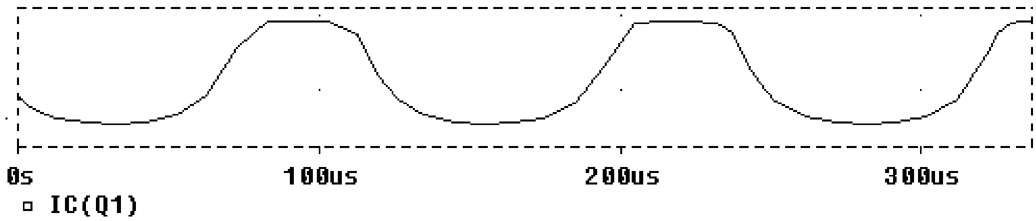
(la curva superiore riporta evidentemente la tensione di alimentazione, mentre la curva inferiore riporta la $i_C(t)$, che è notoriamente invertita di fase rispetto all'ingresso). Come si nota, non c'è alcuna distorsione, a conferma del fatto che il transistor funziona linearmente: la tensione di uscita è il prodotto della corrente per la resistenza R_L.

Adesso aumentiamo l’ampiezza della sinusoide a 20mV. La simulazione fornisce la seguente corrente di collettore:



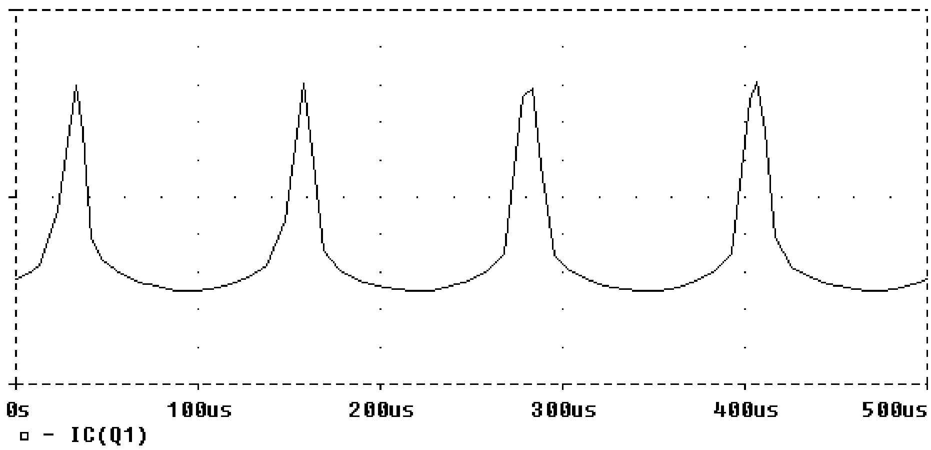
Come si vede, comincia a subentrare una distorsione, dovuta al fatto che la decrescita della I_C con la V_{BE} è meno veloce dell’aumento di I_C con la V_{BE}.

Aumentando ulteriormente la tensione di ingresso a 50mV, la distorsione è ancora più accentuata:

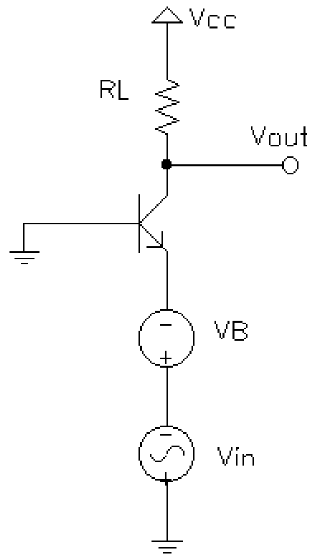


Subentra, in questo caso, anche il fenomeno del taglio della forma d’onda sui picchi, dovuto al fatto che il transistor va in saturazione (il valore minimo della $v_{CE}(t)$ risulta essere infatti di 32mV e si ottiene, ovviamente, in corrispondenza dei picchi negativi dell’ingresso, che poi corrispondono ai picchi positivi della V_B date le polarità considerate).

Adesso vediamo come cambiano le cose modificando la tensione V_B di polarizzazione, portandola ad esempio da 0.7V a 0.75V. La corrispondente corrente di collettore risulta essere di 5.72mA. Applicando in ingresso un segnale di ampiezza 20mV, l’andamento della corrente di collettore risulta essere il seguente (cambiato di segno, per evidenziarne le caratteristiche):



Quindi, a parità di segnale di ingresso, aumentando il valore di V_B otteniamo una corrente di collettore che tende ad assumere sempre più l’andamento di una successione di impulsi. Tali impulsi si ottengono in corrispondenza dei picchi positivi di $v_{IN}(t)$ se poniamo a massa il morsetto a potenziale minore di questo generatore oppure in corrispondenza dei picchi negativi se facciamo il contrario, ossia se disponiamo $v_{IN}(t)$ con la stessa polarità di V_B:

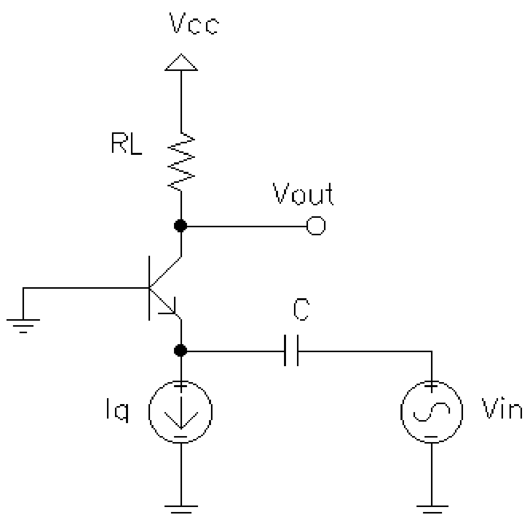


Deduciamo dunque che, all'aumentare di V_B, la corrente di collettore (e quindi l'uscita dell'amplificatore nel caso di carico puramente resistivo) tende ad essere una successione di impulsi che si susseguono a frequenza f_0, che quindi sono distanziati nel tempo di $1/f_0$.

E' ovvio che i picchi non possono essere troppo alti, in quanto è possibile che la giunzione base-collettore vada anch'essa in conduzione, portando il transistor in saturazione.

Si capisce dunque quanto possa essere pericolosa la scelta di polarizzare il transistor imponendo la tensione base-emettitore, dato appunto che il legame tra la I_C e la V_{BE} è particolarmente sensibile: la dipendenza esponenziale di I_C da V_{BE} è infatti tale che una minima variazione della tensione provochi una rilevante variazione della corrente.

Si preferisce allora polarizzare il transistor imponendo direttamente la corrente di collettore, mediante un generatore di corrente di valore I_Q prefissato:



Come detto, il generatore di corrente impone la polarizzazione, fissando il valore della I_E e quindi della I_C. E' chiaramente scomparso il generatore di tensione V_B. Il condensatore di blocco C serve ad evitare che la corrente I_Q venga cortocircuitata dal generatore di segnale $v_{in}(t)$.

In assenza di segnale, in conseguenza dell'imposizione di una corrente $I_Q = I_E^Q \cong I_C^Q$, ai capi della giunzione risulta localizzata una tensione

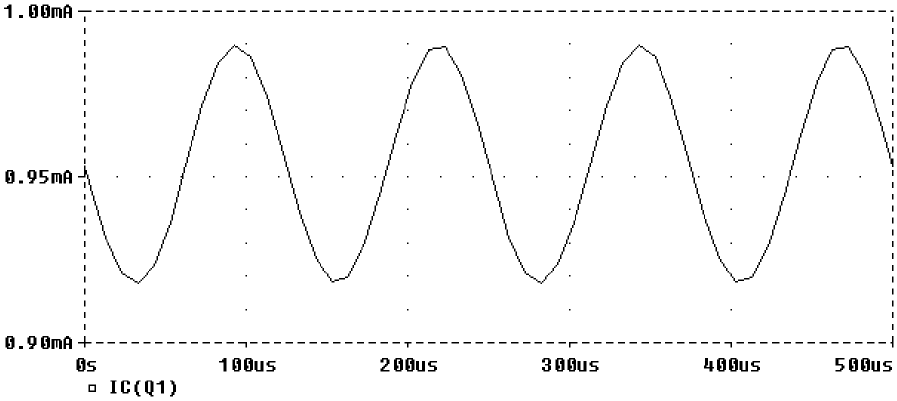
$$V_{BE}^Q = V_T \ln \frac{I_Q}{I_{CS}}$$

Essendo la base del transistor a massa, è evidente che questa stessa tensione (continua) è applicata ai capi del condensatore C, in assenza di segnale (l'emettitore è il nodo a potenziale inferiore): indicando allora con V_C la tensione ai capi di C, abbiamo che $V_C^Q = V_T \ln \frac{I_Q}{I_{CS}}$.

Nel momento in cui applichiamo il segnale $v_{in}(t)$, la tensione V_C, a rigore, non è più continua, ma possiamo invece supporre che lo sia: dobbiamo cioè supporre che il valore di C sia abbastanza grande che il passaggio di corrente alternata faccia variare la V_C in modo impercettibile rispetto al valore V_C^Q in continua.

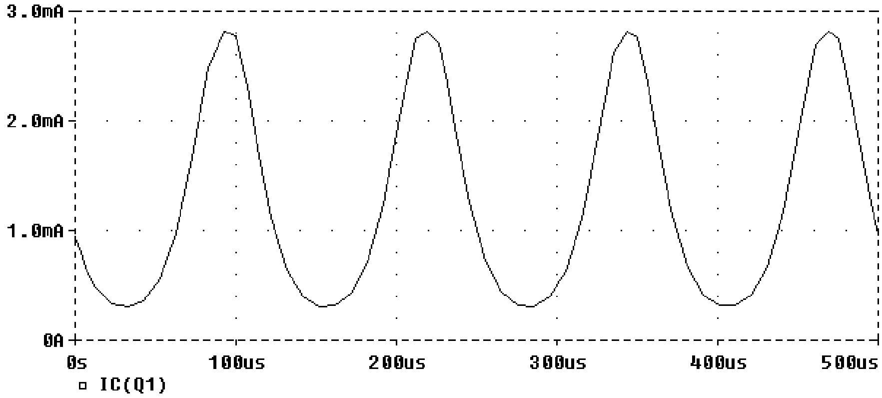
A questo punto, in presenza di segnale, possiamo ripetere discorsi analoghi a quelli fatti prima: dato l'elevato valore della capacità, il condensatore C presenta una notevole inerzia, per cui, all'aumentare della tensione applicata, abbiamo una corrente di emettitore, e quindi di collettore, fortemente deformata, che quindi tende ad assumere un andamento del tipo visto prima. Rispetto a prima, però, c'è questa volta una differenza fondamentale, legata al fatto che la polarizzazione è imposta dal generatore di corrente. Aiutiamoci ancora una volta con una simulazione P-Spice, nella quale imponiamo al transistor una corrente I_Q=1mA. In corrispondenza di questa corrente, la simulazione fornisce $I_C^Q = 0.95\text{mA}$ come corrente di polarizzazione. Fissata la polarizzazione, andiamo ad applicare il segnale:

- finché le escursioni di $v_{in}(t)$ sono piccole (non oltre i soliti 10 mV), la tensione su C praticamente non varia e si ottengono escursioni simmetriche della I_C rispetto al valore I_Q in continua, per cui I_Q rappresenta proprio il valore medio della corrente; ad esempio, con un segnale di ampiezza 1mV, la simulazione fornisce il seguente andamento della I_C:



  Come si vede, il valore medio è proprio 0.95mA, cioè il valore nel punto di lavoro;

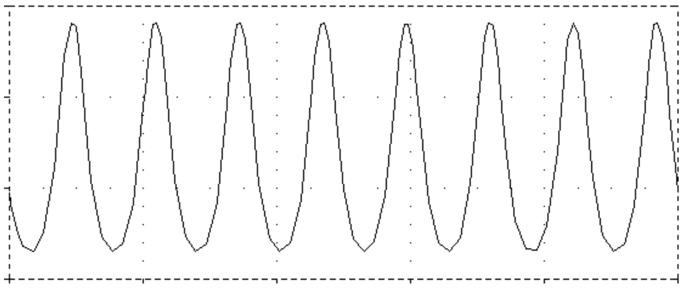
- quando, invece, le escursioni di $v_{in}(t)$ aumentano, per cui l'andamento della I_C è fortemente asimmetrico rispetto a I_Q , il valore medio non è più I_Q, ma è più grande, a causa sempre della caratteristica esponenziale che lega I_C a V_{BE}; ad esempio, con un segnale da 30 mV, l'andamento di I_C è il seguente:



E' evidente la maggiore distorsione rispetto a prima, come è anche evidente che il valore medio è salito. L'aumento del valor medio non può certo provenire dal generatore di corrente, il quale fornisce sempre I_Q, per cui deve provenire necessariamente dal condensatore. Tuttavia, questo condensatore, in risposta al flusso di corrente che lo attraversa, localizza ai suoi capi una tensione che si oppone alla variazione: dato che, all'aumentare dell'ampiezza di $v_{in}(t)$, aumenta l'ampiezza degli impulsi I_C, il condensatore viene attraversato da una corrente (che va dall'emettitore verso terra) in cui gli impulsi in salita sono dominanti rispetto a quelli in discesa; di conseguenza, il condensatore aumenta la tensione dell'armatura di sinistra rispetto a quella di destra, determinando così una diminuzione della tensione V_C ai suoi capi. Ad una diminuzione della V_C corrisponde una diminuzione della I_C, per cui l'azione del condensatore è quella di spostare il valore medio della corrente verso valori più piccoli, anche se la tensione applicata resta la stessa, finché il suddetto valor medio non torna ad essere I_Q.

In definitiva, il condensatore produce una autoregolazione della corrente in modo tale che, a regime instaurato, la corrente di emettitore sia sempre I_Q. *In corrispondenza di questa nuova situazione, la tensione ai capi del condensatore è cambiata: essa tende a diminuire quanto più cresce la tensione $v_{in}(t)$.*

In seguito a questo meccanismo, quanto più aumenta l'ampiezza di $v_{in}(t)$, tanto più stretti diventano gli impulsi della corrente di collettore:

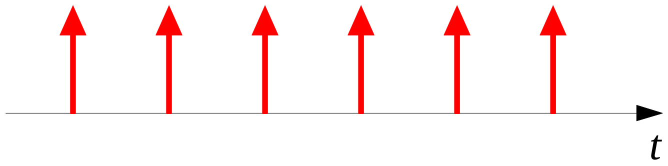


Il valore minimo della corrente è 0, corrispondente a transistor interdetto, mentre il valore massimo deve essere tale che il transistor, in conduzione, fornisca un valor medio di corrente pari ad I_Q.

In pratica, *il condensatore fa si che l'intervallo di conduzione del transistor, in ogni periodo della sinusoide $v_{in}(t)$ di ingresso, dipenda interamente dall'ampiezza V_1 della sinusoide stessa.*

Quando V_1 tende ad essere molto grande, gli impulsi diventano molto stretti e si ripetono a distanza pari al periodo della sinusoide $T=2\pi/\omega_0$. E' bene però sottolineare che l'ampiezza di V_1 non deve essere tale da mandare in saturazione l'amplificatore, in quanto, in caso contrario, gli impulsi verrebbero tagliati sui rispettivi picchi.

Questo discorso ci consente di adottare un buon modello matematico per la corrente di collettore. Infatti, avendo detto che si tratta di impulsi intercalati di T, possiamo pensare, in modo molto approssimato, ad una successione di impulsi matematici δ a distanza T uno dall'altro:



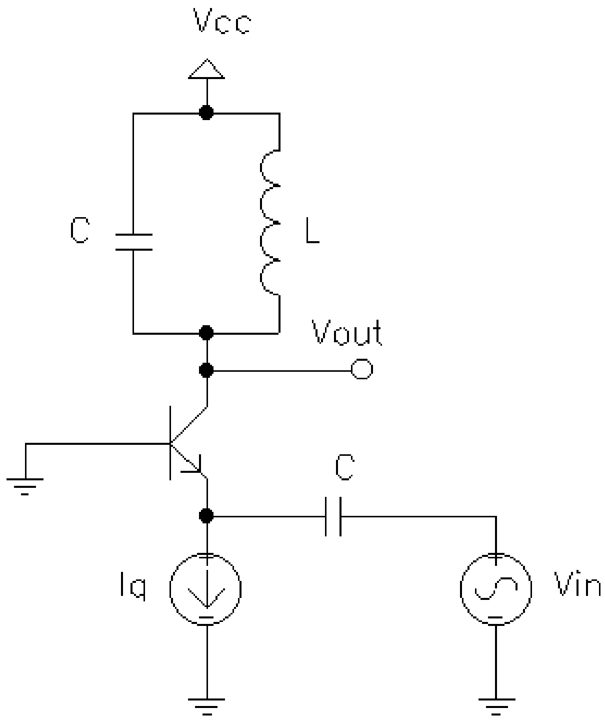
Con questo modello, supponendo che ogni impulso porti una carica Q, deve risultare

$$I_Q = Q/T$$

Infatti, la carica Q deve essere tale che il valore medio di corrente sia pari ad I_Q.

Con questo modello, quindi, ad una sollecitazione sinusoidale corrisponde una corrente di collettore impulsiva. Si tratta quindi di una forma d'onda periodica, il cui spettro sarà a sua volta costituito da una successione di impulsi (quindi di armoniche), a distanza 1/T tra di loro: avremo perciò la componente continua, corrispondente al valor medio I_Q, la prima armonica a frequenza 1/T, la seconda armonica a frequenza 2/T e così via.

Dato che sul collettore andiamo a sistemare un carico risonante, possiamo dimensionare la frequenza di risonanza in modo che coincida con una sola delle armoniche da cui è costituita la corrente di collettore. Per esempio, supponiamo di utilizzare un carico risonante LC, nel quale assumiamo trascurabili le perdite:

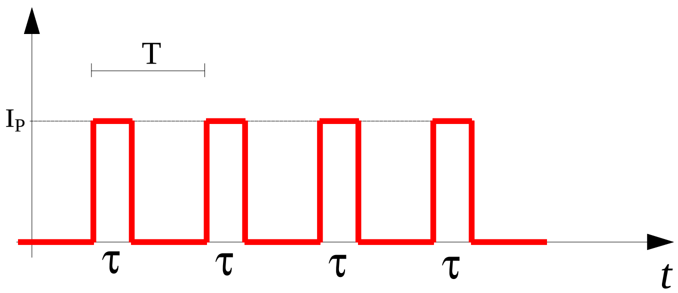


Scegliamo L e C in modo che la frequenza di risonanza $\omega_R = 1/\sqrt{LC}$ del risonante coincida con una delle armoniche di I_C. Questo significa che il carico ha una impedenza diversa da zero solo nell'intorno di ω_R, per cui la caduta di tensione ai suoi capi è una sinusoide a frequenza ω_R: tutti gli altri termini di corrente attraversano il circuito risonante, il quale, però, a quelle frequenze ha una impedenza praticamente nulla, per cui è nulla anche la corrispondente caduta di tensione).

Il risulta ottenuto con questo circuito è dunque il seguente: avendo in ingresso una tensione sinusoidale $v_{in}(t)$, all'uscita riusciamo ad ottenere una tensione sinusoidale alla stessa frequenza anche in presenza di una forte distorsione, a patto, ovviamente, di avere una selettività abbastanza alta (dovuta al risonante). In altre parole ancora, abbiamo trovato che, se il segnale in ingresso è una sinusoide, possiamo spingere la giunzione B-E ben oltre la zona in cui è lecita la linearizzazione, in quanto, all'uscita, continuiamo comunque ad ottenere una sinusoide.

Questo risultato è ottenuto in conseguenza del fatto che, a fronte di una sinusoide in ingresso, otteniamo una corrente di collettore non più sinusoidale, ma ancora periodica (la successione di impulsi visti prima), della quale quindi possiamo ancora sfruttare l'armonica che ci interessa.

E' evidente, comunque, che il modello della corrente di collettore come successione temporale di impulsi matematici è molto spinto, in quanto gli impulsi reali sono tutt'altro che ideale. Diventa già migliore l'uso di impulsi di durata τ finita:

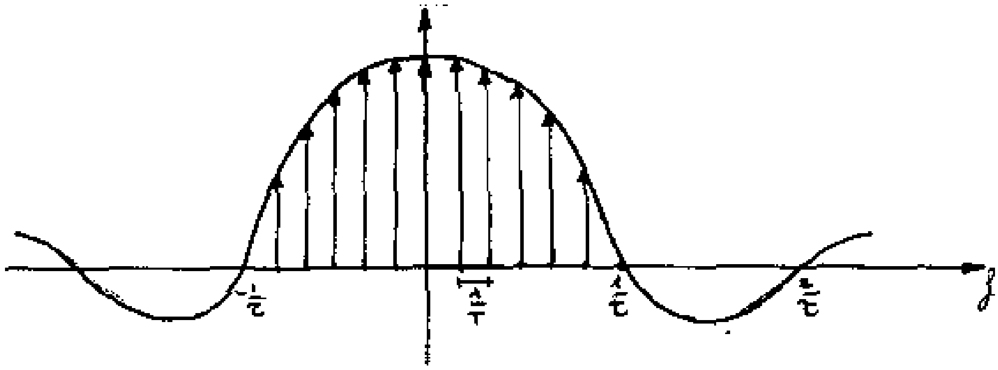


Ovviamente, ricordando sempre che questi impulsini devono convogliare una carica Q tale che $I_Q = Q/T$ e osservando che tale carica vale in questo caso $Q=I_P\tau$, deduciamo che

$$I_Q = \frac{Q}{T} = \frac{I_P \tau}{2\pi/\omega_0} = \frac{I_P \omega_0 \tau}{2\pi}$$

Sulla base di questo modello, possiamo andare a calcolare l'ampiezza della componente di prima armonica della corrente di collettore (cioè la componente a frequenza 1/T), in presenza di pilotaggio molto forte, cioè con ampiezza V_1 del segnale di ingresso molto elevata.

Dobbiamo semplicemente capire come sia fatto lo spettro della forma d'onda riportata nell'ultima figura. Trattandosi di una forma d'onda periodica, di periodo T, il suo spettro sarà una successione di impulsi, a partire dalla frequenza nulla, a distanza 1/T uno dall'altro. Possiamo caratterizzare questa successione di impulsi ricordando come si calcola lo spettro di un segnale periodico: si considera il segnale troncato al primo periodo, se ne calcola lo spettro, lo si scala di un fattore T e lo si campiona (in frequenza) a distanza 1/T. Nel nostro caso, il segnale troncato è un rettangolo di durata τ ed ampiezza I_P, per cui il suo spettro è del tipo $\sin(\pi f \tau)/\pi f \tau$. Campionandolo a passo 1/T, otteniamo lo spettro della corrente di collettore:



L’inviluppo dei vari impulsi è dunque del tipo $\sin(\pi f\tau)/\pi f\tau$, per cui esso dipende dal valore di τ: se gli impulsi di $i_C(t)$ sono molto stretti (ossia τ<<T), il primo zero della funzione $\sin(\pi f\tau)/\pi f\tau$, situato in 1/τ, si trova a notevole distanza dallo zero, il che significa che il lobo principale è abbastanza esteso e quindi anche che l’ampiezza della componente a frequenza 0 è praticamente uguale all’ampiezza della prima armonica positiva e negativa. Dato che la componente a frequenza 0 ha ampiezza pari al valor medio della corrente, ossia I_Q, deduciamo quindi che gli impulsi a frequenze -1/T e 1/T hanno entrambi ampiezza I_Q. Per ottenere l’ampiezza della prima armonica (passando cioè dagli esponenziali complessi alle funzioni sinusoidali reali), dobbiamo sommare le due ampiezze, per cui concludiamo che <u>l’ampiezza della prima armonica è $\boxed{I_{C1} = 2I_Q}$, pari al doppio della corrente media</u>.

Questo vale, come detto, fin quando la durata τ di conduzione del transistor è piccola rispetto al periodo di ripetizione T degli impulsi, ossia al periodo della sinusoide in ingresso

Questo risultato rappresenta una fondamentale differenza tra il circuito con la polarizzazione imposta dal generatore di corrente I_Q ed il circuito con la polarizzazione imposta dal generatore di tensione V_B: infatti, anche se in entrambi i casi risulta $I_{C1} = 2I_{C,medio}$, mentre nel primo caso $I_{C,medio}$ è un dato del problema (è pari ad I_Q) in quanto è imposto da noi, nel secondo caso (quello con V_B) non c’è alcun elemento che imponga un determinato valor medio della corrente di collettore, per cui tale valor medio non è a noi noto.

Guadagno dell’amplificatore

Andiamo adesso a calcolare con maggior rigore analitico il guadagno dell’amplificatore. Nel fare questo calcolo, facciamo due ipotesi:

- la prima, già citata, è quella di assumere un modello esponenziale per il legame tra I_C e V_{BE};
- la seconda è invece, contrariamente a quanto fatto nei precedenti paragrafi, quella di trascurare gli effetti capacitivi intrinseci al transistor, ossia la C_π e la C_μ: trascuriamo cioè gli accumuli di carica presenti nel transistor, ossia anche la “memoria” del dispositivo.

Detto questo, passiamo ai conti. In presenza di solo segnale di polarizzazione, abbiamo detto che la corrente di polarizzazione è

$$I_Q = I_C^Q = I_{CS} e^{\frac{V_{BE}^Q}{V_T}} = I_{CS} e^{\frac{V_C^Q}{V_T}}$$

dove ricordiamo che V_C^Q è la tensione continua ai capi del condensatore.

Aggiungendo il segnale $v_{in}(t) = V_1 \cos(\omega_0 t)$, la tensione ai capi della giunzione base-emettitore diventa $v_{BE}(t) = V_C^Q + V_1 \cos(\omega_0 t)$, per cui la corrispondente corrente di collettore è

$$i_C(t) = I_{CS} e^{\frac{v_{BE}(t)}{V_T}} = I_{CS} e^{\frac{V_C^Q + V_1 \cos(\omega_0 t)}{V_T}} = I_{CS} e^{\frac{V_C^Q}{V_T}} e^{\frac{V_1 \cos(\omega_0 t)}{V_T}} =$$

Abbiamo cioè il termine costante $I_{CS} e^{\frac{V_C^Q}{V_T}}$ ed il termine $e^{\frac{V_1 \cos(\omega_0 t)}{V_T}}$ chiaramente periodico, dato che si tratta di una sinusoide che passa attraverso una non linearità di tipo esponenziale. Possiamo allora sviluppare questo termine periodico in serie di Fourier:

$$i_C(t) = I_{CS} e^{\frac{V_C^Q}{V_T}} \sum_{k=0}^{\infty} A_k \cos((k\omega_0)t + \varphi_k)$$

L'espressione analitica dello sviluppo in serie è stata dimostrata in passato ed è la seguente (i coefficienti dello sviluppo sono le cosiddette funzioni di Bessel modificate):

$$i_C(t) = I_{CS} e^{\frac{V_C^Q}{V_T}} \left[I_0\left(\frac{V_1}{V_T}\right) + 2\sum_{k=1}^{\infty} I_k\left(\frac{V_1}{V_T}\right) \cos(k\omega_0 t) \right]$$

Si tratta cioè di armoniche, a frequenza multipla di ω_0, di ampiezza $I_k\left(\frac{V_1}{V_T}\right)$ che è funzione di V_1 e della tensione termica V_T. Tra queste armoniche abbiamo evidentemente separato quella a frequenza 0 (la continua). Se mettiamo in evidenza proprio il termine in continua, otteniamo

$$i_C(t) = I_{CS} e^{\frac{V_C^Q}{V_T}} I_0 \left[1 + 2\sum_{k=1}^{\infty} \frac{I_k}{I_0} \cos(k\omega_0 t) \right]$$

dove, per semplicità, abbiamo omesso di indicare la dipendenza da V_1/V_T.

Il termine a fattore è pari chiaramente alla corrente media di collettore, la quale, per quanto detto in precedenza, vale I_Q:

$$i_C(t) = I_Q \left[1 + 2\sum_{k=1}^{\infty} \frac{I_k}{I_0} \cos(k\omega_0 t) \right]$$

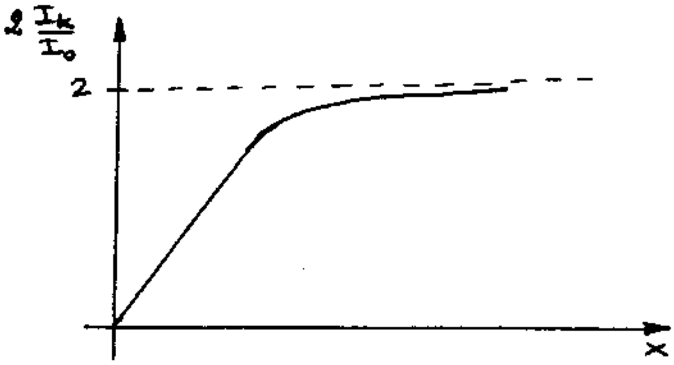
E' bene precisare che i passaggi analitici appena effettuati valgono sia per il circuito con polarizzazione imposta dal generatore di corrente I_Q sia per il circuito con la polarizzazione imposta dalla batteria V_B: la differenza tra i due casi, come già sottolineato in precedenza, è che nel circuito con la V_B, il valore medio I_Q non è noto, mentre lo è nell'altro circuito, dove è imposto dall'esterno[16].

[16] E' bene ricordare anche un'altra cosa: quando parliamo di valor medio di una forma d'onda, in teoria ci riferiamo alla media effettuata da t=-∞ a t=+∞. Nella pratica, invece, non potendo estendere la nostra osservazione da -∞ a +∞, il valor medio potrà essere calcolato solo integrando in un intervallo di tempo finito, per quanto grande. Sceglieremo perciò di integrare per un intervallo di tempo pari ad un numero sufficientemente elevato di periodi della sinusoide in ingresso.

E' evidente, a questo punto, che, per valutare l'ampiezza delle singole componenti armoniche, dobbiamo semplicemente valutare il rapporto $2\frac{I_k}{I_0}\left(\frac{V_1}{V_T}\right)$ tra la funzione di Bessel modificata di ordine k e quella di ordine 0, entrambe valutate per un valore dell'argomento pari a V_1/V_T (c'è poi da considerare che il suddetto rapporto va moltiplicato per I_Q, in base all'espressione prima ricavata) Per semplicità, nei discorsi seguenti poniamo x= V_1/V_T:

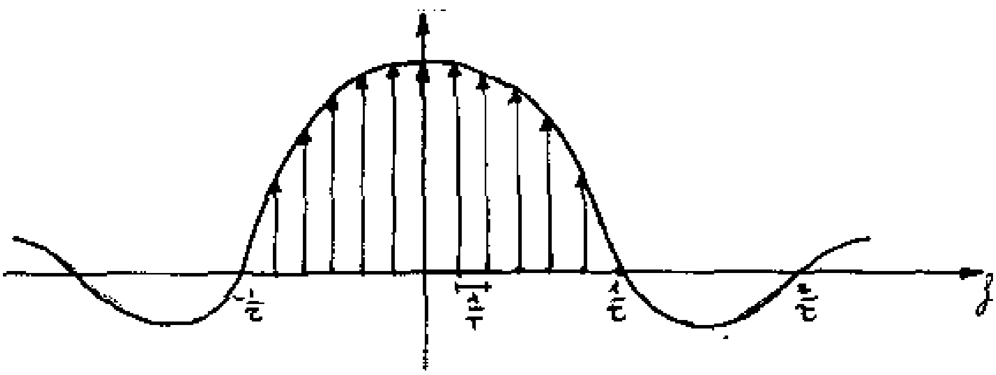
$$i_C(t) = I_Q + \sum_{k=1}^{\infty} 2I_Q \frac{I_k}{I_0}(x)\cos(k\omega_0 t)$$

Senza riportare l'espressione analitica delle funzioni di Bessel modificate, riportiamo direttamente l'andamento di $2\frac{I_k}{I_0}(x)$ in funzione di x:



La curva parte dal valore 0 con un andamento che è praticamente rettilineo. A partire da un certo valore di x (e quindi di V_1), l'andamento tende invece asintoticamente al valore 2, il che significa che I_k tende ad I_0.

Considerando il fattore moltiplicativo I_Q, troviamo dunque conferma di quanto detto in precedenza a livello qualitativo: per x>>1, cioè per V_1>>V_T, l'ampiezza della componente armonica di ordine k generico tende ad un valore $2I_Q$. In realtà, date le approssimazioni con cui si perviene alle funzioni di Bessel modificate, questo risultato è molto ben approssimato per la prima armonica (k=1), mentre lo è sempre meno per le armoniche di ordine superiore.



In generale, dunque, il discorso fatto per la prima armonica vale, con buona approssimazione, per tutte le armoniche vicine al massimo di $\sin(\pi f\tau)/\pi f\tau$. Come già visto in precedenza, quanto minore è τ, tanto più piatto è l'andamento del $\sin(\pi f\tau)/\pi f\tau$ e quindi tante più sono le armoniche con ampiezza pari pressappoco al doppio della continua. Nel caso ideale di impulsi matematici (cioè $\tau=0$), tutte le armoniche sarebbero di ampiezza $2I_Q$.

Il discorso fatto si presta ad un'altra considerazione importante: dal grafico di $2\frac{I_k}{I_0}(x)$ si deduce che l'ampiezza della generica armonica di $i_C(t)$ cresce con x fin quando x è sufficiente piccolo, cioè fin quando V_1 è sufficientemente piccolo. Quando invece x cresce, si giunge alla saturazione, per cui l'ampiezza della generica armonica tende ad un valore fisso $2I_Q$: questo valore è indipendente da V_1, in quanto dipende solo dalla polarizzazione. Quindi, *con il circuito che stiamo considerando (stadio a base comune con corrente di emettitore imposta con un generatore), l'uscita è una sinusoide sincrona con l'ingresso (nell'ipotesi che il carico risonante sia accordato alla frequenza w_0 del segnale ingresso), ma la sua ampiezza, da un certo punto in poi, non dipende più dall'ampiezza dell'ingresso:* allora, se vogliamo che l'ampiezza dell'uscita sul collettore sia linearmente legata all'ampiezza V_1 del segnale iniettato in ingresso, dobbiamo rendere piccolo il valore di V_1 stesso (i soliti 10mV massimi).

Questa considerazione ha una implicazione fondamentale: il circuito è comunque un amplificatore, in quanto può fornire in uscita una sinusoide ad un livello di potenza ben maggiore di quello con cui la stessa sinusoide entra in ingresso, ma questa amplificazione è utile solo se la sinusoide in ingresso non porta informazione legata alla sua ampiezza: infatti, se l'ampiezza dell'ingresso porta informazione (in quanto si tratta di una portante modulata di ampiezza), ma è tale da saturare il guadagno, l'uscita non contiene più tale informazione. Questo circuito, dunque, si presta tipicamente ad amplificare segnali passa-banda in cui l'ampiezza non porta informazione. Per esempio, può amplificare segnali modulati angolarmente (in fase o in frequenza): in questo caso, infatti, la legge di modulazione fa semplicemente spostare gli istanti in cui si trovano i picchi della sinusoide e questo si traduce in un uguale spostamento degli impulsi della $i_C(t)$; di conseguenza, la $i_C(t)$ non presenta più delle sinusoidi a frequenza fissa e in relazione armonica tra loro, ma delle componenti sinusoidali modulate angolarmente, sempre in relazione armonica tra loro.

Tornando ancora al diagramma di $2\frac{I_k}{I_0}(x)$, è evidente che l'amplificatore ha un guadagno variabile con l'ampiezza V_1 della tensione in ingresso: per V_1 piccola, il guadagno cresce linearmente, fino a diventare costante quando V_1 è talmente alto da far saturare il rapporto $2\frac{I_k}{I_0}(x)$.

Transconduttanza di prima armonica

A questo punto, anziché interessarci delle generica componente armonica della corrente di collettore, consideriamo solo la prima armonica, cioè quella alla stessa frequenza ω_0 della sinusoide in ingresso (uguale anche alla frequenza di risonanza del risonante usato come carico): per quanto visto, l'ampiezza di tale componente, per un valore generico di x (e cioè di V_1) è

$$I_{C1} = 2I_Q \frac{I_1}{I_0}(x)$$

In realtà, a voler essere rigorosi, questa non è la prima armonica della corrente di collettore, ma la prima armonica della corrente di emettitore, dato che, sempre a rigore, I_Q è la corrente media di emettitore. Per passare alla corrente di collettore, dovremmo semplicemente tener conto che $I_C=\alpha_F I_E$, dove $\alpha_F=\beta_F/(\beta_F+1)$. D'altra parte, se consideriamo un transistor con $\beta_F \cong 100$, è chiaro che $\alpha_F \cong 1$ e quindi che $I_C \cong I_E$.

Nota I_{C1}, si definisce **transconduttanza di prima armonica** il rapporto tra I_{C1} stessa e l’ampiezza V_1 della tensione forzante:

$$\boxed{G_{m1} = \frac{I_{C1}}{V_1}}$$

Dobbiamo fare una alcune osservazioni su questa definizione. In primo luogo, il parametro G_{m1}, nonostante la somiglianza nel nome, è qualcosa di ben diverso dalla *transconduttanza di piccolo segnale* (simbolo: g_{mQ}) che si definisce nel modello per piccoli segnali di un transistor: infatti, la g_{mQ} è definita dalla relazione $g_{mQ} = \frac{\partial I_C}{\partial V_{BE}}$, cioè come rapporto tra la variazione della corrente di collettore e la variazione della tensione base-emettitore che l’ha determinata; questa definizione è valida qualunque sia la forma d’onda applicata in ingresso al transistor, in quanto si ipotizza che esso venga sollecitato in modo da rendere comunque lecita la linearizzazione della caratteristica esponenziale I_C-V_{BE}. Al contrario, non avendo più l’ipotesi di comportamento lineare su cui basarci, abbiamo il problema per cui la corrente di segnale sul collettore e la tensione applicata hanno un andamento ben diverso nel tempo: la tensione applicata è sinusoidale a frequenza ω_0, mentre la corrente di segnale di collettore è solo periodica, ma con la prima armonica fondamentale a frequenza ω_0. Proprio per questo motivo, <u>si considera solo l’ampiezza della prima armonica, che viene confrontata con l’ampiezza della tensione forzante, che ha la stessa frequenza</u>.

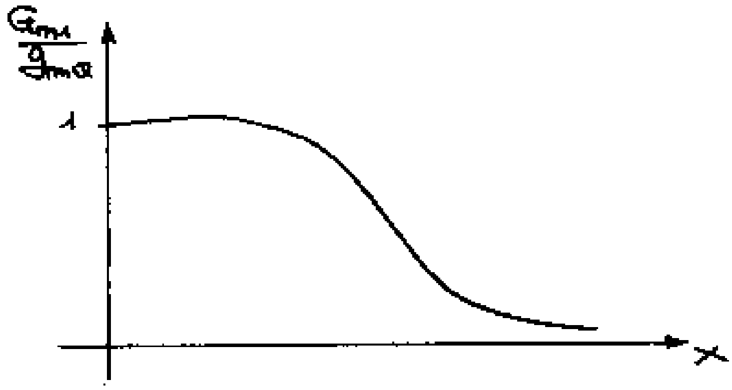
Da questa differenza sostanziale tra g_{mQ} e G_{m1} scaturisce una importante conseguenza: mentre la g_m dipende notoriamente solo dalle condizioni di polarizzazione (ricordiamo a tal proposito che $g_{mQ} = I_C^Q / V_T$ per un transistor bipolare), la Gm1 deve necessariamente dipendere anche dall’ampiezza V_1 della tensione forzante. Ce ne possiamo accorgere con qualche semplice passaggio matematico: infatti, ricordando che l’espressione generale della corrente di collettore era

$$i_C(t) = I_Q + \sum_{k=1}^{\infty} 2I_Q \frac{I_k}{I_0}(x)\cos(k\omega_0 t)$$

è evidente che la prima armonica è $i_{C1}(t) = 2I_Q \frac{I_1}{I_0}(x)\cos(\omega_0 t)$, per cui, sostituendo nell’espressione di G_{m1}, otteniamo che

$$G_{m1} = \frac{I_{C1}}{V_1} = \frac{2I_Q \frac{I_1(x)}{I_0(x)}}{V_1} = \frac{I_Q}{V_1} 2\frac{I_1(x)}{I_0(x)} = \frac{I_Q}{V_T}\frac{V_T}{V_1} 2\frac{I_1(x)}{I_0(x)} = g_{mQ}\frac{V_T}{V_1} 2\frac{I_1(x)}{I_0(x)} = g_{mQ} 2\frac{I_1(x)}{xI_0(x)}$$

La Gm1 è dunque data dal prodotto di g_{mQ} per il termine $2\frac{I_1(x)}{xI_0(x)}$, che varia con V_1. Inoltre, noi conosciamo l’andamento del termine $2\frac{I_1(x)}{I_0(x)}$ con x, per cui possiamo facilmente dedurre anche l’andamento di $2\frac{I_1(x)}{xI_0(x)}$ con x, ossia l’andamento di Gm1 con x. Tracciando, in particolare, l’andamento di G_{m1} normalizzato al valore di g_{mQ}, si ottiene quanto segue:



Si distinguono tipicamente 3 regioni in questo diagramma:

- quando x è piccolo, si può dimostrare che $\lim_{x\to 0} 2\frac{I_1(x)}{I_0(x)} = x$, da cui consegue che

$$\lim_{x\to 0}\frac{G_{m1}}{g_{mQ}} = \lim_{x\to 0} 2\frac{I_1(x)}{xI_0(x)} = 1$$

- quando, invece, x è grande, si può dimostrare che $\lim_{x\to\infty} 2\frac{I_1(x)}{I_0(x)} = 2$, da cui consegue anche che

$$\lim_{x\to\infty}\frac{G_{m1}}{g_{mQ}} = \lim_{x\to\infty} 2\frac{I_1(x)}{xI_0(x)} = 0$$

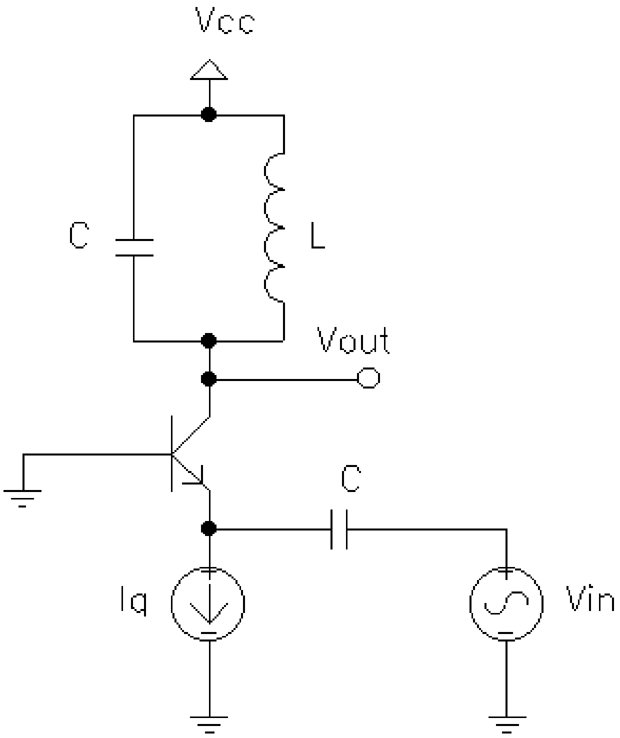
- nella regione intermedia, invece, G_{m1}/g_{mQ} decresce in modo non lineare con l'aumentare di x, ossia dell'ampiezza della tensione applicata. In questo intervallo, cioè, l'amplificatore ha un guadagno che decresce non linearmente all'aumentare della tensione applicata.

Questi risultati analitici sono congruenti con quanto ci aspettavamo: quando x è piccolo, è lecita la linearizzazione, per cui il guadagno G_{m1} viene a coincidere con quello g_{mQ} di piccolo segnale; all'aumentare di x, il guadagno scende rispetto a gmQ, in conseguenza del fatto che interessiamo una regione della caratteristica esponenziale che non può più essere approssimata con una retta; per x che cresce indefinitamente, infine, portiamo il transistor dalla saturazione all'interdizione, per cui praticamente non abbiamo alcun guadagno.

Dire che G_{m1}/g_{mQ} decresce a 0 non significa dire che, all'aumentare del segnale applicato, la tensione di uscita diminuisce. Significa invece che, a partire da un certo valore di V_1, l'ampiezza dell'uscita non varia più all'aumentare di V_1, per cui il guadagno tende asintoticamente a 0.

Tensione di uscita di prima armonica

Noto l'andamento della corrente di collettore in termini di armoniche, possiamo anche calcolare quanto vale la tensione di segnale sul risonante, cioè la tensione di uscita V_{out} che ci interessa prelevare dal circuito:



Se l'ingresso forzante è $V_{in}(t) = V_1 \cos(\omega_0 t)$, avendo detto che $G_{m1}(x) = \frac{I_{C1}}{V_1}$, deduciamo che la componente di prima armonica (a frequenza ω_0) della corrente di collettore è

$$i_{C1}(t) = G_{m1}(x) V_1 \cos(\omega_0 t)$$

Questa è dunque la corrente che alimenta il risonante LC posto sul collettore del transistor. Tale risonante è accordato proprio alla frequenza ω_0 (cioè risulta $\sqrt{LC} = 1/\omega_0$), per cui la tensione che genera, in risposta alla corrente $i_{C1}(t)$, è il prodotto di tale corrente per il valore (resistivo) della sua impedenza in corrispondenza di ω_0. Quanto vale questa impedenza? Se il risonante fosse ideale (privo cioè di dissipazione) e non ci fossero effetti di carico provenienti da altri stadi in cascata, l'impedenza varrebbe ∞. Al contrario, date le perdite nell'induttore, data la parte resistiva dell'impedenza di uscita del collettore e dati eventuali effetti di carico degli stadi successivi, il risonante presenterà, alla frequenza ω_0, una certa resistenza R_{tot}, per cui

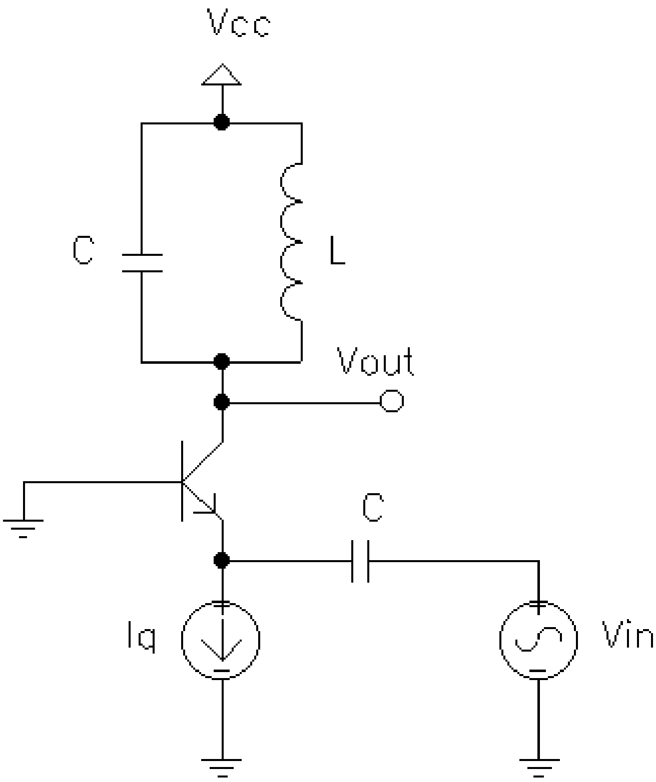
$$V_{out,1}(t) = R_{tot} i_{C1}(t) = G_{m1}(x) R_{tot} V_1 \cos(\omega_0 t)$$

Abbiamo indicato questa tensione con $V_{out,1}(t)$ per specificare che è la tensione relativa solo alla prima armonica della corrente di collettore. Oltre a questo, c'è anche da considerare un'altra cosa: scrivendo che $i_{C1}(t) = G_{m1}(x) V_1 \cos(\omega_0 t)$, <u>abbiamo implicitamente assunto che la corrente di collettore di prima armonica sia in fase con l'ingresso</u>. *Questo dipende dal fatto che <u>il transistor, pur essendo non lineare, è senza memoria</u>* (dato che abbiamo scelto di trascurare gli effetti derivanti da C_π e C_μ): essendo senza memoria, gli impulsi della corrente $i_C(t)$ sono sganciati esattamente sui picchi della tensione forzante, senza sfasamenti in avanti o indietro, il che comporta che $i_{C1}(t)$ sia in fase con l'ingresso e quindi che lo sia anche $V_{out,1}(t)$.

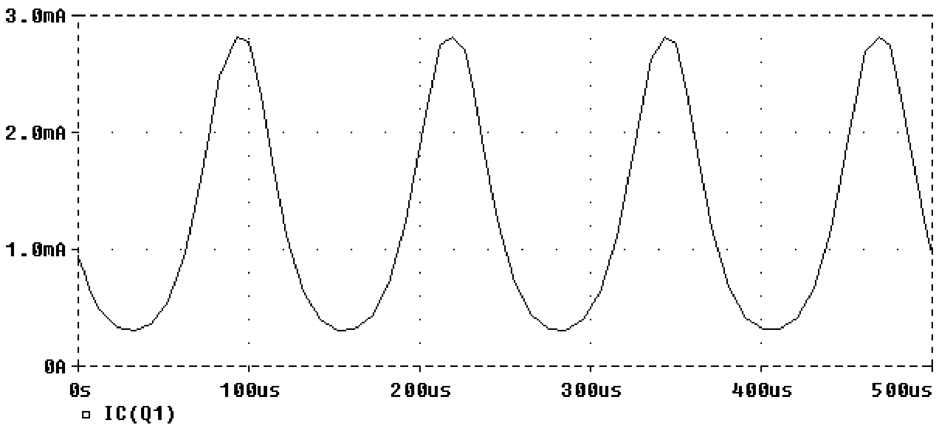
Tutto questo vale, ovviamente, nell'ipotesi che il risonante sia accordato esattamente alla frequenza ω_0: in caso contrario, l'impedenza vista da $i_{C1}(t)$ non sarebbe puramente resistiva, ma avrebbe anche un comportamento reattivo, che darebbe origine a sfasamenti.

Comportamento del condensatore di blocco: effetto di clamping

Riprendiamo ancora l'amplificatore a base comune con risonante LC sul collettore, con polarizzazione imposta dal generatore di corrente I_Q e con il segnale forzante $v_{in}(t)$ portato sull'emettitore del transistor attraverso il condensatore di blocco C:

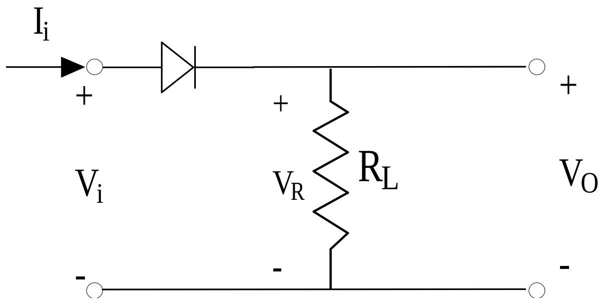


Siamo interessati, in particolare, al comportamento del condensatore di blocco: a questo proposito, abbiamo già osservato che esso, all'aumentare dell'ampiezza V_1 della tensione sinusoidale applicata, fa si che il transistor conduca per frazioni sempre più piccole del periodo della sinusoide. Abbiamo infatti visto che, all'aumentare di V_1, la corrente di collettore tende a diventare una successione di impulsi equispaziati di $T=2\pi/\omega_0$, dove ω_0 è sia la frequenza della sinusoide applicata in ingresso sia la frequenza di risonante del carico LC sul collettore:

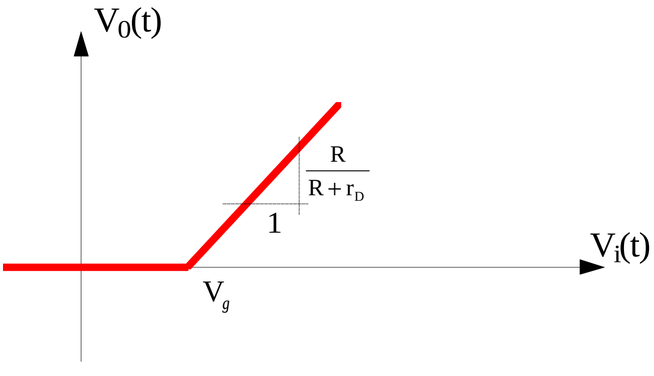


I picchi della corrente di collettore si hanno esattamente in corrispondenza dei picchi positivi della sinusoide in ingresso, per cui, da questo punto di vista, l'amplificatore ha l'effetto di amplificare solo i picchi positivi della suddetta sinusoide. Questo comportamento va sotto il nome di **clamping** e si presenta molto spesso.

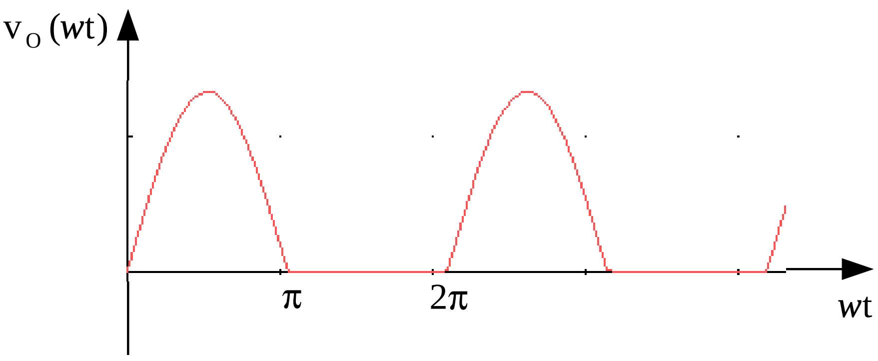
Consideriamo ad esempio un semplice circuito per il raddrizzamento a singola semionda di una forma d'onda sinusoidale:



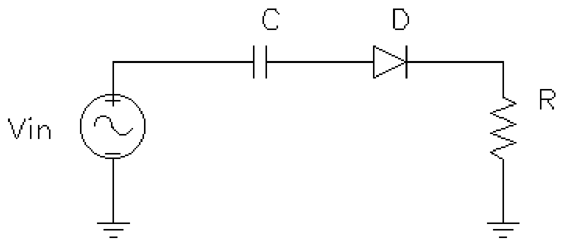
Questo circuito, sfruttando il fatto che il diodo non conduce quando la tensione di ingresso è negativa (in particolare, esso non conduce fin quando la tensione di ingresso non supera il valore V_γ di accensione) ha l'effetto di eliminare i tratti negativi della tensione in ingresso, in accordo ad una caratteristica di trasferimento in tensione fatta nel modo seguente:



Considerando, ad esempio, una tensione sinusoidale in ingresso, l'uscita sarà uguale all'ingresso, ma senza le semionde negative:

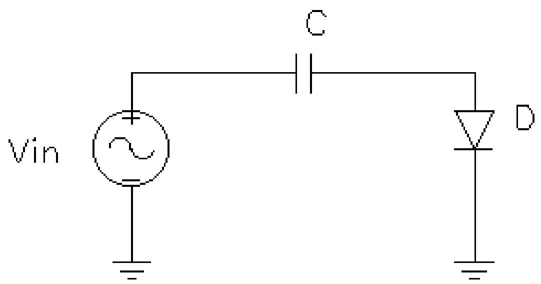


Adesso supponiamo che la tensione di ingresso considerata sia la tensione di segnale applicata, tramite il condensatore di blocco, sull'emettitore dello stadio a base comune prima riportato. Possiamo schematizzare la situazione con il seguente circuito:

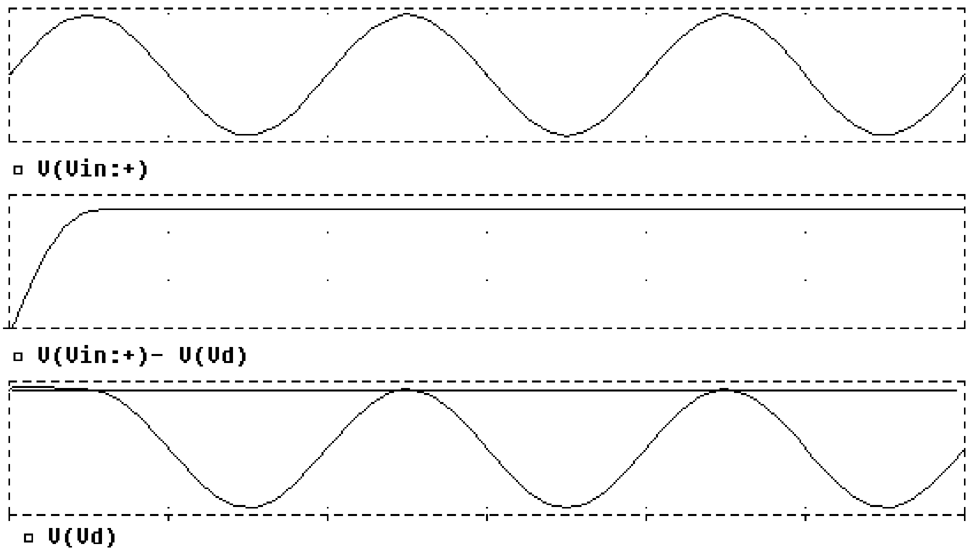


Questo circuito differisce dal raddrizzatore per la presenza del condensatore C, il quale serve ad evitare di caricare la sorgente con la corrente continua dovuta alla polarizzazione. In pratica, il condensatore fa si che la corrente nella maglia abbia valor medio nullo.

Eliminiamo per semplicità la resistenza R, riconducendoci così ad un circuito con un elemento non lineare (il diodo) in serie ad un elemento di memoria (il condensatore):



Supponiamo che il condensatore sia inizialmente scarico (per esempio, supponiamo che il circuito sia stato lasciato a se stante da tempo immemore, con $v_{in}(t)=0$ per $t<0$). All'istante $t=0$, il generatore applica una tensione sinusoidale. Ritenendo il diodo ideale[17], il comportamento del circuito è il seguente:



Inizialmente, con la tensione applicata positiva, il diodo si comporta da cortocircuito, per cui lascia passare corrente e fa si che la tensione ai capi del condensatore segua perfettamente l'ingresso.

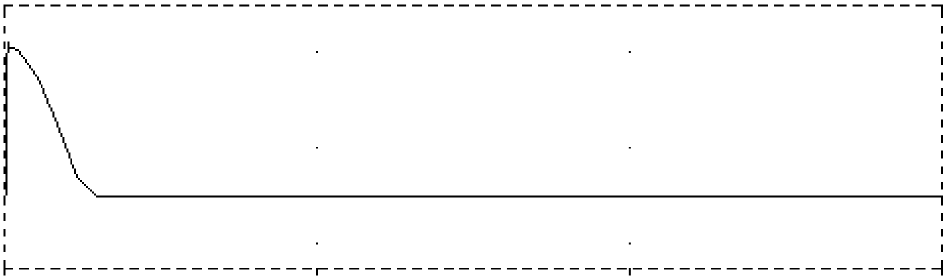
Nel momento in cui l'ingresso raggiunge il suo valore massimo E, anche la tensione ai capi del condensatore vale E; quando, allora, l'ingresso prende a scendere, il condensatore rimane alla

[17] ossia ritenendo che si comporti da cortocircuito quando è in conduzione (tensione positiva applicata ai suoi capi) e da circuito aperto quando è interdetto (tensione negativa ai suoi capi)

tensione E, per cui la tensione ai capi del diodo diventa negativa: il diodo allora interrompe il flusso di corrente, lasciando una tensione E ai capi della condensatore[18].

Quando la tensione di ingresso raggiunge il suo valore massimo negativo (-E), la tensione ai capi del condensatore è ancora E, per cui la tensione ai capi del diodo è -2E.

Ad ogni modo, dato che la tensione sul condensatore non scende più al di sotto di E, il diodo non conduce più, per cui non circola più alcuna corrente nel circuito, come indicato dal grafico seguente:

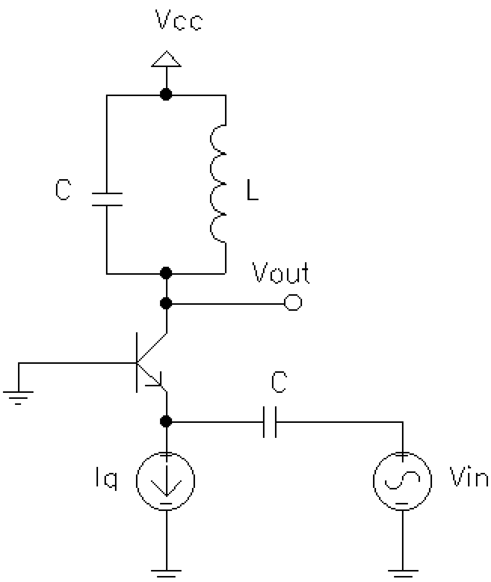


Come si nota, la corrente scorre solo inizialmente, quanto basta per caricare il condensatore alla tensione E, dopo di che il diodo è perennemente interdetto.

Cosa ci interessa osservare nel comportamento di questo circuito? Ci interessa il fatto per cui la tensione ai capi del diodo è uguale a quella forzante, ma con valor medio diverso da zero: infatti, mentre l'ingresso è una sinusoide di ampiezza E, ma con valor medio 0, la tensione ai capi del diodo è ancora una sinusoide di ampiezza E, isofrequenziale con l'ingresso, ma di valor medio pari proprio ad E. Abbiamo cioè recuperato la continua, il che va sotto il nome di **clamping**.

GENERATORE FORZANTE NON IDEALE

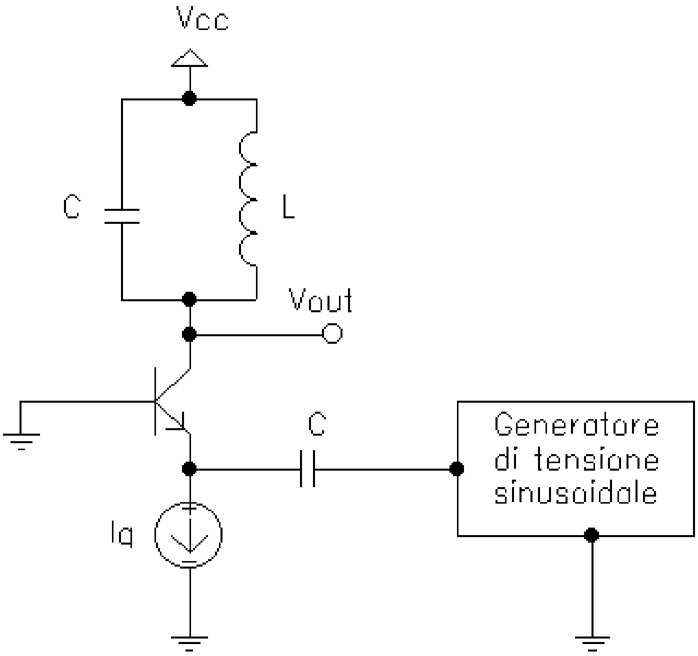
Riprendiamo ancora una volta l'amplificatore a base comune con risonante LC sul collettore, con polarizzazione imposta dal generatore di corrente I_Q e con il segnale forzante $v_{in}(t)$ portato sull'emettitore del transistor attraverso il condensatore di blocco C:



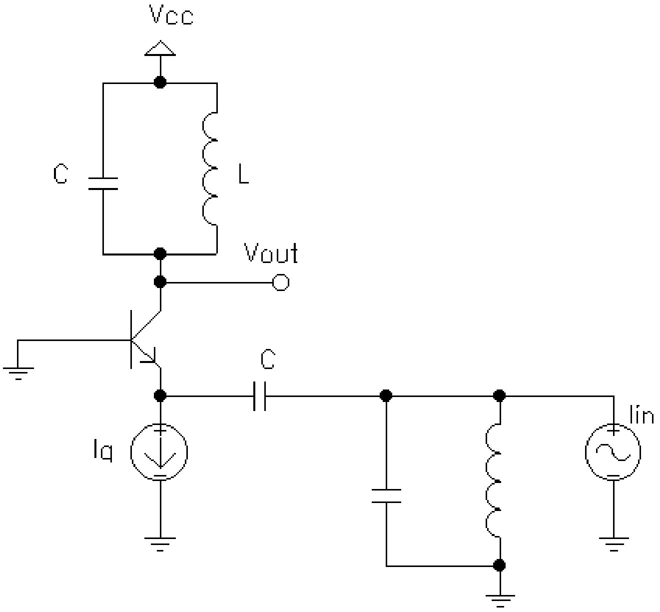
[18] In modo alternativo, si può vedere la cosa dicendo che il diodo resta in conduzione fino all'istante in cui la tensione forzante determina l'inversione del segno della corrente: questo avviene quando la tensione forzante raggiunge il primo picco e poi prende a scendere.

In tutti i discorsi fatti fino ad ora, abbiamo sempre considerato ideale il generatore di tensione, cioè con resistenza serie nulla. In realtà, un generatore ha sempre una resistenza interna non nulla, per quanto bassa: in generale, affinché lo si possa considerare un buon generatore di tensione, è sufficiente che esso presenti una impedenza interna bassa rispetto a quella di carico. Nel nostro caso, supponendo che, alle frequenze di interesse, il condensatore di blocco C sia un cortocircuito e supponendo anche trascurabili le capacità intrinseche del transistor, la resistenza di ingresso vista dal generatore è quella vista guardando dentro l'emettitore del transistor, ossia la nota $1/g_m$. Questa resistenza è dell'ordine di 10 Ω, il che significa che il generatore di tensione dovrebbe avere una resistenza interna ancora più bassa. Per i normali generatori di tensione, questo non è possibile, per cui dobbiamo necessariamente pensare a qualcos'altro per il pilotaggio del circuito.

In linea di principio, possiamo considerare un qualsiasi dispositivo che sia in grado di fornire ai suoi capi una tensione sinusoidale:



Se il dispositivo fornisce una tensione sinusoidale, presentando al contempo una resistenza serie praticamente nulla (deve essere trascurabile rispetto a 10Ω) il circuito si comporterà comunque nel modo finora descritto. Intuitivamente, il dispositivo che fa al caso nostro è ancora una volta un risonante LC, alimentato da una corrente sinusoidale:



Un risonante, in base a quanto già detto in precedenza, è tale che, anche se attraversato da una corrente (periodica) ricca di armoniche, localizza ai suoi capi una tensione praticamente sinusoidale, alla frequenza di risonanza. Questo perché, se il fattore di merito Q è sufficientemente elevato, il modulo dell'impedenza che il risonante presenta alle armoniche della frequenza di risonanza è praticamente infinito: allora, anche se pilotato da più armoniche, esso dà una tensione di valore sensibile solo in corrispondenza della frequenza di risonanza ω_0, mentre da una tensione di valore molto più basso, praticamente trascurabile, in corrispondenza delle armoniche di ω_0.

In base a questo discorso, per sapere se il nostro risonante funzioni bene come generatore di tensione sinusoidale, ossia rispetti i vincoli appena citati, dobbiamo calcolare il suo fattore di merito Q.

Il fattore di merito di un risonante LC parallelo ha espressione

$$Q = R\sqrt{\frac{C}{L}} = R\frac{1}{L}\sqrt{LC} = \frac{R}{L\omega_0}$$

Il termine R che compare in questa espressione tiene conto di tutti i fenomeni che contribuiscono a dissipare quella energia che, in un LC ideale, sarebbe continuamente scambiata tra campo magnetico nell'induttore e campo elettrico nel condensatore[19]. Nel nostro caso, tali perdite sono dovute a vari fattori:

- sicuramente, <u>ci sono perdite nell'induttore</u>: infatti, una parte della corrente che attraversa un induttore reale finisce sotto forma di riscaldamento e viene dissipata per effetto Joule;
- in secondo luogo, <u>non sarà ideale nemmeno il generatore di corrente</u>, che presenterà una resistenza di Norton piuttosto alta: ad ogni modo, questo contributo è generalmente trascurabile, in quanto si tratta di una resistenza molto elevata (in generale, quella di uscita di un transistor in uno specchio di corrente) in parallelo a resistenze decisamente più piccole, che quindi hanno il sopravvento;
- in terzo luogo, c'è la <u>resistenza di ingresso dell'emettitore del transistor</u>: in regime di piccolo segnale, tale resistenza interna è $r_\pi/(\beta+1)\cong 1/g_m$.

Quindi, nel migliore dei casi, la resistenza R da considerare per il calcolo di Q è il parallelo della resistenza r_L dell'induttore e della resistenza $1/g_m$:

$$Q = \left(r_L \,//\, \frac{1}{g_m}\right)\sqrt{\frac{C}{L}}$$

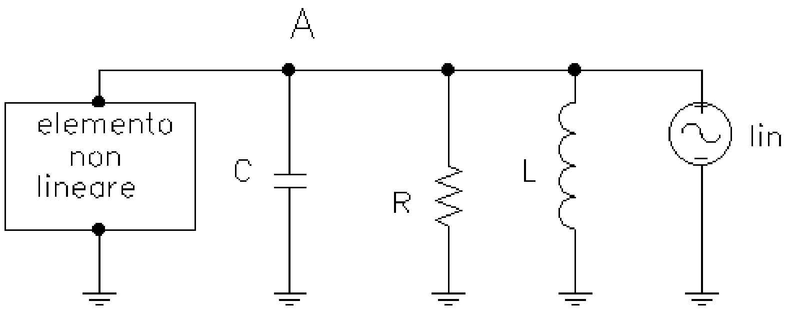
Si nota immediatamente che, se non considerassimo il termine $1/g_m$, il fattore di merito risulterebbe quello proprio del risonante. Aggiungendo invece l'impedenza del carico in parallelo al risonante, allora si parla di **fattore di merito a carico**, indicato con Q_C:

$$Q_C = \left(r_L \,//\, \frac{1}{g_m}\right)\sqrt{\frac{C}{L}}$$

[19] Ricordiamo che un risonante LC ideale funziona da oscillatore perfetto in quanto c'è un continuo scambio di energia tra induttore e condensatore. In presenza, invece, di effetti dissipativi, il risonante si dice **smorzato**, in quanto, proprio a causa della dissipazione, esso può mantenere l'oscillazione solo se alimentato dall'esterno, al contrario del risonante non smorzato (appunto l' LC ideale), dove non serve alcuna sollecitazione esterna.

A noi interessa che questo fattore di merito sia molto maggiore di 1, in modo da autorizzarci a ritenere trascurabili i termini di distorsione dovuti alle armoniche di tensione.

La formula appena ricavata vale però in regime di piccolo segnale. In presenza, invece, di sollecitazioni più ampie, pur continuando a considerare le perdite nell'induttore e la dissipazione dovuta al transistor, dobbiamo anche tener conto che quest'ultimo non lavora più in modo lineare, bensì presenta un comportamento fortemente non lineare. Dobbiamo allora fare riferimento ad un modello generale del tipo seguente:



Nella resistenza R abbiamo incluso, questa volta, solo le perdite dovute all'induttore, mentre nel blocco non lineare abbiamo incluso tutti gli effetti dovuti al transistor. Il pilotaggio del circuito avviene mediante una corrente sinusoidale del tipo $I_{in}(t) = K\cos(\omega_0 t)$.

E' evidente che questo circuito verrà usato solo se si comporta in modo per noi utile, ossia se, a fronte della corrente $I_{in}(t) = K\cos(\omega_0 t)$, riesce a produrre, al nodo indicato con A, una tensione sinusoidale. Perché questo sia possibile, deve succedere che lo smorzamento complessivo del risonante, dovuto alle perdite negli elementi costituenti (C ma soprattutto L) e alla dissipazione nel carico non lineare, sia piccolo: in questo modo, infatti, possiamo ragionevolmente affermare che la risposta alla corrente $K\cos(\omega_0 t)$ è ancora approssimativamente una sinusoide a frequenza ω_0.

Facciamo allora l'ipotesi che il fattore di merito a carico Q_C del risonante sia molto maggiore di 1:

$$Q_C = R_{tot}\sqrt{\frac{C}{L}} >> 1$$

Sotto questa ipotesi, possiamo affermare che la tensione al nodo A è, con ottima approssimazione, una tensione sinusoidale, nonostante la presenza del carico non lineare: sarà perciò una tensione del tipo $V_1\cos(\omega_0 t)$.

A questo punto, dobbiamo verificare che l'ipotesi fatta sia vera, ossia che effettivamente risulti $Q_C>>1$: per fare questa verifica, dobbiamo calcolare tutti gli effetti di dissipazione di potenza (quantificabili tramite una resistenza R_{tot}) del circuito alla frequenza di risonanza ω_0.

Essendo ω_0 la frequenza di risonanza, in corrispondenza di questa frequenza il risonante ha comportamento puramente resistivo (L e C si compensano a vicenda), quantificato dalla resistenza R, per cui la componente della corrente di prima armonica nel risonante è

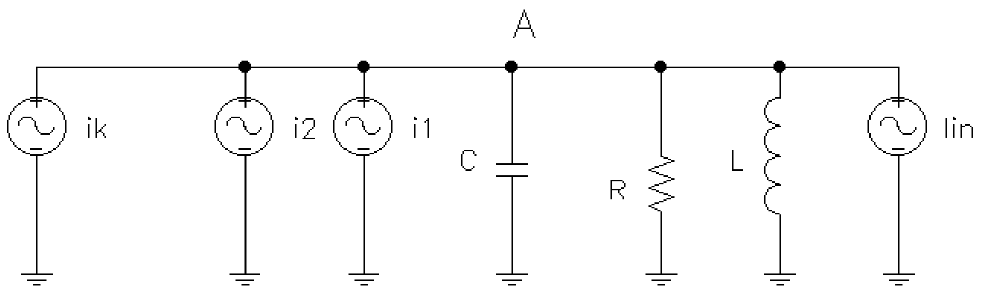
$$i_{R1}(t) = \frac{V_1}{R}\cos(\omega_0 t)$$

D'altra parte, la tensione sinusoidale $V_1\cos(\omega_0 t)$, applicata ai morsetti dell'elemento non lineare, genera l'assorbimento di una corrente i(t) che non è sinusoidale, ma, visto che l'elemento è senza memoria, certamente periodica e in fase con la tensione forzante. Possiamo allora sviluppare in serie di Fourier la corrente i(t), rappresentandola come somma di armoniche:

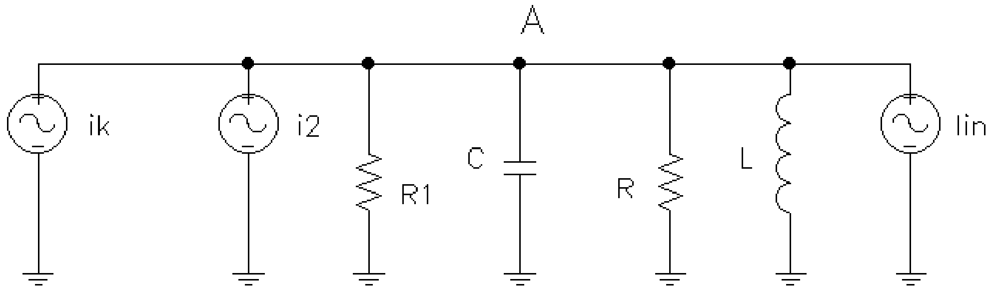
$$i(t) = \sum_{k=1}^{\infty} I_k \cos(k\omega_0 t)$$

Le ampiezze I_k di queste armoniche dipendono sia dalla caratteristica non lineare del dispositivo sia anche dall'ampiezza V_1 della tensione applicata. Il motivo della dipendenza da V_1 è abbastanza intuitivo: se V_1 fosse molto piccola, sarebbe lecita la solita approssimazione di linearità (cioè il fatto di approssimare la caratteristica I-V del dispositivo tramite la tangente nel punto di lavoro), da cui scaturirebbe che l'unica componente non nulla della corrente sarebbe quella della prima armonica (cioè a frequenza ω_0). Al contrario, all'aumentare di V_1 l'approssimazione di linearità diventa sempre meno lecita, da cui appunto la dipendenza di I_k da V_1.

Siamo dunque pervenuti alla situazione in cui l'elemento non lineare è sollecitato ai suoi morsetti dalla tensione $V_1\cos(\omega_0 t)$ e risponde ad essa con infinite armoniche di corrente. Dobbiamo allora chiederci se e quali armoniche di corrente danno origine ad una dissipazione di potenza nel carico non lineare, in quanto ricordiamo ancora che il nostro scopo è quello di calcolare tutti i fattori di dissipazione nel circuito in esame. A tale scopo, possiamo considerare il comportamento del carico non lineare semplicemente ai suoi morsetti: applicando una tensione $V_1\cos(\omega_0 t)$, il carico risponde assorbendo infinite armoniche di corrente. Ciò significa che possiamo sicuramente modellare il dispositivo non lineare con infiniti generatori di corrente in parallelo, ciascuno dei quali produce la propria armonica di corrente:

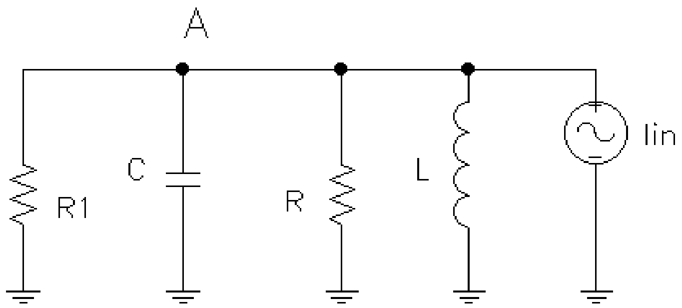


Questi generatori in parallelo forniscono dunque le varie armoniche di corrente. Tra queste armoniche, però, l'unica che da dissipazione di potenza è la prima, ossia $I_1\cos(\omega_0 t)$: infatti, un generatore di corrente che dà una corrente sinusoidale $I_1\cos(\omega_0 t)$ e ai capi del quale è applicata una tensione sinusoidale $V_1\cos(\omega_0 t)$ alla stessa frequenza è del tutto equivalente ad una resistenza di valore R_1 tale che $V_1 = R_1 I_1$:



Questo discorso non vale per gli altri generatori, i quali danno una corrente sempre sinusoidale, ma a frequenza diversa dalla tensione ai loro capi, per cui essi rappresentano dei veri e propri generatori di corrente, i quali non dissipano potenza. Quindi, anche se gli altri termini di corrente non sono nulli, essi non danno luogo a dissipazione di potenza.

In definitiva, *ai fini della valutazione della dissipazione di potenza che avviene all'interno del carico non lineare, possiamo limitarci a considerare la componente di corrente di prima armonica, ritenendola responsabile di una dissipazione di potenza pari a quella di una resistenza di valore* $R_1 = V_1 / I_1$ *sottoposta ad una tensione* $V_1 \cos(\omega_0 t)$*:*



Quindi, in modo del tutto analogo a quanto fatto per la transconduttanza di prima armonica, definiamo anche in questo caso una **conduttanza equivalente di prima armonica**, che rappresenta semplicemente il rapporto tra il termine di prima armonica della corrente (I_1), che circola nel carico non lineare in conseguenza dell'applicazione di una tensione sinusoidale ai suoi morsetti, e il valore di picco (V_1) di tale tensione sinusoidale:

$$G_1 = \frac{1}{R_1} = \frac{I_1}{V_1}$$

Applicando adesso la LKC, deduciamo che la corrente nella resistenza R vale

$$i_R(t) = K\cos(\omega_0 t) - I_1 \cos(\omega_0 t)$$

per cui la tensione al nodo A vale[20]

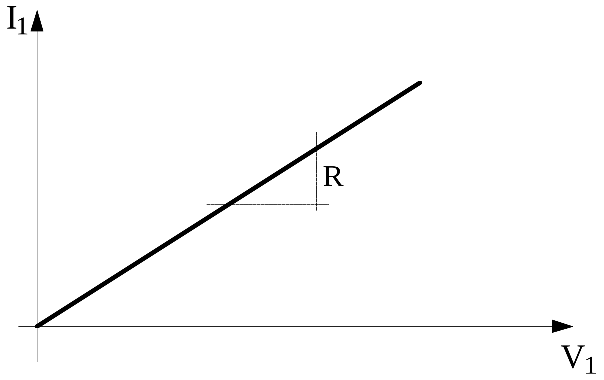
$$V_1 \cos(\omega_0 t) = V_A(t) = Ri_R(t) = R(K - I_1)\cos(\omega_0 t)$$

L'ampiezza K della corrente forzante è nota, per cui sono da determinare V_1 (tensione ai capi del carico non lineare) ed I_1 (ampiezza della corrente di prima armonica assorbita dal carico non lineare). Dobbiamo cioè determinare la coppia corrente-tensione ai morsetti dell'elemento lineare. A tal fine, ci basta intersecare graficamente la caratteristica tensione-corrente dell'elemento non lineare con quella del resto del circuito. Quale caratteristica dobbiamo considerare? Ci ricordiamo un concetto fondamentale dell'Elettrotecnica: quando c'è da trovare il punto di lavoro di un circuito con 1 solo blocco non lineare, si trovano le caratteristiche I-V statiche del blocco non lineare e di quello lineare[21] e le si interseca. Nel nostro caso, invece, noi stiamo considerando una situazione diversa: non siamo infatti un regime stazionario, ma in regime sinusoidale, per cui non sono le caratteristiche statiche che ci interessano, ma quelle dinamiche: in particolare, avendo stabilito che le componenti armoniche di correnti diverse da quella ad ω_0 non diano alcun contributo significativo alla tensione e avendo quindi detto che tale tensione è sinusoidale a frequenza ω_0, <u>ci interessa solo il legame tra l'ampiezza I_1 della prima armonica della corrente e l'ampiezza V_1 della tensione</u>.

[20] Ricordiamo che, alla frequenza di risonanza, induttanza e capacità si annullano a vicenda (le correnti sono uguali ed opposte), per cui è come se non ci fossero).

[21] La caratteristica del blocco lineare, in base al teorema di Thevenin, sarà sempre una retta di pendenza negativa, le cui intercette sono la tensione di Thevenin V_{TH} (asse delle ascisse) e la corrente di cortocircuito V_{TH}/R_{TH} (asse delle ordinate).

Per quanto riguarda la parte lineare del circuito, la caratteristica è immediata, in quanto sappiamo che un risonante RLC si comporta in modo puramente resistivo quando sollecitato a frequenza ω_0:



La cosa si fa invece più complicata per il carico non lineare, del quale, in generale, potremo conoscere, in partenza, solo il legame istantaneo i-v tra corrente e tensione. Ciò significa che, in generale, la caratteristica I_1-V_1 relativo all'elemento non lineare va ricavata applicando una tensione sinusoidali di ampiezza V_1 crescente e misurando la componente armonica fondamentale della corrente conseguentemente assorbita.

Nel caso che noi stiamo considerando, questo passaggio è stato già fatto, in quanto il carico non lineare in esame è la giunzione base-emettitore del transistor: sappiamo che la corrente di emettitore, in presenza di una tensione forzante sinusoidale di ampiezza V_1 e frequenza ω_0, ha espressione

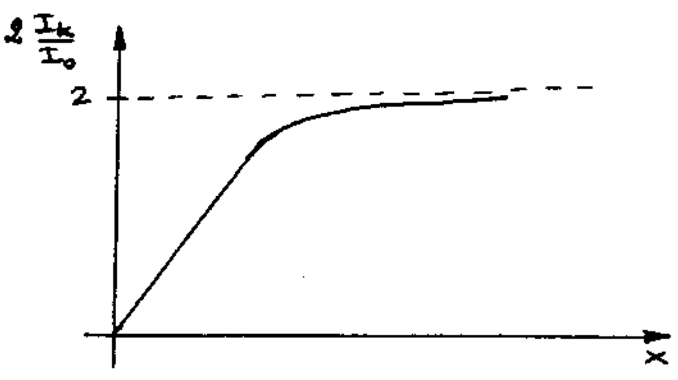
$$i_E(t) = I_Q + \sum_{k=1}^{\infty} 2I_Q \frac{I_k}{I_0}(x)\cos(k\omega_0 t)$$

dove ricordiamo che $x=V_1/V_T$.

Da qui deduciamo che la prima armonica della corrente assorbita è

$$i_{E1}(t) = 2I_Q \frac{I_1(x)}{I_0(x)}\cos(\omega_0 t)$$

Questa relazione lega dunque l'ampiezza $2I_Q \frac{I_1(x)}{I_0(x)}$ della prima armonica di corrente assorbita all'ampiezza V_1 della tensione forzante. Conosciamo anche l'andamento grafico di questo legame:

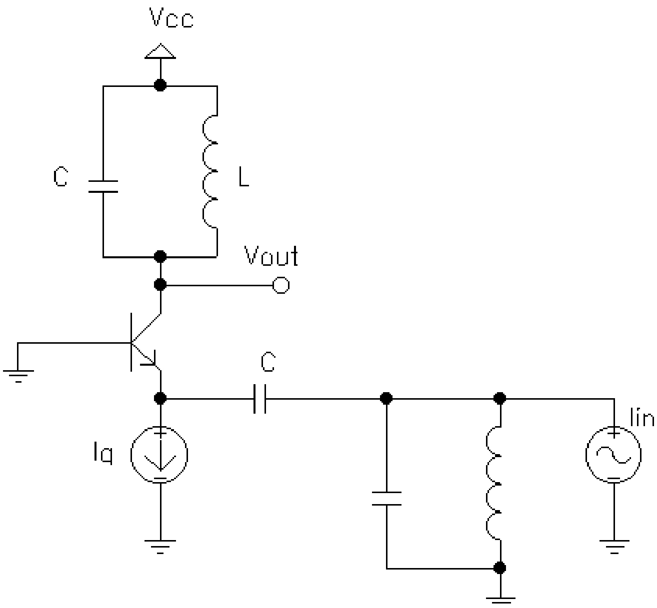


Siamo a questo punto in grado di intersecare graficamente le caratteristiche I-V del risonante e del carico non lineare, il che ci consente di ricavare l'ampiezza sia della corrente di prima armonica effettivamente assorbita dal carico sia della tensione ai capi del risonante.

Indicata brevemente con $\left(V_1^L, I_1^L\right)$ questa coppia tensione-corrente, abbiamo praticamente completato il nostro compito (in quanto sappiamo tutto quello che succede tra circuito di pilotaggio e carico non lineare), salvo un ultimo passaggio: dobbiamo tornare indietro e verificare che, con tale coppia $\left(V_1^L, I_1^L\right)$, risulti effettivamente Q_C>>1: se questo accade, le varie ipotesi fatte sono tutte ragionevoli ed i risultati congruenti, altrimenti dovremmo seguire altre strade.

Calcolo della resistenza di emettitore

Torniamo ancora al seguente circuito:



Siamo giunti ad esso allo scopo di imporre una tensione sinusoidale (con ottima approssimazione) tra l'emettitore del transistor e la massa. Nei paragrafi precedenti, abbiamo inoltre studiato come si comporta il risonante in ingresso data la presenza di un carico non lineare, che, in questo caso, è semplicemente la giunzione base-emettitore del transistor: abbiamo visto che il risonante continua a svolgere i suoi compiti (cioè continua a generare ai suoi capi una tensione praticamente sinusoidale nonostante il carico assorba una corrente con notevole contenuto armonico) solo nell'ipotesi che il fattore di merito complessivo Q_C sia sufficientemente maggiore di 1:

$$Q_C = R_{tot}\sqrt{\frac{C}{L}} = \frac{1}{G_{tot}}\sqrt{\frac{C}{L}}$$

Questo fattore di merito Q_C tiene conto di tutti gli effetti dissipativi presenti nel circuito, che contribuiscono allo smorzamento del risonante: tali termini sono racchiusi nella resistenza R_{tot}, che è il parallelo della resistenza R dell'induttore e della resistenza equivalente presentata dal carico (intesa come rapporto tra l'ampiezza V_1 della tensione sinusoidale applicata e l'ampiezza I_1 della corrente di prima armonica assorbita):

$$Q_C = (R//R_1)\sqrt{\frac{C}{L}} = \frac{1}{G+G_1}\sqrt{\frac{C}{L}} = \frac{1}{G+\frac{I_1}{V_1}}\sqrt{\frac{C}{L}}$$

Il nostro scopo è allora adesso quello di ricavare il valore di G_1. I calcoli sono stati già fatti in precedenza, in quanto abbiamo trovato che l'espressione della corrente di emettitore del transistor, in presenza di una tensione forzante sinusoidale di ampiezza V_1 e frequenza ω_0, ha espressione

$$i_E(t) = I_Q + \sum_{k=1}^{\infty} 2I_Q \frac{I_k}{I_0}(x)\cos(k\omega_0 t)$$

dove ricordiamo che $x=V_1/V_T$.

Da qui deduciamo che la prima armonica della corrente assorbita è

$$i_{E1}(t) = 2I_Q \frac{I_1(x)}{I_0(x)}\cos(\omega_0 t)$$

Possiamo allora scrivere che la conduttanza equivalente è

$$G_1 = \frac{I_{E1}}{V_1} = \frac{2I_Q \dfrac{I_1(x)}{I_0(x)}}{V_1}$$

dove ricordiamo che I_Q è la corrente di polarizzazione imposta da un apposito generatore di corrente.

Questa formula è del tutto analoga a quella trovata, a suo tempo, per lo stesso circuito, per la corrente di collettore: avevamo infatti definito, in quel caso, una transconduttanza di prima armonica mediante la relazione

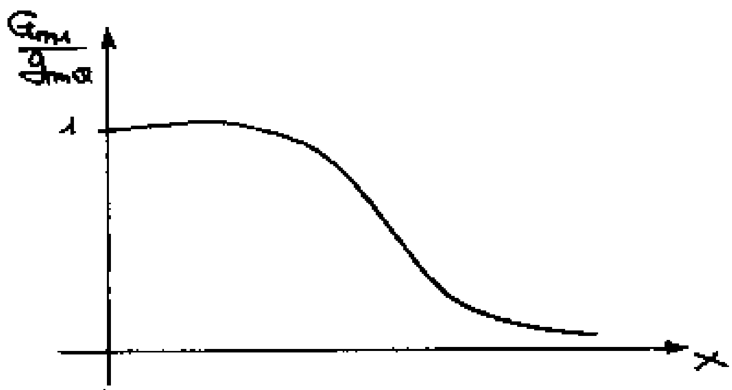
$$G_{m1} = \frac{I_{C1}}{V_1} = g_{mQ} 2 \frac{I_1(x)}{xI_0(x)}$$

E' intuitivo rendersi conto del fatto che G_1 e G_{m1} sono legate tra di loro: infatti, sappiamo che, fin quando il transistor è in zona attiva diretta oppure in interdizione, la corrente di emettitore e la corrente di collettore sono legate dalla relazione $I_C=\alpha_F I_E$, da cui deduciamo quindi che

$$G_1 = \frac{I_{E1}}{V_1} = \frac{I_{C1}/\alpha_F}{V_1} = \frac{1}{\alpha_F}G_{m1} = \frac{1}{\alpha_F}\left(g_{mQ} 2\frac{I_1(x)}{xI_0(x)}\right) = \frac{g_{mQ}}{\alpha_F} 2\frac{I_1(x)}{xI_0(x)}$$

Il termine $\dfrac{g_{mQ}}{\alpha_F}$ non è altro che la transconduttanza (incrementale) dell'emettitore del transistor.

Ad ogni modo, ricordando il diagramma



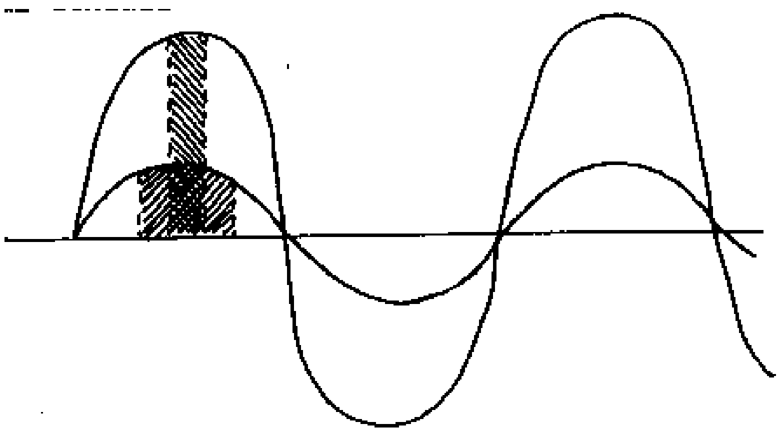
deduciamo quanto segue:

- quando x tende a 0 (cioè per piccoli segnali), si ottiene $G_1 \cong \dfrac{g_{mQ}}{\alpha_F}$, come ci si aspetta che sia;
- quando, invece, x tende ad infinito, ossia quando l'ampiezza V_1 della tensione di pilotaggio aumenta, la conduttanza equivalente G_1 presentata dall'emettitore del transistor tende a diminuire. In termini di resistenza, possiamo cioè affermare che, all'aumentare di V_1, aumenta la resistenza $R_1=1/G_1$ da considerare in parallelo alla resistenza R del risonante vero e proprio.

Mettere in parallelo ad un risonante una resistenza il cui valore aumenta significa ridurre lo smorzamento del risonante, in quanto, a parità di tensione applicata, diminuisce la potenza dissipata (pari al rapporto tra la tensione al quadrato e la resistenza). Questo significa che quanto maggiore è il pilotaggio del transistor, ossia quanto più spingiamo il transistor a comportarsi in modo non lineare, tanto meno esso tende a smorzare il risonante.

Apparentemente, questo è un risultato anomalo, ma in realtà è assolutamente plausibile: infatti, all'aumentare di V_1, il condensatore di blocco contribuisce ad aumentare la tensione di polarizzazione inversa, in modo che la giunzione venga portata in conduzione semplicemente in corrispondenza dei picchi della tensione sinusoidale, il che equivale ad opporsi ad un aumento del valor medio della corrente. Quindi, quanto maggiore è la tensione applicata, tanto più stretti sono gli intervalli di conduzione del diodo, a cavallo dei picchi della sinusoide.

Questa situazione è rappresentata nella figura seguente, dove le zone annerite indicano gli intervalli di conduzione del diodo:



All'aumentare della tensione sinusoidale di pilotaggio, la giunzione base-emettitore conduce sempre meno e quindi anche la dissipazione di potenza nel carico non lineare è sempre minore.

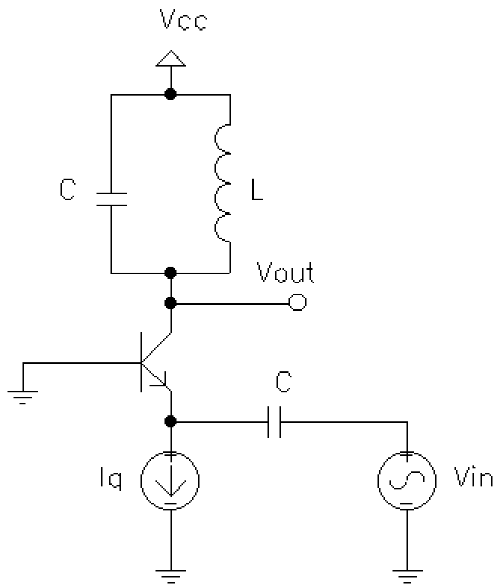
Tutto questo grazie appunto all'azione del condensatore di blocco, che, come detto più volte, opera una autoregolazione del pilotaggio del transistor.

Amplificatori in classe A,B e C

Ricordiamo che un amplificatore il cui transistor conduce solo per una minima frazione del periodo della sinusoide è un amplificatore in **classe C**. Questi amplificatori si differenziano da quelli in **classe A**, nei quali i transistor conducono per l’intero periodo, e da quelli in **classe B**, nei quali i transistor conducono per mezzo periodo.
Gli amplificatori in classe A sono i peggiori in termini di rendimento[22], che ha un valore massimo teorico del 25% quando il segnale in ingresso è sinusoidale; nel caso di un segnale non sinusoidale, come ad esempio un segnale musicale, il rendimento è ancora più basso. Sicuramente migliori sono gli amplificatori in classe B, i cui valori del rendimento possono teoricamente giungere al 78%. Valori ancora più alti si ottengono per gli amplificatori in classe C, con uno svantaggio: rispetto a quelli in classe B, gli amplificatori in classe C presentano una maggiore distorsione dell’uscita, ossia perdono la fedeltà d’ampiezza della forma d’onda. Questo significa che un amplificatore in classe C non può essere usato per amplificatore un segnale modulato d’ampiezza, in quanto l’informazione contenuta appunto nell’ampiezza verrebbe inevitabilmente eliminata dalla distorsione.

Osservazione: problema della saturazione del transistor

Facciamo un ultima osservazione. Consideriamo sempre il circuito seguente:



(dove abbiamo ormai capito che il generatore forzante può essere sostituito da un risonante LC alimentato in corrente).

[22] Il rendimento o l’efficienza di un amplificatore è il rapporto tra la potenza media trasferita al carico e la potenza media assorbita dall’ alimentazione: maggiore è l’efficienza, migliore è il comportamento dell’amplificatore.

Abbiamo trovato in precedenza che, data una tensione sinusoidale $V_{in}(t)=V_1\cos(\omega_0 t)$ applicata tra l'emettitore e la massa, il transistor risponde con una corrente periodica la cui prima armonica ha espressione

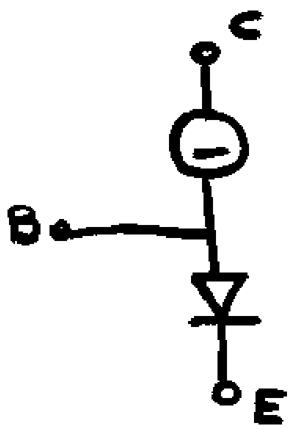
$$i_{C1}(t)=G_{m1}(x)V_1\cos(\omega_0 t)$$

A questa corrente corrisponde quindi una transconduttanza di prima armonica $G_{m1}(x)=\frac{I_{C1}}{V_1}$ e quindi anche una tensione di uscita

$$V_{out,1}(t)=R_{tot}i_{C1}(t)=G_{m1}(x)R_{tot}V_1\cos(\omega_0 t)$$

dove la resistenza R_{tot} tiene conto del carico di collettore alla frequenza di risonanza (perdite nell'induttore, parte resistiva dell'impedenza di uscita del collettore, parte resistiva dell'impedenza del risonante, eventuali effetti di carico degli stadi successivi). Abbiamo inoltre indicato questa tensione di uscita con $V_{out,1}(t)$ per specificare che è la tensione relativa solo alla prima armonica della corrente di collettore.

Questo risultato non è valido sempre, ma solo a patto che sia lecito utilizzare per il transistor il seguente modello di Ebers-Moll:



Si deve cioè vedere dal collettore un generatore di corrente, il che avviene solo se la giunzione collettore-base (n-p) è polarizzazione inversamente. Bisogna dunque accertarsi, con i calcoli, che la tensione applicata alla giunzione collettore-base sia di ampiezza non sufficiente da portare la giunzione stessa in conduzione (cioè da portare il transistor in saturazione): se i calcoli non danno questo, allora il risultato ottenuto non può ritenersi valido.

Partiamo da una condizione in cui non c'è nessun segnale applicato: in questo caso, l'induttore del risonante di carico cortocircuita il collettore all'alimentazione, per cui la tensione sul collettore è V_{CC}. In presenza di segnale, invece, a questa tensione continua è sovrapposta una oscillazione praticamente sinusoidale:

- le escursioni positive di questa tensione sinusoidale non danno problemi, in quanto fanno salire la tensione del collettore e quindi garantiscono un aumento della tensione V_{CE}, ossia un aumento della V_{CB}, ossia un aumento della polarizzazione inversa della giunzione C-B;
- possono invece dare problemi le escursioni negative, che abbassano la V_{CE}, diminuendo la polarizzazione inversa: se la V_{CE} scende al valore $V_{CE,sat}$, la giunzione C-B si polarizza direttamente (con una tensione applicata $V_{CB,sat}=V_{CE,sat}-V_{BE}<0$) ed il transistor va in saturazione.

Quindi, possiamo affermare quanto segue:

- se calcoliamo la tensione di uscita V_{out} dell’amplificatore e troviamo un valore inferiore a V_{CC}, possiamo concludere che, in qualunque istante, la giunzione C-B è sicuramente polarizzata inversamente, per cui è valido il modello circuitale riportato prima e quindi vale anche la formula $V_{out,1}(t) = G_{m1}(x) R_{tot} V_1 \cos(\omega_0 t)$;
- al contrario, se troviamo che $V_{out} > V_{CC} + V_{CB,sat}$, allora la giunzione C-B finisce sicuramente in conduzione sui picchi negativi della tensione di segnale: in questi casi, la giunzione va a “sfogare” a massa (la corrente va dal collettore alla base, che è appunto a massa) l’eccesso di energia accumulata dal risonante, imponendo così come massimo valore della tensione ai capi del risonante il valore che c’era quando è entrata in conduzione, ossia appunto $V_{CC} + V_{CB,sat}$.

Un amplificatore pilotato in questo modo fa quindi localizzare ai capi del carico di collettore una tensione sinusoidale di ampiezza pari approssimativamente al valore $V_{CC} + V_{CB,sat}$. La giunzione C-B agisce perciò da limitatore d’ampiezza, aumentando, però, il contenuto complessivo di armoniche della tensione di uscita desiderata.

Autore: **SANDRO PETRIZZELLI**
e-mail: sandry@iol.it
sito personale: http://users.iol.it/sandry
succursale: http://digilander.iol.it/sandry1