Premier cycle

Année L 1

SCIENCES ET TECHNIQUES DE L'INFORMATION

DU SYSTEME A LA FONCTION

Alain FROMENTEL

v3 - 2006

EFREI

Premier cycle - L1

SCIENCES ET TECHNIQUES DE L'INFORMATION

Du système à la fonction

PREAMBULE

Introduction

De quoi ai-je besoin ? (définition du système)
Comment le faire ? (identification des fonctions)
Avec quels moyens, avec quels objets et comment les fabriquer ? (les composants)

Les technologies de l'information

Elles reposent sur deux technologies essentielles inventées et développées au cours du 20ème siècle : l'électronique et l'informatique
Elles consistent en quatre domaines majeurs :

La représentation de l'information : capteurs, codages, syntaxes, …
Le traitement de l'information : tri, calculs, filtrages, …
La communication de l'information : télécommunications, télécommandes, réseaux …
Le stockage de l'information : mémoires, disques, bases de données, …

Un peu d'histoire

Electricité : énergie facilement productible (centrales électriques avec des principes multiples), assez facilement stockable (mais pour de petites énergies : piles) et très facilement transportable (conducteurs électriques). Le moteur électrique a très rapidement remplacé les machines à vapeur dans les usines, l'électricité a permis de réaliser l'électrolyse, la galvanoplastie, la soudure, …

Electronique : est née par nécessité (!) : besoin de communiquer à (longue) distance, télégraphie, téléphonie, téléphonie sans fil (pour les bateaux !), télévision, téléinformatique, télécommande, …

Electronique : s'est emparée des domaines de la mesure physique (avec capteurs et afficheurs), de la régulation (avec capteurs et actionneurs, « l'automatique »), de l'audiovisuel (son, image), du calcul (calculateurs), du tri de données (ordinateurs), …

Quelques applications actuelles

Le téléphone portable, le « sans fil » → transmission de la parole, émission et réception

La « Hi-Fi », le son « surround » → l'amplification du son, filtrage et codage

La télévision par satellite → analyse, transmission et synthèse de l'image

La télécommande d'engins spatiaux → asservissements, systèmes bouclés

Les machines informatiques → ordinateurs, traitement de données, communication via internet

Fil conducteur du cours « du système à la fonction » : **la téléphonie**

1. Une brève histoire de la communication à distance

- le langage parlé, le messager, le « avis à la population », le porte voix, …
- le courrier écrit, le coursier, le service postal, la poste maritime, la poste aérienne, …
- les messages codés : signes, fanions, télégraphe* Chappe, télégraphe optique, …
- le télégraphe électrique, le code Morse
- le téléphone** électrique « analogique »
- le téléscripteur, la télécopie
- le réseau téléphonique
- la téléinformatique
- la téléphonie numérique
- la téléphonie sans fil
- les « SMS » et « MMS »
- … internet

* télégraphe : graphie (écriture) à distance
** téléphone : phonie (son vocal) à distance

2. La télégraphie et l'électricité : les prémices

2.1 La télégraphie avec l'électricité

2.1.1 L'électricité : un vecteur de communication

Quand on a besoin de transmettre un message à distance, on peut utiliser le service postal, à condition, au préalable d'écrire ce message, c'est-à-dire d'utiliser une représentation de l'information (l'écriture) utilisant un code (dictionnaire des mots), ce dernier utilisant lui-même un ensemble de symboles (les lettres de l'alphabet) arrangées suivant des règles (orthographe puis grammaire).

Un tel service postal révèle l'inconvénient du délai entre l'émission du message par son auteur et sa réception par le destinataire, même avec un service « rapide ».
Ce délai est inhérent aux moyens utilisés pour le transport de l'information : la marche à pied, la voiture, le train, le bateau, l'avion, … Tous ces systèmes possèdent des vitesses de déplacement qui vont de 1 à 1000 kilomètre/heure.

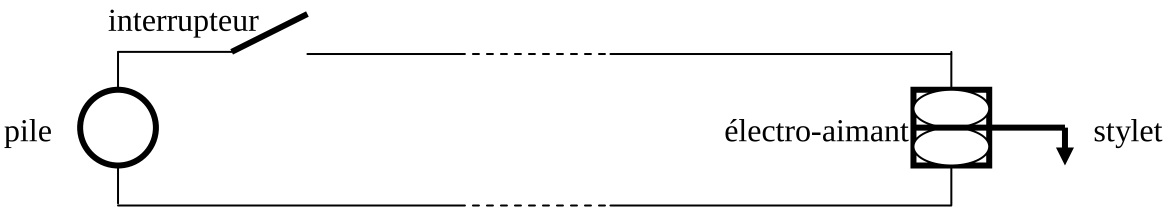
Quand l'électricité fut découverte, une de ses propriétés était l'impression que la vitesse de propagation de l'électricité était infinie (en réalité, elle vaut 300 000 km/s, soit environ 1000 millions de kilomètre/heure !).
Ainsi, l'électricité, déjà fort utile pour les appareils d'éclairage, pour les machines industrielles (moteurs) et pour l'électrochimie, est entrée dans le monde des transmissions de message (les « télécommunications !) … dès lors qu'un système de représentation électrique d'un message fût imaginé.
Ce système de représentation électrique nécessitait un code, utilisant des symboles.

L'électricité (ou plutôt l'électrocinétique) se manifeste par l'existence d'un courant électrique, lui-même du à un mouvement des électrons dans un conducteur métallique : il faut noter d'ailleurs que ce mouvement est « lent » (environ 300 mètre/heure), mais que c'est la « propagation » de ce mouvement de proche en proche qui est très rapide (300 000 km/s).

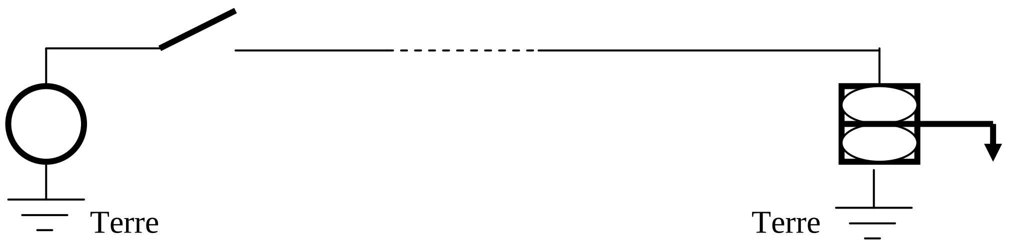
Comment créer un courant électrique ? L'émetteur du message doit disposer d'un générateur électrique (une pile par exemple) : une « source » et de conducteurs reliés à cette pile.
Comment mettre en évidence un courant électrique ? Le destinataire doit disposer d'un moyen permettant de percevoir ce message : un « afficheur ».
Quel code utiliser ? D'abord, quels symboles a-t-on à disposition ? Les plus simples se traduisent par l'existence ou non du courant électrique, ceux-ci se réalisent très facilement par un dispositif « interrupteur ».

2.1.2 <u>Un premier système de communication électrique</u>

Nous sommes donc en mesure de concevoir un système permettant la transmission quasi-instantanée (vitesse = 300 000 km/s) d'une information codée à l'aide de deux symboles que l'on peut noter « 0 » (pour l'absence de courant) et « I » (pour l'existence d'un courant de valeur I).
Ce système comprend :
1 pile, 1 interrupteur (« manipulateur », « key »), des conducteurs électriques, un afficheur (ampoule électrique, vibreur sonore, électro-aimant avec stylet d'écriture, …).



Ce système peut être simplifié (et surtout rendu moins onéreux) en raisonnant sur les potentiels électriques, la Terre étant sensiblement (et sur des distances raisonnables) équipotentielle (la valeur « 0 volt » étant prise en général par convention) ; on peut ainsi faire l'économie d'un des deux conducteurs (économie financière due au coût du cuivre, économie de temps pour la pose, besoin d'une seule bobine de fil pour la pose, pas de nécessité d'isolation si le fil est posé sur des poteaux, …).



Pour réaliser une communication d'information, nous avons à présent le moyen, il nous manque la méthode d'expression des messages à l'aide des « signaux » que nous avons définis : « courant nul 0 » ou « courant non nul I ».
Nous avons besoin de définir une fonction (relation) entre l'ensemble des mots et expressions constituant les messages et un ensemble de séquences de « 0 » et « I ». Cette relation doit être une bijection : à chaque mot ou expression doit correspondre une séquence et une seule, de manière à éviter toute équivoque.

2.1.3 Un codage de l'information

Samuel Morse, inventeur du télégraphe électrique (1843) a aussi inventé un « code » (le « code Morse ») qui réalise une telle bijection d'un ensemble formé de lettres, de chiffres et de signes spéciaux dans un ensemble formé de quatre signes : le point (« short mark »), le trait (« long mark »), le petit espace (« short space ») et le grand espace (« long space »). On peut remarquer qu'il n'y a en réalité que deux signaux (« mark » et « space »), ceux-ci étant chacun de deux durées différentes. Ce code permet donc d'utiliser un système de transmission électrique à deux états « 0 » et « I », tel qeu celui décrit plus haut : la télégraphie « écriture à distance » était née à la condition d'avoir, côté émission, un opérateur capable de transcrire (coder) les mots en signes et, côté réception, un autre opérateur capable de retranscrire (décoder) les signes en mots : des opérateurs de télégraphie !

Le code Morse pour les lettres

A .-	B -...	C -.-.	D -..	E .	F ..-.	G --.	H	I ..
J .---	K -.-	L .-..	M --	N -.	O ---	P .--.	Q --.-	R .-.
S ...	T -	U ..-	V ...-	W .--	X -..-	Y -.--	Z --..	

Il est complété par des codes pour les chiffres, la ponctuation et des informations particulières (début et fin de transmission, erreur, souligné, …).

On pourra remarquer que ce codage est « à longueur variable » : le nombre de signes utilisés dépend de la lettre à coder. Ces longueurs n'ont pas été prises arbitrairement, mais sont fonction de la probabilité d'occurrence d'un caractère : plus un caractère est probable, plus petite sera la longueur du code représentant ce caractère (ainsi la lettre « E » est-elle codée par un seul « point ») : ce type de codage est un exemple de « codage compact », principe toujours utilisé dans les processus de compression de données.

2.1.4 Communication à grande distance

Le transport de l'information est donc assuré par la propagation d'un courant électrique dans un conducteur. La matière qui constitue ce conducteur électrique (cuivre le plus souvent) est caractérisée par sa résistivité électrique ρ. Chaque élément de ce conducteur oppose ainsi une résistance au passage du courant (genre de « frottement » des électrons dans la structure de la matière) qui se traduit par un échauffement de ce conducteur ; ce phénomène est connu sous le nom d'effet Joule. Il y a donc perte d'énergie, ou plutôt conversion d'énergie électrique en énergie thermique qui se rayonne dans l'espace environnant.
La loi de Joule permet de chiffrer la puissance P rayonnée grâce à la relation suivante :

$$P = R\,I^2 \quad \text{exprimée en watt (W)}$$

I étant l'intensité du courant circulant dans le conducteur (en ampères) et R le degré de « résistance » du conducteur, la puissance mesurant la quantité d'énergie par unité de temps.

Pour un conducteur cylindrique plein (« fil électrique ordinaire »), cette « résistance « vaut :

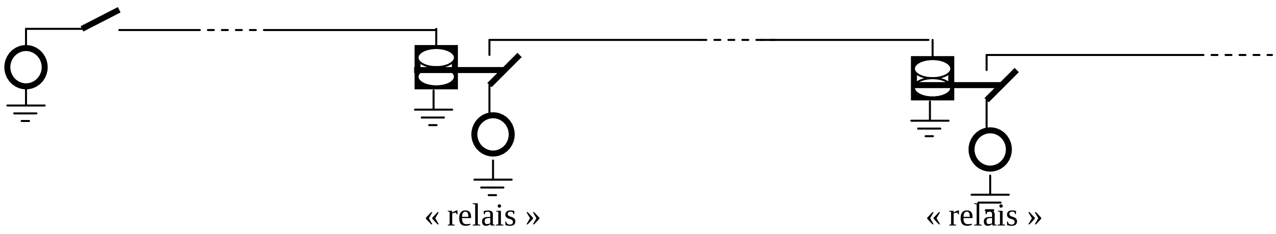
$$R = \rho\,L / S \quad \text{exprimée en ohm } (\Omega)$$

L étant la longueur du conducteur considéré et S sa section.

Ainsi, la puissance perdue en chaleur vaut : $P = I^2\,\rho\,L / S$, directement proportionnelle à la longueur L.

On comprend aisément que, pour des distances importantes de transmission, et même en augmentant la section du conducteur, la puissance reçue puisse devenir faible par rapport à la puissance émise (par la source), peut-être même trop faible pour que l'électroaimant puisse fonctionner !
On a donc imaginé un dispositif dit « relais » (bien nommé) destiné à « régénérer » la puissance initiale ou, plus exactement, destiné à utiliser une puissance faible (affaiblie par la distance) pour commander une puissance plus importante.
CECI CONSTITUE LES PREMICES DE LA NOTION D'AMPLIFICATION ELECTRIQUE.
<u>Ce dispositif est donc à la base des amplificateurs.</u>
<u>De plus, ce même dispositif est exactement à la base de la régénération des futurs signaux numériques !</u>

Le dispositif consiste à utiliser un électroaimant très sensible (car fonctionnant avec peu de courant) qui déclenche un interrupteur mettant ou non en service une nouvelle pile génératrice. Ainsi, on répète automatiquement l'action du manipulateur.
L'ensemble « électroaimant – interrupteur déclenché » est toujours appelé « relais » de nos jours !



et ainsi de suite …

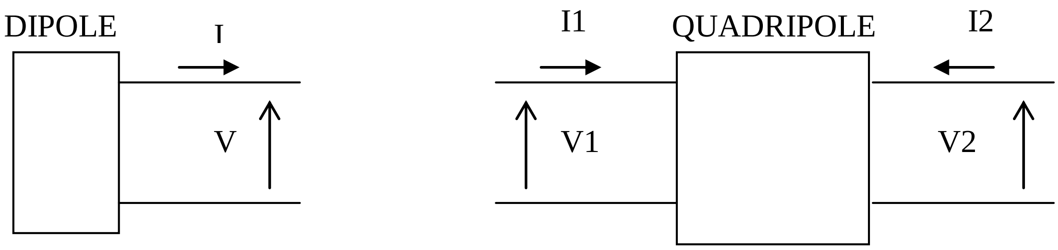
2.2 La modélisation des circuits électriques

2.2.1 Dipôles et quadripôles

Nous avons examiné jusqu'à présent des circuits expérimentaux avec une approche résolument empirique.
Il apparaît nettement un besoin de « chiffrer » les pertes, de manière à pouvoir estimer la distance au delà de laquelle un « relais » sera nécessaire.
Pour cela, nous avons besoin de MODELISER les circuits électriques.
Les modèles élémentaires vont s'appliquer aux trois éléments fondamentaux suivants :

- la source
- l'utilisation
- le moyen de transmission de la source vers l'utilisation.

Les deux premiers sont des DIPOLES électriques (un seul branchement), le troisième est un QUADRIPOLE électrique (une « entrée » et une sortie ») :



A chaque branchement électrique (2 conducteurs) est associé deux grandeurs :

- la différence de potentiel (ddp) ou « tension électrique » V, cette ddp exprime un déséquilibre de charges au niveau d'un dipôle, ce qui pourra donc conduire à un déplacement de ces charges pour tendre à revenir à l'équilibre ! Le travail (au sens énergétique) fourni par les charges Q qui se déplacent vaut $W = Q V$. Cette énergie électrique est naturellement utilisable par des « machines » (c'est le principe de l'utilisation de l'énergie électrique …)

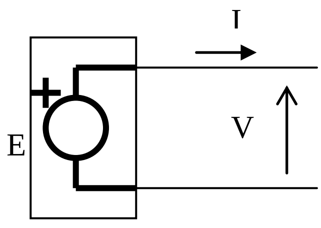
- le courant ou « intensité électrique » I

Le courant exprime le déplacement des charges ou, en un endroit donné du conducteur, la variation de la charge au cours du temps : $I = dQ/dt$.
Il faut noter que, pour le quadripôle, une convention conduit au choix des courants « entrants » dans le quadripôle.

2.2.2 Sources idéales, sources réelles

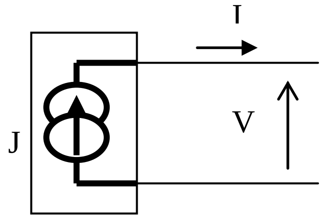
Dipôle représentatif d'une source idéale :

E : force électromotrice (fem)
$V = E \quad \forall I$
C'est une « source de tension »
Puissance délivrée $P = VI$



ou encore :

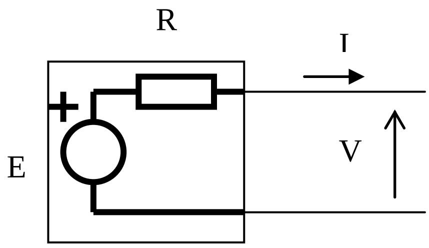
J : courant électromoteur (cem)
$I = J \quad \forall\, V$
C'est une « source de courant »
Puissance délivrée $P = VI$



Dipôle représentatif d'une source réelle :

Les relations ci-dessus ne sont pas exactes en réalité. L'expérience montre que, dans le cas d'une source de tension par exemple, la ddp V n'est égale à E que si I est nul (« source à vide »). Ainsi, lorsque cette source est effectivement utilisée (donc puissance délivrée non nulle), le courant ne peut pas être nul et la conséquence est que la ddp devient inférieure à E. Ce comportement est modélisé grâce à une résistance électrique placée en série avec la source E. En effet, la relation liant E à V s'écrit alors :

$$V = E - RI \qquad \text{(si } I = 0 \text{ alors : } V = E\text{)}$$



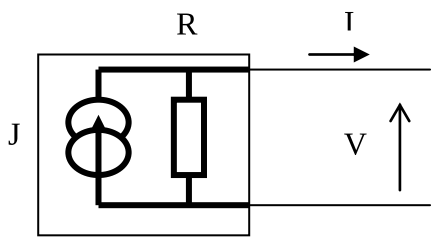
Ce modèle est appelé MODELE DE THEVENIN.

Il faut bien noter deux choses :

- la première, est que cette représentation est un « modèle » comportemental, qui ne prétend pas à décrire la structure « interne » de la source et qui, encore moins, représenterait une « véritable » résistance câblée à l' intérieur de la source !
- la deuxième, est que cette représentation ainsi que le relation consécutive sont linéaires, en particulier, que l'élément « R » ne dépend ni de I, ni de V. S'il y a des cas où cette linéarité est assurée, il y en a d'autres où la relation (V, I) est quelconque, toutefois assimilable localement (petit domaine de V et de I) à une relation linéaire.

Il en est de même pour la source de courant réelle, un raisonnement similaire conduit à la relation suivante et au modèle qui suit :

$$I = J - V/R \qquad \text{(si } V = 0 \text{ alors : } I = J\text{)}$$



Ce modèle est appelé MODELE DE NORTON.

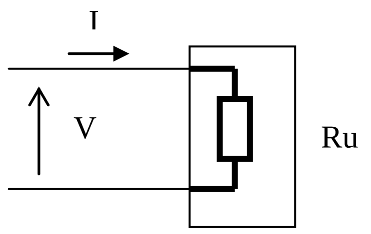
2.2.3 <u>Utilisations ou « charges »</u>

L'aboutissement d'un circuit électrique (surtout dans le contexte d'une transmission) est une utilisation (destinataire de l'information).

Cette utilisation va naturellement consommer de la puissance, en effet, la transmission d'une information se fait matériellement par propagation d'une énergie (électrique, électromagnétique ou optique) ; celle-ci est produite par la source et se dépense dans l'utilisation, dite aussi « charge ».
Ici encore, nous allons avoir recours à un modèle de dipôle consommant de la puissance : la résistance électrique est le dipôle le plus simple. Encore une fois, il est évident que cette résistance n'est qu'un modèle : un système de transmission qui serait terminé par une véritable résistance qui convertit l'énergie électrique intégralement en chaleur serait étrange !

Mieux, cette résistance va même représenter « tout ce qui peut être situé en aval du plan de branchement de cette résistance : on parle alors de « dipôle équivalent ». Ainsi, une simple résistance peut elle représenter tout un système électrique, vu de son entrée (on parle alors de « schéma équivalent »).

Dipôle représentatif d'une utilisation :

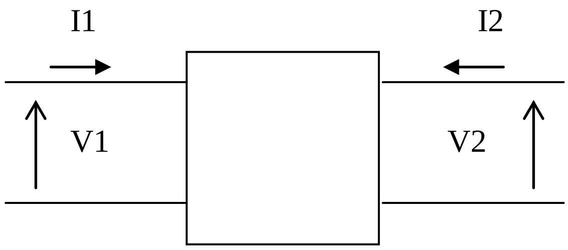


La puissance consommée (ou puissance utile reçue) vaut donc :

$$P = VI = RI^2 = V^2/R$$

2.2.4 <u>Quadripôles</u>

Un quadripôle est un système électrique possédant une entrée et une sortie. Il est donc caractérisé en premier lieu par une « fonction » décrivant son fonctionnement.
En règle générale, la tension et le courant à la sortie seront des fonctions de la tension et du courant à l'entrée. Ces relations peuvent être quelconques.
Le quadripôle sera dit « linéaire » (et le système qu'il représente aussi) si ces relations sont elles-mêmes linéaires (polynôme du premier degré).



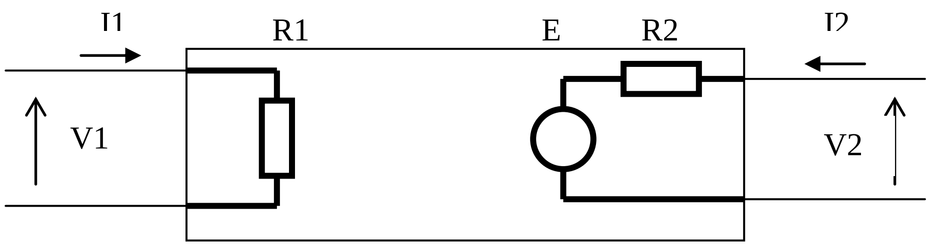
Une écriture matricielle peut être utilisée dans le cas général :

$$\begin{pmatrix} V\,2 \\ \\ I\,2 \end{pmatrix} = \begin{pmatrix} A \end{pmatrix} \begin{pmatrix} V\,1 \\ \\ I\,1 \end{pmatrix}$$

Le choix a été fait d'exprimer les grandeurs de sortie en fonction de celles d'entrée (matrice de « chaîne »), on peut aussi exprimer les relations différemment (les « V » en fonction des « I », l'inverse, des relations « hybrides », etc …)

Un quadripôle est ainsi une « boîte noire » (vu de l'extérieur) qu'il faut néanmoins caractériser si l'on a besoin de l'intégrer dans un système plus grand (ne serait-ce que d'y brancher une source et une charge !).

Pour cela, nul n'est besoin de définir d'autres modèles que ceux imaginés pour les sources et charges. En effet, une « entrée » de quadripôle reçoit l'information, donc consomme de l'énergie, donc est représentable par une résistance modélisant une utilisation. De même, une « sortie » de quadripôle fournit de l'énergie et, à ce titre, est représentable par une source.



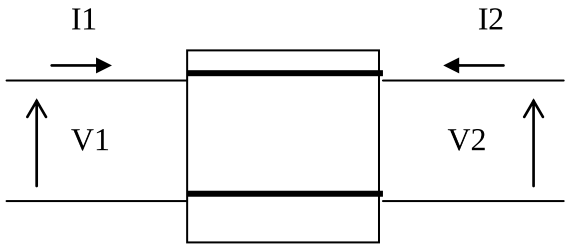
R1 est appelée « résistance d'entrée » ; R2 est appelée « résistance de sortie ».
Les différentes valeurs R1, E et R2 sont exprimables à partir

On obtient ainsi la « caractérisation externe » de ce quadripôle. Celle-ci est suffisante pour pouvoir intégrer ce quadripôle dans un système plus grand, essentiellement vis-à-vis du comportement de ce quadripôle envers les tensions et courants.

Bien entendu, on peut avoir besoin d'une caractérisation plus « poussée » ou « plus fine », décrivant « l'intérieur » de ce quadripôle, soit en décrivant de façon exacte cet intérieur, soit en le décrivant partiellement à l'aide de schémas équivalents.
Enfin, à partir de la caractérisation interne, on peut en déduire la caractérisation externe.

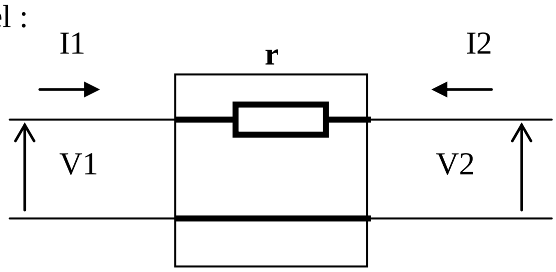
Exemples :
- quadripôle représentatif d'un conducteur idéal (2 fils électriques de résistivité nulle) :

$V2 = V1$
$I2 = - I1$



- quadripôle représentatif d'un conducteur réel :

$V2 = V1 - r\,I1$
$I2 = - I1$



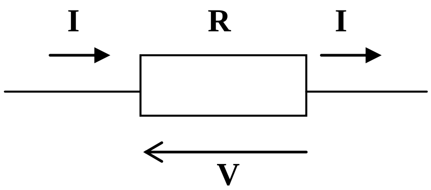
« **r** » représente les pertes par effet Joule des conducteurs. Bien entendu, aucune résistance n'est câblée (!), cette résistance équivalente ne fait que représenter la perte d'énergie par effet Joule cumulée sur toute la longueur des 2 fils électriques (bien que cette résistance ne soit représentée que « sur le fil du haut »).

2.2.5 Théorèmes généraux

- Convention de signes entre tensions et courants :

On peut adopter un sens de courant arbitraire (pas nécessairement celui qui paraît « logique » ou « physique », on rappelle à ce sujet que le sens conventionnel du courant est opposé au sens de déplacement physique des électrons …).
Dès lors que le sens du courant est choisi, celui de la tension aux bornes d'un dipôle passif (c'est-à-dire ne contenant pas de source) en découle :

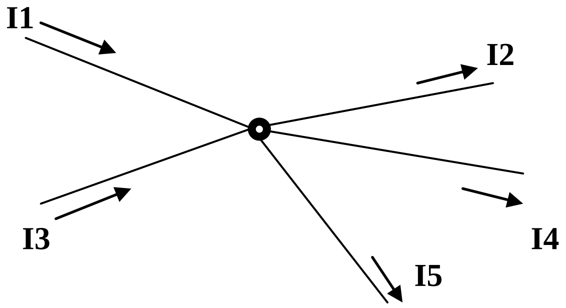
Loi d'OHM
$V = R\,I$
(expression linéaire)



Ce choix est dit « tension et courant en sens inverse ». Ce choix exprime la perte de potentiel lorsque le courant passe dans un élément qui utilise l'énergie (ce qui est naturel, car ce même potentiel exprime une énergie « transportée » par des charges électriques). Un autre choix entraînerait une signe négatif dans l'expression de la loi d'Ohm.

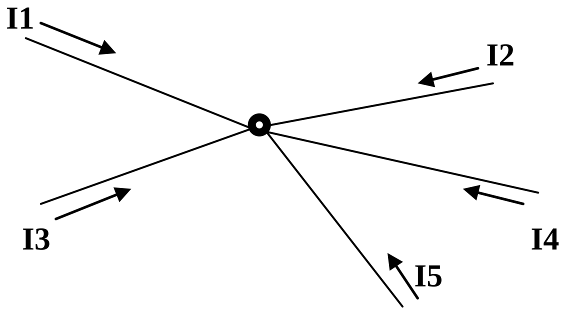
- Loi des noeuds :

Cette loi s'applique lorsqu'il y a une dérivation dans un circuit. Les charges se répartissant, les courants se répartissent également ; étant donné que les charges ne disparaissent pas (!), la loi des nœuds exprime simplement que les courants « entrants » sur un nœud est égale à la somme des courants « sortants » de ce nœud.



I1 + I3 = I2 + I4 + I5

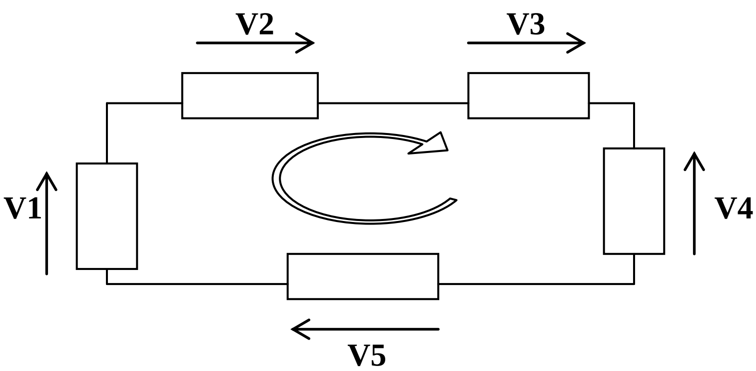
ou, en général, pour des courants tous choisi « entrants » (ou tous « sortants »)



I1 + I2 + I3 + I4 + I5 = 0

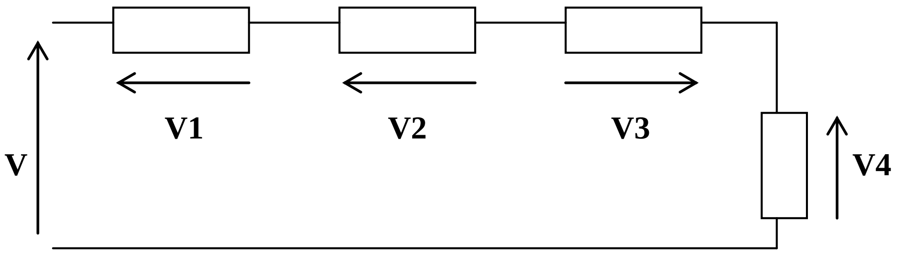
- Loi des mailles :

Cette loi s'applique lorsqu'il y a une boucle dans un circuit ou encore lorsque l'on s'intéresse à une tension aux bornes d'un circuit comprenant plusieurs dipôles.



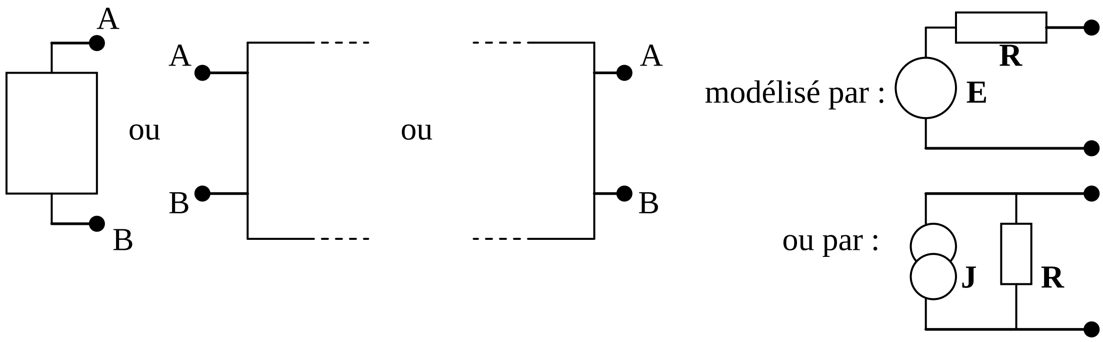
V1 + V2 + V3 – V4 + V5 = 0

La règle est la suivante : on choisit un sens de parcours (arbitraire), on affecte un signe « + » si la tension est dans le sens du parcours, un signe « - » dans le cas contraire.



V -V1 -V2 +V3 -V4 = 0 qui s'écrit aussi : **V = V1 + V2 -V3 + V4 = 0**

- Théorèmes de Thévenin et de Norton :
Ces théorèmes permettent de modéliser tout dipôle ou toute entrée ou sortie de quadripôle sous forme d'un modèle de Thévenin ou de Norton, ces deux modèles étant eux-mêmes équivalents.



Détermination de **E** : $E = V_{AB}$ « à vide », c'est à dire en ne branchant rien entre A et B
Détermination de **J** : $J = I_{AB}$ « en court-circuit », c'est à dire en branchant un fil entre A et B

S'il n'y a pas de sources à l'intérieur du dipôle ou du quadripôle, E et J sont nulles.

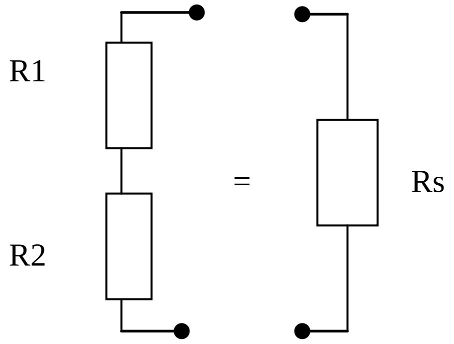
Détermination de **R** : $R = V_{AB} / I_{AB}$ (la valeur est la même pour les deux modèles)

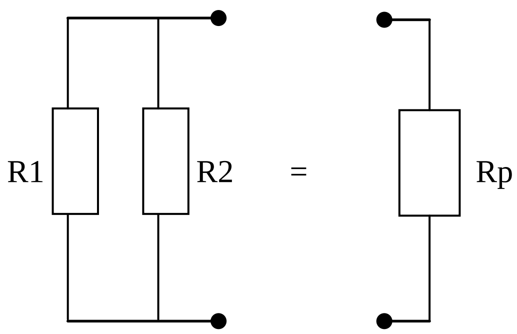
On a enfin : $E = R J$.

- Théorème de superposition :
Ce théorème s'applique exclusivement à des systèmes linéaires (circuits et composants linéaires).
Dans un dipôle ou un quadripôle comportant plusieurs sources indépendantes (c'est à dire n'ayant aucune relation entre elles), chacune des tensions et chacun des courants (en particulier en entrée et en sortie) peuvent être exprimées respectivement en des sommes de tensions et courants partiels, chacune des valeurs partielles étant obtenue en « éteignant » toutes les sources sauf une (éteindre une source de tension, c'est la remplacer par un fil, car V = 0, éteindre une source de courant, c'est la remplacer par rien, car I = 0).

Ce théorème permet de simplifier les calculs sur des circuits complexes et permet également d'isoler un comportement relatif à une source donnée.

- Association de dipôles :
On peut associer des dipôles en série (même courant dans les deux) ou en parallèle (même tension sur les deux) :



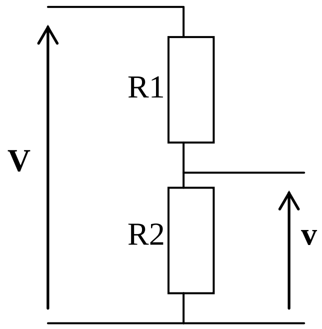


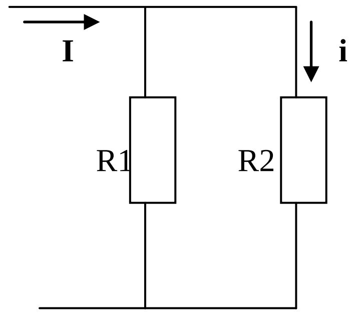
Branchement en série
Rs = R1 + R2

Branchement en parallèle
Rp = R1 R2 / (R1 + R2)

- Diviseur de tension et diviseur de courant :

Ce principe s'applique lorsque l'on a deux (ou plus) dipôles placés en série ou en parallèle:



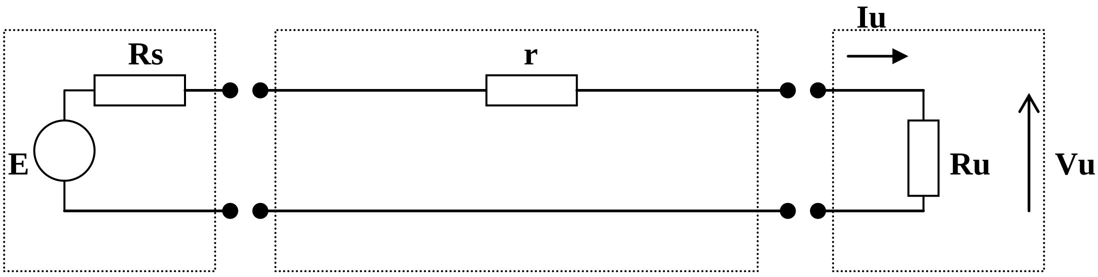


v = V R2 / (R1+R2))

i = I R1 / (R1 + R2)

2.3 Modélisation de la transmission télégraphique

Une transmission télégraphique peut donc être représentée de la manière suivante :



2.3.1 <u>Transfert en tension</u> :

La tension utile, reçue effectivement par le destinataire vaut donc (principe du diviseur de tension) :

$$Vu = E\ Ru / (Rs + r + Ru)$$

Cette expression met en évidence plusieurs choses dont la portée est générale :
- un bon transfert en tension le long de la ligne se fera si **r** est la plus petite possible (évident !), ce qui justifie l'emploi de répéteurs permettant d'utiliser des tronçons de ligne courts (donc ayant une résistance plus faible)
- un bon transfert en tension entre source et utilisation se fera si Ru est grande devant Rs.

De cette manière, si **r** est petite et si Ru est grande devant Rs, on obtient :

$$Vu \cong Vs$$

A la limite, un bon transfert en tension se fera sur une résistance d'utilisation infinie, mais alors le courant sera nul et la puissance utile aussi !

2.3.2 <u>Transfert en courant</u> :

Nous n'avons pas ici de division de courant (pas de nœud), toutefois, en supposant Rs fixée, il est clair que le courant dans Ru sera d'autant plus grand que Ru est faible.
A la limite, un bon transfert en courant se fera sur une résistance d'utilisation nulle, mais alors la tension sera nulle et la puissance utile aussi !

2.3.3 <u>Transfert en puissance</u> :

On peut également raisonner sur la puissance effectivement transmise.
Pour l'instant, nous allons négliger **r**.

$Vu = E\ Ru / (Ru + Rs)$ et $Iu = E / (Ru + Rs)$ entraîne : $Pu = E^2\ Ru / (Ru + Rs)^2$

Cette valeur de Pu sera maximale si et seulement si Ru = Rs.
Ce concept se nomme « adaptation » : on dit que la charge est adaptée (en puissance) à la source.

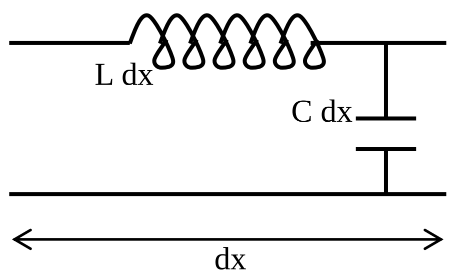
Ce concept dépasse naturellement celui de transfert en puissance, en effet, il est vain d'assurer un transfert maximal en tension si, en échange, le courant Iu est tellement faible qu'il ne permettra pas d'actionner le dispositif placé en réception. Il en est de même pour le transfert maximal en courant.

Ainsi, en l'absence d'un système d'amplification (augmentation de l'énergie), le transfert maximal en puissance doit toujours être recherché, d'autant plus que la présence de **r** dissipe une part de l'énergie émise par la source !

2.3.4 Mesure des puissances et des atténuations :

Les puissances se mesurent naturellement en watt (W).
La téléphonie a normalisé un niveau de puissance émise sur la ligne (appelé d'ailleurs « niveau de ligne » ou *« line level »*) égal à 1 mW.
La téléphonie a également normalisé une résistance de valeur égale à 600 Ω.

Cette valeur correspond à ce que l'on nomme « impédance caractéristique de la ligne téléphonique normalisée » : cette impédance est égale au rapport V / I le long de la ligne, on montre en effet que ce rapport est constant. Ceci vient du fait qu'une ligne, constituée de deux conducteurs séparés par un isolant et parcourus par un courant variable est modélisable localement par le circuit suivant, pour chaque tronçon élémentaire de longueur dx :



La bobine L dx traduit l'effet d'auto induction produit par le courant variable qui circule dans le conducteur (ce courant crée un champ magnétique autour de ce conducteur qui, à son tour crée une fcem d'auto induction sur ce même conducteur : ainsi le conducteur a tendance à s'opposer au passage du courant.
Le condensateur C dx traduit la capacité qu'il y a entre les deux conducteurs séparés par un isolant.

Comme il n'y a pas de pertes en dehors de celle occasionné par l'effet Joule (on n'a pas représenté la résistance R dx qui traduirait ces pertes), l'énergie est conservée par passage dans ce tronçon dx :
Energie emmagasinée par la bobine : $dWL = 1/2 \; L \, dx \, I^2$
Energie emmagasinée par le condensateur : $dWC = 1/2 \; C \, dx \, V^2$
La conservation de l'énergie s'écrit : $dWL = dWC$ d'où : $L I^2 = C V^2$ soit : $\mathbf{(V / I)^2 = L / C}$

Ceci montre donc que le rapport V / I est constant le long de la ligne, ce qui est remarquable. Ce rapport étant homogène à une résistance, on nomme « résistance caractéristique » (de la ligne) ce rapport, noté « Rc ».
Cette résistance caractéristique n'a bien entendu rien à voir avec la résistance (réelle) due à la résistivité du conducteur qui, elle seule, contribue aux pertes.

Cette double normalisation de la puissance et d'une résistance amène naturellement à une valeur particulière de tension électrique : $Uo^2 = Po \,.\, Rc = 10^{-3} \,.\, 600 =$ **775 mV**.

Tous les autres niveaux devant se référer à ce niveau normalisé (puissance ou tension), il a été décidé d'adopter une échelle sans dimension calculée à partir d'un rapport dont le dénominateur serait toujours Po (1 mW) ou Uo (775 mV).

En outre, afin de simplifier les calculs (simples par ailleurs), mais aussi de tenir compte d'une particularité auditive (« loi de Fechner ») énonçant que la sensation de puissance sonore est sensiblement proportionnelle non pas à la puissance acoustique physique, mais au logarithme de celle ci (il s'agit d'une propriété purement physiologique), il a été décidé d'adopter une échelle logarithmique pour représenter les niveaux acoustiques :

$$L = \log (P / Po)$$

mesure ainsi le niveau acoustique P, par rapport à Po = 1mW, en adoptant une échelle logarithmique (logarithme décimal).

En hommage à Graham Bell, cette échelle est associé à une unité (sans dimensions) : le Bel.

Le même raisonnement fut tenu quand il s'agit de mesurer les sensations perçues, mais cette fois par rapport à un niveau physiologique : le choix fut fait de la limite d'audition à la fréquence de 1000 Hz (environ la note « do5 », fréquence correspondant environ à la meilleure acuité auditive : perception du son le plus faible en intensité). Cette valeur minimale égale à 2.10^{-5} Pascal (pression acoustique sur le tympan) correspond à une densité de puissance de 10^{-12} W/m² au voisinage du même tympan (membrane de l'ordre de 1 cm², soit une puissance de l'ordre de 10^{-16} W). A l'opposé, la puissance la plus élevée avant lésion est de l'ordre de 200 Pa, soit 100 W/m², il faut rappeler ici que ce sont des puissances acoustiques, ne surtout pas comparer avec des puissances électriques d'amplificateurs, … car il ne faut pas oublier le rendement très mauvais des haut parleurs et surtout l'affaiblissement en 1/d² de propagation entre ceux-ci et le tympan !).
La connaissance de ces deux valeurs minimale et maximale permet d'accéder à l'étendue dynamique de l'audition humaine : rapport de 10^{14} entre la puissance la plus faible et la puissance la plus élevée. La valeur très élevée de ce rapport ainsi que la loi de Fechner déjà évoquée ont fait utiliser une échelle logarithmique avec une valeur de référence égale à la valeur du seuil d'audition :

$S = \log (p / pmin) = \log (p / 10^{-12})$ p représentant une densité de puissance (W/m²)

Cette échelle se mesure à l'aide d'une unité sans dimension : le « Bel acoustique » ou *« SPL : sound pressure level)*. Deux degrés de cette échelle sont séparés de 1 Bel, l'expérience montre que cette différence de niveau est très grande au sens de la perception, alors que le dixième de ce degré (donc le déciBel) représente une différence perceptible certes, mais faiblement. Il a donc été décidé d'utiliser comme unité courante le déciBel (dB), d'où un facteur « 10 » supplémentaire dans l'expression de S :

$S = 10 \log (p / pmin)$ exprimé en dB SPL

Les puissances électriques seront donc exprimées en dB par rapport à Po = 1mW, afin de rappeler cette valeur de référence de 1 mW, l'unité sera accompagnée du suffixe « m » :

$L = 10 \log (P/ Po)$ exprimée en dBm si Po = 1mW

Par analogie, on peut alors également exprimer des atténuations ou des amplifications avec le même type de calcul :

$A = 10 \log (P2 / P1)$

exprimé en dB (sans suffixe), mesure l'atténuation (si A est négatif) ou l'amplification (si A est positif) d'un quadripôle dont l'entrée est soumise à la puissance P1 et qui fournit en sortie une puissance P2.

Un intérêt de cette expression est que les valeurs sont à présent additives :

Exemple :
Puissance à l'entrée = - 10 dBm (par exemple)
Atténuation de 5 dB (on devrait dire : - 5 dB)
Amplification de 20 dB
Puissance en sortie : -10 - 5 + 20 = + 5 dBm

3. La téléphonie et l'électronique

3.1 La téléphonie point à point câblée

3.1.1 Du télégraphe au téléphone

La transmission proprement dite de la parole ne soulève pas de problème particulier dès lors que la parole sera représentée sous forme d'une tension électrique.

La parole engendre un son, phénomène mécanique dont l'origine est la vibration des cordes vocales. L'air émis par les poumons est ainsi « modulé » par l'ensemble larynx – pharynx – cavité bucco nasale pour engendrer une pression mécanique sur l'air qui est dans la bouche.
L'air étant élastique, cette pression se propage alors de proche en proche pour former le son qui arrive à nos oreilles dont les tympans se mettent à vibrer.
Cette pression «dite « acoustique » vient naturellement s'ajouter à la pression atmosphérique de l'air ambiant. Il faut noter que les amplitudes de pressions acoustiques créant des sons audibles varient entre 2.10^{-5} Pa et 200 Pa, alors que la pression atmosphérique normale vaut environ 2.10^{5} Pa !
Nous sommes donc en face d'un processus essentiellement mécanique.

Pour pouvoir utiliser un « circuit télégraphique », il est indispensable de disposer de deux dispositifs « électromécaniques » :
- le premier, côté source : un objet appelé « microphone », placé à proximité de la bouche et capable de transformer les vibrations mécaniques sonores en vibrations électriques ;
- le deuxième, coté destinataire : on objet appelé « écouteur », placé à proximité de l'oreille et capable de retransformer les vibrations électriques en vibrations mécanique sonores.

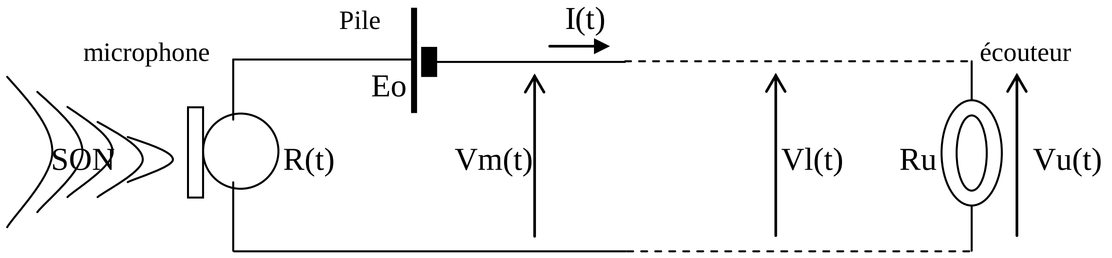
On obtiendra ainsi des vibrations électriques (c'est à dire une fonction variable du temps) « analogues » aux vibrations acoustiques (l'adjectif « analogue » a donné naissance à toute une technologie « analogique » pour laquelle les différentes fonctions sont toutes proportionnelles les unes aux autres, éventuellement à une translation près.

Historiquement, le microphone était fabriqué à l'aide d'un dispositif à résistance variable, la variation de cette résistance étant assurée par la variation de pression mécanique sur un agglomérat résistif, cette variation de pression étant elle-même créée par la vibration d'une membrane souple et élastique sur laquelle arrive le son.
Ce dispositif est entièrement passif en ce sens où il est incapable d'engendrer seul une énergie électrique. Il doit être complété par une source de tension constante (une « pile ») de manière à constituer un dipôle fait de la mise en série de la source avec la résistance variable, cette dernière permettant alors d'obtenir un courant variable analogue aux variations de pression.
Le microphone fait partie des « capteurs » (transformation de grandeur physique en grandeur électrique).

Historiquement également, l'écouteur était réalisé à l'aide d'un électroaimant mettant en mouvement une membrane élastique qui, par sa vibration engendrera un son.
L'écouteur (comme le haut parleur) fait partie des « moteurs » (transformation de grandeur électrique en grandeur physique).

3.1.2 Un système téléphonique

Nous sommes à présent en mesure de réaliser un système de transmission téléphonique rudimentaire.



Vm : tension fournie par l'ensemble microphone – pile
Vl : tension sur la ligne téléphonique (2 conducteurs)
Vu : tension reçue par le destinataire, destinée à l'écouteur
I : courant de ligne
Soit « p(t) » la fonction du temps qui décrit la pression qui s'exerce sur la membrane du microphone.
Cette pression a deux composantes : la pression atmosphérique po et la pression acoustique qui s'ajoute à la première et qui est due au son

$$P(t) = po + pson(t)$$

Compte tenu de la constitution du microphone « capteur de pression » à résistance variable, cette résistance est sensiblement inversement proportionnelle à la pression et devient donc une fonction du temps décrite par :

$$R(t) = R(p(t)) = Ro / p(t)$$
$$I(t) = Eo / (R(t) + Ru) = Eo / (Ru + Ro/p(t))$$

Cette relation est non-linéaire, sauf si on peut garantir que Ru reste toujours très faible devant le terme Ro/p(t). Dans ce cas, on peut négliger la valeur de Ru et l'expression du courant de ligne I(t) devient :

$$I(t) = Eo\ p(t) / Ro$$

qui est une relation proportionnelle.
Soit, en faisant apparaître la composition de p(t) :

$$I(t) = Eo\ po / Ro + Eo\ pson(t) / Ro = Io + i(t)$$

Nous constatons ainsi l'existence de deux termes Io et i(t) :

- le premier « Io » est un courant constant, indépendant du son, dit « <u>courant de polarisation</u> » ou « <u>courant de repos</u> » ou « <u>courant d'alimentation</u> » ;

- le deuxième « i(t) » est le courant variable proportionnel à l'information sonore.

On obtient donc une tension transmise qui est bien une fonction de « son(t) ».

Un tel système a été mis au point la première fois par Graham Bell en mars 1876. *"Mr. Watson, come here. I want you!"* fut la première phrase transmise par « téléphone » entre Mr Bell et Mr Watson, son assistant, dans son laboratoire.

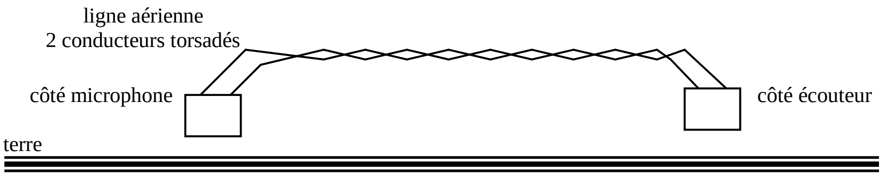
3.1.3 La ligne téléphonique

Ce système peut alors être déployé entre deux interlocuteurs distants grâce à une ligne électrique deux conducteurs.

Rappelons nous à ce sujet que le télégraphe fonctionnait avec un seul conducteur (!), l'autre étant remplacé par « la terre ». Lors de l'utilisation des lignes télégraphiques pour assurer une communication téléphonique (même dans un seul sens), on s'est aperçu que la qualité du son reçu était très médiocre, à cause des « bruits » qui étaient manifestement engendrés sur la ligne. Ces bruits étaient dus à l'induction électromagnétique provenant de sources naturelles ou artificielles extérieures qui s'exerçait dans la grande boucle constituée de la ligne, des équipements d'extrémités et de la terre. Cette aire de boucle étant grande, le flux du champ magnétique se trouvait être grand aussi et donc également la fem parasite induite (loi de Lenz : $e = - d\Phi/dt$).

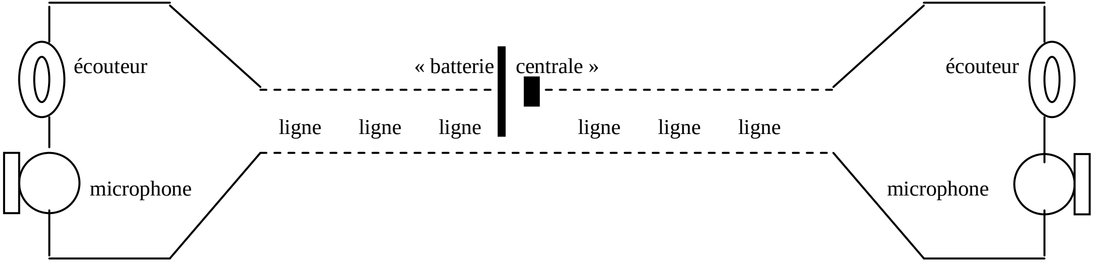
Autant ces bruits n'étaient pas gênants pour la télégraphie qui fonctionne en mode binaire, autant ils étaient préjudiciables à la transmission téléphonique qui fonctionne en mode « de fonction continue », ces bruits étant alors transformés par l'écouteur en sons parasites audibles !

La seule manière d'éviter (ou de limiter cette induction) est d'annuler (ou diminuer considérablement) l'aire de la boucle en adoptant une ligne à deux conducteurs isolés de la terre, ces deux conducteurs étant parallèles et très proches l'un de l'autre et mieux, torsadés pour assurer une auto-annulation de proche en proche d'un résidu possible d'induction (la torsade créant géométriquement des boucles successivement opposées deux à deux, donc créant des fem induites opposées également).



Cet ensemble doit naturellement être « doublé » de manière à pouvoir tenir une conversation (microphone et écouteur à chaque extrémité).
Le doublement strict aurait conduit à une ligne de transmission à 4 conducteurs (2 pour chaque sens de conversation).

Afin d'éviter cela, un montage astucieux fut mis en place avec l'inconvénient de répartir la puissance du microphone en service sur deux écouteurs ; toutefois, cet inconvénient se révèle être un avantage : celui de pouvoir « entendre ce que l'on dit » (le « retour ») ce qui, psychologiquement, « rassure » sur le bon fonctionnement présumé de l'appareil ! (De la même façon qu'une longue période de silence de l'interlocuteur peut faire penser que la liaison est coupée …).



Le concept de « batterie centrale » a fait délocaliser la « pile » de chez les utilisateurs pour la placer chez l'opérateur, celui qui pose et entretient la ligne ! et qui à présent « alimente » la ligne ».

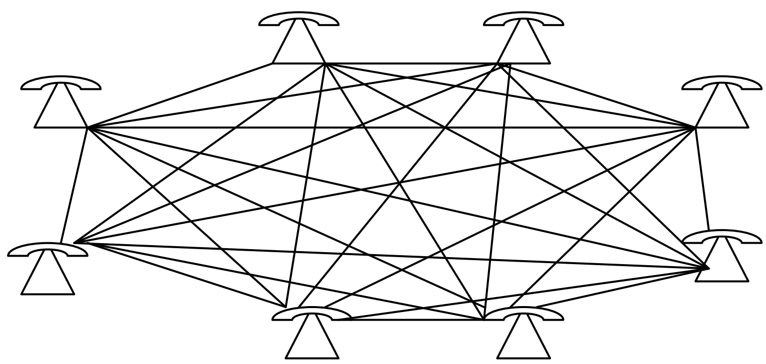
Ce procédé est toujours en vigueur de nos jours :
- nous disposons bien d'un « combiné » (ensemble microphone et écouteur)
- la ligne téléphonique est constituée de 2 conducteurs isolés de la terre
- cette ligne est munie d'une « batterie centrale » (48 volts) qui se trouve au niveau du « central téléphonique ».

En revanche, un poste téléphonique a bénéficié d'un certain nombre de modernisations et d'accessoires : intervention de l'électronique !

3.1.4 La téléphonie publique, le réseau téléphonique

Le système précédent met en relation fixe deux correspondants localisés dans l'espace.
Un service de téléphonie publique doit être capable de mettre en relation, le temps d'une communication, deux correspondants quelconques, « abonnés » de ce service de téléphonie.

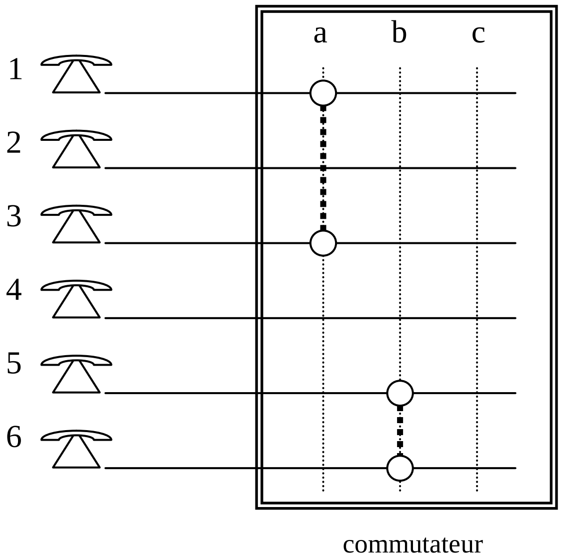
Relier tous les postes téléphoniques entre eux est illusoire et inutile : beaucoup de liaisons seraient parfaitement inutiles car jamais utilisées ! De plus, au delà d'un certain nombre N de postes, chaque poste devant être relié a priori avec les (N-1) autres, le câblage serait inconcevable !



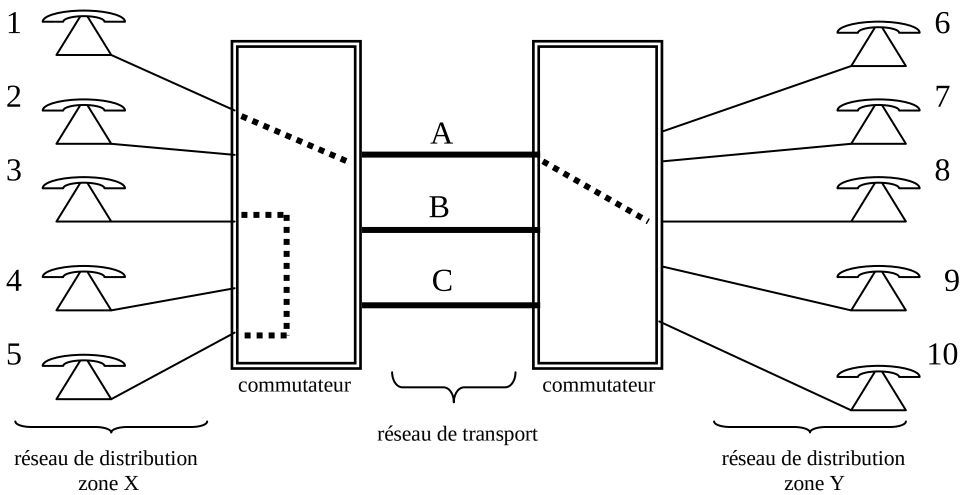
Chaque ligne correspond à deux conducteurs et chaque poste devrait avoir un système de sélection « 1 parmi N-1 ».

On a donc créé deux entités distinctes :
- **les postes téléphoniques** et leur « liaison d'abonné » propre d'une part, l'ensemble de ces liaisons se nomme « **réseau de distribution** » qui converge vers un « commutateur » ;

- **le réseau de transmission commun** (« public ») d'autre part, comprenant un certain nombre de lignes mises à disposition pour effectuer les communications, l'ensemble de ces lignes se nomme soit « circuits de connexion » (interne à un commutateur) soit « **réseau de transport** » (liaisons entre commutateurs).



Exemple d’un service à 6 abonnés et un commutateur à 3 circuits de connexion (abc). Ce service permet 3 communications simultanées.
Le dessin représente une communication entre les abonnés « 1 » et « 3» via le circuit « a » en même temps qu’une autre communication entre les abonnés « 5 » et « 6 ».



Exemple d’un service à 10 abonnés avec 2 zones géographiques X et Y et 3 lignes de réseau de transport inter-zones (ABC). Ce service permet une capacité de 3 communications téléphoniques simultanées en inter zones.
Le dessin représente une communication inter zones entre les abonnés « 1 » et « 8 » via la ligne de transport « A » en même temps qu’une communication « locale » entre les abonnés « 3 » et « 5 ».

Le réseau est hiérarchisé géographiquement : quartier (ou entreprise), secteur, ville, arrondissement, département, région, pays, continent, monde. Ceci pour des raisons de « trafic » : les communications inter-continents sont moins nombreuses que les communications locales pour un instant donné. On peut donc optimiser le nombre de liaisons de transport.

Les connexions dans les commutateurs et inter commutateurs se sont effectuées d'abord manuellement (opérateurs humains avec des câbles de connexion), puis automatiquement (systèmes électromécaniques, puis électroniques, voire « informatiques »)

La numérotation permet le repérage des postes, elle est unique pour chacun des postes dans le monde ! Elle est basée sur un système « à préfixe » permettant d'opérer rapidement à un « routage » (mise en place des diverses connexions au fur et à mesure de la numérotation).

3.1.5 **Des accessoires** (fort utiles)

- tonalité (invitation à numéroter) : synthèse d'un son (la première !)
- sonnerie d'appel : dispositif électrique « basique »
- dispositif de numérotation : « impulsions » (coupures de ligne, codage élémentaire), codage multifréquences (fréquences « vocales », codage un peu plus évolué)
- présentation du numéro : nécessité de transmission de l'information de l'appelant, transmission de la « signalisation »
- répondeur téléphonique : dispositif local, annexe à la téléphonie
- …

3.1.6 **La télécopie** (façon de « détourner » le système pour en faire un autre)

Une feuille de papier imprimée, un éclairage, un capteur optique « tout ou rien », un « codeur phonique » nommé « modulateur » dont le rôle est d'associer un son audible à un état de lumière reçue par le capteur (lumière = « pip », noir = « pup ») et … l'inverse à l'arrivée, la fonction impression (thermique, à aiguilles, à jet d'encre, laser, …) remplaçant la fonction capteur.

3.1.7 **Le « télex » et la téléinformatique**

Télex (télé-texte) : un clavier, un codeur binaire permettant d'attribuer un code différent pour chaque caractère y compris ceux « non-imprimables » (le code « ASCII »), un « codeur phonique » (1 = « pip », 0 = « pup ») et … l'inverse à l'arrivée, la fonction impression (électromécanique au début) remplaçant le clavier.

Téléinformatique : même principe, sauf que la source d'information est déjà un fichier binaire et que, pour augmenter la vitesse de transmission d'un fichier, le « codeur phonique » est plus perfectionné, le terme « modem » pour « modulateur – démodulateur » s'est imposé (00 = pap », 01 = « pep », 10 = « pip », 11 = « pop » ou bien plus encore en définissant une tonalité unique « pup » modulée en amplitude et en phase : modulateurs « QAM » modernes)

3.2 Distances, atténuations et amplification

3.2.1 Le répéteur télégraphique électromagnétique

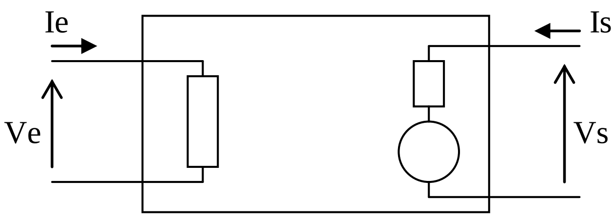
Son défaut est de ne fonctionner qu'en « tout ou rien » (donc en binaire) … mais c'est peut être une idée (!)… car il fonctionne remarquablement bien et est quasi-insensible aux perturbations … c'est justement cette remarque qui fait que la très grande majorités des transmissions électriques d'aujourd'hui est réalisée avec des codages « numériques » (représentation discrète de l'information, de quelque nature que ce soit), dont le plus élémentaire est le codage « binaire » (soit « tout ou rien » !)

3.2.2 Le répéteur téléphonique

Le téléphone électrique, nous l'avons vu, fonctionne de manière « analogique » : la ligne téléphonique véhicule des courants et des tensions qui sont des fonctions continues (et à variations continues) du temps, soit des fonctions du type $x(t)$.
On a besoin d'un système du type $y = kx$ avec $k > 1$
amplificateur (« qui augmente l'amplitude »)
mais attention : augmenter l'amplitude … ou la puissance ?
En effet, augmenter l'amplitude de la tension électrique ne provoque pas nécessairement une augmentation de la puissance (exemple : si le courant délivré diminue proportionnellement à l'augmentation de la tension délivrée, la puissance ne change pas : c'est le cas du transformateur électrique pour lequel le produit « VI » est invariable !).
Un amplificateur est justement caractérisé par la fait que la puissance délivrée est supérieure à la puissance fournie.

3.2.3 Les modèles des « amplificateurs »

Ce sont donc des quadripôles qui vont recevoir une puissance faible (car affaiblie) et qui doivent fournir une puissance plus élevée, avec un facteur de proportionnalité (ou au moins une relation linéaire) entre une grandeur d'entrée (V ou I) et une grandeur de sortie (V ou I).
Ces quadripôles peuvent toujours être représentés par un schéma équivalent externe :



Un amplificateur est, par définition, tel que Ps est supérieure, voire très supérieure, à Pe.
De plus, la puissance d'entrée Pe peut être très faible, voire quasi-nulle.

On peut alors distinguer deux types d'entrées :

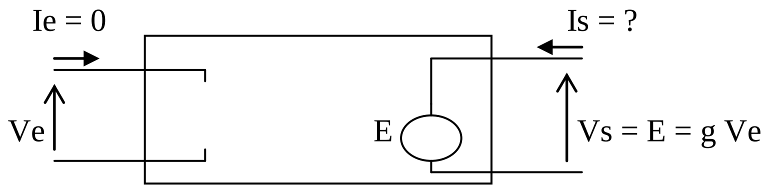
- Pe très faible car Ie est très faible, ce qui conduit à Re grande, l'information utile étant donc contenue dans Ve ;
- Pe très faible car Ve est très faible, ce qui conduit à Re petite, l'information utile étant donc contenue dans Ie.

De la même façon, la sortie peut être vue sous deux aspects :

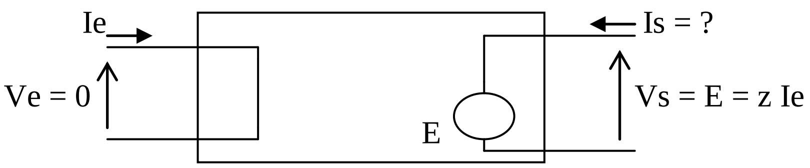
- une source de tension E, mais alors avec une résistance Rs en série petite (pour ne pas consommer inutilement à « l'intérieur » du quadripôle, car cette puissance consommée serait perdue pour la sortie !)
- une source de courant J, mais alors avec une résistance Rs parallèle grande (même raison !)

On peut enfin idéaliser ces propos (résistance faible → résistance nulle ou résistance grande → résistance infinie). On arrive alors à 4 configurations possibles pour la modélisation d'un « amplificateur » ; dans chacun des cas, la puissance d'entrée est nulle et la puissance de sortie est arbitraire, donc « aussi grande que l'on souhaite ».

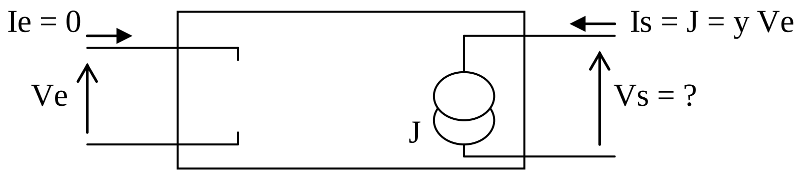
Conversion tension – tension (CVV) de relation : E = g Ve



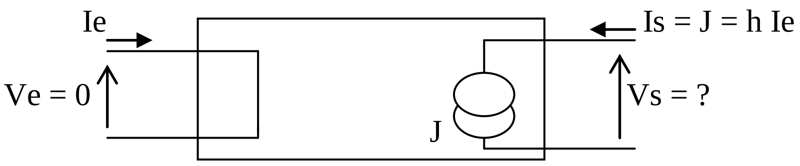
Conversion courant – tension (CIV) de relation : E = z Ie



Conversion tension – courant (CVI) de relation : J = y Ve



Conversion courant – courant (CII) de relation : J = h Ie



Les grandeurs g, z, y et h sont des constantes de proportionnalité, ce sont ces grandeurs qui assurent la « transmission » de l'information à l'intérieur du quadripôle.
Dans tous les cas, on vérifie bien que Pe = 0 et que Ps est arbitraire, elle sera déterminée par la valeur de la résistance d'utilisation.

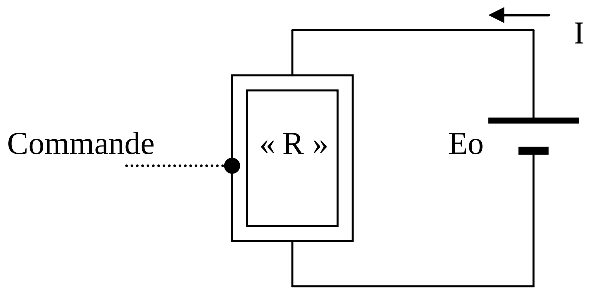
Il apparaît donc que ces quadripôles « fabriquent » de l'énergie, donc contiennent obligatoirement une ou plusieurs sources d'énergie : on les appelle pour cela des quadripôles « actifs ».
En réalité, on peut parfaitement imaginer de tels quadripôles vérifiant les propriétés précédentes, ne contenant pas explicitement de sources d'énergie, à condition qu'il y ait des sources d'énergie externes qui viennent « alimenter » ce quadripôle en énergie (d'où le nom « alimentations » donné à ces sources externes).

Nous sommes ici au niveau du modèle, nous connaissons les propriétés que doit avoir un tel type de quadripôle « amplificateur », il nous faut à présent trouver un composant électrique capable de réaliser ces propriétés.
Les composants électriques traditionnels (R, L, C) ne donnent pas satisfaction car le premier est un dipôle qui consomme de l'énergie électrique en la convertissant de manière irréversible en chaleur (l'inverse de ce qui est attendu !) et les deux suivants ne font que stocker de l'énergie. De même, le transformateur électrique (2 bobines parfaitement couplées magnétiquement) est conservateur de l'énergie (dans le meilleur des cas).

3.2.4 A la recherche du composant « amplificateur »

- Une idée : à partir d'une puissance faible (puissance Pe), utiliser une source externe d'énergie (« alimentation ») dont la tension ou le courant pourra être commandé de façon linéaire par cette puissance faible.

Il peut donc s'agir d'un composant dipôle dont la « résistance » pourrait être commandée, à condition que cette commande se fasse avec une énergie faible.



Si « R » est une fonction de la commande : « R » = R(com),alors le courant dans le circuit vaut : I = Eo / R(com).
La puissance disponible à la sortie vaut Ps = Eo , aussi grande que l'on veut car fonction de la valeur de Eo.

Nous recherchons donc un dispositif de ce type.

- Lee de Forest (1907) : le tube à vide

Un flux électronique dans le vide créé entre une cathode et une anode est contrôlé par un champ électrique provoqué par une différence de potentiel sans consommation de courant (principe de la cathode chauffée par un filament qui émet des électrons et principe de la « grille » qui laisse passer, freine ou repousse ces électrons).

- Bardeen Brattain Shockley (1948) : le transistor à semiconducteur (matériau dont une propriété est que sa résistivité décroît quand la température augmente)

Un courant électrique est contrôlé par un champ électrique provoqué par une différence de potentiel appliquée sur une jonction semiconductrice « en direct » (transistor bipolaire) ou en « en inverse » (transistor à effet de champ). D'une manière comme d'une autre, le courant appliqué est très faible.

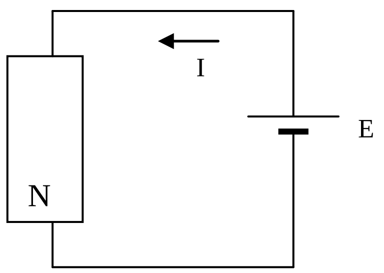
- ? ? ? (à venir) : les « pistes » sont : le transistor magnétique (« spintronique »), le transistor quantique (« à un électron », « nanoélectronique »), le transistor optique (contrôle d'un flux lumineux), …

3.2.5 Le transistor bipolaire

Principe sommaire :

Le transistor bipolaire NPN : composant formé d'un « **semiconducteur** » (silicium en général, atome de la colonne « IV » du tableau de Mendeleiev) que l'on « dope » (accroissement de la conduction électrique grâce à l'introduction d'atomes ayant un électron en plus (N) ou en moins (P) que le semiconducteur (atomes des colonnes respectivement « V » ou « III » du tableau de Mendeleiev).
Un atome ayant un électron en plus est dit « donneur » (d'électron), un atome avec un électron en moins est dit « accepteur » ou « trou ».
Les semiconducteurs ont la propriété d'avoir une conductivité électrique nulle à 0 K, mais croissante avec la température.

Ainsi, un semiconducteur dopé N par exemple se comporte exactement comme une résistance, à la différence près que cette résistance dépend de la température.



$I = E / R$

$R = R(T)$

L'assemblage « **transistor bipolaire** » est réalisé suivant le profil suivant (transistor de type « NPN »)

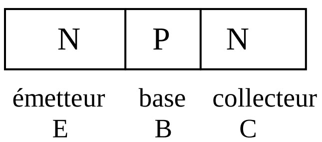
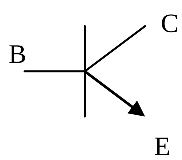
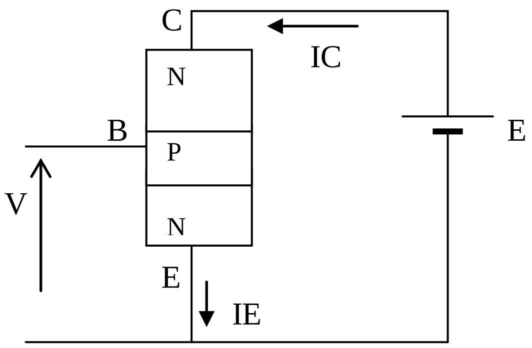


Schéma électrique :



L'idée est d'interposer entre deux électrodes « N » (donc conductrices), une troisième électrode dopée « P » destinée à contrôler le flux d'électrons sortant de l'émetteur et allant vers le collecteur.
Pour cela, l'émetteur, semiconducteur dopé « N » dispose naturellement d'électrons disponibles que l'on va extraire et accélérer vers la base grâce à un champ électrique orienté dans le sens « B→E ». Pour que ces électrons puissent arriver jusqu'au collecteur, un autre champ électrique est placé dans le sens « C→E ».



Si la base était également dopée « N », l'ensemble formerait un barreau conducteur dont la résistance serait fonction du dopage (et de la température). Mais la base est dopée « P », ce qui fait que des électrons issus de l'émetteur peuvent se recombiner avec des « trous » de cette zone « P ». Si le dopage « P » est faible, si la zone « base » est très mince et si enfin le champ électrique est très accélérateur entre C et B, la probabilité pour qu'un électron de E se recombine avec un trou de B est très faible et, par conséquent, il y a une grande probabilité pour que cet électron de E arrive jusqu'à C : le transistor est dit « **passant** ».

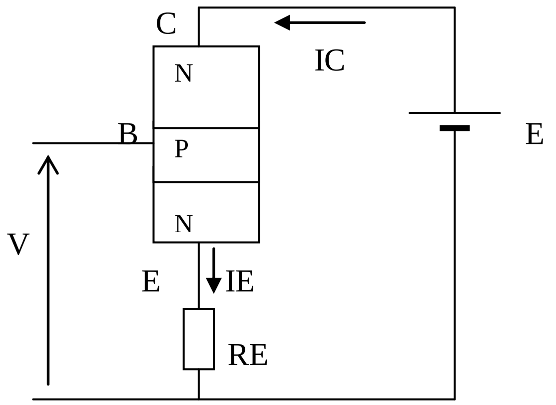
Traduit en termes de tensions et de courants électriques : avec une source E positive et une ddp appliquée V positive également, IE sera positif, IB sera faible et IC sera proche de IE : on est dans un fonctionnement de type « interrupteur fermé », à la différence près que le semiconducteur oppose une résistance. Toutefois, si l'on peut assurer une ddp nulle (ou presque nulle) entre C et E, alors, l'interrupteur fermé devient réellement équivalent à un fil. Ceci peut se faire en plaçant une résistance RC en série avec le collecteur, on a alors :
VCE (= VC – VE) = E – RC IC, il suffit de choisir la valeur de RC telle que VCE s'annule. Le transistor est dit alors « **saturé** ».

En revanche, si le champ électrique entre B et E est nul ou négatif dans le sens E→B, alors aucun électron de E ne pourra aller dans la zone B et par conséquent, aucun électron n'arrivera en C : le transistor est dit « **bloqué** ».

Traduit en termes de tensions et de courants électriques : avec une source E positive et une ddp appliquée V négative, IE sera nul et IC sera nul aussi : on est dans un fonctionnement de type « interrupteur ouvert ».

Nous avons dès lors un fonctionnement de type « interrupteur commandé » : interrupteur entre E et C, commandé par la tension électrique entre B et E.

Nous recherchons à présent un fonctionnement linéaire, c'est-à-dire un courant dans le transistor (IC) qui serait commandé par la tension V de manière proportionnelle (ou tout au moins linéaire).
Pour cela, il est d'abord nécessaire de transformer les variations de V en variations de IE, cette transformation s'effectue grâce à une simple résistance intercalée en série avec l'émetteur du transistor.



Nous avons alors :

$$V = VBE + RE\ IE$$

L'étude détaillée du transistor (voir après) nous montre que, tant que V est positive, IE aussi et la conséquence est que VBE est alors quasi invariable (fonction du semiconducteur, du dopage et de la température), égale à 0,65 V (silicium, dopage « normal », 300 K). Donc :

$$V = 0,65 + RE\ IE \quad \text{soit} \quad IE = (V - 0,65) / RE$$

Le courant IC est très proche du courant IE : $IC = \alpha\ IE$ α étant un facteur intrinsèque du transistor (α = 0,995 à 0,998 pour un transistor « moderne », dépend bien sûr du dopage et de l'épaisseur de la base).

On obtient donc enfin :

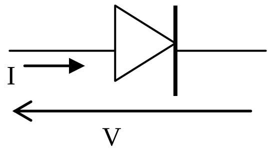
$$IC = (\alpha / RE) . (V - 0,65)$$

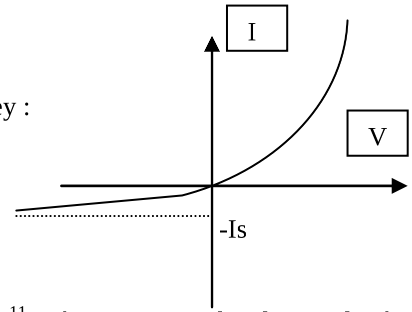
Nous avons dès lors un fonctionnement linéaire, qui peut donc être utilisé pour réaliser une quadripôle amplificateur.

Le fait que le courant de base soit très petit met en évidence une commande en tension et le fait que ce soit le courant collecteur qui soit contrôlé met en évidence une sortie sous forme d'une source de courant : un transistor (muni de sa résistance RE et de sa source E) va donc réaliser un quadripôle amplificateur de type « CVI ».

Etude détaillée :
Ce composant est un assemblage de deux « diodes », l'une apparaît entre émetteur et base, l'autre entre base et collecteur :

La diode semiconductrice :





C'est un assemblage PN (« jonction PN ») qui est décrit par la loi de Shockley :

$I = Is\ (\exp(qV/kT) - 1)$

q : charge de l'électron = $1{,}6\ 10^{-19}$ C
k : constante de Boltzmann = $1{,}38\ 10^{-23}$ J. K^{-1}
T : température thermodynamique = 300 K (ambiante moyenne)
$kT/q = U_T = 26$ mV (à 300 K)
Is = dépend du matériau, du dopage, de la température (10 pA = 10^{-11} A à 300 K pour du silicium dopé « normalement »

Compte tenu de la loi exponentielle d'une part et de la valeur très faible du courant Is d'autre part, on constate que l'on peut distinguer trois zones :

$V < 0$: $-Is < I < 0$ donc $I \approx 0$

$0 < V < Vd$: $0 < I < Imax$ (Vd de l'ordre de 0.7 V pour une diode au silicium)

$V > Vd$: I prend des valeurs extrêmement importantes et conduit à la destruction de la diode !

Pour la deuxième zone, le calcul de la pente de l'exponentielle amène à : $dV/dI = U_T / I$ (homogène à une résistance, dite « résistance de la diode rd ». Pour des valeurs non nulles ou très petites de I, cette valeur de résistance est très faible (26 Ω pour 1 mA et 0.26 Ω pour 100 mA !).
Pour la même raison, on atteint très rapidement des valeurs proches de Vd, dès que I est positif.

Aussi, le schéma équivalent de la diode pour les deux premières zones est-il très simple :

V < 0

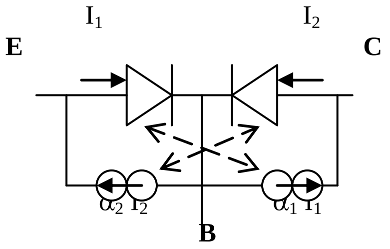
I > 0

Selon les valeurs de résistances et de tensions existantes dans une maille contenant la diode, les valeurs rd ou Vd ou les deux peuvent être négligées, on arrive alors à une diode équivalente à un simple fil, pour I > 0.

Condition si …	**Conséquence** … alors	**Etat de la diode**
V < 0	**I = 0**	**Diode « bloquée »**
I > 0	**V = 0**	**Diode « passante »**

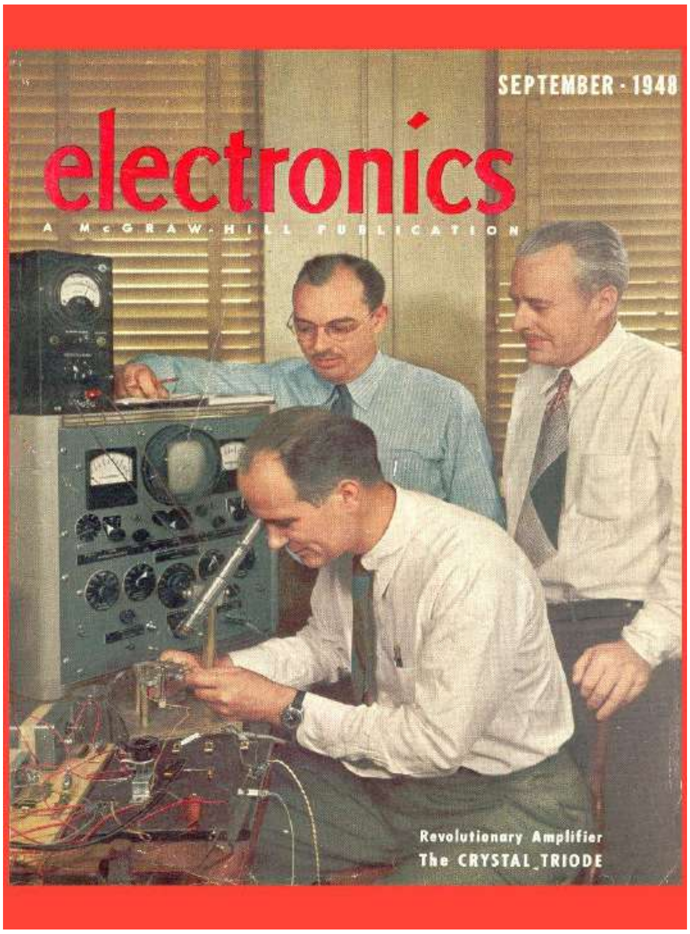
L’effet transistor :

La découverte de « l’effet transistor » (Bardeen – Brattain – Schokley en 1948) fait apparaître deux courants additionnels liés aux courants directs (c’est-à-dire dans le sens de la conduction de la diode) respectifs dans chacune des deux diodes (dès lors que la base est « mince ») :



α_1 et α_2 sont inférieurs mais proches de 1.

Ce nouveau composant, initialement appelé « crystal triode » pour « composant à trois électrodes fabriqué à partir d’un cristal (de germanium à l’époque) », a permis de réaliser un amplificateur « révolutionnaire » … par rapport au tube à vide.

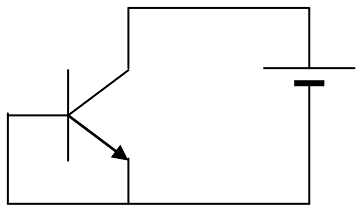


3.2.6 Utilisation pratique du transistor bipolaire en interrupteur

« Régime bloqué »
On place la diode BE en régime bloqué (tension nulle de B vers E) et la diode BC en régime bloqué également (tension positive de B vers C). Ces conditions s'écrivent :

$$V_{BE} = 0 \text{ et } V_{CB} > 0.$$

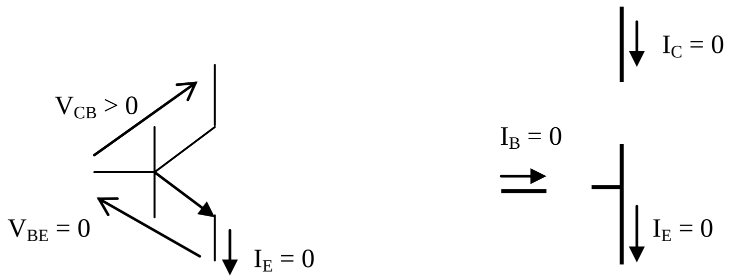
Ceci se réalise aisément en adoptant la configuration suivante :



Sous ces conditions, le transistor présente les propriétés suivantes :

$$I_C = 0 \quad \text{et } I_B = 0 \ \ (\text{donc } I_E = 0)$$

Le transistor est équivalent à un circuit ouvert.

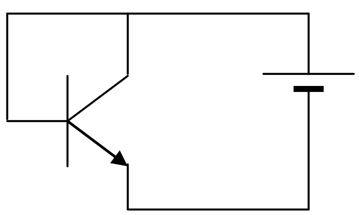


Le transistor se comporte alors comme un interrupteur ouvert entre émetteur et collecteur.

« Régime saturé »
On place la diode BE en régime passant (courant circulant de B vers E) et la diode BC en régime passant également (courant circulant de C vers B). Ces conditions s'écrivent :

$$I_E > 0 \text{ et } I_C > 0.$$

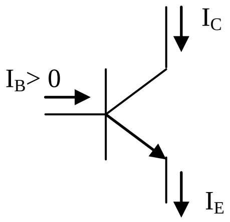
Ceci se réalise aisément en adoptant la configuration suivante :

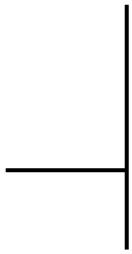


Sous ces conditions, le transistor présente les propriétés suivantes :

$$V_{BE} \cong 0 \quad \text{et} \quad V_{CB} \cong 0$$

Le transistor est équivalent à un circuit électrique fermé.





Le transistor se comporte alors comme un interrupteur fermé entre émetteur et collecteur.

« Régimes bloqué et saturé »
Ces deux régimes sont à la base de la réalisation de fonctions logiques : réalisation d'un interrupteur commandé ouvert (transistor bloqué) ou fermé (transistor saturé).
Des associations de transistors permettent de réaliser des opérations logiques élémentaires (non, et, ou) puis des fonctions évoluées.
Ces fonctions évoluées permettent à leur tour de réaliser des systèmes logiques câblés, puis programmés.
En outre, ces mêmes transistors utilisés en interrupteur permettent de réaliser la commutation de circuits téléphoniques.

3.2.7 <u>Utilisation pratique du transistor bipolaire en amplificateur linéaire</u>

« Régime normal ou passant ou linéaire »

On utilise ce composant en plaçant la diode BE en régime passant (courant circulant de B vers E) et la diode BC en régime bloqué (tension positive de B vers C). Ces conditions s'écrivent :

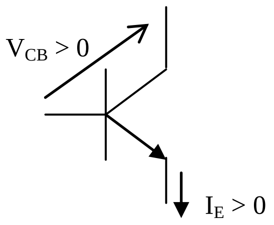
$$I_E > 0 \text{ et } V_{CB} > 0.$$

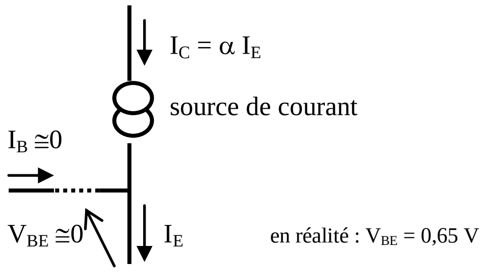
Ces conditions correspondent à établir un champ électrique dans le sens B$\rightarrow$E (diode BE conductrice, les électrons de E vont vers B) et un autre champ électrique dans le sens C$\rightarrow$B (diode BC bloquée, mais les électrons non recombinés dans la base sont bien attirés vers C)

Sous ces conditions, le transistor présente les propriétés suivantes :

$$I_C = \alpha\, I_E \quad (\alpha \text{ inférieur mais très proche de 1) et } V_{BE} \cong 0 \text{ (positive).}$$

La conséquence est que I_B est presque nul ($I_B = I_E - I_C$). La résistance d'entrée de la base est ainsi très grande, assimilable à l'infini. Le collecteur délivre un courant proportionnel à celui de l'émetteur et se comporte comme une source de courant presque idéale. Un schéma équivalent du transistor ainsi configuré (« polarisé ») est donc :

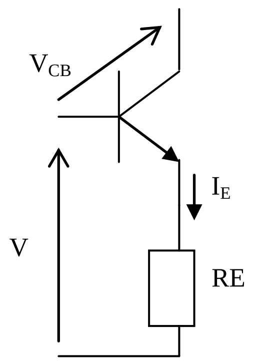




Les propriétés à l'entrée (base) et à la sortie (collecteur) font naturellement penser à des performances d'amplification dans la mesure où la puissance à l'entrée est quasi-nulle ($I_B \cong 0$ et $V_{BE} \cong 0$) alors que la puissance disponible à la sortie ne l'est pas, mieux, « semble » être aussi grande que l'on veut car elle est fournie par une source de courant pouvant – a priori – délivrer une puissance quelconque de la forme : αI_C . Ru, si Ru est la résistance d'utilisation branchée en parallèle avec cette source de courant.

Le circuit ainsi réalisé se comporte bien comme un quadripôle de type CVI.

On peut donc construire un système « amplificateur » qui serait constitué progressivement de la façon suivante :



La résistance RE permet de créer un courant I_E proportionnel à V1 :

$$I_E = V / RE$$

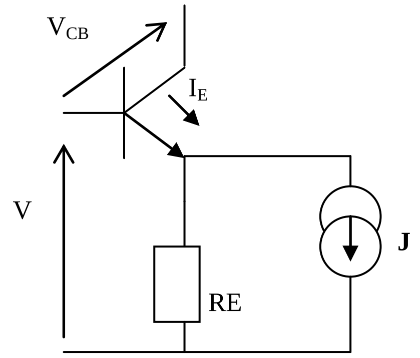
à condition bien sûr que la diode BE soit passante. Pour cela, I_E doit toujours être positif, quelque soit la valeur de V. Afin d'assurer cette condition, il est nécessaire d'ajouter une source de courant J, en effet cette condition est manifestement fausse dès que V est négative !

En ajoutant une source de courant J, on obtient : $I_E = V / RE + J$

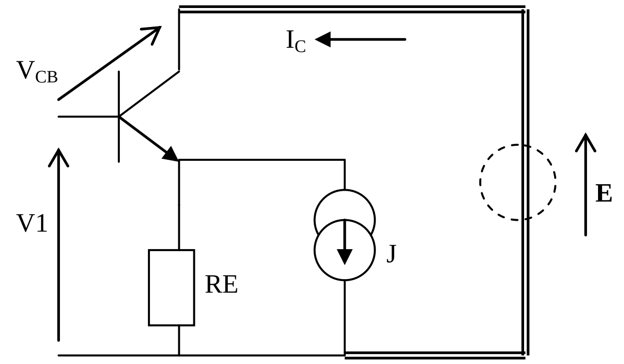
La condition $I_E > 0$ devient donc :

$$V / RE + J > 0, \text{ soit } V > - J\, RE$$

Donc V peut à présent être positive (sans limites), nulle, ou négative (limite égale à – J RE)



Pour permettre « l'effet transistor », il est nécessaire à présent d'assurer $V_{CB} > 0$ (donc le blocage de la diode BC), toujours pour toute valeur de V. Ceci est possible en considérant une maille {V – V_{CB} – fermeture du circuit} :



La fermeture du circuit est représentée par le double trait

Il apparaît clairement la relation : $V + V_{CB} = 0$ d'où$V_{CB} = - V$
Le transistor ne fonctionnera donc pas pour les valeurs positives de V (car alors V_{CB} serait négative !)
On peut remédier à ce problème en insérant dans la maille une source de tension E, comme indiqué en pointillés sur la figure ci-dessus.

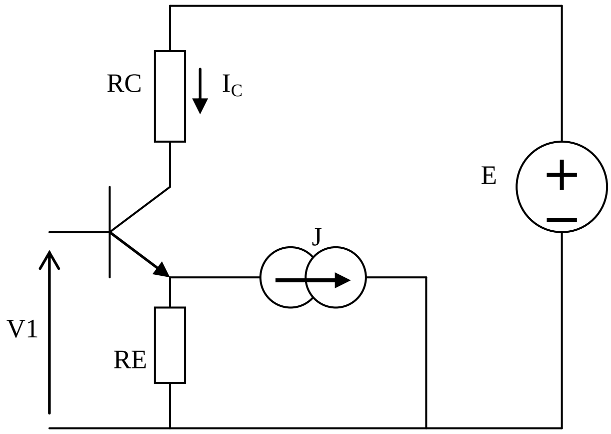
On obtient alors la relation suivante : $V + V_{CB} – E = 0$ d'où $V_{CB} = E – V$. Le transistor fonctionnera donc tant que V restera inférieur à E (on choisira E en conséquence).

L'amplificateur CVI fonctionne à présent, avec une relation entre V et I_C :

$$I_C = \alpha\, I_E = \alpha\, (V / RE + J) = \alpha/RE\ V + \alpha J$$

Cette relation est bien linéaire avec le coefficient $\boldsymbol{\alpha}$**/RE.**

Enfin, on peut « récupérer » l'information sur une résistance d'utilisation placée en série dans le circuit du collecteur (dans lequel circule I_C) :



Nous retrouvons bien l'appel à deux sources d'énergie :
E est une source d'énergie permettant d'assurer la condition $V_{CB} > 0$
J est une source d'énergie permettant d'assurer la condition $I_E > 0$

L'amplificateur est terminé !

3.3 Distances, perturbations et filtrage

3.3.1 Origine et effet des perturbations : bruits, interférences, diaphonie

Lors d'une transmission réelle (c'est-à-dire avec une ligne téléphonique d'une longueur importante), il est apparu que le signal reçu était entaché de **perturbations**, donnant lieu à des « bruits » parasites audibles venant donc gêner la transmission de l'information, cela allant de la simple gêne (altération du confort) à l'impossibilité de compréhension du message.

Une étude sommaire de cette perturbation montre que celle-ci est de caractère **additif** : il s'agit d'un « courant perturbateur » qui vient s'ajouter, dans la ligne, au courant utile représentant la parole :

$$\mathbf{I = I_S + I_B}$$

I_S représentant le courant « signal » et I_B représentant le courant « bruit »

Quelle peut-être l'explication de ce courant I_B ?

Un courant électrique naît d'un mouvement électronique. Ce dernier naît lui-même d'une sollicitation « électrique » volontaire (force électromotrice d'origine chimique, magnétique, mécanique, optique, …).

Mais il peut naître également de **l'état thermique d'un conducteur** (la température mesurant précisément l'état d'agitation des atomes et molécules, température permettant de définir une énergie interne propre à pouvoir participer à la libération d'électrons), c'est le cas du « bruit thermique », ce **bruit thermique** naît ainsi dans la ligne téléphonique elle-même (bruit intrinsèque). Ce bruit est incontournable car pour l'annuler il faudrait que la température thermodynamique (absolue) soit nulle (0 K = - 273 °C) ou, en tous cas, très faible (rappel : la température ambiante moyenne est de l'ordre de 300 K)

Il peut naître aussi du **phénomène d'induction électromagnétique** mutuelle entre conducteurs parcourus par des courants variables dans le temps (ce sont des « interférences »). Ces interférences résultent néanmoins de deux processus : le premier est la **production d'un champ électromagnétique** par un courant variable (incontournable), le deuxième est le résultat de **l'induction de ce champ** sur les conducteurs transportant le signal utile (on peut théoriquement supprimer cette induction en entourant complètement ces conducteurs d'un autre conducteur porté à un potentiel constant, celui de la terre par exemple : c'est le principe de la « cage de Faraday ».

Enfin, s'il s'agit de conducteurs transportant le même type d'information, on parle plutôt de « **diaphonie** » (exemple : un câble téléphonique « multipaires » rassemblant, comme son nom l'indique, plusieurs lignes téléphoniques, 1 ligne = 1 « paire » de conducteurs)

Ces différents phénomènes physiques contribuent à ce courant I_B.

L'expression de celui-ci est naturellement impossible à déterminer (beaucoup trop de facteurs, surtout pour le bruit thermique !), I_B est pour cela qualifié de « **courant aléatoire** ».
De ce fait, il est impossible de retrancher ce courant au courant I reçu et par conséquent, le courant utile I_S est impossible à retrouver : **l'information est définitivement corrompue** !

3.3.2 Notion de représentation fréquentielle, théorème de Fourier

Bien entendu, il est inacceptable que l'information ne puisse pas être reçue correctement.
Il faut donc rechercher une manière de pouvoir « séparer » le signal du bruit.

On peut d'abord examiner les fonctions I_S (t) et I_B(t).
Cet examen nous montre que **ces deux fonctions se « ressemblent »** par les caractères suivants :

- toutes les deux sont de valeur moyenne nulle
- toutes les deux sont des fonctions continues à variations continues
- aucune des deux n'est périodique ou ne révèle des caractéristiques particulières : en effet, le signal représentant la parole est « quelconque » en ce sens où il n'y a pas de modèles particuliers pour les sons parlés (en fait si, ce sont les « phonèmes », mais ils sont très nombreux et, à l'époque du développement de la téléphonie, trop difficiles à identifier) ; par ailleurs, comme nous l'avons vu, le bruit est de caractère « aléatoire ».

Il est donc nécessaire de rechercher une autre représentation de ces signaux pour tenter de trouver des éléments discriminants.

Le **théorème de Fourier** montre qu'une vibration périodique peut toujours être décomposée en une somme discrète de vibrations sinusoïdales d'amplitudes et de phases variées mais de périodes égale et sous-muliples de la période de la vibration périodique (analyse de Fourier) ; les fréquences associées étant alors toutes multiples de la fréquence fondamentale dont la valeur est égale à l'inverse de la période, les intervalles entre fréquences successives ayant cette même valeur. Ces fréquences sont dites « harmoniques ».

Soit s(t) une fonction périodique de période T :

- si cette fonction est sinusoïdale, elle s'écrit :

$$s(t) = C \cos (2 \pi F t + \varphi) \quad \text{avec } F = 1/T$$

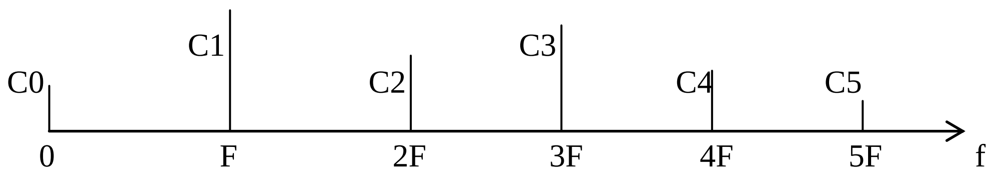
ou encore : $$s(t) = A \cos (2 \pi F t) + B \sin (2 \pi F t)$$

- si cette fonction n'est pas sinusoïdale, le théorème de Fourier prouve que cette fonction peut s'écrire :

$$\mathbf{s(t) = C_0 + C_1 \cos (2 \pi F t + \varphi_1) + C_2 \cos (4 \pi F t + \varphi_2) + C_3 \cos (6 \pi F t + \varphi_3) + \ldots}$$

ou encore, la même chose en faisant ainsi apparaître les termes en « F », en « 2F », en « 3F », etc …
On appelle « **SPECTRE** » de la fonction s(t) la donnée de l'ensemble des couples $\{C_i \ \ \varphi_i \}$ pour chaque fréquence $f_i = i\ f$.

Ce « spectre » peut être représenté symboliquement (et partiellement, car module seul, sans la phase) par des segments de droite parallèles à l'axe des ordonnées, dont la longueur vaut C_i et l'abscisse $f_i = i\ f$.



Le calcul des « **coefficients de Fourier** » se réalise à l'aide des relation suivantes (qui sont des produits scalaires, donc des « projections » de la fonction sur des « axes f_i ») :

$A_i = 1/T \int_0^T s(t) . \cos (2\pi\ i\ F\ t)\ dt$ $\qquad$ $B_i = 1/T \int_0^T s(t) . \sin (2\pi\ i\ F\ t)\ dt$

Puis : $C_i^2 = A_i^2 + B_i^2$ $\qquad$ $\varphi_i = \arctan (B_i / A_i)$

En appliquant un tel théorème à la voix (qui est une vibration, mais très rarement périodique), on passe **de la notion de somme discrète à la notion de somme continue** : en effet, une vibration non périodique peut toujours être considérée comme étant périodique, mais de période infinie (!) , les intervalles de fréquences associés étant alors nuls (!).

On a alors recours à une transformation d'une fonction s(t) en une autre fonction S(f), le passage de l'une à l'autre se faisant grâce à la « **transformée de Fourier** » qui, à partir d'une fonction réelle quelconque du temps (mais « de carré intégrable ») associe une fonction de la variable complexe :

$$S(f) = \int_{-\infty}^{+\infty} s(t)\ \exp (- 2\ \pi\ j\ f\ t)\ dt \qquad \text{avec } j^2 = -1$$

Cette fonction S(f) admet un module |S(f)| et une phase arg (S(f)).
C'est le module (voire le module au carré) qui est en général qualifié de « spectre ».

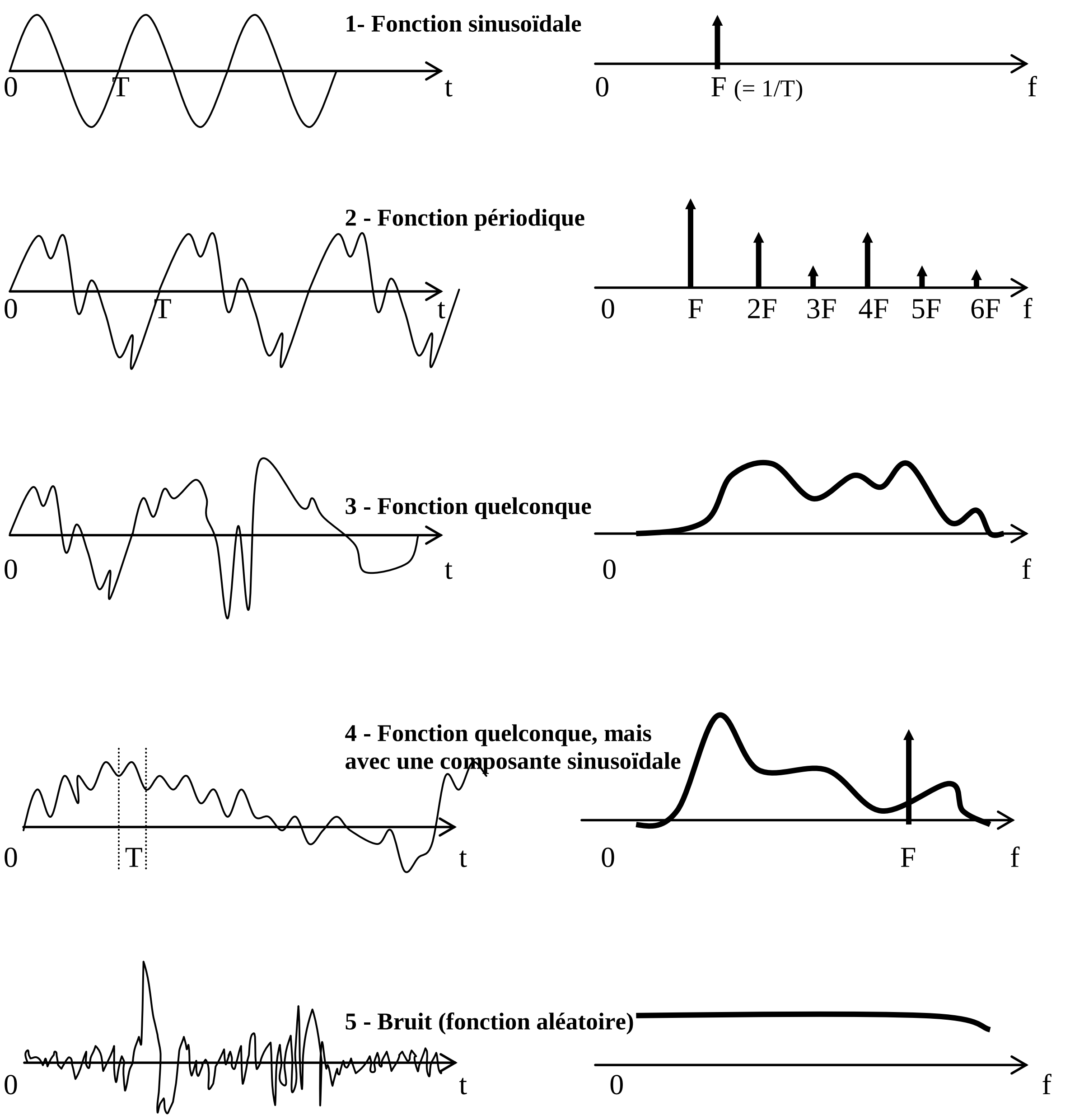
Ce module n'est pas borné en fréquences en règle générale, toutefois, on remarque que, pour les signaux réalistes (signaux physiques, la parole et les sons en particulier),au delà d'une certaine fréquence et parfois, en deçà d'une certaine fréquence, ce module décroît régulièrement pour quasiment s'annuler.
Dans ces conditions, on peut définir un « intervalle de fréquences » dit « bande » dans laquelle les valeurs spectrales sont « significatives ».

En règle générale également, les spectres sont des fonctions continues de la fréquence, toutefois, pour des fonctions périodiques du temps, ou des fonctions du temps ayant une ou plusieurs composantes périodiques, on retrouve naturellement les « raies » (segments de droites) du théorème de Fourier : ces raies représentent les valeurs ponctuelles C_i correspondantes aux fonctions sinusoïdales de la décomposition de Fourier de (ou des) composantes périodique(s).

En réalité ces « raies » (le nom vient de l'optique : raies d'absorption dans un spectre lumineux, utilisées en astrophysique pour contribuer à déterminer des types de molécules dans une étoile) sont des fonctions mathématiques très particulières, appartenant à la famille dites des « distributions ». Ce sont des fonctions qui ne s'appliquent qu'à d'autres pour donner finalement une valeur. Les « raies » sont des « distributions de Dirac », notées δ et définies par : $\delta(f)\ S(f) = \delta(f)\ S(0)$ donnant ainsi la valeur en f = 0 de la fonction S(f). On a par ailleurs, pour tout fo : $\delta(f\text{-}fo)\ S(f) = \delta(f\text{-}fo)\ S(fo)$ donnant de même la valeur en f = fo de la fonction S(f). Ces valeurs sont des valeurs ponctuelles, la représentation par un segment de droite est purement conventionnelle. On représente souvent ces « raies » par une « flèche » plutôt que par un segment, la valeur S(0) ou S(fo) apparaissant à coté de la « flèche » (cette « flèche » n'a rien à voir avec un vecteur !)

Signaux ou bruits
Fonctions du temps

Spectres S(f)
Fonctions de la fréquence



On remarque, sur ces illustrations que les spectres sont très différentes suivant les types de fonctions, ce qui permet de penser qu'une opération de « tri » serait plus abordable dans le domaine de la représentation fréquentielle que dans la représentation temporelle : ce « tri » se nomme « filtrage ».

3.3.3 <u>Notion de filtrage</u>

Le filtrage est donc une opération permettant de **« trier » des informations utiles** dans un signal à plusieurs composantes, dont certaines sont des perturbations.

Le filtrage est aussi, par extension, un moyen de **modifier l'information** en modifiant son spectre.
Le filtrage est également enfin un moyen que l'on peut employer dans le but de **rechercher une caractéristique particulière** (comme une périodicité « cachée » par exemple

Le filtrage est donc une opération très riche quant à ses applications, quelques exemples :

- pour une image, quelle est la luminosité moyenne ? (caractéristique scalaire)
- pour une image, où sont situés les contours ? (caractéristique géométrique)
- pour une image, comment réaliser (ou combattre) un flou ? (modification)
- pour un texte, quelle est la fréquence de la lettre « E » ? (caractéristique scalaire)
- pour un texte, remplacer les minuscules par des majuscules (modification)
- pour un son, quelle est la répartition entre les « basses » et les « aiguës » ? (caractéristique scalaire)
- pour un son, renforcer les « basses » (modification)
- …

Tous ces exemples sont des cas de filtrage.

Nous allons nous intéresser plus spécialement au « filtrage fréquentiel » car, dans le cas de la téléphonie, il s'agit de « trier » l'information utile du signal reçu, entaché de perturbations.

La **parole** définit une famille de signaux qui, au sens de s(t) n'a pas de propriété particulière sauf d'avoir une valeur moyenne nulle !
En revanche, au sens de S(f), on remarque (par la mesure, mais aussi en analysant le processus vocal) que le spectre de la voix parlé présente des valeurs très significatives dans la bande [300 Hz – 3400 Hz], alors, qu'en dehors de cette bande, les valeurs spectrales sont faible voire nulles.

Une normalisation internationale a ainsi fixé la « bande téléphonique » à cette valeur de bande [300 Hz – 3400 Hz].

Désormais, notre système de transmission ne doit plus transmettre que des signaux correspondant à cette bande. De ce fait, toute fonction ayant des valeurs spectrales hors de cette bande peut être considérée comme étant une perturbation et doit être éliminée à la réception et bien sûr à chaque amplification.

Nous avons vu que les perturbations étaient de plusieurs types :

- **<u>le bruit thermique intrinsèque et incontournable</u>**, son spectre a une caractéristique remarquable : c'est une fonction (théoriquement) constante : on dit que c'est un « **bruit blanc »,** terminologie qui n'est qu'une simple analogie avec le cas de la lumière blanche pour laquelle (pour l'œil humain), toutes les couleurs (fréquences) sont représentées avec la même contribution.

Un corollaire de cette définition du bruit blanc est que sa **puissance est infinie** : en effet, toutes les composantes fréquentielles de 0 à l'infini étant représentées, la composition de Fourier en ferait une fonction d'amplitude infinie (physiquement impossible).
On comprend aisément que, dans ces conditions, le signal téléphonique (traduisant la parole) puisse rapidement être « noyé » dans le bruit (une bande infinie étant largement plus grande que la bande téléphonique …).

Toutefois, le remède est clair, même s'il est imparfait : un filtrage efficace de la fonction reçue dans la bande [300 Hz – 3400 Hz] élimine une puissance partielle énorme de bruit : toute celle contenue dans la bande [0 – 300 Hz] $\cup$ [3400 Hz - ∞].

Il ne subsistera que la puissance partielle de bruit dans la bande [300 Hz – 3400 Hz] qui vaut : $N = kT\,B$

$k = 1{,}38\ 10^{-23}\ JK^{-1}$ constante de Boltzman
$T = 300\ K$
$B = 3100\ Hz$ largeur de la bande considérée

Il faut remarquer que ce résultat est indépendant de la ligne et de sa longueur, mais ne dépend (en première approximation) que de la température à laquelle cette ligne (mais aussi l'écouteur) est soumise et surtout que de la largeur de bande de fréquences considérée.

On obtient alors une puissance de bruit N (« N » comme « *noise* ») environ égale à 10^{-17} W, soit 10^{14} fois plus faible que le niveau « ligne » téléphonique (10^{-3} W).

Il est clair que ce n'est pas ce type de perturbation qui soit la plus gênante.

Le calcul sommaire précédent appelle toutefois à deux remarques :
- la ligne n'intervient pas directement pour la puissance de bruit, mais indirectement car elle provoque une atténuation du signal utile si bien qu'au cours des kilomètres parcourus, le rapport puissance de signal sur puissance de bruit décroît ;
- lorsque l'on passe au travers d'un amplificateur, le bruit à l'entrée est naturellement amplifié également, mais, en plus, il faut ajouter le bruit propre engendré par les composants qui réalisent l'amplificateur. On appelle « facteur de bruit » le rapport entre le bruit réel en sortie et le bruit incontournable dû à l'amplification du bruit en entrée ; un amplificateur performant devrait avoir un facteru de bruit proche de l'unité.

- **les perturbations dues à l'induction électromagnétique** engendrée par d'autres courants étrangers à l'application de téléphonie, leurs spectres sont très divers car dépendent des autres applications telles que transport d'énergie électrique (essentiellement 50 Hz et les premières harmoniques), machines électriques, radiodiffusion hertzienne, … Il n'y a aucune raison pour qu'il y ait une densité de puissance particulièrement importante dans la bande [300 Hz – 3400 Hz].

De plus, on peut notablement atténuer ces effets d'induction en adoptant :
- soit un écran magnétique (« cage de Faraday ») consistant à enfermer la ligne dans une gaine conductrice équipotentielle (« écran », souvent confondu avec le « blindage », ce dernier permettant une bonne tenue mécanique, résistance à l'arrachement et/ou résistance au percement) ;
- soit l'adoption de la « torsade » des deux conducteurs de la ligne de manière à réduire au maximum la surface dans laquelle s'exerce le flux du champ magnétique (adopté pour des liaisons courtes, chez l'abonné par exemple) ;
- soit les deux (adopté pour des liaisons longues).

Ici encore, le filtrage permet d'éliminer naturellement une très grande partie de la puissance des perturbations qui se trouvent « hors bande ».

- la **diaphonie**, phénomène identique au précédent, à la différence près qu'il s'agit d'une ligne téléphonique perturbée par d'autres lignes téléphoniques, surtout quand elles sont situées dans le même câble « multipaires » !

Là encore, la technique de la torsade des paires permet d'affaiblir la diaphonie. En plus, une torsade globale de l'ensemble des paires entre elles permet de ne pas favoriser une diaphonie entre deux paires qui seraient toujours proches, en effet, cette torsade globale engendre une somme de diaphonies partielles, créant ainsi un genre de « brouhaha » pouvant être assimilé à un « bruit de fond » … finalement plus acceptable (au sens auditif comme au sens de la confidentialité) que si une autre conversation était omniprésente !

Cette fois, bien entendu, le filtrage n'est d'aucun recours !

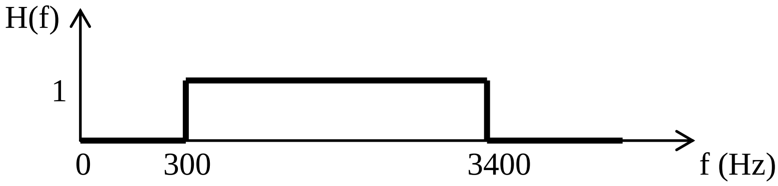
3.3.4 <u>Caractérisation des filtres, aspects fréquentiel et temporel :</u>

Un filtre est un dispositif de type « quadripôle » qui va s'insérer dans une chaîne de transmission (dans notre cas) dans le but de permettre le « tri » de l'information utile parmi les différentes composantes du signal reçu et perturbé. Nous nous limiterons ici à cette application des filtres, compte tenu de nos besoins pour notre système téléphonique.

Le fait que la bande de fréquences doive être limitée à la valeur [300 Hz – 3400 Hz], fait qu'un filtre peut être modélisé par un système qui assurerait une multiplication du spectre du signal d'entrée par une fonction égale à 1 dans la bande [300 Hz – 3400 Hz] et égale à 0 partout ailleurs.

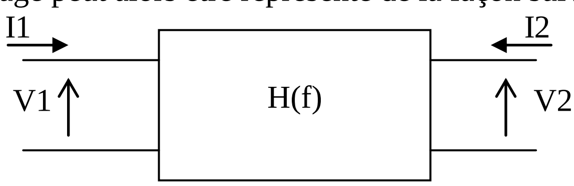
L'expression et le graphe d'une telle fonction de la fréquence, que l'on notera H(f), sont immédiates :

$$H(f) = \left| \begin{array}{ll} 1 \quad \text{pour} & 300\ \text{Hz} < f < 3400\ \text{Hz} \\ 0 \quad \text{ailleurs} & \end{array} \right.$$



On dit aussi qu'il s'agit d'une fonction « porte » ou encore fonction « fenêtre »

Le système de filtrage peut alors être représenté de la façon suivante :



Une difficulté est que la fonction H est définie en fonction de f, alors que les courants et les tensions sont des fonctions du temps ! (rappelons qu'il s'agit de la représentation électrique de la parole !)

La caractérisation d'un filtre nécessite donc une cohérence entre les variables des fonctions.

On peut alors avancer deux méthodes :

- la première est la **méthode fréquentielle**, où la variable choisie est la fréquence. C'est la méthode historique, elle a l'avantage de bien « visualiser » l'effet fréquentiel du filtre. Elle nécessite de traduire les signaux d'entrée et de sortie en termes de spectres, donc rigoureusement, de réaliser la transformation de Fourier du signal d'entrée, d'appliquer (par multiplication) la fonction de filtrage pour obtenir le spectre de sortie et, enfin, de réaliser la transformée de Fourier inverse pour retrouver le signal de sortie.

Ces calculs sont en général longs et compliquées pour des signaux quelconques, en revanche, ils sont élémentaires pour des fonctions d'entrée et de sortie sinusoïdaux.
De ce fait, la caractérisation d'un filtre par la méthode fréquentielle se bornera à la donnée de la fonction H(f) ainsi qu'au comportement de ce filtre en régime sinusoïdal.
Tous les autres comportements pourront s'en déduire par application du théorème de Fourier étendu (« toute fonction peut se représenter sous forme d'une somme discrète ou continue, finie ou infinie de fonctions sinusoïdales »).

Cette fonction H(f) se nomme « **fonction de transfert** ».

$$x(t) \Leftrightarrow X(f) \Longrightarrow \boxed{H(f)} \Longrightarrow Y(f) \Leftrightarrow y(t)$$

$$\mathbf{Y(f) = X(f) \, . \, H(f)}$$

Les expressions de X(f) et de Y(f) n'ont pas besoin d'être explicitées. La caractérisation se faisant en régime sinusoïdal, chaque composante de fréquence fo donnée subit une multiplication par la valeur H(fo) correspondante.

Remarque : de la même façon que les spectres sont des nombres complexes (dû à la transformée de Fourier), la fonction H(f) aussi :

$$H(f) = |H(f)| \exp (j \arg (H(f)) = |H(f)| e^{j\varphi(f)}$$

Ainsi, une composante d'entrée $x(t) = A \cos (2\pi \, fo \, t + \theta)$ deviendra, par passage dans le filtre :

$$y(t) = B \cos (2\pi \, fo \, t + \theta') = A \, |H(f)| \cos (2\pi \, fo \, t + \theta + \varphi(fo))$$

L'amplitude a bien été multipliée par un facteur dépendant de la fréquence |H(fo)| et il y a eu rotation de phase d'un angle φ(fo).

Une fonction de transfert est donc une fonction complexe, toutefois, pour les applications courantes, la donnée du seul module est suffisante pour caractériser, de façon pratique, un filtre.

- la deuxième est la **méthode temporelle**, où la variable choisie est le temps. C'est une méthode plus « moderne » et particulièrement bien adaptée au filtrage « numérique » (c'est-à-dire quand les signaux sont exprimés sous forme de suites numériques : cas d'une image numérisée, d'une suite de nombres, d'un texte codé, …).
Cette méthode consiste à exprimer H sous forme d'une fonction du temps : h(t)

La relation de filtrage $Y(f) = H(f)\,X(f)$ s'écrit alors (en faisant la transformée de Fourier inverse des deux membres) :

$$y(t) = x(t) * h(t)$$

l'opérateur « étoile * » désignant ici un « produit de convolution » et non un produit simple.
Le produit de convolution est une « nouvelle opération » qu'a fait apparaître la transformation de Fourier.
Le produit de convolution exprime précisément l'opération que subissent les valeurs de x(t) par passage dans le filtre caractérisé par h(t). Mais que représente h(t) ?

Ce produit de convolution, comme tout produit, admet un élément neutre, notons le provisoirement $\sigma(t)$, cette fonction sera donc telle que :

$$\sigma(t) * h(t) = h(t)$$

Si l'on rapproche cette expression de la précédente, on peut donc dire qu'en appliquant un « signal » représenté par $\sigma(t)$ à l'entrée d'un filtre, celui-ci sortirait une fonction égale à sa propre fonction h(t) le caractérisant.
Il apparaît donc que cette fonction $\sigma(t)$ soit très précieuse, mais quelle est cette fonction ?

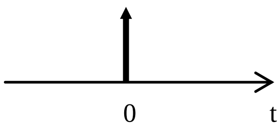
Pour répondre à cela, il faut expliciter le produit de convolution :

$$y(t) = \int_{-\infty}^{+\infty} x(u) \,.\, h(t-u) \,.\, du$$

La recherche de l'élément neutre est la recherche de la fonction $\sigma(t)$ telle que :

$$f(t) = \int_{-\infty}^{+\infty} f(u) \,.\, \sigma(t-u) \,.\, du \qquad \text{pour toute fonction } f(t)$$

La réponse est que $\sigma(t)$ n'est autre que la fonction (ou plutôt distribution) déjà rencontrée : $\delta(t)$, fonction de Dirac (représentée symboliquement par une « flèche »).
Il s'agit « physiquement » d'une « impulsion » localisée en t = 0, c'est-à-dire une fonction que l'on pourrait représenter (non rigoureusement) par une multiplication par 1 d'une autre fonction, mais seulement en t = 0 et par une multiplication par 0 partout ailleurs.



La fonction h(t) est pour cela appelée « réponse impulsionnelle » du filtre : c'est le signal que fournira le filtre si on lui place une impulsion (de Dirac) à l'entrée.

La réponse impulsionnelle h(t) est une fonction réelle, elle peut aussi être une suite de nombres.

Physiquement, une fonction d'entrée x(t) (ou respectivement une suite numérique) peut toujours être interprétée comme une suite continue (respectivement discrète) de valeurs ponctuelles, représentables chacune par une impulsion de Dirac pondérée par la valeur de la fonction en chaque point (respectivement par la valeur numérique considérée).
Ceci s'écrit :

$$x(t) = \int_{-\infty}^{+\infty} x(u) \,.\, \delta(t-u) \,.\, du$$

et se lit : « une fonction x(t) résulte de la somme (somme au sens de la réunion) de tous ses points x(u) pour toutes les valeurs de u ». C'est le « théorème de convolution, qui ne fait qu'exprimer la relation de l'élément neutre !

De ce fait, la fonction de sortie peut alors être exprimée à l'aide des relations successives suivantes, dans lesquelles le symbole « → » exprime le passage à travers le filtre :

$\delta(t)$	$\rightarrow$	$h(t)$	car h(t) est la réponse impulsionnelle du filtre
$\delta(t-u)$	$\rightarrow$	$h(t-u)$	changement de variable
$c \,.\, \delta(t-u)$	$\rightarrow$	$c.\, h(t-u)$	« c » est un facteur constant
$x(u) \,.\, \delta(t-u)$	$\rightarrow$	$x(u) \,.\, h(t-u)$	x(u) est aussi un facteur constant, pour u donné
$\int_{-\infty}^{+\infty} x(u) \,.\, \delta(t-u) \,.du$	$\rightarrow$	$\int_{-\infty}^{+\infty} x(u) \,.\, h(t-u) \,.\, du$	en faisant la somme pour tous les u
$x(t) * \delta(t)$	$\rightarrow$	$x(t) * h(t)$	en écrivant avec l'opérateur de convolution
$x(t)$	$\rightarrow$	$y(t)$	résultat

On trouve bien ici justifié que la sortie y(t) est le produit de convolution de l'entrée x(t) par la réponse impulsionnelle du filtre h(t).

3.3.5 Caractérisation des filtres, fonction de transfert et tracé de Bode :

Nous allons nous intéresser ici à la seule fonction de transfert H(f) (représentation fréquentielle).

Nous avons vu que H(f) était un nombre complexe : $H(f) = |H(f)|\, e^{j\varphi}$

La caractérisation fréquentielle passe par le comportement en régime sinusoïdal, pour lequel il est légitime d'adopter également la représentation complexe :

$$x(t) = A \cos(2\pi f t + \varphi) \rightarrow \underline{X} = A\, e^{j\omega t}\, e^{j\varphi} \rightarrow \underline{\underline{X}} = A\, e^{j\varphi}$$

La dernière représentation est « abrégée », c'est-à-dire ne comprend pas le terme $e^{j\omega t}$, dans la mesure où celui-ci sera présent dans tous les facteurs, donc factorisable dans les expressions et simplifiable dans les équations.

Par la suite, nous omettrons également le soulignement simple ou double (qui rappelle l'usage de la notation complexe) tant qu'il n'y a pas d'ambiguïté possible (nous emploierons en revanche la lettre majuscule X pour désigner la notation complexe abrégée associée à x(t)).

Les circuits réalisant le filtre sont des circuits électriques pour lesquels nous adopterons aussi la notation complexe (impédances complexe). De ce fait, nous noterons désormais H(jω) la fonction de transfert :

$$H(j\omega) = |H(j\omega)|\, e^{j\varphi(\omega)}$$

H(jω) est une fonction a priori quelconque (dépend du type de filtre à réaliser), en réalité, et en particulier si ce filtre est réalisé à partir d'éléments de circuits électriques (dipôles R L C et sources, y compris sources liées par une relation), on montre aisément que la fonction de transfert est toujours sous la forme d'une fraction rationnelle en jω :

$$H(j\omega) = \frac{a0 + a1\,j\omega + a2\,(j\omega)^2 + a3\,(j\omega)^3 + a4\,(j\omega)^4 + \ldots + am\,(j\omega)^m}{b0 + b1\,j\omega + b2\,(j\omega)^2 + b3\,(j\omega)^3 + b4\,(j\omega)^4 + \ldots + bn\,(j\omega)^n}$$

avec les propriétés suivantes :
- tous les coefficients du dénominateur sont positifs ou nuls
  (c'est aussi le cas pour le numérateur pour des filtres « usuels »)
- le degré du dénominateur (n) est toujours supérieur ou égal à celui du numérateur (m)
- le numérateur et dénominateur sont factorisables en binômes du premier degré et en trinômes du second degré

On emploie couramment, dans un but de simplification, la notation dite « opérateur de Laplace » consistant à remplacer le terme (jω) par le symbole « **p** » (parfois, on rencontre également le symbole « s » avec la même signification).

En réalité, cet opérateur est issue d'une opération de transformation, relativement similaire à la transformée de Fourier, qui consiste à transformer une fonction du temps en une « fonction en p », permettant de réaliser du « calcul symbolique » dont un des buts est précisément d'éviter le produit de convolution en le remplaçant par un produit simple (comme le fait aussi la transformée de Fourier). On obtient bien la correspondance $p = j\omega$, mais seulement dans le cas où la fonction est sinusoïdale, ce qui est le cas ici.

On obtient alors l'écriture suivante :

$$H(p) = \frac{a0 + a1\,p + a2\,p^2 + a3\,p^3 + a4\,p^4 + \ldots + am\,p^m}{b0 + b1\,p + b2\,p^2 + b3\,p^3 + b4\,p^4 + \ldots + bn\,p^n}$$

On appelle « ordre du filtre » la valeur de n (le degré du dénominateur), on montre en effet que le comportement essentiel de ce filtre se traduit grâce à cet ordre.

Le fait que le numérateur et le dénominateur soient factorisables en binômes du premier degré et en trinômes du second degré permet d'écrire H(p) sous la forme d'un produit de facteurs dont les termes sont au maximum de second degré en p :

$$H(p) = Ha(p)\,.\,Hb(p)\,.\,Hc(p)\,.\,\ldots\ldots\;Hq(p)$$

Le filtre H(p) peut alors être considéré comme étant la mise en cascade de q filtres indépendants et chaque filtre Hi(p) peut alors s'écrire :

$$Hi(p) = \frac{a2\,p^2 + a1\,p + a0}{b2\,p^2 + b1\,p + b0}$$ certains coefficients pouvant être nuls

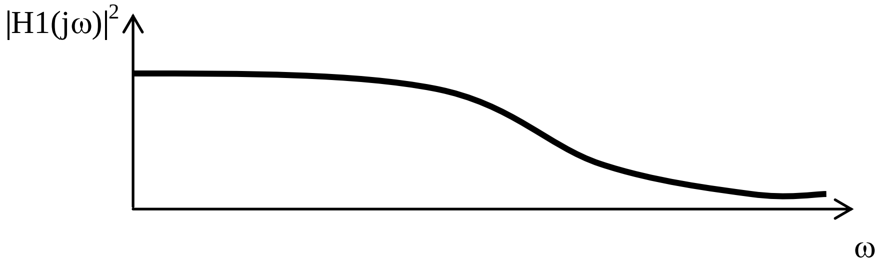
Il ne reste donc que ce type de filtre à caractériser.
On peut les passer tous en revue selon les coefficients non nuls.

Prenons par exemple, un filtre tel que seuls a0 ; b1 et b0 soient non nuls :

$$H1(p) = \frac{a0}{b1\,p + b0}$$ soit, en (jω) : $$H1(j\omega) = \frac{a0}{b1\,j\omega + b0}$$

Le module vaut : $|H1(j\omega)|^2 = \dfrac{a0^2}{b0^2 + b1^2\,\omega^2}$

Cette fonction peut être tracée directement :



On obtient ainsi une courbe monotone décroissante, avec un module de fonction de transfert qui vaut (a0/b0) pour une fréquence nulle, qui décroît pour des fréquences croissantes et qui tend vers 0 pour une fréquence tendant vers l'infini.

Ce filtre peut être qualifié de « PASSE – BAS » au regard de son comportement, en particulier si l'on choisit a0 = b0 (transfert égal à 1 pour f = 0).

Précisément, si a0 = b0, l'expression se met sous la forme :

$$|H1(j\omega)|^2 = \frac{1}{1 + b^2\,\omega^2}$$

La valeur ω_0 telle que le terme $b^2\omega^2$ vaut 1 est remarquable car alors : $|H1(j\omega_0)|^2 = 1/2$, valeur juste intermédiaire entre 1 et 0 d'une part, mais aussi atténuation d'un facteur 2 de la puissance d'entrée : en effet, $|Y(j\omega)|^2 = |H1(j\omega)|^2 \cdot |X(j\omega)|^2$ (expression du filtrage, convertie en modules au carré), ainsi $|H1(j\omega_0)|^2 = 1/2$ signifie bien que la puissance de sortie est moitié de celle d'entrée (la puissance étant proportionnelle au carré de la tension ou du courant électrique).

On peut employer ici une échelle logarithmique pour mesurer les puissances et les atténuations (*cf chap. 2.3.4*) : un facteur 1/2 correspond alors à :

$10 \log (1/2) \approx -3$ dB (- 3 décibels)

Cette valeur fait que le point $\{\omega_0, -3 \text{ dB}\}$ est un point remarquable.

On peut avantageusement généraliser l'emploi de l'échelle logarithmique en adoptant l'écriture suivante :

$$10 \log|H1(j\omega)|^2 = 20 \log |H1(j\omega)| = 10 \log \frac{1}{1 + b^2\,\omega^2} = 10 \log (1) - 10 \log (1 + b^2\,\omega^2)$$

soit : $10 \log|H1(j\omega)|^2 = -10 \log (1 + b^2\,\omega^2)$

Mais … la simplification ne peut pas se poursuivre car on arrive sur le logarithme d'une somme qui n'est pas remarquable …

On emploie alors une technique dite « asymptotique », consistant, comme son nom l'indique, à rechercher des fonctions équivalentes pour $\omega \to 0$ d'une part et $\omega \to \infty$ d'autre part :

$\omega \to 0$ $\quad$ $10 \log|H1(j\omega)|^2 \equiv - 10 \log (1) = 0$

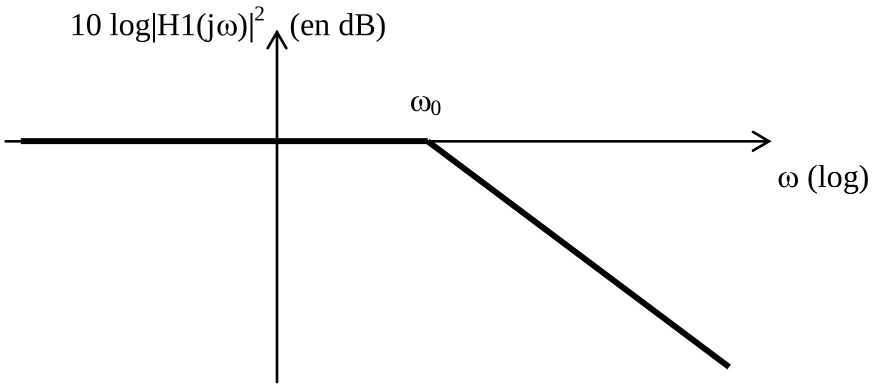
$\omega \to \infty$ $\quad$ $10 \log|H1(j\omega)|^2 \equiv - 10 \log (b^2\omega^2) = - 20 \log (b\omega) = - 20 \log b - 20 \log \omega$

Nous obtenons alors des expressions du premier degré maximum en ($\log \omega$).

Ceci a donné l'idée d'adopter également une échelle logarithmique pour ω (de manière à transformer ω en $\log \omega$), ce qui permet, en adoptant une échelle « log – log », d'avoir finalement des tracés qui se résument à des droites !

Cette méthode est connue sous le nom de « DIAGRAMME DE BODE » ou « TRACE DE BODE » :

Cela consiste à tracer les asymptotes et à les faire se réunir en ω_0



On peut apporter une précision supplémentaire sur la pente de la droite oblique .Son expression est :

$10 \log|H1(j\omega)|^2 \equiv - 20 \log b - 20 \log \omega$

ainsi :

quand ω est multipliée par 2, un facteur $-20 \log 2 = - 6$ dB apparaît

quand ω est multipliée par 10, un facteur $-20 \log 10 = - 20$ dB apparaît

On dit alors que la pente vaut « - 6dB par octave » ou « -20 dB par décade ».

Il faut enfin se rappeler que l'atténuation réelle de ce filtre pour ω_0 vaut 3 dB

On appelle « fréquence de coupure » la fréquence qui correspond à ω_0.

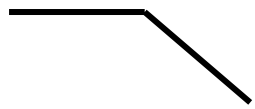
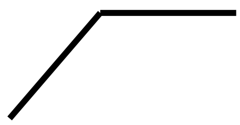
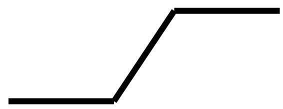
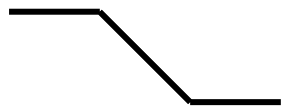
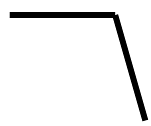
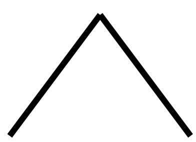
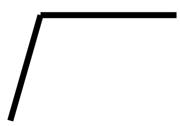
Toutes ces données font que le tracé de Bode permet d'avoir une estimation rapide (et souvent suffisante) du comportement du filtre.

Le tracé réel de la fonction (quand il est nécessaire) est facilité par ce tracé de Bode grâce aux asymptotes qui sont déjà tracées et aussi grâce au point $\{\omega_0, - 3 \text{ dB}\}$.

On peut naturellement étendre ce procédé à tous les autres filtres de manière à se constituer une « bibliothèque » de filtres élémentaires ; pour cela, on reprend l'expression générale du filtre d'ordre 2 :

$$H(p) = \frac{a2\ p^2 + a1\ p + a0}{b2\ p^2 + b1\ p + b0}$$

certains coefficients pouvant être nuls

Expression de H(p)	Type de filtre réalisé	Allure du Tracé de Bode
$\frac{1}{1 + b\ p}$	PASSE BAS premier ordre	
$\frac{a\ p}{1 + b\ p}$	PASSE HAUT premier ordre	
$\frac{1 + a\ p}{1 + b\ p}$ $a > b$	ACCENTUATEUR	
$\frac{1 + a\ p}{1 + b\ p}$ $a < b$	DEACCENTUATEUR	
$\frac{1}{1 + b\ p + c\ p^2}$	PASSE BAS second ordre	
$\frac{a\ p}{1 + b\ p + c\ p^2}$	PASSE BANDE second ordre	
$\frac{a\ p^2}{1 + b\ p + c\ p^2}$	PASSE HAUT second ordre	

3.3.6 **Réalisation des filtres**

Principes généraux :

Un filtre, relativement à son comportement fréquentiel, doit réaliser une atténuation ou une amplification dont le facteur est une fonction de la fréquence.
Nous devons donc rechercher des moyens permettant de réaliser une telle discrimination de fréquences.

Pour cela, nous pouvons opérer d'abord à partir d'une fonction sinusoïdale (qui, rappelons-le est la base du théorème de Fourier).
Prenons précisément un signal qui serait composé de deux fonction sinusoïdales d'amplitudes comparables A1 et A2, mais de fréquences très différentes l'une de l'autre :

$$x(t) = A1 \cos(\omega 1\, t + \varphi 1) + A2 \cos(\omega 2\, t + \varphi 2) = x1 + x2 \qquad \text{avec } \omega 2 >> \omega 1$$

Supposons que l'on veuille renforcer la composante x2 et atténuer la composante x1.
Dérivons l'expression de x(t) :

$$\begin{aligned} dx(t)/dt &= - \omega 1\, A1 \sin(\omega 1\, t + \varphi 1) - \omega 2\, A2 \sin(\omega 2\, t + \varphi 2) \\ &= - B1 \sin(\omega 1\, t + \varphi 1) - B2 \sin(\omega 2\, t + \varphi 2) \end{aligned}$$

$$B1 = \omega 1\, A1 \quad \text{et} \quad B2 = \omega 2\, A2$$

On obtient alors B2 >> B1 car $\omega 2 >> \omega 1$: la composante x2 a bien été renforcée au détriment de la composante x1.

Supposons à présent que l'on veuille l'inverse : renforcer la composante x1 et atténuer la composante x2.
Intégrons l'expression de x(t) :

$$\begin{aligned} \int x(t)\, dt &= A1/\omega 1 \sin(\omega 1\, t + \varphi 1) + A2/\omega 2 \sin(\omega 2\, t + \varphi 2) \\ &= C1 \sin(\omega 1\, t + \varphi 1) + C2 \sin(\omega 2\, t + \varphi 2) \end{aligned}$$

$$C1 = A1 / \omega 1 \quad \text{et} \quad C2 = A2 / \omega 2$$

On obtient alors C1 >> C2 car $\omega 2 >> \omega 1$: la composante x1 a bien été renforcée au détriment de la composante x2.

Ceci montre que des comportements de « filtrage » peuvent être obtenus par de simples opérations de dérivation et d'intégration : ces opérations sont en effet à la base de tout filtre.
Plus généralement, un filtre est régi par une équation intégro-différentielle en ω.

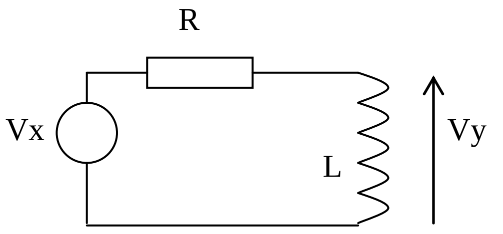
Réalisations avec des circuits électriques :

Il faut d'abord identifier des composants (ou des moyens) permettant de réaliser les deux opérations de dérivation et d'intégration.
Les composants élémentaires bobine L et condensateur C font l'affaire, en effet :

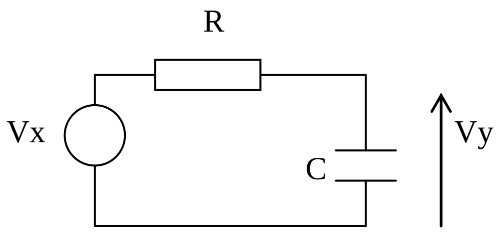
- une bobine L est décrite par la relation : $V = L\, dI/dt$
une bobine permet donc de réaliser (idéalement) un CIV dérivateur

- un condensateur C est décrit par la relation : $V = 1/C \int I\, dt$
un condensateur permet donc de réaliser (idéalement) un CIV intégrateur

Des circuits d'utilisation élémentaires s'en déduisent :



$$Vy = Vx\,(jL\omega / R + jL\omega) = Vx\ H1\,(j\omega)$$



$$Vy = Vx\,(1 / 1 + jRC\omega) = Vx\ H2\,(j\omega)$$

En utilisant l'opérateur de Laplace :

$$H1\,(p) = \frac{Lp}{R + Lp}$$ c'est donc un « filtre passe –haut du premier ordre »

$$H2\,(p) = \frac{1}{1 + RCp}$$ c'est donc un « filtre passe –bas du premier ordre »

Nous voyons donc que des circuits aussi simples que ceux-là permettent de réaliser des filtres, certes rudimentaires, car peu « sélectifs » (l'atténuation ne croît que de 20 dB par décade).

Des associations de tels circuits permettent d'augmenter la sélectivité de ces filtres. Et de réaliser des filtres d'autres types, selon le besoin.

<u>Réalisations avec des circuits électroniques :</u>

On montrer que l'emploi de circuits électroniques, en particulier de CVV idéaux de gain 1 ou de CVV idéaux de gain infini (très grand en pratique), permet de s'affranchir totalement d'un des deux composants L ou C ; dans la pratique, on utilise que le condensateur car celui-ci est relativement parfait à l'opposé de le bobine pour laquelle on ne peut jamais faire abstraction de la résistance série (résistance du conducteur de la bobine !).

Réalisation avec des fonctions numériques (logiques) :

La réalisation de filtres « numériques », contrairement aux précédents, ne s'appuie pas sur des considérations harmoniques (comportement du filtre en régime sinusoïdal), mais sur sa caractérisation temporelle, c'est-à-dire, la relation de convolution qui existe entre l'entrée x(t) et la sortie y(t) du filtre :

$y(t) = x(t) * h(t)$ h(t) étant la réponse impulsionnelle du filtre

La réponse impulsionnelle peut bien sûr être calculée à partir de la réponse fréquentielle (transformée de Fourier inverse). Aussi, en règle générale et notamment pour les signaux temporels, on détermine d'abord la réponse fréquentielle souhaitée, puis on en déduit la réponse impulsionnelle.

L'opération de convolution est une opération qui agit sur des fonctions continues. Une numérisation impose une **discrétisation** de ces trois fonctions x(t), y(t) et h(t).

Dans le cas de filtrage de signaux analogiques (donc donnant des fonctions continues du temps), la première opération à réaliser est la numérisation de ces signaux, c'est-à-dire (sommairement) une discrétisation du temps (opération appelée « **échantillonnage** ») suivie d'une quantification des valeurs (opération appelée « **quantification** »).

Echantillonnage : $s(t) \rightarrow s(k\,Te)$

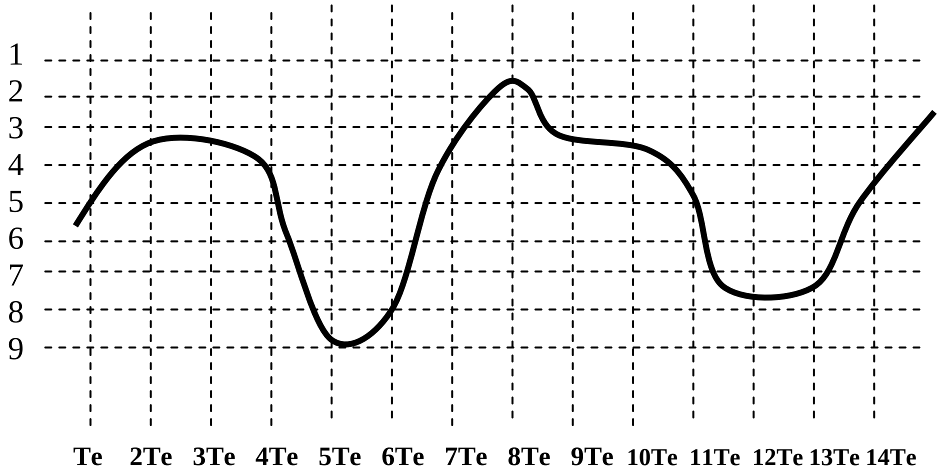
La variable temps « t » devient un indice progressif « k » (k est un entier)

Te est la période d'échantillonnage, un théorème (Nyquist – Shannon) énonce que cette opération est sans perte d'information si Fe = 1 / Te (fréquence d'échantillonnage) est supérieure ou égale à la fréquence maximale significative du spectre du signal considéré (dans le cas de la téléphonie : Fmax = 3400 Hz, donc Fe doit être supérieure ou égale à 6800 Hz, on choisit 800 Hz dans la pratique)

Quantification : $s(k\,Te) \rightarrow s_i\,(k)$

La valeur réelle « s » à l'instant « kTe » est approchée par une valeur entière « s_i » , la fonction « s (k Te) » devient alors une suite numérique de valeurs « s_i », indexées par le paramètre « k »

Ces deux opérations sont schématisées ci-dessous :



La fonction s(t) devient alors la suite **: {… 5 3 3 4 9 8 4 2 3 3 5 8 7 5 … },** indexée par la suite {k Te}.

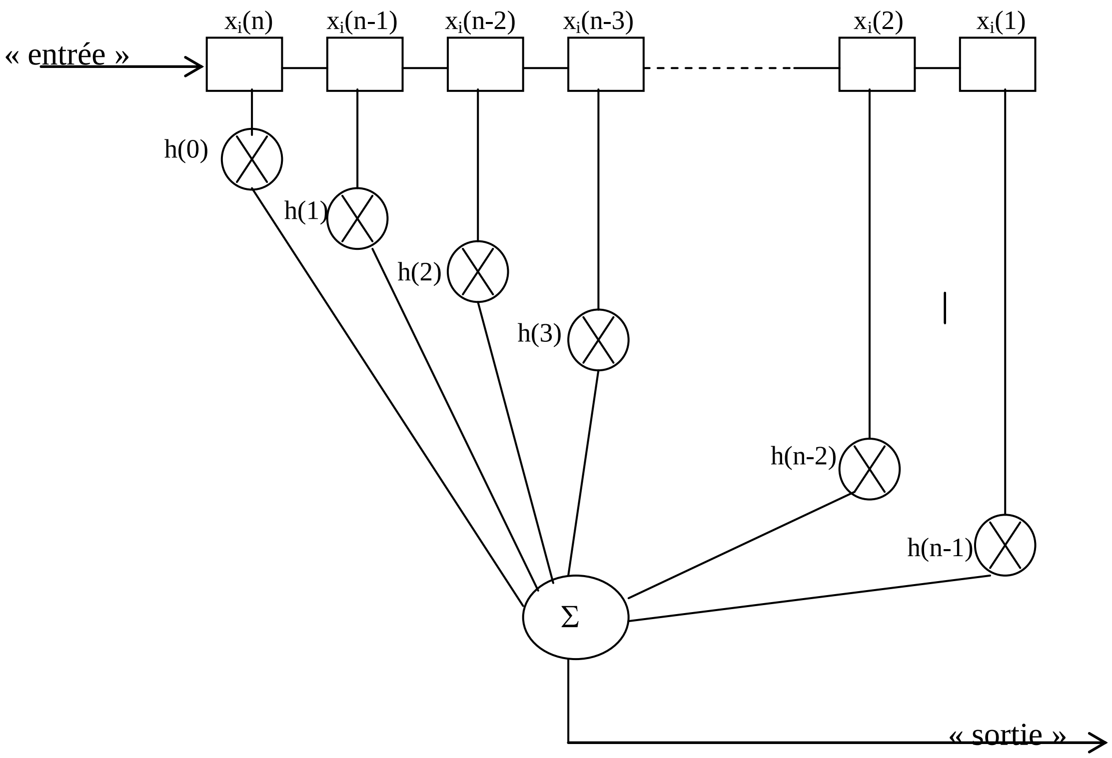
Il y a bien sûr une **erreur systématique de quantification**, mais celle-ci est majorée par le demi-intervalle qui sépare deux valeurs entières consécutives, et peut être rendue aussi petite que l'on veut en adoptant un intervalle plus faible (ce qui, par ailleurs, multiplie le nombre de valeurs entières nécessaires à la représentation de la fonction).

L'ensemble de ces deux opérations se nomme « numérisation » ou « conversion analogique – numérique ».

Une fois cette opération effectuée sur x(t) et sur h(t), on peut réaliser la convolution, sous sa forme discrète, l'expression :

$$y(t) = \int_{-\infty}^{+\infty} x(u) \,.\, h(t-u) \,.\, du \qquad \text{devenant} \qquad y_j(k) = \sum x_i(u) \,.\, h(k-u)$$

La structure élémentaire d'un « filtre numérique » se déduit de l'expression précédente :



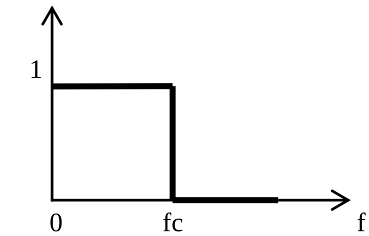
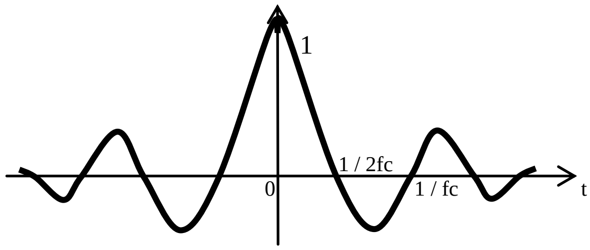
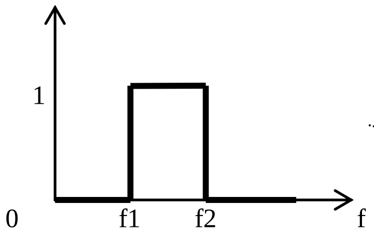
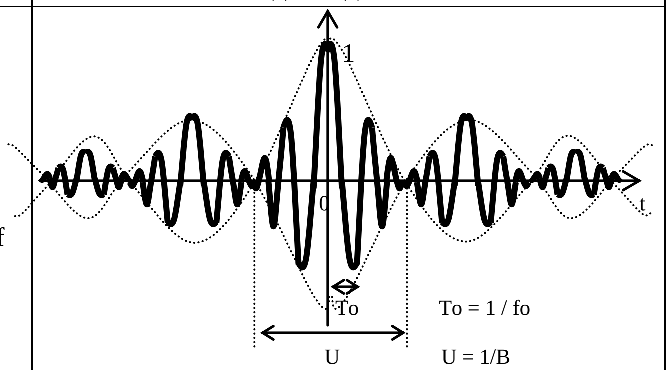
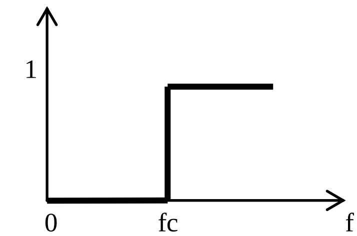
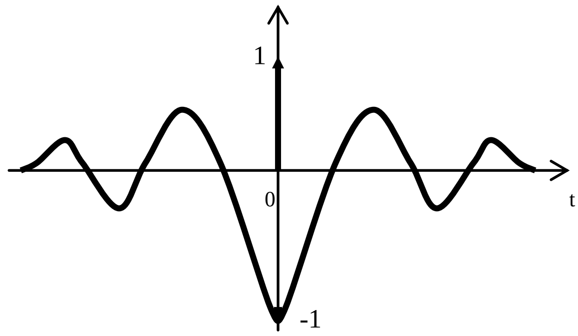
Le filtre présenté ci-dessus à une « profondeur de calcul » de n valeurs (il agit donc sur n valeurs consécutives de la fonction numérique d'entrée).

Ce filtre est donc constitué d'un « registre à décalage » permettant d'avoir ces n valeurs successives de x_i, que l'on multiplie (multiplication ordinaire de deux nombres) par les coefficients « h » (ces coefficients sont caractéristiques du filtre) pour enfin faire la somme de tous les termes et obtenir UNE valeur, ici $y_j(n)$.
Une fois ce calcul fait (quasi-immédiat, au temps de calcul des opérations près), une nouvelle valeur x_i (n+1) entre dans le registre et « pousse » les autres : il y a alors, dans le registre, les valeurs x_i(n+1)jusqu'à x_i(2), ce qui va donner en sortie y_i(n+1) et ainsi de suite … Une nouvelle valeur « sortira » donc à chaque nouvelle valeur en entrée , ce qui créera la suite $\{y_j\}$, donc l'information de sortie, résultat du filtrage de la suite $\{x_i\}$, information d'entrée.

Cette structure particulière de filtre, issue directement de l'expression de convolution est appelée « filtre transversal ».

On rappelle bien sûr, que si l'information est initialement représentée sous forme d'une fonction continue du temps x(t), il est nécessaire d'opérer préalablement à l'échantillonnage et à la quantification de cette fonction.
De même, si la fonction de sortie doit également être une fonction y(t) (ce qui est le cas pour la téléphonie bien entendu !), le filtre numérique doit être suivi de l'opération inverse de la numérisation, dite « restitution » ou « conversion numérique analogique ».

A titre d'information, voici les <u>allures</u> des réponses imulsionnelles pour les trois familles principales de filtres :

filtre	Allure du tracé de Bode (filtre idéal)	Allure de la réponse impulsionnelle
Passe bas	 H(f) = 1 pour f < fc = 0 pour f > fc	 $h(t) = \sin(2\pi f_c t) / \pi t = 2 f_c \ \mathrm{sinc}(2\pi f_c t)$ sinc : sinus cardinal ; sinc (x) = sin(x) / x
Passe bande	 $f_o^2 = f_1 f_2$ B = f2 – f1	
Passe haut		

4. La téléphonie numérique

4.1 Critique de la téléphonie analogique

- sensibilité à l'atténuation
- sensibilité à la non-linéarité (déformation des signaux)
- sensibilité aux perturbations
- une ligne = une seule conversation téléphonique …

… donc nécessité de multiplication des lignes pour augmenter la capacité de transmission (augmentation du nombre de communication simultanées).

4.2 Retour à la télégraphie et au répéteur télégraphique, mais … quel est le rapport ?

La télégraphie fonctionne en « tout ou rien » et transmet une information
« le tout ou rien » est :

- peu sensible à l'atténuation (sauf si elle devient considérable)
- peu sensible aux non-linéarités
- peu sensible aux perturbations
- une ligne = une information …

… mais, un procédé nommé « télégraphie harmonique » a permis d'envisager plusieurs transmissions simultanées et indépendantes sur une même ligne, ce procédé se nomme aujourd'hui « multiplexage fréquentiel »

Ce procédé consistait à associer une fonction sinusoïdale de fréquence donnée à une transmission. L'information « tout ou rien » se transformait alors en une amplitude de cette fonction sinusoïdale « tout ou rien » : il s'agit d'un procédé dit « modulation d'amplitude par tout ou rien » : amplitude non nulle ou nulle.
Ensuite, on peut opérer à une somme de ces différentes fonctions et transmettre l'ensemble (soit N conversations sur N fréquences) sur une seule ligne.
A l'arrivée, la sélection de la conversation se réalise aisément par une opération de filtrage sélectif (sélection d'une bande étroite autour de la fréquence considérée).

Ce procédé fut peu employé pour la télégraphie, mais fût à la base de deux systèmes majeurs :

- la radiodiffusion hertzienne qui utilisa le même principe de fréquences différentes pour des informations différentes (toujours en vigueur de nos jours !) ;
- le multiplexage qui consiste, pour un « canal » de transmission donné, à transmettre plusieurs informations distinctes sans qu'elles se perturbent mutuellement.

Mais … la téléphonie n'est pas la télégraphie … sauf à trouver une représentation de l'information « voix » sous forme symbolique, à l'aide d'un « codage » approprié.

En ce sens, la méthode « analogique » employée historiquement pour la téléphonie peut être qualifiée de très sommaire et très simpliste. En effet, le problème n'est pas de transmettre, comme on le fait dans la méthode analogique » un signal représentant la pression acoustique au cours du temps, mais l'information produite par la voix.
Ainsi, formellement, la pression acoustique n'est elle-même qu'un vecteur de transmission, l'écriture en est un autre d'ailleurs.

4.3 Un codage numérique de la voix ? Une discrétisation d'une fonction continue

Une première approche du codage fut également relativement sommaire puisque basé sur le signal acoustique existant. Le raisonnement fut le suivant, en partant du but à atteindre :

- la télégraphie ne fonctionne qu'en « tout ou rien », donc ne peut adopter qu'un alphabet de base binaire,
- n'importe quel nombre peut s'exprimer sous n'importe quelle base, binaire en particulier,
- une fonction réelle du temps décrit un ensemble de nombres réels (rangés en fonction du temps)
- un signal électrique analogique représentant la parole est une fonction réelle du temps.

Les transformations successives apparaissent clairement en faisant le chemin inverse (de la parole à la télégraphie).

Toutefois, deux difficultés se présentent :

- la première due au fait que la fonction du temps étant continue, le nombre de valeurs à transmettre est infini par unité de temps ;
- la deuxième due au fait que les valeurs sont réelles, donc exprimables en général par un nombre infini de chiffres, quelle que soit la base de numération choisie.

Ces deux difficultés amènent à rechercher la pertinence d'une double discrétisation : temps et valeurs.
Nous retrouvons bien évidemment la problématique déjà abordée lors du filtrage numérique : celle de la numérisation (*cf chap. 3.3.6*).

4.4 Numérisation : avec quels moyens et quels choix ?

4.4.1 Echantillonnage

On rappelle que la discrétisation du temps se nomme « échantillonnage » et que cette opération est régi par le théorème de Nyquist – Shannon de manière à ne pas perdre d'information (un corollaire est que cette opération est parfaitement réversible).

L'échantillonnage doit donc se réaliser avec une cadence $Fe = 1/ Te$ supérieure ou égale à la fréquence maximale significative du spectre du signal téléphonique fmax, soit 3400 Hz.
La condition est donc : $Fe > 2 \text{ fmax} = 6800$ Hz.

On montre que l'opération inverse (restitution) peut être tout simplement réalisée à l'aide d'un filtre passe-bas idéal dont la fréquence de coupure fc est égale précisément à $Fe / 2$.
Le problème est que le filtre passe bas idéal étant irréalisable physiquement, la restitution exacte est impossible.
On peut toutefois s'approcher de la restitution exacte sous deux conditions (complémentaires) : la première est de choisir Fe la plus grande possible devant 2 fmax, la deuxième est d'utiliser un filtre passe-bas d'ordre le plus élevé possible.

En téléphonie, la fréquence d'échantillonnage est normalisée (pour le monde entier) à : Fe = 8000 Hz.

4.4.2 Quantification

On rappelle que la discrétisation des valeurs se nomme « quantification » et que cette opération apporte une erreur systématique due à l'arrondi effectué lors de l'approche d'une valeur réelle par une valeur entière. Cette erreur est majorée par le demi-intervalle de quantification (*cf chap. 3.3.6*).

On rappelle également que le niveau « ligne » téléphonique est normalisé à Lo = 1mW (niveau « 0 dBm »), ce qui correspond à Uo = 775 mV sur une résistance de 600 Ω (*cf chap. 2.3.4*). Ce niveau correspond à la « valeur efficace moyenne d'une conversation normale ».
On appelle en outre, « facteur de crête » (noté « cr »), le facteur multiplicatif qu'il faut appliquer à une valeur efficace pour obtenir la valeur maximale atteinte par le signal, enfin il faut toujours avoir en tête que le signal téléphonique est alternatif (valeurs positives et négatives, valeur moyenne nulle) : l'intervalle total de l'ensemble des valeurs atteintes par le signal téléphonique vaut donc : 2 cr Uo.

Ainsi, le choix d'un intervalle de quantification, noté « q », et la connaissance de l'intervalle total des valeurs atteintes (soit 2 cr Uo) par le signal permettent de déterminer le nombre de valeurs quantifiées :

$$N_Q = 2\ cr\ Uo / q$$

A son tour, cette valeur de N_Q permet de connaître le format des nombres entiers une fois exprimés en base 2, ce format binaire B (nombre de bits) est en effet égal à :

$$B = \log_2 N_Q \quad (N_Q = 2^B)$$

En récapitulant : le nombre de bits (format) par échantillon, pour un niveau ligne donné, va permettre de connaître l'intervalle de quantification, donc l'erreur de quantification.
Cette erreur de quantification sera perçue par l'utilisateur comme une **perturbation** (de la famille des « bruits ») et donc, doit être suffisamment faible pour qu'elle ne soit pas gênante.

On remarque que la gêne n'est pas directement liée à la puissance de la perturbation mais au **rapport entre la puissance utile et la puissance de la perturbation** . Ce rapport se nomme « rapport signal sur bruit » et se note habituellement « S/N » (« *Signal to Noise ratio* »).

Le calcul montre que : **S / N ≈ 6 B** S / N étant exprimé en dB

En téléphonie (objectif : transmission d'un message par l'intermédiaire de la parole), on considère qu'une communication est possible dès que S / N est supérieur à 45 dB, le choix de B = 8 est donc suffisant.

En téléphonie, le format binaire est normalisé (pour le monde entier) à : B = 8 bits.

4.4.3 Débit série de la téléphonie numérique

La transmission de type « télégraphique » de la téléphonie numérisée consiste à envoyer les valeurs binaires les unes derrière les autres, ceci pour chaque échantillon.

Le débit binaire (nombre de bits par seconde) vaut donc : D = B . Fe = 8 . 8000 = 64 kbit / s

En téléphonie, le débit binaire normalisé vaut : 64 kbit / s

5. La téléphonie sans fil

5.1 Critique de la téléphonie filaire

Cette critique est évidente : la téléphonie nécessite une paire de conducteurs de bout en bout entre les interlocuteurs.

Elle interdit donc la mobilité, sauf avec un rayon d'action limité par le « cordon téléphonique » du poste d'abonné (mobilité : déplacement de l'interlocuteur durant une communication).
Elle interdit l'itinérance, sauf dans un secteur desservi apr la même ligne d'abonné, car un numéro (donc une tarification) est associé à une ligne d'abonné (itinérance : établissement ou réception d'une communication à quelque endroit que ce soit).

Elle interdit bien entendu toute liaison téléphonique avec les navires ou les avions ou … les voitures !

5.2 L'expérience de la « radio »

Parallèlement au développement de la téléphonie (filaire) s'est produit le développement de la radiodiffusion hertzienne terrestre.
Il s'agit d'un autre service, parfaitement distinct qui consiste à diffuser des informations (sonores ou visuelles) vers le public en utilisant cette fois des moyens électromagnétiques, c'est-à-dire un principe d'action à distance basé sur l'induction électromagnétique : un courant variable crée un « champ électromagnétique », celui-ci se propage naturellement (sans support matériel, même dans le vide), enfin ce champ provoque une induction (fem) dans un circuit électrique, l'ensemble des grandeurs mises en jeu étant proportionnelles entre elles, donc propres à transporter une information.

5.3 La téléphonie « sans fil »

Ce procédé résulte naturellement de la combinaison de deux procédés : celui de la téléphonie et celui de la radiodiffusion.

En quelque sorte, la ligne téléphonique d'abonné est remplacée par une liaison électromagnétique entre le poste d'abonné (portable) et une « station de base ». Le réseau de distribution est donc virtuel.
En revanche, le réseau de transport ainsi que les systèmes de commutation sont ceux de la téléphonie filaire.

6. La téléinformatique : quand le réseau téléphonique transporte des données

6.1 voix → « données » et données → « voix » : la convergence

La téléphonie, qu'elle soit filaire ou sans fil, est aujourd'hui majoritairement numérique, au moins au niveau du réseau de transport et de la commutation.

Nous avons vu que les choix de téléphonie faisaient que les échantillons étaient cadencés à la fréquence de 8000 Hz et étaient codés sur 8 bits (un octet). La voix est donc représentée par une suite d'octets binaires.

Un octet peut aussi représenter un « caractère » codé à l'aide du code « ASCII ».
Un octet peut aussi représenter une « composante couleur » d'un point d'une image.

Ainsi, un réseau de transport et de commutation transporte en réalité des « octets ».
Un « observateur » de ce réseau pourrait alors légitimement se poser la question suivante :
« les octets binaires représentent-ils les caractères alphabétiques successifs de mots constituant une phrase, ou les composantes couleurs des points successifs constituant une image , ou les valeurs successives des échantillons constituant un son, ou des octets appartenant à une des structures de données d'un programme informatique, ou … ».

Ceci signifie, qu'à l'inverse, un réseau téléphonique numérique est à même de transporter des données numériques, de quelque nature que ce soit : c'est la « convergence » de plusieurs applications (néanmoins toutes liées à l'information) vers un moyen unique de communication (le « réseau »).

6.2 Du réseau de voix au réseau de données numériques : qu'est ce que cela change ?

Comme nous venons de le voir, la réponse est « rien » !
Toutefois, il subsiste une différence majeure : le facteur « temps ».

La parole, la musique, le son en général, de même que l'image animée (cinéma, télévision) sont des fonctions du temps. La représentation binaire « porte » ce temps par l'intermédiaire de la fréquence d'échantillonnage (aussi bien pour le son : 8000 échantillons par seconde, que pour l'image animée : 25 images par seconde).

En revanche, une suite de caractères formant un texte n'est associée à aucune échelle de temps, pas plus d'ailleurs que les points formant une image fixe ni les octets d'un programme.

De même (les notions se rejoignent) :
- le débit binaire associé à la transmission d'un son ou d'une image animée sont déterminés par le son ou l'image eux-mêmes (la détermination se fait à cause du théorème de Nyquist – Shannon), le réseau doit pouvoir assurer un tel débit ;

- le débit binaire associé à la transmission d'un texte ou d'une image fixe peuvent être déterminé au contraire par le réseau, le débit fixant alors la vitesse de transmission, donc, finalement le temps de transmission global du texte ou de l'image fixe.

Enfin :
- les échantillons successifs d'un son ou les images successives d'une image animée ne peuvent pas subir d'interruption de transmission sous peine d'un « blanc » (pour le son) ou d'un « noir » (pour l'image) ou encore d'une image qui se fige, on dit que l'information doit être transmise avec « fluidité » ;
- les octets d'un texte ou d'une image fixe peuvent être transmis par « paquets » (de façon, par exemple, à « laisser la place » à d'autres paquets appartenant à d'autres communications »), les intervalles de temps entre les paquets successifs étant arbitraires, mis à part que ces intervalles contribuent à allonger le temps global de transmission.

Pour toutes ces raisons, on distingue deux modes de communication :
- le mode « circuits », adapté en particulier au son et à l'image animé, consistant à établir une liaison continue entre deux points du réseau, réservée uniquement à cette communication ;

- le mode « paquets », adapté en particulier aux « données informatiques » en général (y compris textes et images fixes), consistant en l'acheminement de données groupées en paquets sur une liaison partagé entre plusieurs communications « pseudo - simultanées » (les paquets sont quand même envoyés les une derrière les autres). C'est ce mode qu'utilise « Internet » avec des protocoles de la famille « IP » (*Internet Protocol*).

6.3 « L'inversion » : quand la parole est considérée comme une donnée parmi d'autres (« la fin du téléphone de Graham Alexander Bell ? »)

Néanmoins, la parole est bien « une donnée comme les autres », y compris sur des réseaux en mode « paquets » dès lors que les contraintes de temps et de « fluidité » vues plus haut sont satisfaites.
Ceci est techniquement possible, notamment grâce à une capacité et un débit suffisamment grands pour y satisfaire et pour permettre la coexistence simultanée de l'ensemble des applications de communication : sons, images animées, données informatiques.

Une application de communication dite « voix sur IP » (*VOIP – voice over IP*) est là pour en témoigner.

Alors, ***« la fin du téléphone de Graham Alexander Bell » ?***