

Àudio digital

Francesc Tarrés Ruiz
Javier Melenchón Maldonado

PID_00153332



Els textos i imatges publicats en aquesta obra estan subjectes –llevat que s'indiqui el contrari– a una llicència de Reconeixement-NoComercial-SenseObraDerivada (BY-NC-ND) v.3.0 Espanya de Creative Commons. Podeu copiar-los, distribuir-los i transmetre'ls públicament sempre que en citeu l'autor i la font (FUOC. Fundació per a la Universitat Oberta de Catalunya), no en feu un ús comercial i no en feu obra derivada. La llicència completa es pot consultar a <http://creativecommons.org/licenses/by-nc-nd/3.0/es/legalcode.ca>

Índex

1. Activitat 1. Representació en temps i freqüència de senyals d'àudio.....	7
1.1. Introducció	7
1.1.1. Contingut	7
1.1.2. Metodologia	7
1.1.3. Recursos	7
1.2. Objectius	8
1.3. Guia	9
1.3.1. Introducció	9
1.3.2. Forma d'ona temporal i representació en freqüència	9
1.3.3. Audacity: programa de captura, edició i processament de àudio	20
1.3.4. Audacity: captura de senyals i maneig bàsic de pistes ...	23
1.3.5. Audacity: formes d'ona bàsiques i representació del temps freqüència	27
1.3.6. Superposició de tons	32
1.3.7. Seqüenciació de tons i interpretació de l'espectre	33
1.3.8. Espectrograma	35
1.3.9. Anàlisi d'espectre del senyal de veu	36
1.3.10. Espectrograma d'un xiulet	39
1.3.11. Timbre dels instruments musicals	41
1.4. Problemes/preguntes	43
2. Activitat 2. Digitalització de senyals d'àudio: mostreig.....	46
2.1. Introducció	46
2.1.1. Contingut	46
2.1.2. Metodologia	47
2.1.3. Recursos	47
2.2. Objectius	48
2.3. Guia	49
2.3.1. La retícula de mostra	49
2.3.2. La freqüència de mostreig	53
2.3.3. Anàlisi en freqüència de <i>Panis angelicus</i>	55
2.3.4. Anàlisi en freqüència de <i>Sintonía Ópera</i>	56
2.3.5. Freqüències de mostreig per a senyals de veu	57
2.3.6. Senyals de banda limitada	58
2.3.7. Teorema del mostreig	59
2.3.8. Aliàsing i filtres antialiàsing	60
2.3.9. Alguns exemples d'aliàsing	62
2.3.10. L'aliàsing en vídeo i cinema	66
2.3.11. Selecció de la freqüència de mostreig en sistemes d'àudio	68

2.3.12.	Freqüència de reproducció del senyal d'àudio	70
2.4.	Problemes/preguntes	71
2.5.	Avaluació	72
3.	Activitat 3. Digitalització de senyals d'àudio: quantificació..	75
3.1.	Introducció	75
3.1.1.	Contingut	75
3.1.2.	Metodologia	76
3.1.3.	Recursos	76
3.2.	Objectius	77
3.3.	Guia	78
3.3.1.	La retícula de mostreig i el pas de quantificació	79
3.3.2.	Soroll de quantificació i relació senyal-soroll	85
3.3.3.	La relació senyal-soroll	88
3.3.4.	Codificació PCM (<i>pulse code modulation</i>) del senyal d'àudio	91
3.3.5.	La selecció de les paraules codi en PCM (opcional)	94
3.3.6.	PCM en complement a 2 (CA2) (opcional)	96
3.3.7.	La taxa de bits i l'emmagatzemament de la codificació PCM	97
3.3.8.	Nivells de quantificació i sistemes d'àudio	99
3.3.9.	Avantatges i inconvenients de la digitalització del senyal d'àudio	102
3.4.	Problemes/preguntes	104
3.5.	Avaluació	105
4.	Activitat 4. Filtratge de senyals d'àudio.....	109
4.1.	Introducció	109
4.1.1.	Contingut	109
4.1.2.	Metodologia	109
4.1.3.	Recursos	110
4.2.	Objectius	111
4.3.	Guia	112
4.3.1.	Exemples amb filtratge passabaix	115
4.3.2.	Exemples amb filtratge passaalt	120
4.3.3.	Combinacions de filtres en sèrie i en paral·lel	122
4.3.4.	Filtres de pas de banda	123
4.3.5.	Filtres de <i>shelving</i>	125
4.4.	Problemes/preguntes	127
4.5.	Avaluació	129
5.	Activitat 5. Equalització de senyals d'àudio.....	131
5.1.	Introducció	131
5.1.1.	Contingut	131
5.1.2.	Metodologia	131
5.1.3.	Recursos	132
5.2.	Objectius	133

5.3.	Guia	134
5.3.1.	Equalitzadors paramètrics	135
5.3.2.	El factor de qualitat en els filtres paramètrics	141
5.3.3.	Aplicacions dels equalitzadors paramètrics	142
5.3.4.	Equalitzadors gràfics	146
5.3.5.	Estructura d'un equalitzador gràfic	147
5.4.	Problemes/preguntes	150
5.5.	Avaluació	152
	Bibliografia	155

1. Activitat 1. Representació en temps i freqüència de senyals d'àudio

1.1. Introducció

1.1.1. Contingut

Aquesta activitat s'ha de considerar com una primera aproximació als senyals d'àudio digital i al seu maneig mitjançant programes de captura i edició. L'objectiu principal és que l'estudiant es familiaritzi amb els conceptes de forma d'ona i espectre del senyal d'àudio, i aprengui les relacions que hi ha entre ells. En la primera part de l'activitat es presenten senyals musicals complexos, s'examinen les formes d'ona, i s'aprèn a capturar els senyals des de diferents recursos com el CD o el micròfon. Posteriorment, es treballa amb senyals tonals, conceptualment molt intuïtius, per a introduir el concepte d'anàlisi d'espectre i representació en freqüència del senyal. Una vegada introduïts els conceptes bàsics, s'estén per a senyals de veu i instruments musicals.

1.1.2. Metodologia

Els conceptes s'introdueixen mitjançant exercicis que l'estudiant fa mitjançant una aplicació de captura i edició d'àudio. Els exercicis són molt guiats i cerquen que es vagin comprènent els conceptes bàsics de representació de senyals en el domini del temps i en el domini de la freqüència. Al final de l'activitat es proposa una sèrie de problemes o exercicis addicionals per a consolidar els aprenentatges d'aquesta activitat.

1.1.3. Recursos

Els recursos principals utilitzats en aquesta activitat són:

- Programari de captura i processament d'àudio Audacity (llicència lliure).
- Simulador web audiovisual de tons purs.
- Senyals prèviament capturats i presentats com a materials addicionals de l'activitat.
- Proposta d'activitats de cerca d'informació en webs i recursos d'Internet per a aprofundir en els coneixements adquirits en aquesta activitat (aplicacions de programari d'equalització, fabricants d'equalitzadors, etc.).

- Activitats proposades per a verificar l'adquisició dels coneixements d'aquesta activitat.

1.2. Objectius

Els objectius principals que es persegueixen en aquesta activitat són:

- 1) Introduir el concepte de forma d'ona i representació del senyal en el domini del temps.
- 2) Aprendre el maneig bàsic de programes de captura, edició, processament i reproducció de senyals d'àudio i familiaritzar-s'hi.
- 3) Identificar senyals sinusoidals de diferents freqüències amb el so que produeixen.
- 4) Barrejar diferents senyals sinusoidals i interpretar-ne la forma d'ona i el so que produeixen.
- 5) Veure i interpretar la representació en freqüència de senyals sinusoidals purs.
- 6) Interpretar la representació en freqüència de senyals complexos.
- 7) Aprendre a capturar senyals de veu a partir de micròfons integrats a l'ordinador, i analitzar la forma d'ona i el seu espectre.
- 8) Aprendre a capturar senyals musicals a partir d'una font de CD.
- 9) Introduir el timbre dels instruments musicals.
- 10) Representar el temps-freqüència dels senyals mitjançant l'espectrograma.

Aquests objectius estan relacionats amb les **competències de l'assignatura** següents:

- 1) Capacitat de capturar àudio i imatges de manera eficient i eficaç.
- 2) Capacitat de digitalitzar eficientment i eficaçment continguts d'àudio i imatge.
- 3) Capacitat d'operar de manera digital amb senyals digitals d'àudio i imatge.

I amb les **competències generals del grau** següents:

1) Ser capaç d'analitzar un problema en el nivell d'abstracció adequat a cada situació i aplicar habilitats i coneixements adquirits per a abordar-lo i resoldre'l.

2) Capturar, emmagatzemar i modificar informació d'àudio, imatge i vídeo digitals aplicant principis i mètodes de realització i composició del llenguatge audiovisual.

3) Atendre adequadament consultes sobre projectes, tecnologies i mercat de productes multimèdia, i avaluar de manera precisa l'entorn d'aplicació, els recursos i les alternatives tecnològiques disponibles.

1.3. Guia

1.3.1. Introducció

El senyal d'àudio és una part molt important de qualsevol producció multimèdia, per la qual cosa és fonamental conèixer els sistemes que en permeten la captura, edició, processament i producció. No obstant això, per a poder fer totes aquestes operacions és necessari comprendre les característiques dels senyals d'àudio i els mètodes que hi ha per a representar la informació que contenen.

En aquesta activitat ens centrarem principalment a estudiar dos mètodes alternatius per a representar els senyals d'àudio: la **forma d'ona temporal** i l'**espectre**.

Veurem que aquestes dues formes de representació són complementàries i que podem extreure informació important de cada una. També aprendrem algunes opcions bàsiques d'un programa de captura, edició i processament de senyals d'àudio. Aquesta aplicació de programari s'utilitzarà en totes les activitats d'àudio d'aquesta assignatura. Explicarem els fonaments i eines bàsiques dels programes de captura i edició d'àudio, centrant-nos en concret en l'aplicació que utilitzarem durant l'assignatura.

Iniciarem l'activitat amb una sèrie d'exercicis exploratoris i, a continuació, continuarem amb l'explicació i la pràctica amb una aplicació d'edició i captura d'àudio concret.

1.3.2. Forma d'ona temporal i representació en freqüència

Els sons i les imatges, tant estàtiques com dinàmiques (imatges en moviment o vídeo), arriben als nostres sentits a través de medis físics com l'aire o la llum. L'ésser humà ha anat ideant diferents alternatives al llarg de la seva història per emmagatzemar la impressió que li produeixen aquests estímuls, des de les

primitives pintures rupestres fins als últims reproductors MP3, passant per les fotografies o els discos de vinil. Totes aquestes tecnologies tenen l'objectiu de capturar diferents sèries efímeres d'estímuls, per poder tornar-los a percebre a voluntat.

Durant les darreres dues dècades, els estímuls sonors i visuals estan essent capturats i emmagatzemats en massa a través de les tecnologies de digitalització.

Al llarg d'aquest apartat i dels següents començarem a estudiar amb detall aquest procés de digitalització.

En aquest apartat començarem treballant un concepte clau que intervé en el procés de digitalització de continguts sonors i visuals i que alhora forma part de la naturalesa dels sons (i també de les imatges, tot i que no ho tractarem en aquesta assignatura): el concepte de **freqüència**. Com que relacionarem el so amb la seva freqüència representada sobre una pantalla o paper, també necessitarem representar els diferents sons d'una forma visual; una d'aquestes serà la **forma d'ona**. Per ajudar-nos a desenvolupar i treballar aquests conceptes farem una sèrie d'exercicis exploratoris mitjançant una simulació web audiovisual. Si bé l'explicació podrà seguir-se a través d'aquest text escrit, recomanem a l'estudiant que la treballi a través del web, ja que li permetrà incorporar la informació sonora a les explicacions.

Per poder utilitzar la simulació web amb totes les seves prestacions sonores, és necessari disposar de l'última versió del navegador Firefox. Es poden utilitzar altres navegadors, però només podrà experimentar-se la part visual de la simulació.

Exercici 1. Aspecte d'un so senzill

En aquest exercici aprendrem a distingir les diferents representacions d'un so. Assegureu-vos de tenir els altaveus amb el volum activat. Atesa la naturalesa dels sons implicats, recomanem no tenir un volum alt, especialment si utilitzem auriculars.

- **Situeu el cursor a sobre de la caixa inferior de la simulació web. Després proveu a moure'l dins d'aquesta caixa i responeu a les preguntes següents:**

- 1) Com sentiu el so en desplaçar el cursor a la dreta? I si el desplaceu a l'esquerra?
- 2) Us sembla un so agradable o molest? El trobeu natural o artificial?
- 3) Què creieu que són les animacions que surten a la caixa *Forma d'ona* i *Ones de pressió*?

A continuació expliquem els elements de la simulació i donem resposta a les preguntes formulades.

Resposta:

La caixa de sota de tot, on moveu el ratolí, indica la freqüència seleccionada en l'instant actual mitjançant una línia vermella vertical i una caixa de text amb un nombre seguit de les lletres *Hz*. *Hz* és l'abreviació d'hertz, una unitat física que s'utilitza per a indicar de manera numèrica una freqüència; representa el nombre de voltes, cicles o repeticions per segon. Un nombre alt de Hz indica una freqüència elevada i un nombre petit, una freqüència baixa. En la simulació podeu veure que els valors estan limitats entre 100 Hz i 1.000 Hz, més que suficient per als propòsits d'aquesta. Quan seleccioneu un nombre elevat, movent el ratolí a la dreta, percebreu que el so és més agut que quan seleccioneu un nombre petit, movent el ratolí a l'esquerra.

L'animació que surt sota **ones de pressió** és una representació de la densitat de l'aire per on es desplaça el so. Per desgràcia (o per sort), la pressió a l'aire no té cap incidència en els rajos de llum que la travessen i, per tant, no la podem percebre visualment a la realitat. La representació donada és una proposta del que veuríem si poguéssim percebre visualment els canvis de pressió a l'aire; hem assignat el color cian als punts amb pressió màxima (més gran que l'atmosfèrica) i el color blanc als punts de pressió mínima (més petita que l'atmosfèrica). Observeu que quan un so és més agut, és a dir, la seva freqüència és més elevada, l'aire presenta variacions de pressió més ràpides i menys espaiades. El cas contrari passa amb un so més greu (o menys agut, segons com es vulgui dir). Aquestes ones de pressió, en arribar a les nostres orelles, són les responsables que escoltem els sons tal com els escoltem. Hem associat la descripció d'**agut** a un so de freqüència elevada, com, per exemple, hem associat la descripció de **negre** al color de les línies que formen les vores de les dues caixes que hi ha en la simulació.

Només queda comentar el contingut de la caixa titulada **Forma d'ona**. Aquí podeu veure una línia amb pujades i baixades¹ sempre amb la mateixa cadència (que només canvia quan moveu el ratolí a la caixa de **Freqüències**) que es desplaça amb el temps. Si us hi fixeu, podreu notar que sempre que a les

Heinrich Rudolf Hertz

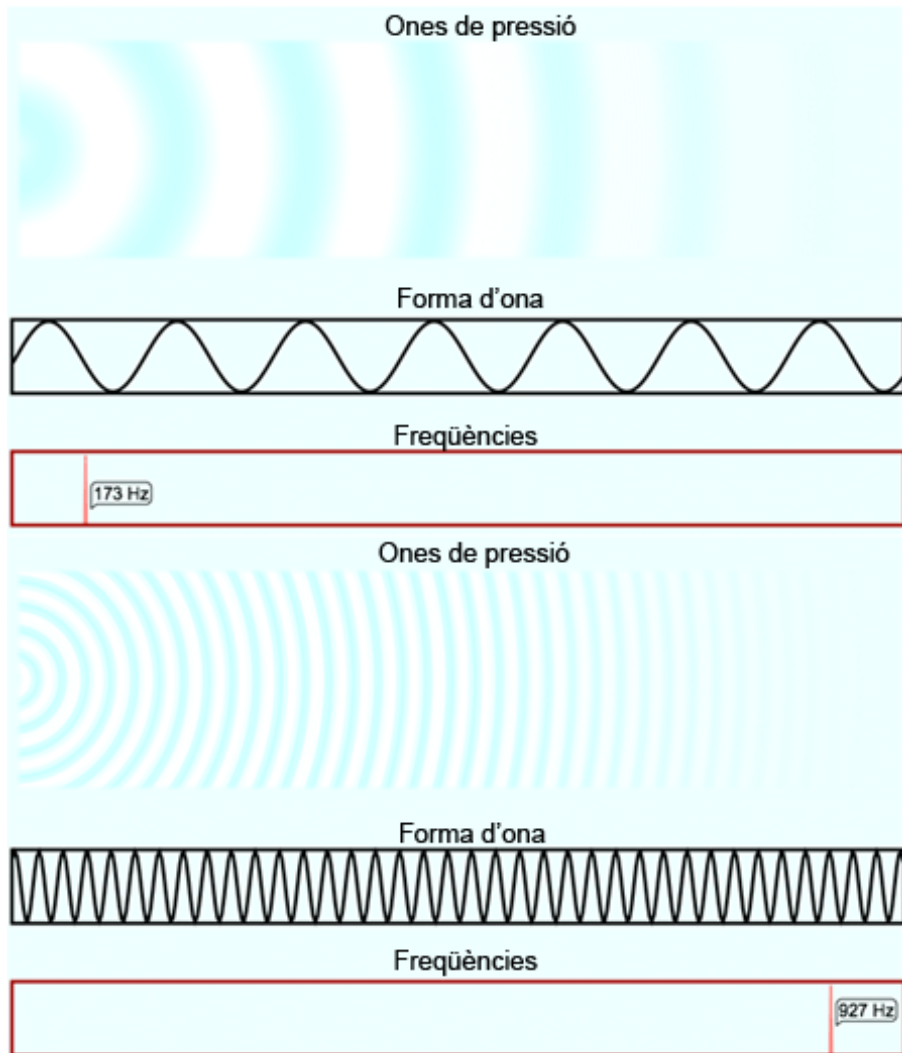
La unitat física de freqüència hertz es posa en honor a Heinrich Rudolf Hertz, físic alemany descobridor de la propagació de les ones electromagnètiques.

⁽¹⁾Una pujada i baixada completes i seguides forma el que s'anomena un *cicle*.

ones de pressió hi ha el color blau, que indica una pressió elevada, la línia de la **forma d'ona** se situa a la part alta de la caixa; al contrari, quan les **ones de pressió** mostren un color blanc, que indica una pressió baixa, la línia de la **forma d'ona** se situa a la part baixa de la caixa (aquestes dues correspondències són més fàcils d'observar per a freqüències petites). La resta de tonalitats de blau tenen també la seva correspondència amb la **forma d'ona**, la qual se situa més a dalt per a tonalitats més properes al blau i més a baix per a tonalitats menys saturades i, per tant, més properes al blanc. La representació en **forma d'ona** d'un so és una manera abstracta de representar amb una línia² les variacions de pressió a l'aire que formen un so i, per tant, el so en si mateix.

⁽²⁾La línia representada per la forma d'ona es pot dibuixar com una sèrie de valors numèrics molt seguits i units mitjançant segments molt curts que donen l'efecte de corba que puja i baixa. Aquesta representació té unes limitacions, com es veurà més endavant quan es parli del mostreig.

Figura 1. Aspecte visual mostrat per la simulació per a una freqüència petita (a dalt) i una freqüència elevada (a baix)



Noteu l'aspecte que pren el so de les freqüències seleccionades, tant en forma d'ones de pressió a l'aire com en la representació abstracta de forma d'ona.

Els sons que reproduïx la simulació i amb els quals heu experimentat en aquest exercici reben el nom de **tons purs**. Aquest tipus de tons estan formats per una única freqüència (d'aquí el qualificatiu de *purs*) i són inexistents a la natura per si sols i de forma aïllada. Haureu notat que són sons molestos, ja que no estem acostumats a percebre'ls a la vida diària, però tenen gran ressonància dins del nostre aparell auditiu. Tot i així, tots els sons es poden obtenir a partir de tons purs (o es poden descompondre en una agregació de tons purs, depèn de com es miri).

Vegeu també

Un tast de la creació de sons a partir de tons purs el podem observar en l'exercici 4; de fet, en l'assignatura de *Tractament i publicació d'àudio* es veurà com crear sons a partir d'aquests elements.

Exercici 2. Aspecte de dos tons purs simultanis

En aquest exercici explorarem l'aspecte visual i sonor que té una agregació de diferents tons purs. La simulació està preparada per a compensar el volum de reproducció de diferents tons purs simultàniament; tot i així, igual que en l'exercici anterior, es recomana partir d'un volum baix dels altaveus o auriculars i apujar-lo si no se sent suficient.

Situeu el cursor a la caixa de freqüències i seleccioneu un to pur d'una freqüència de 120 Hz. Quan la tingueu seleccionada feu clic al damunt amb el ratolí i tot seguit moveu el ratolí fora de la caixa de freqüències. Hi podreu observar que el to escollit ha quedat fixat.

Observació

A la simulació web, si voleu tornar a començar podeu prémer el botó de *Reset*, o el botó de *Pausa* per congelar la reproducció. Tingueu en compte que si torneu a moure el ratolí dins de la caixa de freqüències es desfà la pausa (igual que si torneu a prémer el botó de *Pausa*).

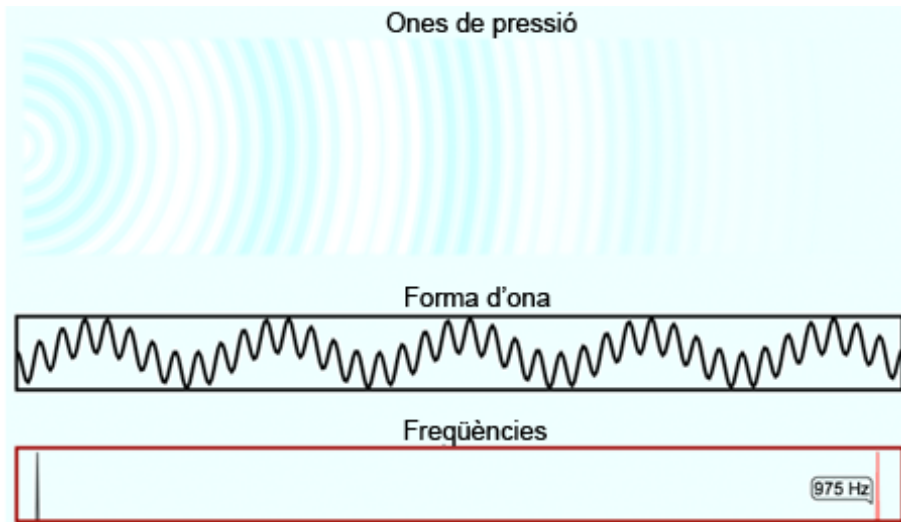
Movent ara el ratolí dins de la zona de freqüències elevades (per sobre dels 800 Hz), responeu a les preguntes següents:

- 1) Què observeu sobre la forma d'ona i ones de pressió? Descriviu el so que sentiu.
- 2) Si moveu el ratolí per sota dels 800 Hz i us aneu apropant a la freqüència seleccionada, continueu veient i sentint el mateix? Si no, descriviu quan noteu un comportament diferent i quin és.

Resposta:

Podreu apreciar que sentiu els dos tons, el que heu fixat a 120 Hz i el que teniu seleccionat en tenir el ratolí dins de la caixa i per sobre dels 800 Hz. Tots dos tons viatgen per l'aire junts, a través d'ones de pressió (i forma d'ona) compostes per tots dos tons. En la forma d'ona es pot veure clarament una vibració ràpida que va pujant i baixant. Aquest aspecte es deu al fet que les dues formes d'ona corresponents a cada to pur s'han agregat en forma de suma. Els diferents tons purs i, en general, sons que viatgen pel mateix medi (normalment aire, a través d'ones de pressió) sempre se sumen. Per tant, el que percebem normalment per l'oïda sempre és una suma de diferents sons.

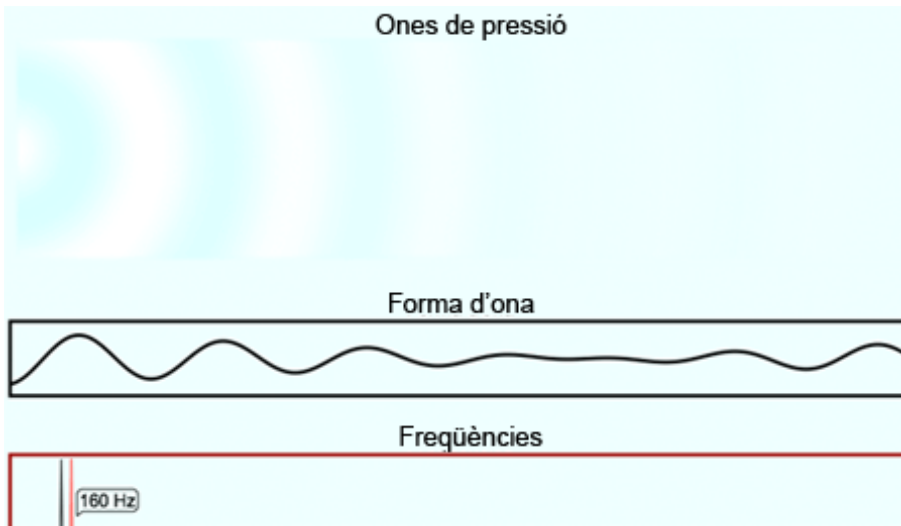
Figura 2. Visualització de dos tons purs simultanis de freqüències molt diferents



Si anem reduint el valor de la freqüència del to seleccionat pel ratolí (disminuint-la progressivament per sota de 800 Hz), podrem veure com la vibració ràpida és cada cop més lenta, però continua pujant i baixant al mateix ritme, ja que la vibració més lenta (la de 120 Hz) està fixada. Hi ha un punt al voltant dels 250 Hz en què visualment no distingim, a partir de la forma d'ona, una vibració ràpida i una altra de més lenta; és més, hi ha un punt al voltant dels 140 Hz en què ja no podem distingir els dos tons. Tots dos hi són, barrejats, però la nostra oïda presenta limitacions físiques, i aquesta és una d'elles, no poder distingir tons de freqüències properes quan sonen simultàniament.

Quan dos tons se situen molt propers, el que es produeix és la percepció d'un sol to pur amb la presència d'increments i decrements de volum més o menys ràpids en funció de com de prop se situen els dos tons. Aquest tipus d'efecte es denomina *efecte psicoacústic* i es veurà amb més detall en properes assignatures. Visualment es pot distingir també, en veure una vibració que augmenta i disminueix de volum lentament (més lentament com més propers són els dos tons). Com que la forma d'ona es representa a una velocitat sensiblement inferior a la del so (altrament no es veuria res), cal esperar-se una mica per a poder observar aquest fenomen sobre la forma d'ona i ones de pressió.

Figura 3. Visualització de dos tons purs molt similars



S'hi pot apreciar (de forma molt local) com la forma d'ona augmenta i disminueix de volum mantenint una sola freqüència d'oscil·lació aparent.

La simulació permet afegir fins a 40 tons purs. Podeu fer proves i experimentar els efectes de l'addició de múltiples tons purs. En propers exercicis explotarem aquesta característica (concretament a l'exercici 4).

Exercici 3. Amplitud i sonoritat

En aquest exercici explorarem una característica més d'una ona sonora, l'amplitud i com afecta aquesta a la sonoritat.

Creeu un to pur d'una freqüència de 240 Hz fent clic sobre la caixa de freqüències i seleccionant la de 240 Hz. A continuació, situeu-vos sobre la freqüència de 320 Hz. Sentireu dos tons diferenciats. Feu clic sobre la freqüència de 320 i no mogueu el ratolí; veureu que el text de la caixa on s'indica la freqüència canvia a vermell³ i apareix *Amplitud: 1*, això significa que teniu seleccionat aquest to pur. Fent clic a sobre n'augmenteu l'amplitud. Veureu que el to de 240 Hz es fa més petit a mesura que augmenteu l'amplitud del de 320 Hz; això es deu a qüestions de representació de la simulació, en què tot s'escala segons l'amplitud més gran. També podeu moure la rodeta del ratolí endavant i endarrere per augmentar o també per disminuir l'amplitud, respectivament. Si se us ha mogut el ratolí, desseleccionant el to de 320 Hz, només heu de tornar a situar-lo a sobre per poder continuar modificant-ne l'amplitud.

⁽³⁾En situar el ratolí sobre un to existent, se selecciona i podem variar-ne l'amplitud.

Responen a les preguntes següents:

1) Sentiu tots dos tons simultàniament amb qualsevol canvi d'amplitud en el to de 320 Hz? Quan deixeu de sentir-ne un o l'altre? Comenteu també la forma d'ona que veieu en fer canvis a l'amplitud.

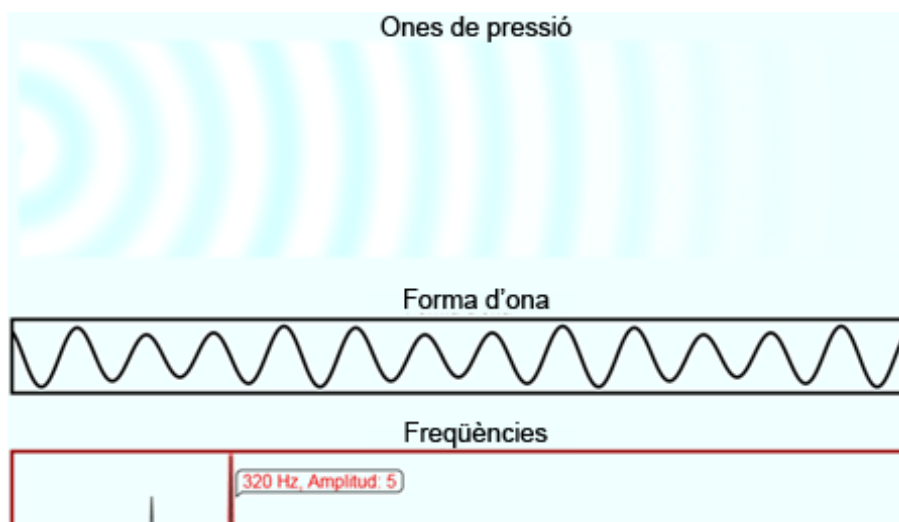
2) Torneu a respondre la pregunta 1 anterior seleccionant un to de 940 Hz en comptes del de 320 Hz (haureu de prémer *Reset* per poder esborrar els tons que hi ha).

Resposta:

Quan augmenteu l'amplitud del to de 320 Hz (o del de 940 Hz) veureu que la seva forma d'ona pren molta importància fins al punt de gairebé anul·lar (comparativament) la del to de 240 Hz; el resultat és una forma d'ona que gairebé sembla un sol to pur a 320 Hz (o 940 Hz), tot i que no ho és mai, ja que el to de 240 Hz sempre hi és present, encara que sigui amb poca importància relativa.

Aquest fet té impacte sobre la nostra oïda. Podreu comprovar que amb una amplitud de 4 (o 5, depenent de la persona i del volum de l'altaveu) el to de 320 Hz anul·la la sonoritat del to de 240 Hz. No obstant això, el to de 940 Hz necessita una amplitud més gran per a fer que deixem de sentir el de 240 Hz. Com més diferència hi ha entre els dos tons, més gran ha de ser un dels dos per a anul·lar acústicament l'altre. Aquest efecte s'anomena *emascarament de freqüències* (sonores) i és la base en què se sustenta l'MP3 per a fer les compressions d'àudio: si una freqüència no se sent, perquè té una freqüència propera molt més forta, no es guarda la que no se sent (o s'elimina), i així s'estalvia espai d'emmagatzemament.

Figura 4. Un to de 240 Hz a amplitud 1 no se sent quan hi ha un to de 320 Hz a amplitud 5



La forma d'ona del to de 320 Hz predomina sobre la de 240 Hz, però visualment es detecta que no hi ha només un sol to pur, encara que només en sentim un.

L'amplitud té a veure amb la sonoritat, tot i que no amb una raó proporcional. Augmentar l'amplitud al doble no fa que sentim el so el doble de fort. Perquè un so soni un punt més alt, cal multiplicar l'amplitud per l'arrel quadrada de 10, o el que és el mateix, per 3,1623 (aproximadament). Així, si volem que el to pur de 320 Hz soni el doble de fort que el de 240, cal donar-li una amplitud de 3; si volem que soni el triple de fort, cal donar-li una amplitud de 10; pera tenir-ho el quàdruple de fort, cal una amplitud de 31; i així successivament. Això es deu al fet que la nostra oïda mesura la sonoritat de forma logarítmica i ens permet augmentar molt el rang d'intensitats a sentir, respecte a fer-ho amb una regla proporcional. La representació visual escollida en la freqüència, és a dir, l'alçària dels pics, reflecteix la sonoritat en comptes de l'amplitud.

Transformada de Fourier

Com a curiositat, la gràfica que dibuixarem a la caixa de freqüències rep el nom de *transformada de Fourier* de la forma d'ona dibuixada a sobre, tot i que en l'aplicació es presenta de manera simplificada.

Exercici 4. Aspecte d'un so complex

En aquest exercici utilitzarem diferents tons purs per a crear un so més complex aprofitant les capacitats de la simulació web. Cal aclarir que, en tractar-se d'una simulació senzilla, s'obindrà una aproximació d'un so conegut, però no una reproducció exacta.

Utilitzant la simulació web, creeu set tons purs amb les components de freqüència i amplitud següents:

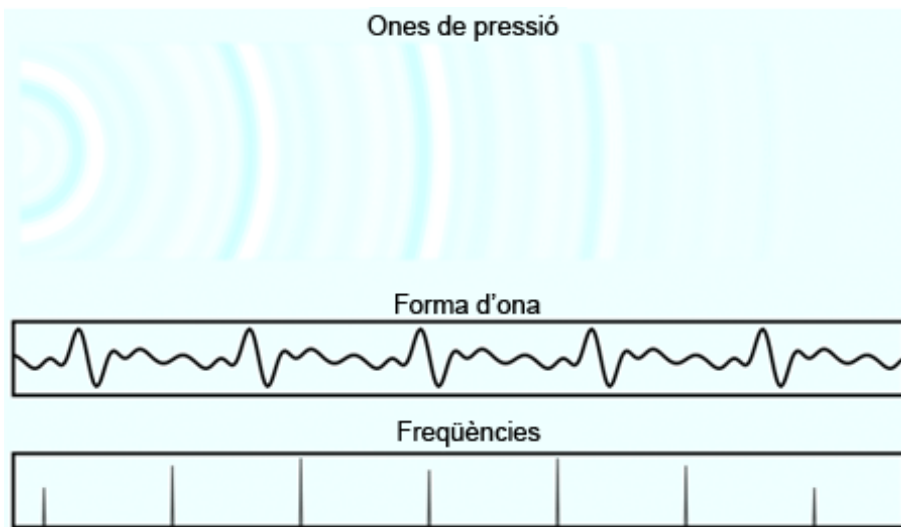
	To pur 1	To pur 2	To pur 3	To pur 4	To pur 5	To pur 6	To pur 7
Freqüència	130 Hz	260 Hz	390 Hz	520 Hz	650 Hz	780 Hz	910 Hz
Amplitud	3	10	16	8	16	10	3

Escolteu el so produït amb atenció. Us recorda algun tipus de so conegut? Aquells amb experiència musical tindreu avantatge en contestar aquesta pregunta.

Resposta:

El so produït pels set tons purs proposats és una aproximació d'un Do2 produït per un fagot. No és un so exacte, ja que es necessitarien molts més tons purs per a recrear-lo amb precisió, però podeu veure que amb un conjunt limitat es pot obtenir una bona aproximació. El fet de crear sons complexos a partir de tons purs forma part de la disciplina coneguda com a *síntesi de so additiva*.

Figura 5. Aproximació de l'aspecte visual del so produït per un fagot tocant un Do2 en forma d'ona, ones de pressió i components freqüencials



Exercici 5. Mesures d'escala de la simulació

En aquest exercici obtindrem les mesures d'escala de les ones de pressió de la simulació web. En tractar-se d'una simulació, i com s'ha pogut experimentar en exercicis anteriors, la forma d'ona es mou a una velocitat molt inferior a la que tindria a la realitat per a poder apreciar-ne visualment l'evolució.

Per fer aquest exercici heu de saber que la finestra que mostra tant la forma d'ona com les ones de pressió va 50 vegades més lenta que la velocitat real. Recordem també que la velocitat del so a l'aire és de 343 metres per segon.

Primer de tot, determineu quina seria la durada real d'una ona que es desplaça per tota la caixa de forma d'ona, és a dir, quanta estona es visualitzaria dins de la caixa de forma d'ona si la simulació la mostrés a temps real i no de forma alentida?

Un cop sabeu la dada anterior, esbrineu quines serien les mides (amplada i alçària), en metres, de la zona de les ones de pressió, en cas de que tingués mides reals.

Resposta:

Per a respondre a la primera pregunta només necessitem la primera dada i mesurar el temps que la simulació mostra una forma d'ona internament. Per a fer-ho, necessitarem un cronòmetre. Obtindrem la mesura experimentalment i de forma aproximada. Si ho poguéssim mesurar amb plena exactitud, veuríem que les ones triguen 2 segons exactes a travessar la caixa de forma d'ona. Com que aquesta velocitat mostrada és 50 vegades inferior a la real, significa que, a la realitat, trigaria 50 vegades menys, és a dir 2 segons entre 50, que

són 0,04 segons. Les ones travessarien la caixa de **forma d'ona** en 4 centèsimes de segon. Està clar que hem d'alentir-ho si hi volem apreciar alguna cosa visualment.

Per a respondre a la segona pregunta, necessitem el resultat de la primera i saber que el so es desplaça a 343 metres per segon. Si de punta a punta de la caixa de **forma d'ona** hi ha 0,04 segons en temps real, en aquest temps, un punt de pressió d'aire "viatja" 13,72 metres ($0,04 \text{ s} \times 343 \text{ m/s}$). Així, l'amplària de la regió simulada d'ones de pressió ocuparia, si es dibuixés a escala 1:1, una distància de 13,72 metres.

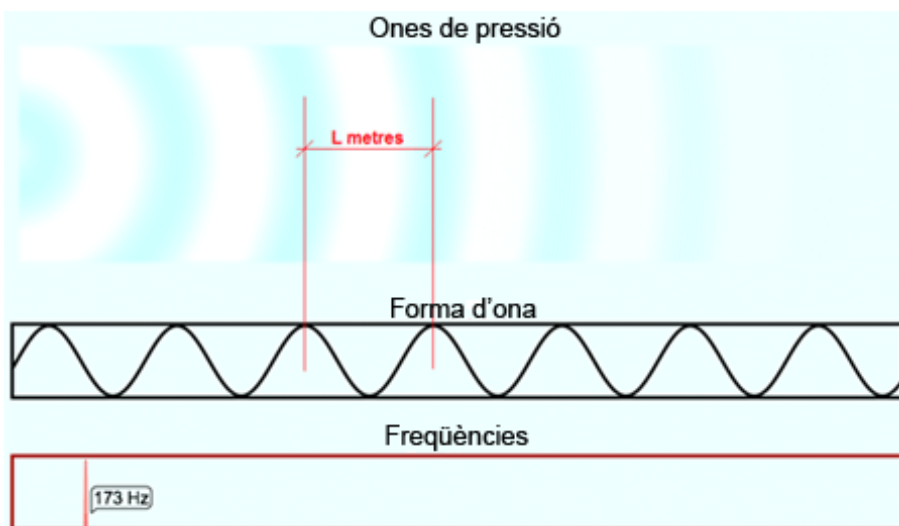
Exercici 6. Període i longitud d'ona

En aquest exercici s'identifiquen dos conceptes associats directament amb la freqüència: la **longitud d'ona** i el **període**.

La **longitud d'ona** es mesura en metres i és la distància que hi ha entre dos punts d'una ona que es repeteixen després de completar un cicle, per exemple, el temps que passa entre dos pics d'un to pur, com es pot veure en la figura 6.

El **període** es mesura en segons i es el temps que triga a completar-se una longitud d'ona, o el que és el mateix, el temps que triga a completar-se un cicle; en el cas d'un to pur, el temps que triga a completar una oscil·lació, o el temps que triga a acabar en el punt més alt des del punt més alt immediatament anterior.

Figura 6. Longitud d'ona d'una ona sonora, mesurada en metres



Sabent la definició anterior i recordant que el so es desplaça a 343 metres per segon, quina és la longitud d'ona d'un to pur de 1.000 Hz? I la d'un to pur de 100 Hz? Quin és el seu període en segons?

De fet, si l'oïda humana capta entre 20 Hz i 20.000 Hz, quina és la longitud d'ona associada a aquests dos límits? I el seu període?

Resposta:

La longitud d'ona d'un to pur de 1.000 Hz és de 0,343 metres i la d'un to de 100 Hz és de 3,43 metres. El període d'un to pur de 1.000 Hz és de 0,001 segons i el d'un to pur de 100 Hz és de 0,01 segons. La longitud d'ona d'un to pur de freqüència de 20 Hz és de 17,15 metres i el seu període és de 0,05 segons. La longitud d'ona d'un to pur de freqüència de 20.000 Hz és de 0,01715 metres (o 17,15 mil·límetres) o té un període de 0,00005 segons (o 0,05 mil·lèsimes de segon).

1.3.3. Audacity: programa de captura, edició i processament de àudio

Hi ha un gran nombre d'aplicacions per a la **captura, edició i processament** del senyal d'àudio entre les quals podem destacar:

- Acoustica,
- Amadeus Pro,
- Creative WaveStudio,
- GoldWave,
- Ardour,
- Qtractor,
- SoundForge,
- SoundBooth,
- Wave Editor,
- etc.

Totes aquestes aplicacions permeten la captura de diverses pistes d'àudio procedents de micròfons, entrades de línia, CD, fitxers digitals, etc. És possible tallar i enganxar fragments, aplicar efectes independents a cada pista i barrejar els resultats en una banda sonora final amb múltiples pistes. Alguns d'aquests programes són d'ús lliure, encara que els més populars (WaveStudio, SoundBooth, SoundForge) només solen proporcionar-se amb la compra de maquinari de digitalització específic o la compra amb el distribuïdor.

El programa que utilitzarem per a l'edició i el registre de senyals d'àudio en aquesta assignatura és **Audacity**, un programari lliure de codi obert que ja porta molts anys en el mercat i que actualment pot competir quant a funcionalitats i qualitat del tractament amb els productes comercials més reconeguts.

Descàrrega

L'Audacity es pot descarregar directament des d'<http://audacity.sourceforge.net>.

Sistemes operatius

Gairebé tots els programes anteriors tenen una versió per al Windows i alguns disposen de versions per al MAC i/o Linux.

Sistemes operatius

L'Audacity disposa de versions per als sistemes operatius més utilitzats: Windows XP, Windows Vista, MAC OS X, GNU/Linux.

Durant la instal·lació pot escollir l'idioma amb què es vol fer. Possiblement us sentireu més còmodes en castellà, encara que si la feu en anglès us ajudarà a familiaritzar-vos amb el tipus de vocabulari utilitzat en sistemes d'àudio.

El **procés d'instal·lació** és molt simple.

- Per al **Windows**, us heu de descarregar el fitxer `audacity-win-unicode-x.x.x.exe` (en què `x.x.x` representen la versió del programa) en un directori des del qual vulgueu fer la instal·lació. En executar el programa us demana permís per acceptar les condicions de la llicència GNU, s'instal·la per defecte en el directori d'Arxius de Programes i crea per defecte les icones del programa en el menú inici i en l'escriptori.
- El procediment d'instal·lació per a **MAC OS X i Linux** és encara una mica més fàcil, ja que en descarregar i descomprimir el programa de l'aplicació es genera directament un fitxer executable en el directori d'aplicacions.

Quan **s'inicia el programa** Audacity, apareix una pantalla d'ajuda que us permet fer diversos programes d'aprenentatge per a reproduir arxius, gravar veu, editar so, exportar i importar formats d'àudio, etc. Si ho voleu podeu elaborar algun d'aquests programes encara que durant aquesta activitat explicarem amb detall alguna de les funcions anteriors. Aquesta **pantalla de benvinguda** es mostra en la figura 1.

Figura 1. Pantalla de benvinguda del programa Audacity



Es recomana seguir algun dels programes d'aprenentatge si es tenen dificultats amb el maneig del programa i el seguiment d'aquesta activitat.

Si premem el botó d'acceptar apareix la **pantalla principal** del programa Audacity, que es mostra en la figura 2.

⁽⁴⁾ Les figures que es mostren poden no coincidir exactament amb les de la versió d'Audacity que estigueu utilitzant, l'idioma en el qual hàgiu fet la instal·lació o el sistema operatiu. En general, les diferències haurien de ser mínimes i l'activitat s'hauria de poder seguir sense gaires dificultats. Les figures que presentem durant aquesta activitat procedeixen de la versió en castellà del sistema operatiu Vista.

Figura 2. Pantalla principal del programa Audacity



El programa disposa de diversos **menús en la part superior** de la finestra, en els quals trobem les funcions habituals:

- **Menú *Archivo***. Disposa d'opcions per a crear nous arxius, obrir i desar arxius, desar projectes, i exportar i importar senyals d'àudio en diferents formats.
- **Menú *Editar***. Permet fer funcions de retallar i enganxar amb les pistes d'àudio, dividir pistes, gestionar silencis, etc.
- **Menú *Ver***. S'utilitza principalment per a determinar la fracció del senyal d'àudio que es mostra en la finestra, i controla per tant l'escala horitzontal (zoom) amb què es presenten les gràfiques.
- **Menú *Control***. Gestiona els enregistraments, reproduccions pauses, barreges, anar al principi o al final de pistes, etc.
- **Menú *Pistas***. Permet afegir noves pistes al projecte, alinear-les, silenciar algunes pistes i afegir etiquetes a les pistes.
- **Menú *Generar***. Es poden generar diferents tipus de senyals artificials (sintètics). Aquest menú s'utilitzarà molt en aquesta activitat per a generar senyals senzills que ens ajudin a interpretar les relacions entre la representació en temps i la representació en freqüència.
- **Menú *Efecto***. En aquest menú es disposa de diferents possibilitats per a generar efectes sobre el senyal d'àudio. Les funcions que inclou depenen molt de la versió del programa, ja que actualment les actualitzacions del programa s'estan centrant principalment en la incorporació d'efectes addicionals. En aquesta assignatura només veurem alguns dels efectes conceptualment més senzills d'interpretar.
- **Menú *Analizar***. S'utilitza per a determinar la representació en freqüència del senyal (espectre). Disposa a més d'altres eines d'anàlisi que permeten

trobar silencis de manera automàtica, determinar el ritme i el compàs musical, etc.

- **Menú *Ayuda*.** Inclou una interfície per a la cerca d'ajuda per temes i paraules clau, l'accés a la pantalla de benvinguda i els programes d'aprenentatge, la informació de versió i un test de rendiment de la màquina.

A més d'aquestes opcions de menú, el programa disposa de diversos **botons** que permeten accedir ràpidament a les funcions més habituals. En la figura 2 es mostren algunes les opcions que s'utilitzen amb més freqüència, com són els botons de reproducció i els controls de volum de la sortida i del micròfon.

En els exercicis següents anirem presentant alguns exemples sobre les característiques dels senyals d'àudio, alhora que s'exposarà com es duen a terme algunes de les funcions d'Audacity.

1.3.4. Audacity: captura de senyals i maneig bàsic de pistes

En aquest exercici aprendrem com es captura un senyal d'àudio procedent d'un CD i un senyal de veu procedent d'un micròfon gestionant la manipulació bàsica de les pistes d'àudio i interpretant les formes d'ona.

- 1) Per començar seleccioneu un CD⁴ d'àudio qualsevol.
- 2) Per capturar el senyal d'àudio d'un CD, introduïu el CD i seleccioneu en el menú *Archivo* l'opció *Importar* i després *Audio*.

Apareixerà l'explorador d'arxius.

- 3) Seleccioneu el CD i la pista d'àudio que voleu capturar.

Aquesta operació resulta més senzilla en uns sistemes operatius que en d'altres.

- Si esteu utilitzant el **MAC OS X** o el **Linux** simplement haureu de seleccionar el tema d'interès, i el sistema operatiu seleccionarà de manera automàtica els convertidors adequats i carregarà l'arxiu a Audacity.
- El sistema operatiu **Windows** us demanarà que primer passeu del format **CDA**, propi del disc compacte d'àudio, al format **WAV**. Hi ha diferents maneres de fer aquest procés. Possiblement, la més senzilla és obrir el reproductor de **Windows Media**, seleccionar *Herramientas\Opciones*, i en la pestanya de *Copiar Música desde CD* seleccionar el format **WAV** (sense pèrdua). Després, es tracta de seleccionar el tema del disc que es vol convertir a WAV i fer la còpia (amb el botó *Iniciar copia desde CD*). Per

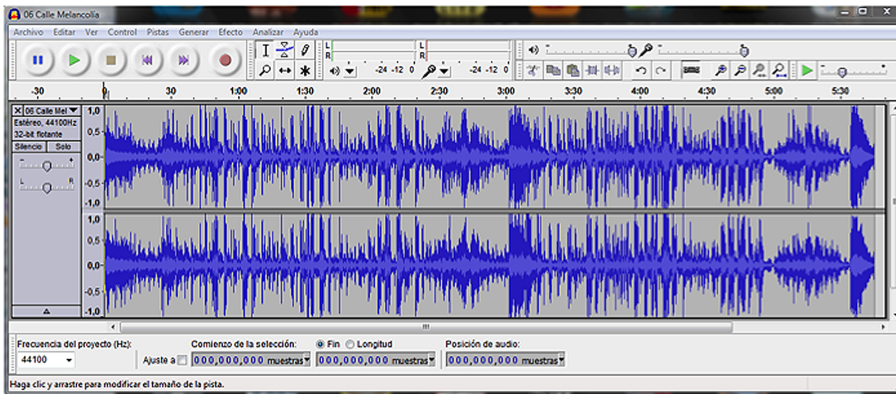
⁽⁵⁾Se suggereix prendre un tema exclusivament musical o amb un vocalista que vulgui acompanyar cantant (recordeu que en l'exercici barrejarem dues pistes, la del CD amb la de la vostra pròpia veu).

defecte, el tema seleccionat es convertirà al format WAV i s'arxivarà en el directori predefinit del Windows *Mi música*.

4) Ara, per a obrir-lo amb Audacity hem de seleccionar l'opció **Importar\Audio** en el menú **Archivo**. S'obrirà la finestra de l'explorador d'arxius i haurem de seleccionar el tema que acabem de convertir a WAV.

En la figura 3 es mostra una captura de la pantalla d'Audacity una vegada s'ha carregat el tema musical. Es tracta d'un arxiu estèreo, per la qual cosa apareixen les formes d'ona del senyal del canal dret i la del canal esquerre. En la part superior de la pista estèreo hi ha representat un eix de temps que en el nostre exemple arriba pràcticament als 6 minuts. Els eixos verticals de les gràfiques estan normalitzats entre -1.0 i 1.0 , la qual cosa significa que el senyal sempre estarà comprès entre aquests dos valors.

Figura 3. Pantalla d'Audacity una vegada carregat un tema musical estèreo complet (aproximadament 6 minuts)



5) Pot reproduir-se el tema musical prement el botó de **Play** (en verd).

A mesura que té lloc la reproducció es pot observar el desplaçament del cursor, que indica el moment de la reproducció en què ens trobem. Fixeu-vos que l'amplitud que prenen els senyals en l'eix vertical depèn de la intensitat sonora. Així, en els moments de poca intensitat sonora veiem que els nivells dels senyals són baixos i que augmenten quan augmenta la intensitat. En efecte, la forma d'ona que es representa en la pantalla es correspon amb la tensió que, degudament amplificada, enviaríem a un altaveu per a la reproducció del senyal d'àudio.

En la línia inferior se'ns indica que la freqüència de mostreig del projecte és de 44.100 Hz. Això significa que treballem en un format d'àudio digital que pren un total de 44.100 mostres per segon del senyal (analitzarem amb més detall aquest paràmetre en l'activitat d'àudio següent). També en la línia inferior se'ns indica la posició del cursor en format numèric. Per defecte es veu el número de mostra en què està situada la selecció del senyal d'àudio.

Per manejar l'arxiu més fàcilment i fer algunes proves sobre edició seleccionarem un fragment d'uns 30 segons.

6) Amb el ratolí marqueu un fragment del senyal d'aproximadament 30 segons (cap a la meitat del tema).

7) Una vegada seleccionats aquests 30 segons, en el menú **Editar**, seleccioneu **Recortar** (Ctrl+T).

Veureu que desapareix tota la informació que no estava seleccionada.

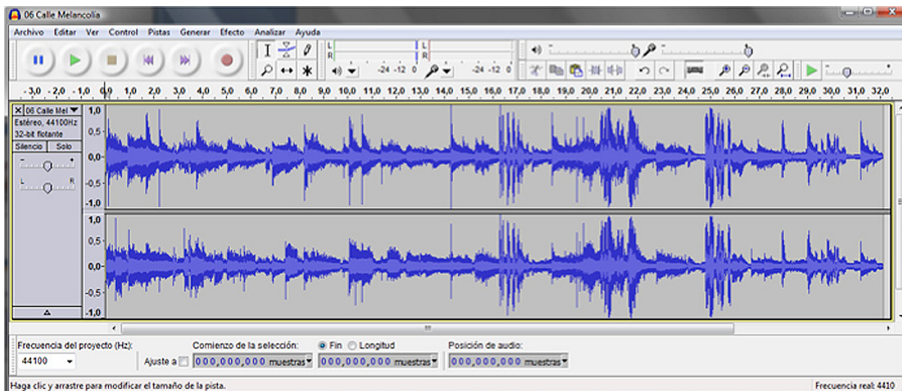
8) Per situar aquest segment al principi de la pista podem fer **Cortar** (Ctrl+X), després situar el cursor al principi de la pista i activar **Pegar** (Ctrl+V).

9) Ara, en el menú **Ver** executeu l'ordre **Ajustar a la ventana** (Ctrl+F) i es mostrarà el detall de les formes d'ona en tota la finestra de l'Audacity.

10) **Reproduïu** l'arxiu i fixeu-vos com només estem treballant amb un fragment de l'arxiu original.

En la figura 4 mostrem el resultat que hem obtingut en el nostre exemple. Ara, l'eix horitzontal té una durada aproximada de 32 segons.

Figura 4. Retallada i selecció d'un fragment d'uns 30 segons



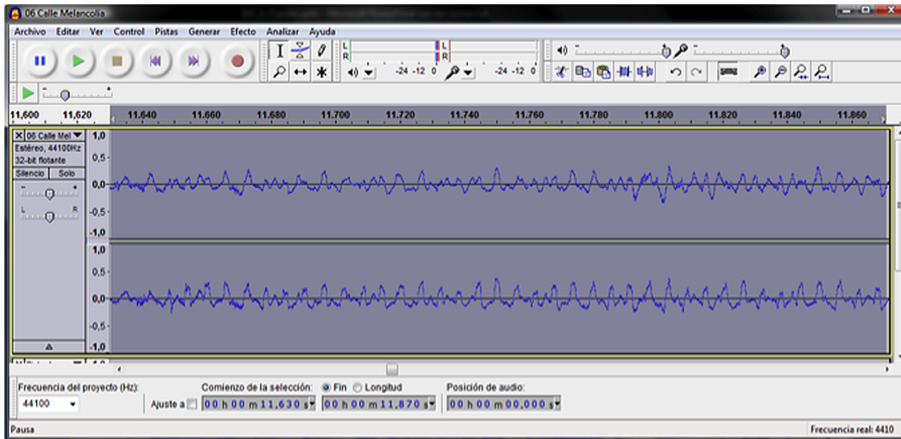
11) Podem veure amb més detall la forma d'ona del senyal seleccionant un fragment amb el ratolí (seleccioneu aproximadament la meitat d'un segon o menys) i mostrant en la pantalla el detall seleccionat mitjançant l'ordre **Ver/Ampliar la selecció** (Ctrl+E). Podeu tornar a visualitzar la pista completa amb Ctrl+F.

En la figura 5 es mostra el resultat d'ampliar la forma d'ona en uns 0,25 segons. Ara és possible veure el detall de la representació del senyal, és a dir, de la tensió que enviaríem a un altaveu per a la seva reproducció.

Observació

Fixeu-vos que el maneig de les ordres Cortar, Pegar i Copiar és anàleg al d'un processador de textos. Simplement, en comptes de moure'ns per un document ens movem en les pistes d'àudio.

Figura 5. Ampliació d'un fragment d'uns 240 ms per a visualitzar el detall de la forma d'ona



Ara registrarem una pista de veu utilitzant el micròfon incorporat a l'ordinador.

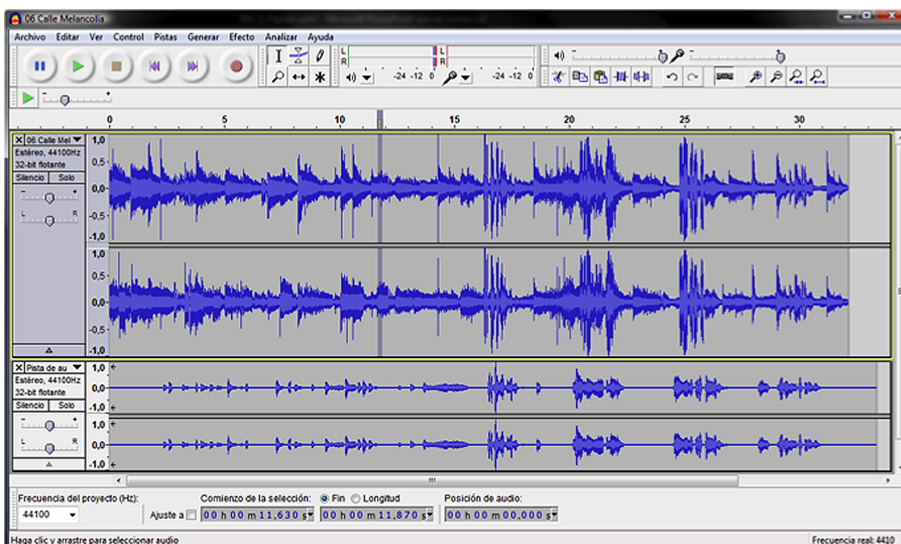
1) Situeu el cursor a l'inici de la pista d'àudio, poseu-vos uns auriculars i disposeu-vos a gravar la vostra veu en una nova pista d'àudio. Per a això, simplement heu de prémer el botó de **Rec** (botó vermell).

Pels auriculars escoltareu la pista original mentre que pel micròfon es registrarà la vostra veu. Acompanyeu el cantant o parleu mentre té lloc l'enregistrament.

2) Ara podeu **reproduir** i veureu com el programa barreja totes les pistes.

El resultat que hem obtingut en el nostre exemple es mostra en la figura 6.

Figura 6. Detall de la pista d'àudio procedent del CD més la pista d'àudio procedent d'un enregistrament de veu



Podeu actuar sobre el resultat de la barreja mitjançant el tauler de control de pistes que apareix al costat de les pistes. Aquest tauler proporciona informació sobre el tipus de senyal (estèreo, 44.100 Hz, 32-bit flotant) i disposa de diversos controls per a la barreja. El botó *Silencio* deixa la pista en silenci durant la

reproducció, el botó *Solo* desactiva la resta de pistes durant la reproducció, i només se sent la pista d'interès. Les dues barres de desplaçament horitzontals són per al control de volum individual de la pista i per al control de balanç entre el canal dret o l'esquerre. Si voleu aplicar efectes en una pista, haureu de seleccionar una pista amb el cursor i aplicar algun dels efectes disponibles en el menú *Efecto* (proveu, per exemple, d'aplicar un efecte d'eco en la pista de veu que acabeu de registrar).

Si us ve de gust, es poden incorporar noves pistes fent més enregistraments. Podeu incloure senyals de línia externs si el vostre ordinador disposa de connectors. Se suggereix que es facin diverses proves amb diverses pistes, manipulant la informació, aplicant funcions de retallar i enganxar fragments d'àudio, modificant els nivells dels senyals, aplicant efectes en les diferents pistes, etc., fins que us trobeu còmodes amb l'aplicació i el maneig de les posades.

3) Finalment, una vegada ajustats tots els paràmetres de les pistes podeu enregistrar el resultat final mitjançant l'ordre *Exportar* que trobareu en el menú *Archivo*.

1.3.5. Audacity: formes d'ona bàsiques i representació del temps freqüència

En aquest exercici generarem algunes formes d'ona bàsiques, n'escoltarem els resultats i n'analitzarem la composició en freqüència.

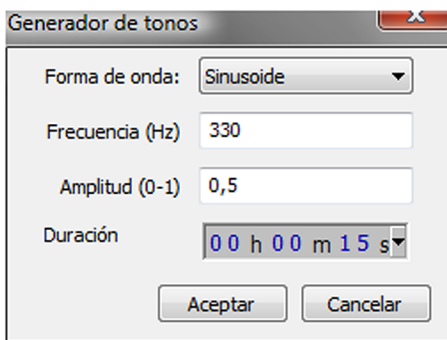
L'objectiu és introduir la representació espectral d'un senyal que com veurem al llarg de l'assignatura pot resultar molt útil per a interpretar la composició dels sons i generar diferents tipus de filtres per a condicionar els senyals.

Començarem un nou projecte d'Audacity.

1) Podem tornar a arrencar l'aplicació o executar l'opció *Nuevo* en el menú *Archivo* (Ctrl+N).

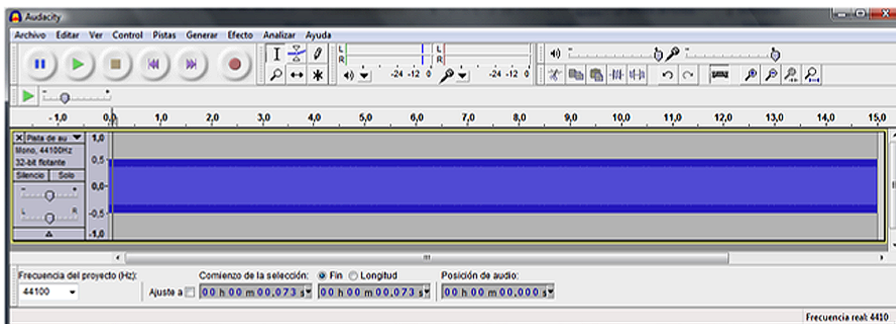
2) Ara anirem al menú *Generar* i seleccionarem l'opció *Tono*, amb la qual cosa apareixerà el menú de generació de tons representat en la figura 7. Seleccionarem una forma d'ona sinusoidal, amb una freqüència de 330 Hz, una amplitud de 0,5 i una durada de 15 segons (els mateixos paràmetres que es mostren en la figura 7).

Figura 7. Quadre de diàleg per a la generació d'un to sinusoidal



Escolteu el senyal. Es tracta d'un to pur que es correspon aproximadament amb la nota musical mi_3^5 . En principi, en generar el senyal es visualitzarà en la finestra la forma d'ona en tota la seva longitud, és a dir, 15 segons. Com el senyal té una freqüència de 330 Hz, això significa que en un segon tindrem un total de 330 períodes de la sinusoide. Així, doncs, estem veient un total de $330 \times 15 = 4.950$ períodes en la finestra que es representa en la figura 8. Evidentment, la resolució de la pantalla gràfica de l'ordinador no ens permet distingir la forma sinusoidal del senyal.

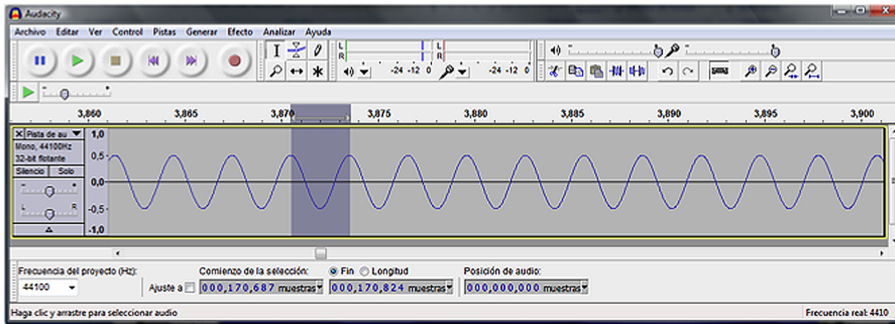
Figura 8. Representació gràfica de la forma d'ona del senyal tonal durant els 15 segons



Per avissualitzar la forma d'ona sinusoidal podem seleccionar un fragment curt del senyal i visualitzar-lo en la finestra d'Audacity (Ctrl+E).

3) Seleccioneu un fragment i amplieu-lo fins a obtenir una gràfica amb unes característiques semblants a la que es mostra en la figura 9, en què pot apreciar-se el caràcter sinusoidal del senyal. En aquesta figura també hem seleccionat un període complet del senyal. Tal com ens indica la barra inferior, el començament de la selecció va des de la mostra 170.687 fins a la mostra 170.824. Així, doncs, podem concloure que el nombre de mostres d'un període és de 137.

Figura 9. Representació d'un fragment del senyal sinusoidal i selecció d'un període bàsic



D'acord amb aquest resultat, podríem estimar la freqüència del senyal a partir del seu nombre de mostres. En efecte, com estem prenent un total de 44.100 mostres per segon. Fixeu-vos que 44.100 mostres per segon és un nombre de mostres molt elevat. Significa que el temps que transcorre entre dues mostres és tan sols d' $1/44.100$ segons, és a dir, 0,00002268 segons, o el que és el mateix, 0,02268 mil·lisegons o 22,68 microsegons. Com que el nombre de mostres que té el període de repetició bàsic del senyal és de 137, el temps equivalent serà de $137/44.100 = 3,106$ mil·lisegons. La freqüència és l'invers del període, per la qual cosa obtindrem una estimació de la freqüència de $44.100/137 = 321,89$ Hz.

El resultat obtingut no coincideix amb la freqüència teòrica, però això és a causa de l'error que cometem en seleccionar un únic període basant-nos en la representació gràfica de la forma d'ona.

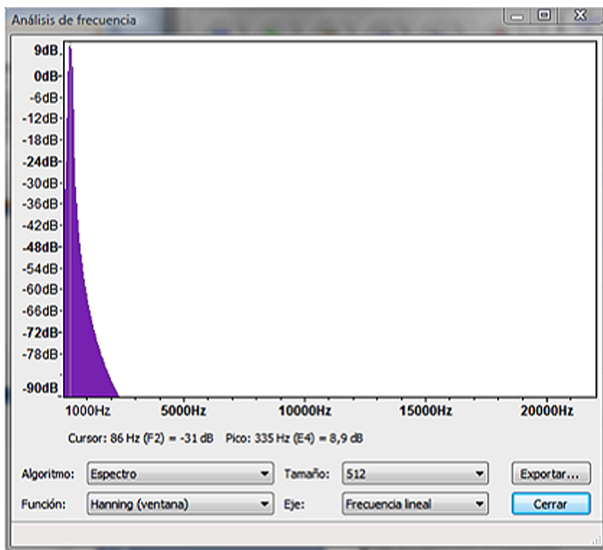
Activitat

Com a exercici es proposa que repetiu la selecció d'un període diverses vegades i observeu com el nombre de mostres estimat varia, ja que és difícil fer una estimació amb molta precisió.

Vegem una manera alternativa d'estimar la freqüència del senyal basant-nos en l'**anàlisi espectral**. Per analitzar el senyal, torneu a visualitzar tot el senyal complet (Ctrl+F).

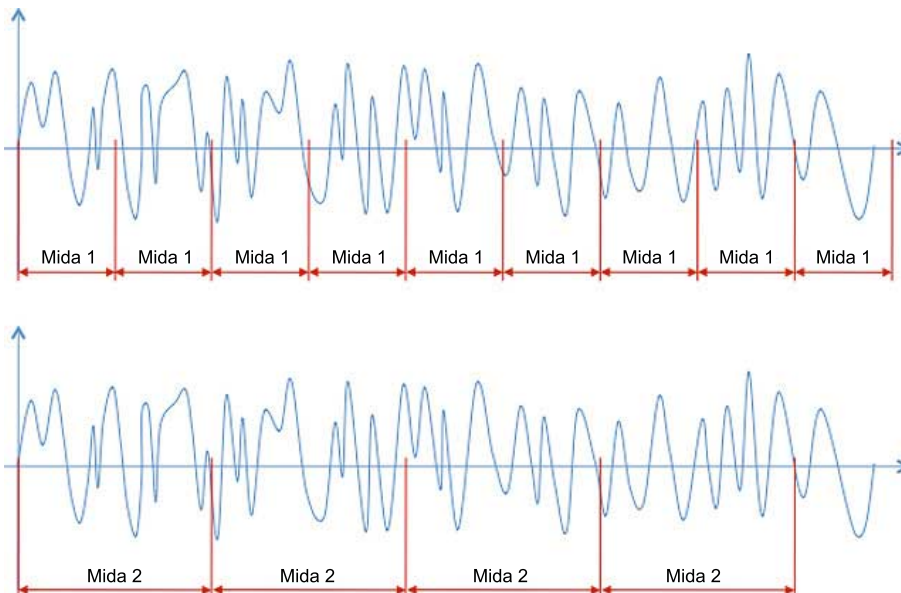
1) Seleccioneu un fragment de diversos segons del senyal, i en el menú **Anali-zar** seleccioneu l'opció **Anàlisis de espectro**. Obtindreu com a resultat la gràfica de la figura 10, en la qual es representa una estimació de les freqüències que componen el nostre senyal. Aquesta estimació es fa amb un mètode matemàtic denominat *transformada de Fourier*. Veiem que l'estimació que hem obtingut per defecte tampoc no és excessivament potent, ja que tal com s'indica en el peu de la figura, es detecta una cresta en la freqüència de 335 Hz (que tampoc no es corresponen amb la real de 330 Hz). En els paràgrafs següents veurem com podem millorar aquesta estimació de la freqüència modificant alguns paràmetres.

Figura 10. Estimació de freqüència obtinguda a partir de l'anàlisi d'espectre d'Audacity



2) En la part inferior de la finestra d'*Análisis de frecuencia* veiem que hi ha quatre caselles amb opcions que podem seleccionar. La primera permet seleccionar diversos algoritmes d'anàlisi. En el nostre cas, sempre utilitzarem l'**anàlisi d'espectre** per a estudiar la composició en freqüència dels senyals. La segona casella de selecció permet ajustar la mida de la finestra d'anàlisi. Aquest és un factor clau per a millorar la resolució de l'anàlisi d'espectre del senyal. Per comprendre el significat d'aquest paràmetre considerarem l'exemple que es presenta en la figura 11.

Figura 11. Anàlisi de Fourier sobre fragments del senyal d'àudio

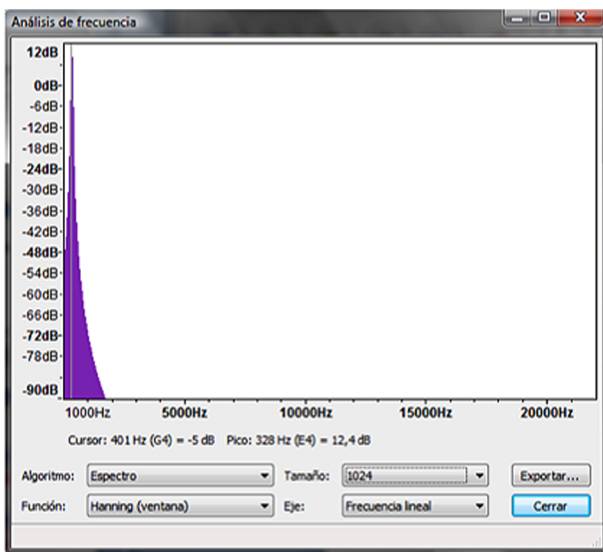


En aquest cas tenim representat un senyal amb una selecció d'un elevat nombre de mostres, des del principi fins al final de la gràfica. L'estimació de l'espectre es fa dividint el senyal en petits fragments (també denominats *finestres*) que tenen una mida prefixada (512 mostres en l'exemple de la figura 10). Per a cada un d'aquests fragments es fa una estimació de l'espectre (transfor-

mada de Fourier), i es fa la mitjana dels resultats obtinguts que donen lloc a una gràfica com la que es veu en la figura 10. Així, doncs, el que visualitzem és la mitjana de tots els espectres calculats per a les diferents finestres. En el nostre exemple, la transformada de Fourier només es fa sobre un nombre prefixat de 512 mostres.

Si el temps durant el qual analitzem el senyal és curt, l'estimació de les freqüències que componen el senyal no serà gaire precisa. En el nostre cas, 512 mostres corresponen a un temps d'anàlisi de $512/44.100 = 11,61$ mil·lèsimes de segon. Si augmentem la mida de la finestra al doble (1.024 mostres) estarem analitzant el senyal durant més temps, per la qual cosa, sense entrar en els detalls dels algorismes matemàtics de la transformada de Fourier, sembla natural que puguem estimar les freqüències que el componen amb més precisió. En efecte, en la figura 12 representem el resultat de l'anàlisi en freqüència que obtenim aplicant una finestra de mida de 1.024 mostres sobre el mateix fragment del senyal. La cresta és molt més aguda i la seva posició és més precisa (el màxim és en 328 Hz). Proveu amb una finestra de 2.048 mostres i veureu com la precisió encara millora més (331 Hz).

Figura 12. Anàlisi en freqüència amb finestres de 1.024 mostres



Precisió en la freqüència del senyal

Fixeu-vos que si posem una finestra prou gran podem determinar amb gran precisió la freqüència del senyal.

El paràgraf anterior sembla suggerir que com més llarga és la finestra millor serà l'estimació de l'espectre. Això és cert si les característiques del senyal no varien, com és el nostre cas. En efecte, tenim un **to pur** que es manté en les mateixes condicions durant 15 segons complets, per tant, en principi, podríem utilitzar tota aquesta informació per a estimar l'espectre del senyal. En canvi, amb senyals musicals o de veu, les característiques varien de manera molt ràpida en el temps, els canvis són dinàmics i de vegades pot interessar que aquestes finestres siguin més petites per a estudiar com evolucionen les característiques de l'espectre al llarg del temps.

Resumint, depenent de les característiques dels senyals és possible que estiguem més interessats a utilitzar fragments de **gran longitud** (senyal estacionari) o fragments de **curta longitud** (senyals variants). En cada cas haurem de seleccionar la longitud apropiada.

3) A més de la mida de la finestra també podem seleccionar la funció, que per defecte té seleccionat el valor Hanning. En principi no modificarem aquest valor per a les diferents estimacions de freqüència que farem en aquesta assignatura. La funció representa la importància que es dóna a les mostres dins de la finestra. Així, amb una finestra rectangular (una de les opcions que podem escollir), totes les mostres de la finestra tenen la mateixa importància. En una finestra de Hanning les mostres centrals tenen més pes en els algoritmes que fan l'estimació dels espectres.

4) Finalment, també podem seleccionar el tipus d'escala que s'aplica a l'eix horitzontal. Per defecte apareix l'escala lineal encara que en àudio, moltes vegades, és habitual utilitzar una escala logarítmica. La raó per la qual s'utilitza una escala logarítmica és que l'oïda es comporta de manera logarítmica, és a dir, tenim molta més resolució per a discriminar senyals de baixa freqüència que per distingir senyals d'alta freqüència. Proveu de modificar l'escalat en l'eix horitzontal a logarítmic. En aquest cas la posició de les freqüències en l'eix horitzontal es correspon amb la percepció real que tenim de les freqüències dels senyals.

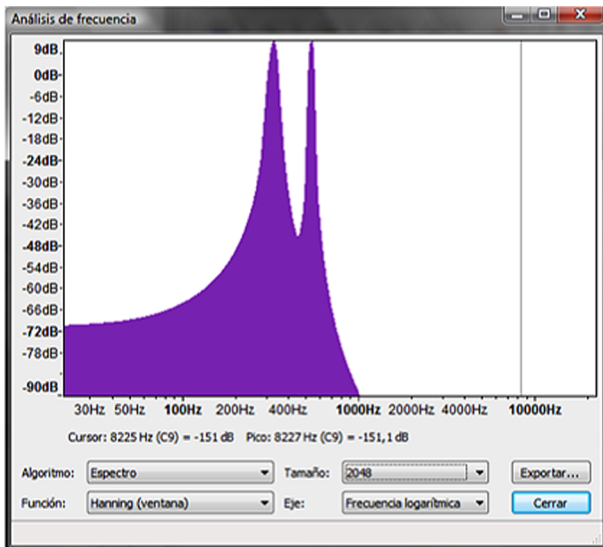
1.3.6. Superposició de tons

Ara generarem un nou to a una freqüència de 540 Hz.

1) Per a això accediu al menú *Pistas* i seleccioneu *Añadir nueva\Pista de audio*. Situeu el cursor a l'inici d'aquesta nova pista i genereu un to sinusoidal de 15 segons, amb una amplitud de 0.5 i una freqüència de 540 Hz. Escolteu el senyal i comproveu com l'oïda pot distingir perfectament entre les dues freqüències.

2) Seleccioneu ara amb el cursor un fragment de diversos segons dels dos senyals i feu l'anàlisi d'espectre (per a seleccionar les dues pistes heu de crear un rectangle amb el ratolí que inclogui les dues pistes –utilitzeu Ctrl+A si voleu seleccionar tots els senyals de totes les pistes). Utilitzeu diversos valors de la **mida** de la finestra d'anàlisi i passeu de freqüència lineal a logarítmica per veure els efectes sobre el senyal. Comproveu que les estimacions de freqüència són correctes, sempre que la mida de la finestra sigui prou gran.

Figura 13. Exemple de superposició de dos tons



Mostra d'un dels resultats dels espectres que hem obtingut amb aquest exemple.

Observació

Observeu com ara l'espectre mostra clarament la superposició dels dos tons.

1.3.7. Seqüenciació de tons i interpretació de l'espectre

Aquest exemple és semblant a l'anterior però en comptes de posar els dos tons simultàniament, en pistes paral·leles, els posarem l'un darrere de l'altre usant la mateixa pista.

1) Partim de l'exercici anterior i esborrem la segona pista (n'hi ha prou de marcar la creu en la casella que té els controls de volum i balanç de la pista).

2) Després posem el cursor al final de la primera pista, la que té un to amb 330 Hz i 15 segons de durada. Generem ara un to de 540 Hz, amb 15 segons de durada. Com a resultat tindrem una única pista amb una durada de 30 segons. En els 15 primers segons tenim un to de 330 Hz, i en els següents un de 540 Hz. Si escoltem el senyal ens adonarem del canvi de freqüència.

3) Seleccionem ara un fragment d'àudio que inclogui els dos tons i aproximadament durant el mateix temps. Per exemple, seleccioneu l'àudio entre el segon 12 i el 18. Feu l'anàlisi d'espectre.

Vegem què passa quan el fragment que seleccionem inclou el doble de temps d'un senyal que de l'altre.

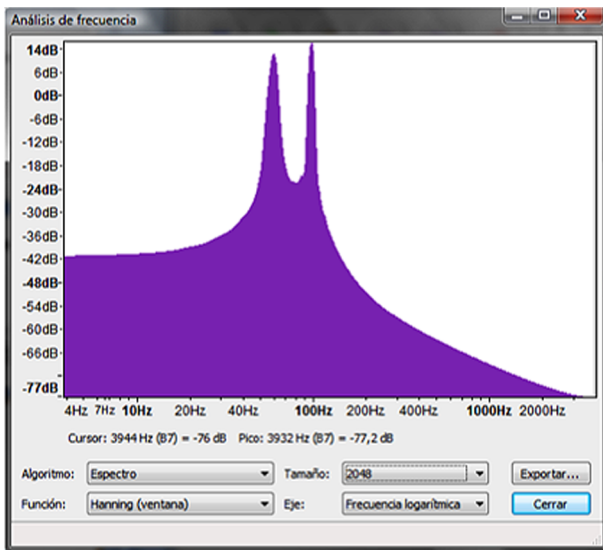
4) Seleccionem, per exemple, entre el segon 12 i el 21. Ara tenim una selecció en què el primer to és present durant 3 segons mentre el segon té una durada de 6 segons.

5) Feu l'anàlisi de freqüència. Ara hauríeu de tenir un resultat com el de la figura 14, en què l'amplitud de la cresta situada en la freqüència de 540 Hz és més gran que la de la cresta situada en 330 Hz. Intenteu raonar per què una de les crestes té més potència que l'altra.

Reflexió

El resultat que hauríeu d'obtenir és el que ja heu obtingut en l'exercici 3. Podeu explicar per què?

Figura 14. Espectre obtingut quan fem l'anàlisi prenent un període de temps més llarg del senyal de 540 Hz



6) Finalment, seleccioneu un fragment de senyal en què només hi hagi un dels tons. Veurem que en l'estimació d'espectre ara només apareix el to que hem seleccionat.

L'explicació d'aquests resultats correspon a la idea que hem explicat en la figura 11 en comentar el procediment de càlcul de l'estimació d'espectre. En efecte, el senyal es divideix en fragments i es calcula l'espectre de cada un dels fragments. Posteriorment es fa la mitjana de tots aquests espectres i es presenten en la pantalla. Això significa:

a) Quan hem pres el mateix temps dels dos tons en la mitjana han participat tants espectres corresponents a un dels tons com a l'altre, per tant, han aparegut les dues crestes amb la mateixa importància.

b) Quan hem un dels tons ha intervingut el doble de temps que l'altre s'ha fet la mitjana de més espectres corresponents al segon to que al primer, per la qual cosa la cresta ha tingut més amplitud.

c) Quan seleccionem un fragment en el qual només tenim un to, només participa en la mitjana aquest to.

Per tant, és important tenir en compte que la gràfica que ens presenta l'anàlisi d'espectre és una gràfica mitjana i que hem d'interpretar-la correctament en cada moment.

Essencialment, l'anàlisi d'espectre ens proporciona una estimació de les característiques freqüencials del senyal durant el temps de durada de la finestra d'anàlisi. Això ens permet veure en una sola gràfica els diferents elements de freqüència que componen el senyal, d'un sol cop d'ull podem conèixer la composició en freqüència d'un senyal complex. No obstant això, si el senyal canvia les seves característiques al llarg del temps, no sabem en quins instants es

Observació

Fixeu-vos, per exemple, que si no examinéssim amb més detall el senyal, podríem concloure que el resultat que hem obtingut en el primer apartat és perquè tenim dos tons que són actius alhora.

produeixen aquests canvis. En general, per a veure amb precisió els instants en què es produeixen els canvis necessitem recórrer a la **representació temporal**. L'espectrograma, que veurem en l'exercici següent, constitueix un compromís excel·lent entre la resolució en temps (que ens proporciona la representació temporal) i la resolució en freqüència (que ens proporciona l'estimació d'espectre). L'espectrograma és una representació en 3D de com els espectres evolucionen al llarg del temps. Veurem els detalls d'aquesta representació en l'exercici següent.

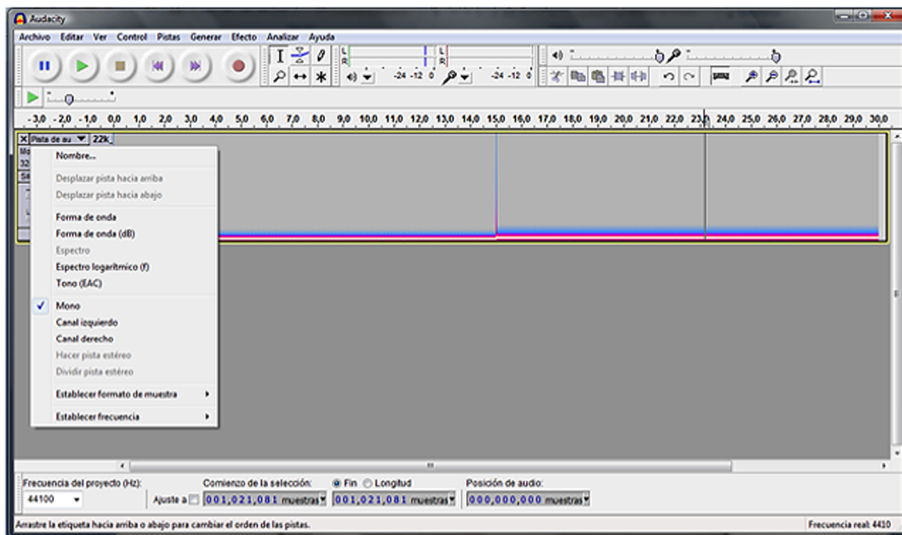
1.3.8. Espectrograma

Una alternativa a l'anàlisi d'espectre que hem fet en l'apartat anterior és l'espectrograma. Essencialment, l'espectrograma consisteix a representar tots els espectres que es van obtenint en analitzar cada un dels fragments d'un senyal àudio. Seguirem aquest primer exemple de l'espectrograma amb el mateix senyal que hem generat en l'apartat anterior.

1) Per a visualitzar l'espectrograma d'un senyal d'àudio amb el programa Audacity, hem d'anar al desplegable de **Pista de audio** (vegeu la figura 15) i seleccionar l'opció *Espectro*.

- La representació que obtenim ha de ser interpretada correctament i amb molta cura. En l'eix horitzontal continuem tenint el temps mentre que en l'eix vertical tenim la freqüència. La freqüència en el nostre exemple va de 0 kHz a 22 kHz. Els espectres es representen en colors, assignant els colors més clars quan en les freqüències en les quals es produeixen amplituds o crestes significatives.
- L'espectrograma també es pot representar en un eix logarítmic de freqüències. Per a això, s'ha de seleccionar l'opció d'*Espectro logarítmico* en el desplegable de la pista d'àudio.

Figura . Representació de l'espectrograma del senyal



Observació

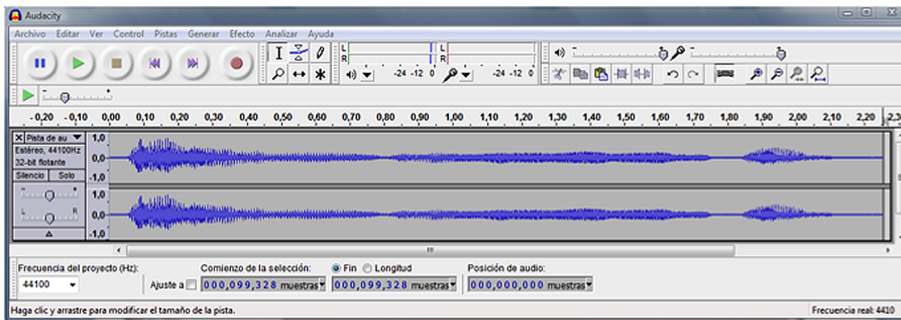
Fixeu-vos, en el nostre exemple, com després dels 15 segons, la ratlla blanca (la cresta de freqüència) es desplaça lleugerament cap amunt a causa del canvi de la freqüència.

1.3.9. Anàlisi d'espectre del senyal de veu

En aquest exercici examinarem l'espectre del senyal de veu. Per a això proposem que registreu la pròpia veu dient la paraula *saludo* (en castellà) lentament. En la figura 16 es mostra una possible forma d'ona obtinguda amb aquesta paraula.

Quan digitalitzeu la vostra pròpia versió veureu que els resultats no coincideixen, ni en forma d'ona ni en estimació de freqüència, amb els que presentem en l'exemple. Això passa perquè cada persona té les seves pròpies característiques de generació de veu. De fet, podem reconèixer les persones per la veu a causa d'aquestes característiques específiques de cadascú.

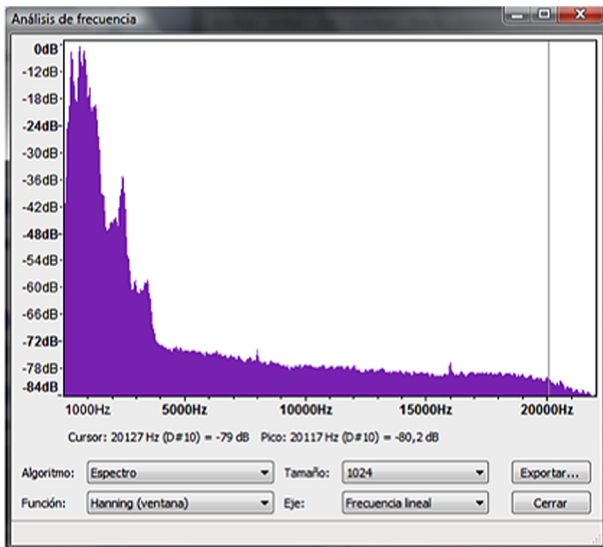
Figura 16. Forma d'ona del senyal de veu amb la frase *saludo*



Es demana que feu les proves següents amb el senyal que heu obtingut:

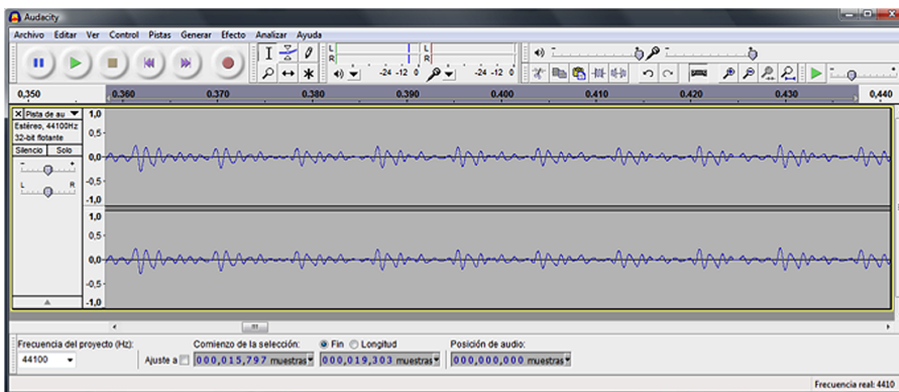
a) Seleccioneu el senyal complet (Ctrl + A) i analitzeu-ne l'espectre. Tingueu en compte que el que visualitzeu és la mitjana de tots els espectres dels fragments analitzats per a aquest senyal. L'espectre que hem obtingut en el nostre exemple es mostra en la figura 17.

Figura 17. Mitjana de l'espectre en tot el senyal de veu



b) Seleccioneu la part de la /a/ i feu un *zoom* sobre la forma d'ona. Hauríeu de veure que la forma d'ona de les vocals és un senyal més o menys complex però que es repeteix de manera periòdica. En la figura 18 veiem la forma d'ona obtinguda per al nostre exemple.

Figura 18. Forma d'ona de la vocal /a/

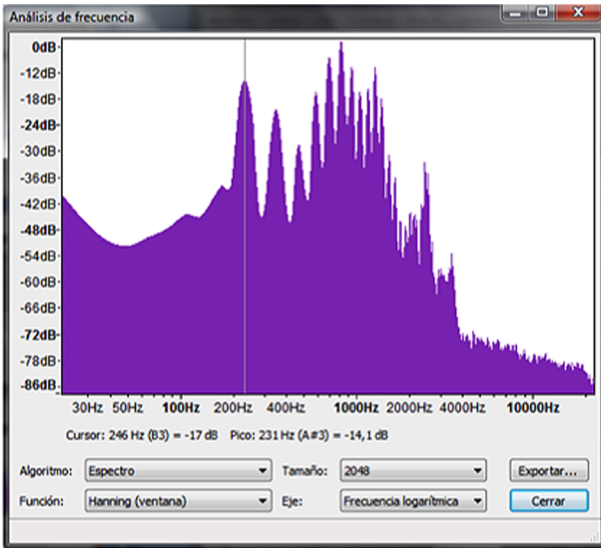


c) Seleccioneu un fragment prou llarg de la vocal /a/ i analitzeu-ne l'espectre. El resultat que hauríeu d'obtenir és que es tracta d'un senyal harmònicament complex, com ho mostra el nostre exemple de la figura 19. En el nostre cas detectem un primera cresta significativa en la freqüència de 231 Hz i posteriorment apareixen una sèrie d'harmònics que són els que hi donen el matis tímbric d'aquesta veu.

Pitch

Fixeu-vos, tal com es marca en el requadre en vermell, que hi ha una forma bàsica que es repeteix de manera periòdica. El període de repetició es denomina *període de pitch*.

Figura 19. Espectre de la vocal /a/



d) Repetiu la mateixa prova que en els apartats b) i c) amb la vocal /o/. Es proporcionen unes gràfiques d'exemple en les figures 20 i 21.

Figura 20. Forma d'ona de la vocal /o/

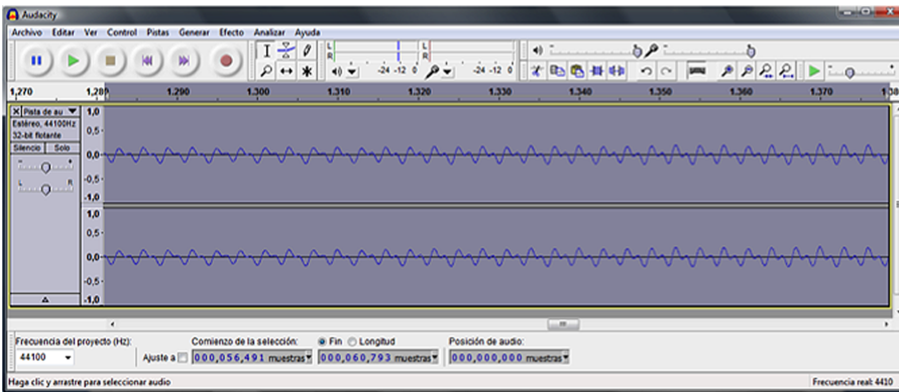
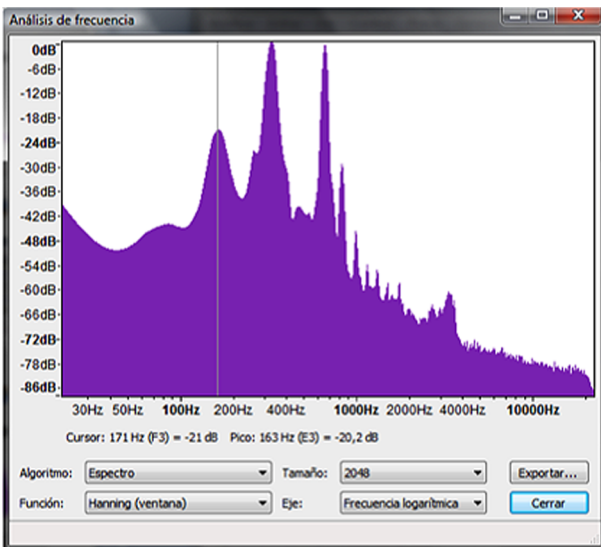


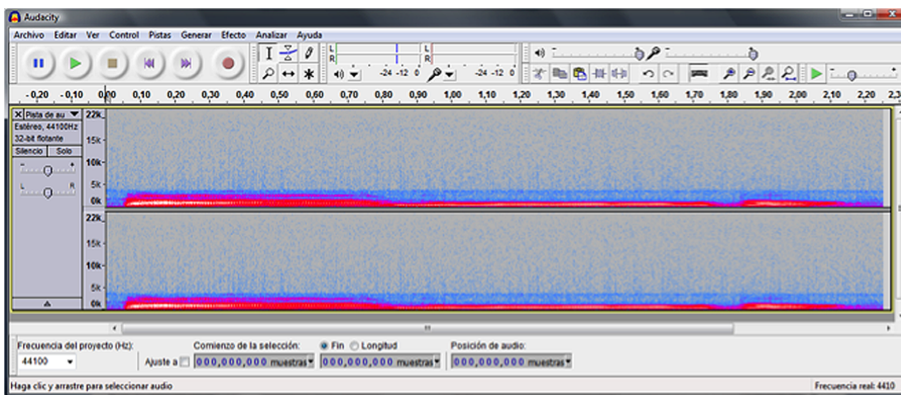
Figura 21. Espectre de la vocal /o/



e) Intenteu veure la forma d'ona de les consonants /s/ i /d/ (és més complex perquè la durada de les consonants és més curta que la de les vocals). Comproveu que no es detecten periodicitats.

f) Finalment, representeu l'espectrograma del senyal d'àudio complet i comproveu com evolucionen els espectres al llarg del temps, modificant-se la freqüència fonamental de la veu i la importància relativa dels harmònics. L'espectrograma se sol considerar com una mesura biomètrica característica de la persona, com si es tractés d'una empremta dactilar. Hi ha mètodes per a identificar si una persona és en un determinat enregistrament a partir de l'anàlisi de l'espectrograma. En la figura 22 es mostra el resultat de l'espectrograma obtingut en el nostre exemple.

Figura 22. Espectrograma de la paraula *saludo*



1.3.10. Espectrograma d'un xiulet

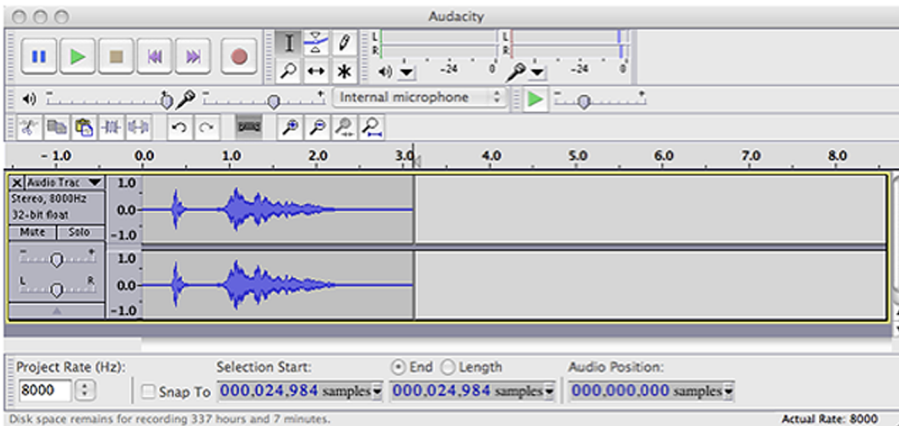
En aquest exercici aprofundirem en l'estudi de l'histograma mitjançant l'anàlisi d'un de simple però que pot resultar molt il·lustratiu, com és un xiulet.

Es proposa que amb el programa Audacity es capturi un xiulet (un fragment relativament curt) en què es produeixi un canvi de freqüència significatiu.

La idea principal d'aquest exercici és comprovar com la representació de l'espectrograma ens ajuda a situar els components de freqüència en el temps, mentre que les representacions en el domini temporal i en el domini freqüencial no són tan clares.

1) En efecte, en la figura 23 es mostra el resultat d'adquirir un xiulet i la seva representació en forma d'ona. Perquè la representació en freqüència i en espectrograma es pugui visualitzar correctament, és convenient modificar la freqüència de mostreig a 8.000 Hz abans d'adquirir el senyal (la freqüència de mostreig es pot modificar en la casella de selecció situada en la part inferior esquerra de l'aplicació).

Figura 23. Representació d'un xiulet en el domini temporal

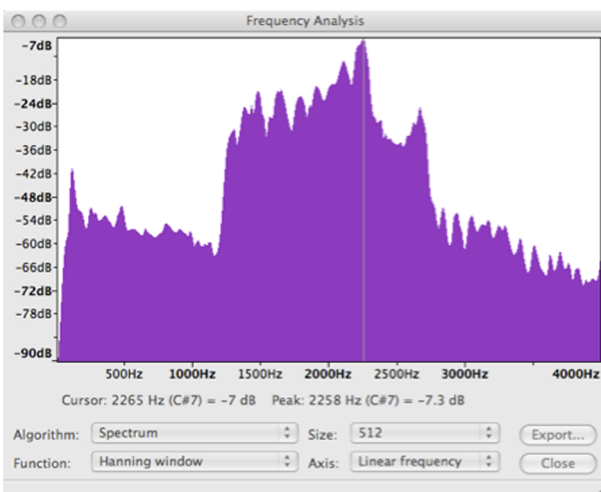


La freqüència de mostreig està modificada a 8.000 Hz

En la representació en temps del xiulet ens adonem que és possible veure que el senyal es compon de dos fragments bàsics, que podem situar-los perfectament en el temps, encara que no en podem determinar les freqüències. Així, doncs, la representació en temps ens permet tenir molta precisió per a saber quan el senyal canvia les característiques.

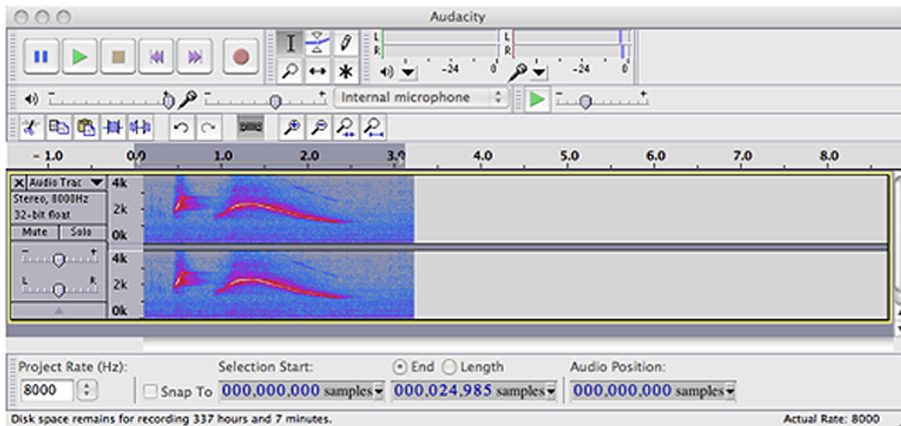
2) En la figura 24 representem l'estimació espectral que hem obtingut seleccionant tot el senyal i analitzant l'espectre. Ara veiem que el senyal és espectralment complex, format per moltes freqüències, que podríem determinar amb precisió. No obstant això, en aquesta representació no sabem els instants en què es produeixen aquests canvis, és a dir, on es fan els canvis de freqüència, com estan disposades les freqüències en el temps.

Figura 24. Representació de l'estimació d'espectre del xiulet de la figura anterior



3) Finalment, en la figura 25, que correspon a l'espectrograma, veiem com obtenim un compromís excel·lent entre **resolució temporal** i **freqüencial**, ja que és possible identificar l'evolució de l'espectre al llarg del temps.

Figura 25. Representació del xiulet mitjançant l'espectrograma



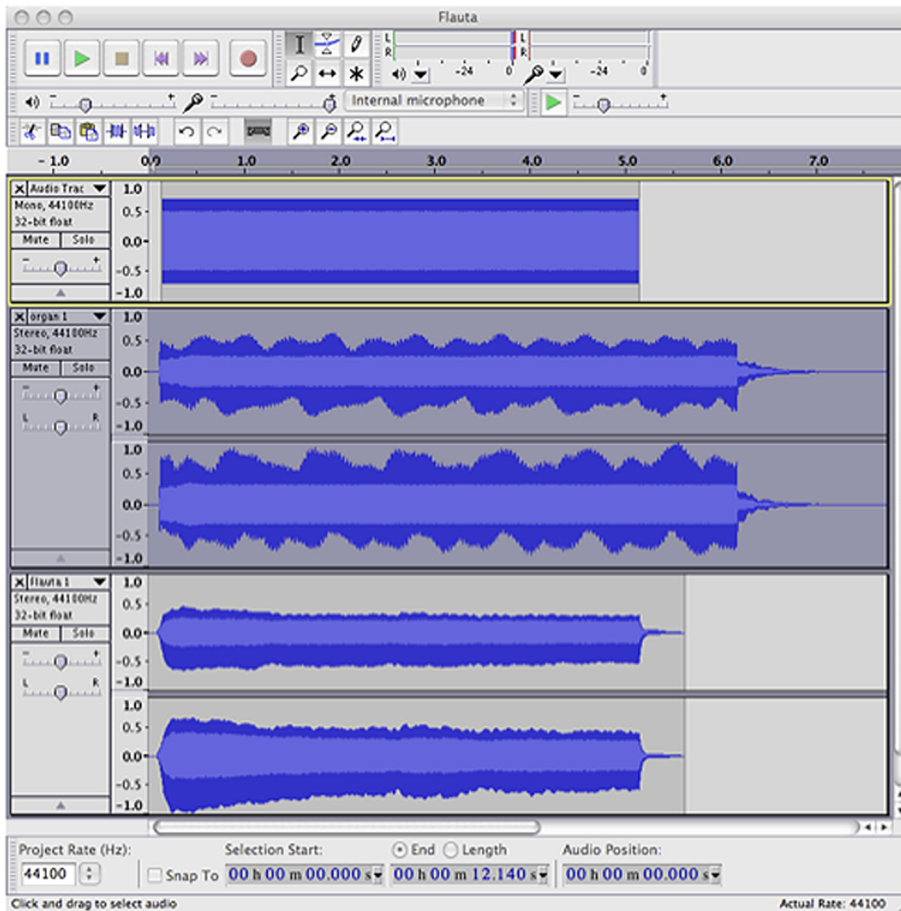
En aquesta gràfica podem apreciar com els diferents components freqüencials evolucionen en el temps

1.3.11. Timbre dels instruments musicals

Per finalitzar amb aquest apartat, analitzarem espectralment instruments musicals que produeixen la mateixa nota.

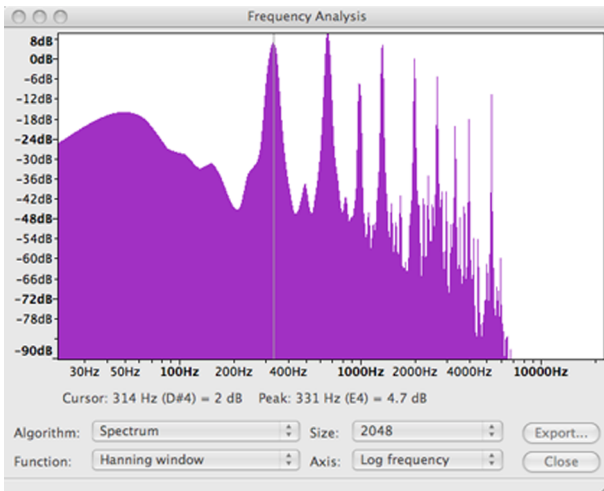
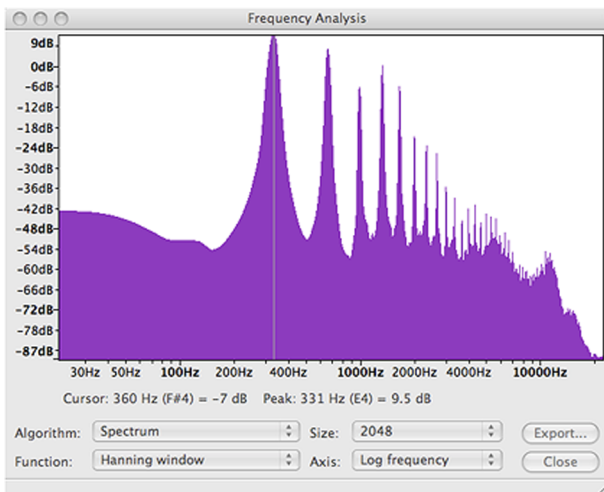
- 1) Inicialitzeu de nou Audacity i genereu un senyal sinusoidal de 330 Hz amb una durada d'uns 5 segons.
- 2) Posteriorment, importeu en una nova pista el senyal organo.wav que es proporciona amb el material d'aquesta pràctica.
- 3) Finalment importeu també el fitxer flauta.wav. La pantalla d'Audacity hauria de disposar dels tres senyals tal com es mostren en la figura 26.

Figura 26. Pantalla d'Audacity amb les tres pistes corresponents al to, l'orgue i la flauta



Es demana que feu les proves següents:

- Reproduïu els tres senyals individualment. Fixeu-vos que tots tres estan produint la mateixa nota musical, encara que l'oïda pot identificar-los clarament com a instruments diferents.
- Seleccioneu el to pur de 330 Hz i comproveu mitjançant l'anàlisi espectral que està compost per una única freqüència.
- Seleccioneu el senyal proporcionat per l'orgue i determineu-ne l'espectre. Observeu la riquesa d'harmònics a freqüències que són múltiples de la freqüència fonamental.
- Repetiu l'apartat anterior per a la flauta. En aquest cas també veiem que hi ha una important riquesa harmònica encara que les amplituds relatives dels harmònics són diferents que per al cas de l'orgue. Són aquestes diferències en les importàncies relatives de cada harmònic les que caracteritzen cada un dels instruments musicals, proporcionant-los un timbre característic a cada un. En les figures 27 i 28 es mostren els espectres obtinguts per a l'orgue i la flauta respectivament.

Figura 27. Anàlisi d'espectre de l'orgue per a una nota mi_3 Figura 28. Anàlisi d'espectre de la flauta per a una nota mi_3 

1.4. Problemes/preguntes

1) Es proposa que captureu el senyal d'àudio d'una pista de CD, tal com ja heu fet en el primer exercici d'aquesta activitat. Podeu utilitzar el mateix tema musical o un altre.

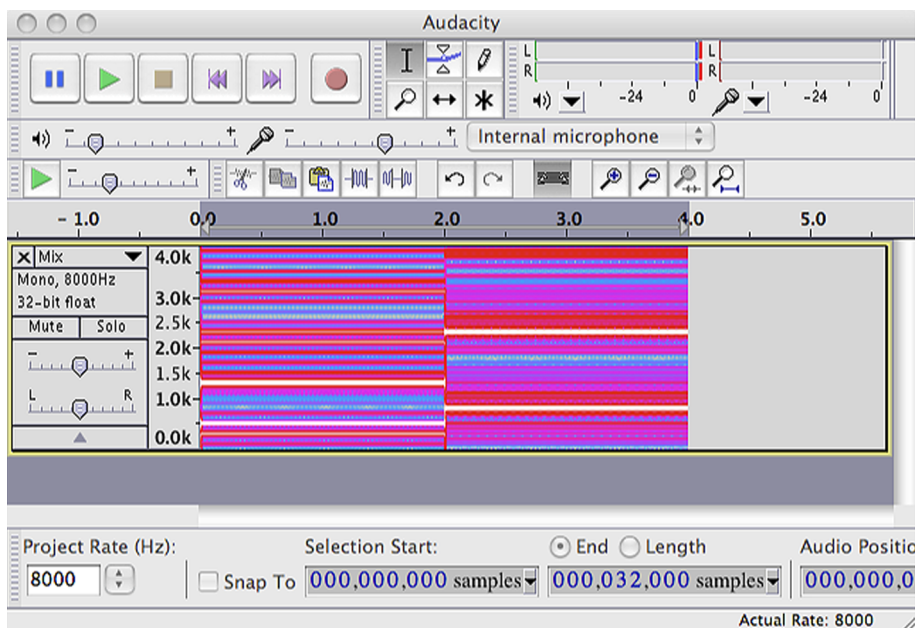
- Tragueu els auriculars de l'ordinador, amplifiqueu el senyal pels altaveus fins a un nivell de conversa normal i activeu el botó de gravació Rec d'Audacity. Amb això, es gravarà una nova pista, capturada pel micròfon, sobre el senyal que estan reproduint els altaveus.
- Una vegada finalitzat l'enregistrament, escolteu el resultat obtingut amb la superposició de les dues pistes. Expliqueu què passa quan superposeu les dues pistes. Podeu estimar quin és el retard entre les dues pistes?

2) Elaboreu una llista, el més completa possible, d'aplicacions de captura i edició d'àudio que pugueu trobar per Internet. Seleccioneu tres d'aquestes aplicacions i intenteu enumerar-ne amb el màxim detall possible les característiques incloent-hi aspectes com:

- Nombre de pistes que es poden processar.
- Característiques de la interfície gràfica (captureu les pantalles).
- Capacitats de tallar, enganxar i copiar.
- Control de volum i balanç individual per pista.
- Representació de la forma d'ona temporal dels senyals de cada pista.
- Representació de l'espectre de les pistes, control de to o equalització de les pistes individuals.
- Capacitat de controlar targetes de captura d'àudio específiques.
- Disponibilitat d'efectes d'àudio (enumerem els efectes més significatius, etc.).

3) Tenim un senyal d'una durada de 30 segons sobre la qual hem fet un espectrograma que es mostra en la figura adjunta. Quines característiques temps-freqüència ens podeu indicar sobre aquest senyal?

Figura 29



Proporcioneu els detalls sobre:

- Quants fragments estacionaris veieu.
- Quina durada té cada fragment estacionari.
- Quins són els dos components freqüencials més importants de cada fragment estacionari del senyal.

4) En l'arxiu **MarcadoTelefónico.wav** que es proporciona juntament amb el material d'aquesta activitat, hem registrat els tons que produeix un telèfon mòbil en marcar un nombre. La qualitat de l'enregistrament no és excel·lent,

de fet, ha estat realment capturat pel micròfon incorporat en un ordinador portàtil mentre es marcaven els nombres al telèfon mòbil. Es demana que determineu els nombres que han estat marcats sabent que la taula de freqüències que utilitzen els telèfons per a marcar els nombres és la de la figura adjunta.

Figura 30

	1209hz	1336hz	1477hz	1633hz
697hz	1	2	3	Un
770hz	4	5	6	B
852hz	7	8	9	C
941hz	*	0	#	D

2. Activitat 2. Digitalització de senyals d'àudio: mostreig

2.1. Introducció

2.1.1. Contingut

La captura i digitalització dels senyals d'àudio té un paper fonamental en tots els sistemes d'edició, efectes i producció de so, ja que els paràmetres amb què s'ha fet l'adquisició determinen directament la **qualitat final del so**.

En aquesta activitat introduïrem el procés de digitalització del senyal d'àudio com un problema d'aproximació dels senyals analògics mitjançant una **retícula**.

Aquesta retícula ens permet identificar dos elements clau en la digitalització:

- la freqüència de mostreig (pas de la retícula en l'eix horitzontal) i
- el pas de quantificació (pas en l'eix vertical).

Els criteris per a seleccionar la freqüència de mostreig i la seva incidència en la qualitat final del senyal es tractaran en aquesta activitat. En l'activitat següent s'analitzaran els efectes del pas de quantificació i s'avaluaran els efectes de tots dos en la digitalització i l'amplada de banda dels senyals.

Per a estudiar la freqüència de mostreig es proposen diferents proves auditives que intentaran familiaritzar-nos amb la relació entre la freqüència de mostreig i la qualitat del senyal. Aquestes proves ens indicaran que hi ha un límit en la freqüència de mostreig més enllà del qual l'oïda humana no és capaç d'apreciar diferències entre els senyals, per la qual cosa no tindrà sentit augmentar excessivament el nombre de mostres per segon. Una vegada introduïts aquests conceptes de manera experimental, intuïtiva i poc rigorosa s'enuncia el **teorema del mostreig** o **teorema de Nyquist**, que ens proporciona les condicions necessàries per a conèixer la freqüència a què s'ha de mostrejar un senyal.

La part final de l'activitat se centra en el **problema de l'aliàsing** definint-lo de manera intuïtiva i proporcionant diferents experiències auditives per a identificar-lo.

2.1.2. Metodologia

L'activitat se centra a definir el concepte de mostreig d'un senyal i establir els criteris per a seleccionar el nombre de mostres per segon que s'han de prendre per a la seva correcta representació.

En la metodologia utilitzada l'estudiant ha de fer diferents proves auditives amb fitxers d'àudio digitalitzats amb diferents freqüències de mostreig i comparar els resultats obtinguts.

Amb això, se cerca que ens familiaritzem amb els efectes que la selecció de la freqüència de mostreig té sobre la qualitat final del senyal d'àudio.

Per a donar suport a l'audició també se suggereix que s'analitzin els senyals des del punt de vista de l'espectre, amb la qual cosa és possible associar l'espectre d'un senyal amb les seves característiques audibles i veure com incideix la freqüència de mostreig en l'espectre del senyal. S'apliquen conceptes de freqüència i espectre que s'han estudiat en activitats prèvies.

Una vegada introduïts els conceptes bàsics des d'un punt de vista purament experimental, es formalitzen els conceptes de teorema de mostreig i aliàsing, i se suggereix que l'estudiant emprengui altres experiències de digitalització i mesura d'espectres. El teorema del mostreig i l'aliàsing es presenten sense demostració, encara que es proporcionen molts exemples auditius perquè quedin consolidats.

2.1.3. Recursos

Els recursos principals utilitzats en aquesta activitat són:

- Arxius d'àudio predigitalitzats amb diferents freqüències de mostreig per a la seva audició.
- Programari de captura i processament d'àudio Audacity (licència lliure).
- Captura i anàlisi de veu a diferents freqüències de mostreig.
- Explicacions dels fonaments de la digitalització i la freqüència de mostreig incloses en el text de la guia de l'activitat.
- Proposta d'activitats de cerca d'informació en webs i recursos d'Internet per a aprofundir en els coneixements adquirits en aquesta activitat (webs de fabricants de targetes de captura, *wiki*, etc.).

- Problemes proposats per a verificar l'adquisició dels coneixements d'aquesta activitat.

2.2. Objectius

Els objectius principals que es persegueixen en aquesta activitat són:

- 1) Introduir el concepte de digitalització d'un senyal.
- 2) Interpretar la funció dels dos elements clau en el procés de digitalització d'un senyal: la freqüència de mostreig i el pas de quantificació.
- 3) Relacionar, mitjançant experiments dirigits, la qualitat subjectiva d'un senyal d'àudio amb la seva freqüència de mostreig.
- 4) Conèixer el teorema del mostreig i els criteris de selecció de la freqüència de mostreig.
- 5) Identificar el fenomen de l'aliàsing en les seves diferents manifestacions, tant en àudio com en un altre tipus de senyals.
- 6) Conèixer les freqüències de mostreig que s'utilitzen en els sistemes principals d'àudio i veu digitals.

Aquests objectius estan relacionats amb les **competències de l'assignatura** següents:

- 1) Capacitat de capturar àudio i imatges de manera eficient i eficaç.
- 2) Capacitat de digitalitzar eficientment i eficaçment continguts d'àudio i imatge.
- 3) Capacitat d'interpretar i modificar informació auditiva de manera freqüencial.

I amb les **competències generals del grau** següents:

- 1) Capturar, emmagatzemar i modificar informació d'àudio, imatge i vídeo digitals aplicant principis i mètodes de realització i composició del llenguatge audiovisual.
- 2) Atendre adequadament consultes sobre projectes, tecnologies i mercat de productes multimèdia avaluant de manera precisa l'entorn d'aplicació, els recursos i les alternatives tecnològiques disponibles.

2.3. Guia

En aquesta activitat ens centrarem en el procés de digitalització del senyal d'àudio i molt especialment en els criteris necessaris per a escollir una freqüència de mostreig adequada.

L'objectiu final que perseguim és determinar el **nombre de mostres** que s'han de prendre en cada segon d'un senyal en funció de les seves característiques. A més, volem familiaritzar-nos amb aquest paràmetre per a mostrar com la seva elecció resulta fonamental en la qualitat final del senyal d'àudio.

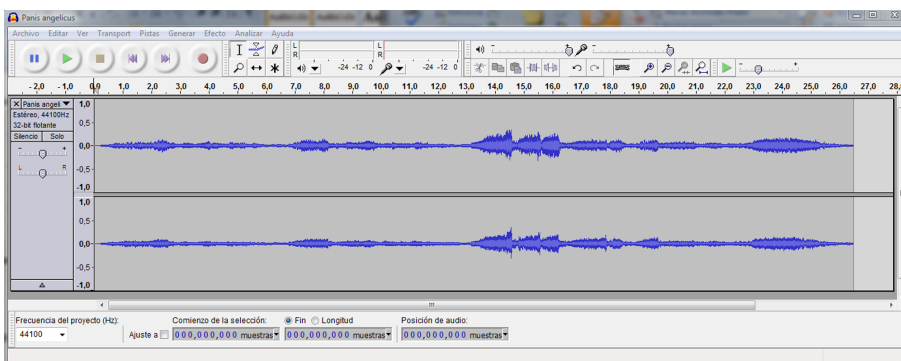
Com en l'activitat anterior, utilitzarem principalment el programa Audacity per a visualitzar, analitzar i escoltar els diferents senyals que ens serviran de guia per a comprendre els conceptes que volem il·lustrar.

2.3.1. La retícula de mostra

1) Per començar obrirem amb el programa Audacity el fitxer **Panis Angelicus.wav**. Aquest fitxer es proporciona com a recurs addicional per a aquesta activitat i consisteix en un breu fragment d'una interpretació del *Panis Angelicus* per Kiri Te Kanawa. En obrir el fitxer observem la forma d'ona dels dos canals estereofònics. Com que l'escala de l'eix temporal inclou tot el fragment musical no es poden veure gaires detalls de la forma d'ona. No obstant això, amb aquesta resolució visual tenim una idea general de com evoluciona en el temps el nivell del senyal.

2) Escolteu l'àudio per familiaritzar-vos amb el seu contingut i la localització en el temps de les frases musicals. En la figura 31 es mostra aquesta pantalla inicial.

Figura 31. Mostra de la pantalla inicial en obrir el fitxer Panis Angelicus.wav amb el programa Audacity



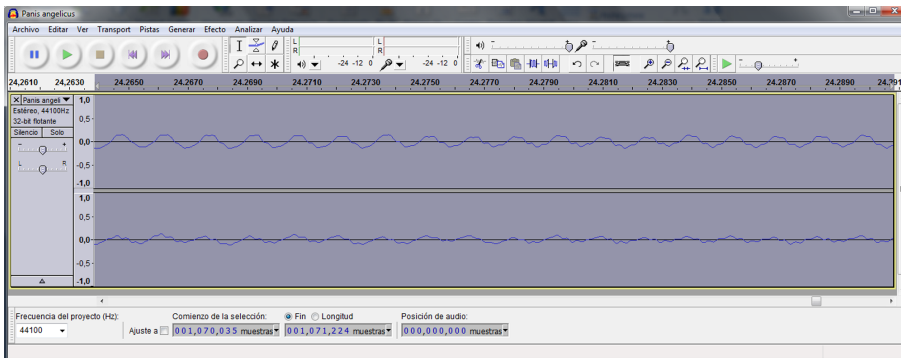
Farem un zoom del senyal per veure els detalls de la forma d'ona.

Per a fer el zoom s'ha de seleccionar un fragment d'àudio i posteriorment executar els menús *Ver\Ampliar la selecció* (Ctrl+E).

3) Observeu en pantalla el resultat i comproveu que, tot i que ara el detall amb què veiem la forma d'ona del senyal ha millorat considerablement, encara no podem apreciar els encreuaments per zero i altres peculiaritats del senyal.

4) Amplieu el factor de zoom fins que pugui visualitzar un fragment de senyal compost per uns 18-20 cicles d'una forma d'ona quasiperiòdica. En la figura 32 es mostra aquesta ampliació del senyal en què es veu que la forma d'ona és una repetició d'una forma d'ona bàsica. En aquest exemple concret es mostren un total de 18 repeticions (cicles) del senyal bàsic. Si proveu d'ampliar altres fragments del senyal original veureu que tenen característiques semblants, canviant la forma d'ona del cicle bàsic i el període de repetició.

Figura 32. Ampliació d'un fragment del senyal



En aquest exemple hem reduït l'escala de visualització fins a poder observar el caràcter quasiperiòdic de la forma d'ona. En el senyal superior podem comptar un total de 18 períodes fonamentals.

Fixeu-vos que amb la resolució actual el senyal té l'aspecte d'un senyal analògic, és a dir, un senyal l'amplitud del qual és contínua i que pren valors per a tot instant de temps. En efecte, el senyal que tenim en pantalla pot representar perfectament l'evolució en el temps de la tensió obtinguda a la sortida d'un micròfon. Al seu torn, aquesta tensió representa les variacions del nivell de pressió acústica que hi ha a la membrana del micròfon.

La característica principal dels senyals analògics és que es poden representar mitjançant funcions que evolucionen de manera contínua en el temps.

5) Continuant amb el nostre exemple, ampliem una mica més el valor del zoom del senyal fins a visualitzar-ne només uns 2 o 3 cicles. Ara veurem que en la forma d'ona apareixen marcades unes ratlles que ens indiquen el caràcter digital del senyal. En efecte, les ratlles es corresponen amb els valors del senyal que realment estan disponibles en l'arxiu wav original que hem introduït.

Suggeriment

Se suggereix fer un zoom de la zona d'àudio entre el segon 24 i el 25.

Senyals analògics

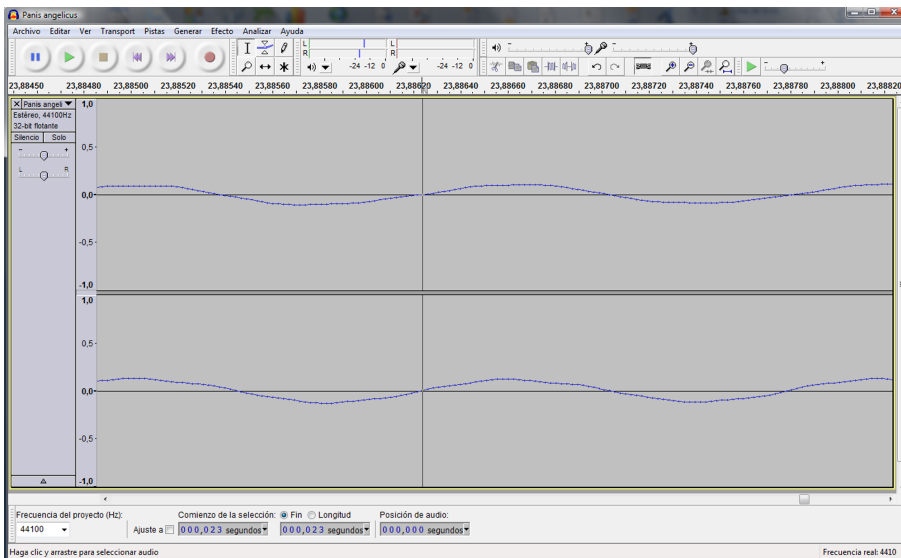
Un matemàtic diria que un senyal analògic és una funció real de variable real, encara que per a nosaltres serà suficient el concepte de **funció contínua** que evoluciona en el temps.

La conseqüència més important d'aquest resultat és que la informació proporcionada a l'arxiu digital original només conté els valors del senyal en uns instants de temps determinats: els **instants de mostreig**.

En la figura 33 es reproduïx un resultat semblant al que hauríem d'haver obtingut fins ara en l'activitat. És important notar que les ratlles estan distribuïdes uniformement al llarg de l'eix temporal (horitzontal), la qual cosa ens indica que el mostreig del senyal es fa a intervals de temps regulars i que denominarem el **període de mostreig**.

L'invers del període de mostreig és la **freqüència de mostreig** i ens indica el nombre de mostres que prenen per segon.

Figura 33. Representació d'un petit fragment de l'arxiu en què s'observa que el senyal digital està format per un conjunt de mostres preses a intervals de temps constants.



Activitat

Determineu el nombre de mostres que s'han pres per a un període de repetició bàsic del senyal d'àudio que teniu representat en la pantalla (per facilitar el compte fixeu-vos en els encreuaments per zero del senyal).

Sabent que aquest senyal s'ha mostrejat a una freqüència de 44.100 Hz, determineu el període i la freqüència de la forma d'ona que teniu representada en la pantalla.

Solució

El resultat depèn del fragment de senyal que s'hagi ampliat. Si el fragment correspon al so entre el segon 24 i 25, les mostres d'un període fonamental haurien de ser unes 67 o 68.

Per al cas anterior (posem un període de 68 mostres) el període de repetició del senyal serà:

$$P = 68 \times \frac{1}{f_m} = 68 \times \frac{1}{44.100} = 0,00154195 = 1,5419 \text{ ms}$$

La freqüència fonamental del senyal en aquest fragment serà l'invers del període:
 $F = 648 \text{ Hz}$.

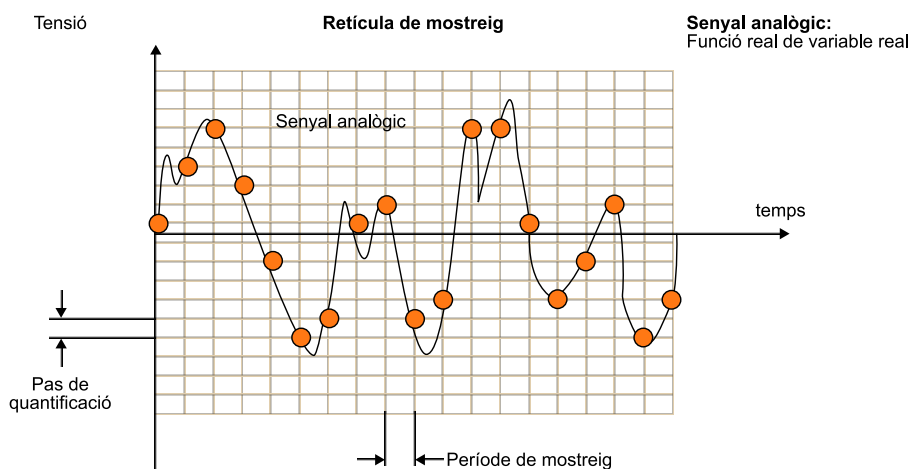
Un senyal digital només conté informació en els instants de mostreig. Aquests instants de mostreig tenen lloc a intervals de temps constants, que es coneixen com el període de mostreig. La freqüència de mostreig és l'invers del període de mostreig i ens indica el nombre total de mostres que es prenen per segon.

D'altra banda, si ens centrem en els valors d'amplitud que el senyal pren en la gràfica, veurem que no tots els valors són possibles sinó que hi ha uns nivells predeterminats.

Això no és senzill de veure, ja que el nombre total de nivells és molt elevat i les diferències entre un nivell i el següent poden ser molt petites. En definitiva, en l'eix horitzontal (temps) el senyal només pren valors en uns instants concrets mentre que en l'eix vertical, el senyal només pot prendre uns nivells d'amplitud predeterminats. Aquesta divisió dels eixos horitzontal i vertical dona lloc al concepte de **retícula de mostreig**.

Per aclarir aquests conceptes representem esquemàticament la retícula de mostreig en la figura 34. En aquesta gràfica es representa un senyal analògic i la retícula de mostreig que acabem de descriure. La retícula de mostreig es caracteritza pels seus passos tant en l'eix horitzontal com en l'eix vertical. Els valors d'aquests passos es coneixen com a **període de mostreig** (horitzontal) i **pas de quantificació** (vertical). El senyal digital són els punts de la retícula de mostreig que més s'aproximen al senyal analògic i que s'han marcat amb cercles vermells en la figura 34. Resulta evident que el senyal digital solament és una aproximació al senyal analògic i que la qualitat d'aquesta aproximació depèn de la densitat de la retícula de mostreig. Així, doncs, com més gran sigui la densitat de mostres i de passos de quantificació millor serà l'aproximació digital al senyal analògic.

Figura 34. Aproximació d'un senyal analògic mitjançant una retícula de mostreig



L'observació anterior suggereix una pregunta:

Existeix un nombre ideal per al període de mostreig i per al pas de quantificació? Aquesta és la pregunta clau que intentarem resoldre en aquesta activitat i en la següent.

De moment, en aquesta activitat ens centrarem a analitzar el període de mostreig o, el que és el mateix, el nombre de mostres per segon que s'han de prendre d'un senyal (freqüència de mostreig). Veurem que la freqüència de mostreig depèn de les característiques en freqüència del senyal i que són aquestes característiques les que determinen el nombre mínim de mostres que s'han de considerar. En l'activitat següent estudiarem els efectes del pas de quantificació en la qualitat del senyal digital.

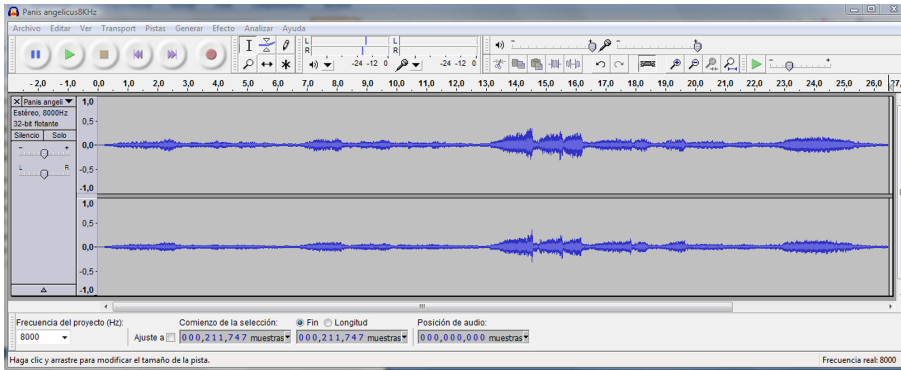
2.3.2. La freqüència de mostreig

En aquest apartat experimentarem amb la freqüència de mostreig i estudiarem quins són els requisits que s'han de complir per a digitalitzar correctament un senyal.

Arribarem a establir el teorema del mostreig que ens permetrà conèixer el nombre mínim de mostres per segon que són necessàries per a digitalitzar correctament un senyal, sense que es puguin observar diferències significatives amb el senyal analògic original. En tot aquest apartat i en la resta d'aquesta activitat suposarem que el nombre de nivells de quantificació és prou elevat perquè no influeixi en la qualitat del senyal.

1) Comencem carregant en l'Audacity els senyals **Panis Anglicus.wav** i **Panis Angelicus8kHz.wav**.

- El primer és un senyal en format WAV (àudio no comprimit, el format del qual s'estudiarà en activitats posteriors) pres directament del CD original i mostrejat a 44.100 Hz.
- El segon senyal s'ha obtingut utilitzant una freqüència de mostreig de 8 kHz. Veiem que no és possible apreciar diferències a simple vista entre les dues formes d'ona. No obstant això, si ajustem els marcadors de posició del cursor a mostres, en la part inferior del menú (vegeu la figura 35), comprovarem que mentre que un dels fragments té un total de 1.168.000 mostres (26,5 s x 44.100 mostres/s) l'altre senyal només té unes 212.000 mostres (26,5 s x 8.000 mostres/s).

Figura 35. Forma d'ona del senyal *Panis angelicus* mostrejada a 8 kHz

En les finestres inferiors s'ha seleccionat el nombre de mostres per a visualitzar la selecció i la posició del cursor.

Ara escoltarem amb certa atenció l'àudio de cada un dels dos senyals.

2) Escolteu els dos fragments i observeu les diferències de qualitat entre les dues versions.

Nota

Les freqüències de mostreig s'han escollit perquè resulti bastant evident la diferència de qualitat entre els dos senyals. No obstant això, en un ordinador portàtil o en equips amb àudio de baixa qualitat és possible que no resulti fàcil discriminar els dos senyals. En aquest cas, se suggereix que s'utilitzin auriculars.

3) Intenteu expressar qualitativament i anoteu en un paper les diferències de qualitat que aprecieu entre els dos senyals.

Repetim aquest experiment amb un altre senyal d'àudio.

1) Ara, obrim amb l'Audacity els fitxers **SintoniaOpera-22050.wav**, **SintoniaOpera-16000.wav** i **SintoniaOpera-8000.wav**. En aquest cas, es tracta de la sintonia d'un programa de ràdio difós per Internet. La qualitat de partida d'aquests arxius és considerablement inferior a l'enregistrament del primer fragment.

2) Com en el cas anterior podem comprovar que les freqüències de mostreig són de 22.050 Hz, 16.000 Hz i 8.000 Hz, i que el nombre total de mostres dels fitxers coincideix amb l'esperat. Escolteu atentament els tres senyals i intenteu expressar qualitativament les diferències que observeu.

Els resultats principals de les audicions anteriors són:

1) La qualitat del senyal *Panis angelicus* mostrejat a 44.100 Hz és considerablement superior a la mostrejada a 8.000 Hz. En aquesta última s'aprecia una pèrdua considerable d'aguts i brillantor.

2) Els dos senyals **SintoniaOpera** mostrejats a 22.050 Hz i 16.000 Hz són indistingibles pel que fa a la qualitat.

3) El senyal **SintoniaOpera** mostrejat a 8.000 Hz representa una pèrdua de qualitat considerable respecte als altres dos.

En resum, un dels efectes que pot tenir la reducció de la freqüència de mostreig és una pèrdua de qualitat que es percep com una disminució dels aguts (casos 1 i 3).

No obstant això, en algunes circumstàncies no és possible observar cap millora o pèrdua quan es canvia la freqüència de mostreig (cas 2). Per a quantificar i analitzar amb més precisió el que passa, per què en alguns casos apreciem diferències i en altres casos no, és convenient que ens centrem en l'anàlisi en freqüència dels senyals.

2.3.3. Anàlisi en freqüència de *Panis angelicus*

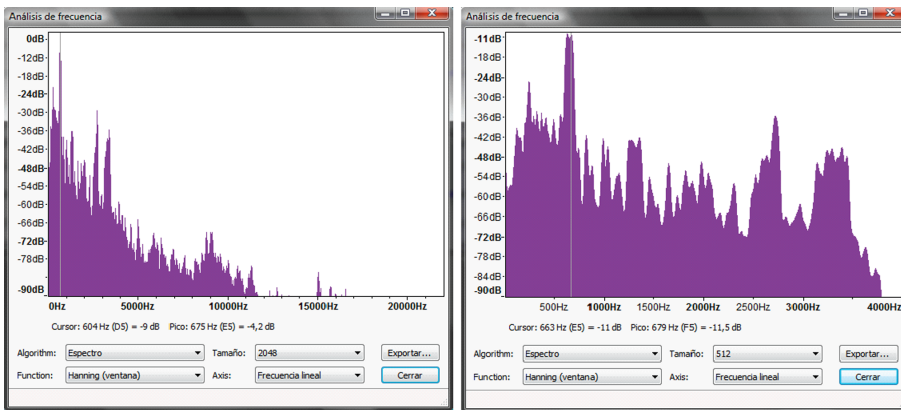
Compararem la representació en freqüència d'un fragment del senyal *Panis angelicus* mostrejat a 44.100 Hz amb la representació del mateix fragment del senyal mostrejat a 8.000 Hz.

1) Per a calcular la representació en freqüència (transformada de Fourier) hem de seleccionar un fragment de senyal. En aquest cas, suggerim que es prengui el fragment entre el segon 23 i el segon 24, que correspon a una nota d'una freqüència fonamental aproximada de 675 Hz.

2) Seleccioneu el fragment d'un segon de durada amb el cursor i executeu les opcions *Analitzar* i *Anàlisis de espectre*. Escolliu una representació de l'espectre, amb una finestra de **Hanning** de **2.048** mostres. Observeu el resultat obtingut i repetiu les mateixes operacions per al fitxer mostrejat a 8.000 Hz. En aquest cas trieu una mida de la finestra de 512 mostres⁶. Els resultats obtinguts es representen en la figura 36, en què s'observa un canvi molt significatiu de l'espectre dels dos senyals. En efecte, l'espectre del senyal mostrejat a 44.100 Hz s'estén més enllà dels 10.000 Hz, fins i tot amb alguns components visibles entorn dels 15.000 Hz. En canvi, l'espectre del senyal de 8.000 Hz queda limitat a components inferiors als 4.000 Hz, per la qual cosa ha perdut gran part dels components d'alta freqüència del primer senyal. També és important observar que el contingut freqüencial dels dos senyals entre 0 kHz i 4 kHz és aproximadament igual.

⁽⁶⁾ Les mides de les dues finestres són diferents perquè volem que l'interval temporal en què es calculen les dues transformades siguin aproximadament iguals. Perquè el període de temps considerat sigui aproximadament el mateix haurem de prendre més mostres del senyal amb una freqüència de mostreig alta que del senyal amb una freqüència de mostreig baixa. Vegeu l'activitat d'àudio 1 per interpretar el significat de la durada de les finestres

Figura 36. Comparativa entre els espectres del senyal *Panis angelicus* mostrejat a 44.100 Hz i el mostrejat a 8.000 Hz



En la imatge de la dreta el contingut spectral està limitat a freqüències inferiors als 4 Hz.

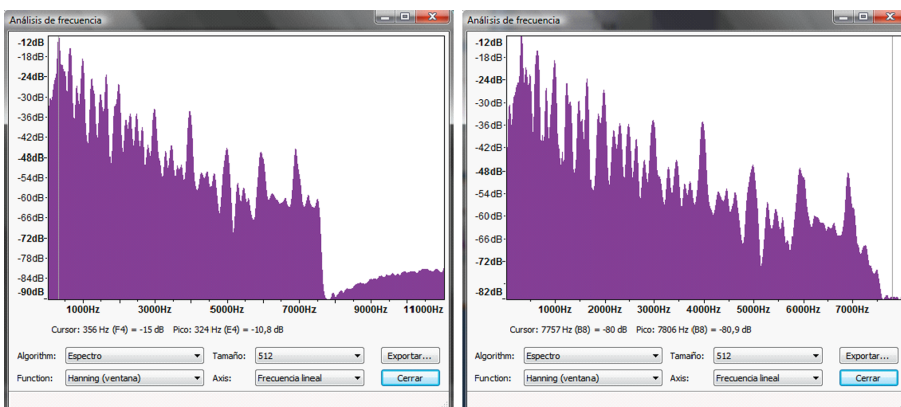
2.3.4. Anàlisi en freqüència de *Sintonía Ópera*

En aquest apartat repetim les proves anteriors però ara amb l'altre senyal. Hem pres el fragment entre 1 segon i 3 segons en tots els casos per fer l'anàlisi en freqüència.

En la figura 37 es pot veure una comparativa entre el senyal mostrejat a 22.050 Hz i el mostrejat a 16.000 Hz. En aquest cas, resulta evident que els dos senyals tenen pràcticament el mateix contingut en freqüència. En efecte, el contingut entre 0 kHz i 8 kHz dels dos senyals és exactament igual. El senyal de 22.050 Hz té uns components freqüencials molt baixos (molts dB per sota de la resta), centrats entorn d'11 kHz. No obstant això, a causa de la seva potència reduïda, aquests components no són audibles⁷.

⁽⁷⁾ Segurament, aquests components de potència reduïda procedeixen d'algun processament fet en l'emissora de ràdio (o en un altre sistema) per a canviar la freqüència de mostreig a 22.050 Hz.

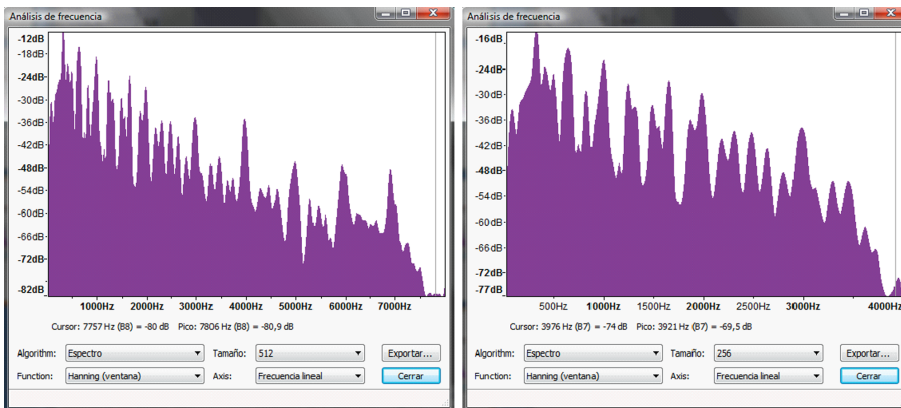
Figura 37. Espectres dels senyals *Sintonía Ópera* mostrejats a 22.050 Hz i 16.000 Hz (fragment de 2 segons de duració)



Els espectres dels dos senyals són idèntics en la banda entre 0 Hz i 8.000 Hz. S'observen alguns components addicionals, entre 9.000 Hz i 11.000 Hz, per al senyal de freqüència de mostreig més alta. No obstant això, a causa de la baixa potència aquests components no resulten audibles.

Finalment, en la figura 38 també es comparen els espectres entre el senyal *Sintonía Ópera* mostrejat a 16.000 Hz i el mostrejat a 8.000 Hz. En aquest cas, es veu que l'últim senyal només té components freqüencials fins a 4.000 Hz, per la qual cosa és lògic que des d'un punt de vista auditiu representi una pèrdua de brillantor i de components d'alta freqüència.

Figura 38. Comparativa entre els espectres del senyal Sintonía Ópera mostrejat a 16.000 Hz (esquerra) i 8.000 Hz (dreta)



El senyal de la dreta ha perdut tots els components freqüencials a partir dels 4.000 Hz.

Així, doncs, des d'un punt de vista experimental i intuïtiu sembla que podem obtenir les conclusions següents sobre la freqüència de mostreig d'un senyal.

Quan un senyal es mostreja a F mostres per segon llavors tots els components freqüencials superiors a $F/2$ no són presents.

Així, per exemple, quan la freqüència de mostreig sigui de 32.000 Hz, tots els components freqüencials superiors a 16.000 Hz no seran presents.

Si un senyal només té components freqüencials significatius fins a una freqüència W , llavors no obtenim cap millora auditiva apreciable si mostregem a freqüències superiors a $2W$.

Per exemple, si un senyal sols té components en freqüències fins als 4.000 Hz, llavors no té sentit mostrear-lo a més de 8 kHz. La qualitat que obtindrem a 16 kHz, 30 kHz o 8 kHz és la mateixa.

2.3.5. Freqüències de mostreig per a senyals de veu

Abans de començar a formalitzar els conceptes del teorema del mostreig es proposa que dugueu a terme l'experiència següent:

1) Ajusteu el volum d'enregistrament del micròfon en el programa Audacity i enregistreu la vostra veu a diferents freqüències de mostreig. Per a fer el registre es recomana que repetiu constantment una mateixa paraula (per exemple, *provant, provant*) unes quatre o cinc vegades i posteriorment compareu la qualitat de l'audició utilitzant diferents freqüències de mostreig.

2) Ajusteu els nivells de la manera correcta perquè el nivell d'enregistrament sigui el correcte, sense que es produeixi distorsió per saturació dels amplificadors.

Repetiu l'enregistrament per a diverses freqüències de mostreig, començant per 48.000 Hz i reduint a 44.100 Hz, 22.050 Hz, 16.000 Hz i 8.000 Hz. Per a cada un dels registres, seleccioneu un fragment d'àudio (la paraula que heu repetit) i determineu-ne l'espectre.

Justifiqueu quina penseu que és la freqüència de mostreig idònia per al senyal de veu⁸.

En les seccions següents formalitzarem alguns conceptes relacionats amb el mostreig de senyals. En concret, definirem el concepte de senyals de banda limitada i establirem el teorema del mostreig. També introduïrem el concepte d'aliàsing.

⁽⁸⁾El resultat final depèn de la qualitat del vostre sistema d'enregistrament: micròfon incorporat a l'estació de treball, micròfon extern, qualitat del micròfon, qualitat de la targeta d'adquisició.

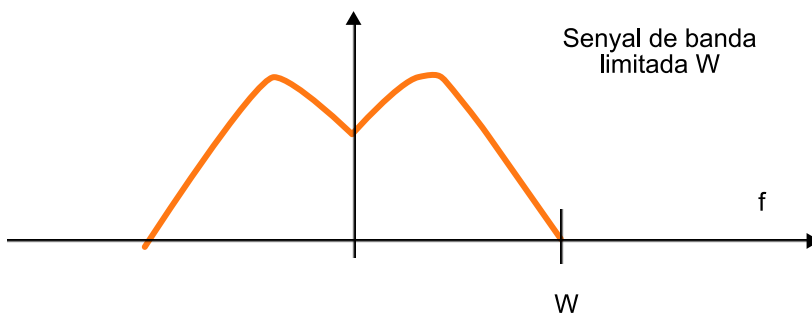
2.3.6. Senyals de banda limitada

Es diu que un senyal és de **banda limitada** quan el seu contingut espectral és nul a partir d'una determinada freqüència W .

En la figura 39 es representa esquemàticament l'espectre d'un senyal de banda limitada.

La freqüència que actua com a límit del contingut espectral del senyal es denomina **amplada de banda**.

Figura 39. Senyal de banda limitada amb una amplada de banda W



Els senyals de banda limitada són molt importants en els sistemes digitals ja que, com veurem, el teorema del mostreig estableix que només es pot dur a terme un mostreig sense pèrdua d'informació quan el senyal és de banda limitada. Els senyals d'àudio i de vídeo són de banda limitada, per la qual cosa admeten representacions digitals.

En el cas dels **senyals d'àudio** és la mateixa oïda humana la que estableix el límit d'amplada de banda del senyal. La capacitat per a la detecció de **freqüències i tons** depèn de cada individu i de l'edat, però està establerta entre 15 kHz i 20 kHz. Els senyals amb una freqüència superior a 20 kHz es poden

considerar inoïbles per a l'oïda. Per això, el senyal d'àudio únicament requereix disposar de contingut espectral fins a aquesta freqüència màxima perquè pugui ser reproduït amb total fidelitat.

En telefonia analògica convencional es considera que l'amplada de banda del senyal de veu és d'uns **3,4 kHz**. Aquesta amplada de banda proporciona informació suficient perquè els missatges telefònics es puguin interpretar correctament i es pugui reconèixer l'interlocutor. El fet que en telefonia no s'utilitzi tota l'amplada de banda del senyal d'àudio està lligat a un problema tecnològic i econòmic per a l'optimització dels recursos d'amplada de banda. En efecte, com més petita és l'amplada de banda de cada senyal individual que s'ha d'encaminar al llarg de la xarxa telefònica, millor es poden aprofitar els recursos, transmetent més canals en paral·lel (multiplexació) per les mateixes connexions físiques.

A la pràctica no resulta evident determinar l'amplada de banda d'un senyal. Quan analitzem l'espectre sempre apareix un **soroll de fons** que dificulta la identificació dels components d'espectre que corresponen al senyal i els que corresponen al soroll.

Es poden elaborar diferents regles genèriques per a calcular l'amplada de banda d'un senyal a partir de la visualització del seu espectre, encara que totes aquestes regles són empíriques i tenen excepcions.

Una regla utilitzada amb certa freqüència és donar com a amplada de banda del senyal la freqüència a partir de la qual tot l'espectre està x dBs per **sota del seu valor màxim**. El valor d' x depèn de les característiques del nostre sistema d'adquisició.

Activitat

Tenint en compte la regla anterior determineu l'amplada de banda dels senyals de veu digitalitzats en l'apartat anterior (considerant que heu utilitzat un sistema de captura domèstic, preneu el valor d' $x = 60$ dBs).

2.3.7. Teorema del mostreig

El teorema del mostreig estableix que quan un senyal és de banda limitada es pot mostrejar sense que es produeixin pèrdues d'informació utilitzant una freqüència de mostreig superior al doble de la seva amplada de banda.

$$f_{\text{muestreo}} = \frac{1}{T_{\text{muestreo}}} > 2 \cdot W$$

Variació d' x dBs

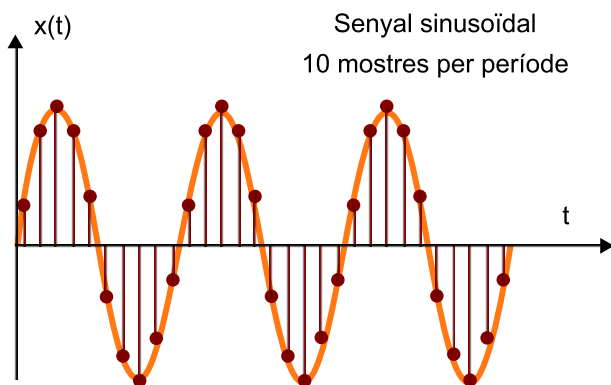
Si tenim un sistema d'adquisició econòmic, amb bastant soroll, x pot prendre el valor de 50 dBs o 60 dBs. En sistemes d'adquisició d'estudis professionals podem estar entorn de 85 dBs i 100 dBs.

Aquesta relació entre la freqüència de mostreig i l'amplada de banda del senyal també es coneix amb el nom de **criteri de Nyquist**.

Així, per exemple, si un senyal té una amplada de banda de 15 kHz, la freqüència mínima a què s'ha de mostrejar és el doble, és a dir, 30 kHz.

D'acord amb aquest teorema, la freqüència de mostreig mínima per a poder treballar amb senyals d'àudio d'alta fidelitat estaria situada per sobre de 40 kHz (20 kHz d'amplada de banda). En la figura 40 s'il·lustra el procés de mostreig sobre un component sinusoidal.

Figura 40. Mostreig d'un component sinusoidal amb una freqüència de mostreig que verifica el criteri de Nyquist.



En aquest exemple es prenen un total de 10 mostres per període, per la qual cosa el teorema del mostreig es verifica sense cap mena de problema. En efecte, en prendre 10 mostres per període, la freqüència de mostreig és 10 vegades superior a la freqüència del senyal, de manera que complim perfectament amb el teorema de Nyquist. En la figura 40 veiem, de manera intuïtiva, que el nombre de mostres és prou elevat per a poder reconstruir la forma del senyal original a partir de les seves mostres. Essencialment, el teorema de Nyquist ens diu que podrem fer aquesta reconstrucció sempre que prenguem més de dues mostres per període (freqüència de mostreig superior al doble de l'amplada de banda del senyal).

2.3.8. Aliàsing i filtres antialiàsing

Per a garantir que el procés de mostreig d'un senyal analògic és correcte, hem d'assegurar-nos que es compleix el criteri de Nyquist, això és, que la freqüència de mostreig que utilitzem és **més alta que el doble de l'amplada de banda**. A la pràctica, per a garantir que això es compleix es fa una mena de doble revisió.

- En efecte, per exemple, si volem mostrejar un senyal d'àudio d'alta fidelitat (20 kHz d'amplada de banda) escollirem una freqüència de mostreig que sigui com a mínim dues vegades més alta que l'amplada de banda.
- La segona part d'aquesta doble revisió és que hem de garantir que en cap circumstància no tenim senyals que no compleixen amb el criteri de Nyquist. Per a això, una vegada hem seleccionat la freqüència de mostreig,

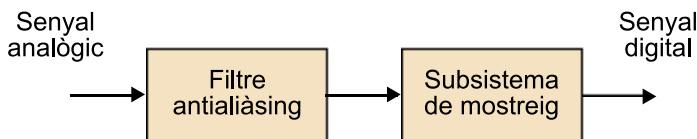
44.100 Hz

En àudio digital, una freqüència de mostreig molt utilitzada són **44.100 Hz**, que compleix amb el teorema de mostreig.

posarem un sistema que s'encarrega d'eliminar en el senyal els possibles components freqüencials que no compleixin amb el teorema de mostreig. Aquest sistema es denomina **filtre antialiàsing** i la seva funció és bàsicament eliminar els possibles components de senyal la freqüència dels quals sigui superior a la meitat de la freqüència de mostreig (22.050 Hz en el nostre exemple).

A la pràctica, si la nostra primera hipòtesi sobre les característiques del senyal és correcta, el filtre antialiàsing no actuarà. No obstant això, sempre és convenient posar-lo perquè en un sistema real podem tenir, a més del senyal desitjat, interferències o sorolls no desitjats les freqüències dels quals poden ser superiors al límit marcat pel teorema de mostreig. En la figura 41 es representa un diagrama de blocs en el qual s'indica la presència del filtre antialiàsing que precedeix el subsistema de mostreig del senyal.

Figura 41. Filtratge passabaix previ al mostreig d'un senyal analògic

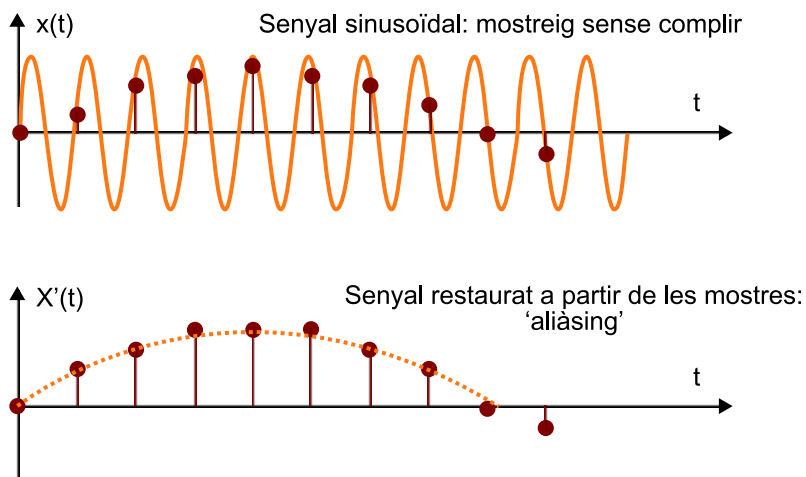


El filtre antialiàsing impedeix l'aparició del fenomen d'**aliàsing**, que consisteix en el fet que els components del senyal amb una freqüència superior a la meitat de la freqüència de mostreig (les que no compleixen el criteri de Nyquist) apareixen, una vegada mostrejats, com a components de senyal amb freqüències "falses", dins de l'amplada de banda del senyal útil.

El fenomen de l'antialiàsing s'il·lustra de manera molt simplificada en la figura 42. En aquest exemple es representa un component sinusoidal i les seves mostres, utilitzant una freqüència de mostreig que no verifica el criteri de Nyquist (hem vist que per a sinusoides hauríem de prendre un mínim de dues mostres per període). De manera intuïtiva, el senyal que podem reconstruir a partir de les mostres és el que obtenim a partir de la unió dels valors (punts) del senyal en els instants de mostreig.

Això dóna lloc a l'aparició d'un component sinusoidal, amb una freqüència molt més baixa, que no hi havia en el senyal original i que per això rep el nom d'**àlies**. El problema és més important del que al principi pot semblar a causa d'aquest canvi de freqüència en els components del senyal, que poden aparèixer com a interferències, és a dir, com a senyals que prèviament no existien.

Figura 42. Aparició del fenomen de l'antialiàsing quan la freqüència de mostreig no compleix el criteri de Nyquist



Podríem pensar que l'estructura de mostreig de la figura 42 solament és necessària quan no podem assegurar que el senyal és de banda limitada. Així, en el cas de senyals d'àudio podríem prescindir del filtre antialiàsing sempre que garantíssim que la freqüència de mostreig supera 40 kHz, amb la qual cosa garantim que els nostres senyals verifiquen el criteri de Nyquist. Tanmateix, el problema no és tan senzill, ja que res no impedeix assegurar que la font de senyal o els mateixos amplificadors rebin interferències de freqüències superiors a 20 kHz.

És important que ens adonem que aquestes freqüències no resulten audibles mentre els senyals es mantinguin en el domini analògic, però que quan fem un mostreig, a causa de l'aliàsing, poden canviar de freqüència i entrar en la banda audible.

Per això, el filtre antialiàsing és estrictament necessari, en tots els casos, abans de procedir al mostreig del senyal.

2.3.9. Alguns exemples d'aliàsing

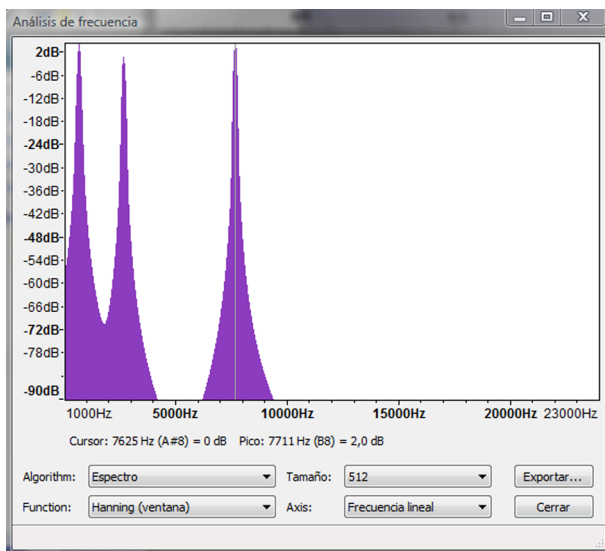
Per a comprendre millor el problema i els efectes de l'aliàsing en senyals d'àudio, sembla aconsellable dur a terme algunes audicions que ens n'evidenciïn els efectes.

En aquesta secció escoltarem i analitzarem amb el programa Audacity alguns fitxers d'àudio que ens intentaran mostrar els efectes de l'aliàsing. Començarem les proves amb senyals generats artificialment, utilitzant tons i modulacions, en els quals l'aliàsing es presenta de manera molt evident. Posteriorment, utilitzarem arxius musicals en què l'aliàsing es presenta d'una manera menys clara.

1) Començarem amb un senyal sintètic format per tons de diferents freqüències i que s'ha digitalitzat amb diferents freqüències de mostreig. El senyal original és una superposició de tons amb modulació d'amplitud i amb una freqüència màxima de 7.700 Hz. El senyal **Orig_Tones.wav** està mostrejat a **48.000 Hz**, de manera que compleix el criteri de Nyquist.

2) Seleccioneu un fragment d'un segon del senyal i feu la vostra anàlisi en freqüència. L'espectre del senyal es mostra en la figura 43, en què s'observa que està compost per tres tons en les freqüències aproximades de 700 Hz, 2.700 Hz i 7.700 Hz. Com que la freqüència de mostreig és considerablement superior al doble de la màxima freqüència, no observariem diferències significatives entre un senyal analògic compost per aquests tres tons i el senyal digital corresponent.

Figura 43. Espectre del senyal sintètic Orig_Tones.wav, compost per tres tons amb freqüències de 700 Hz, 2.700 Hz i 7.700 Hz mostrejat a 48.000 Hz, complint per tant amb el teorema de mostreig



El pas següent és analitzar quin és el senyal que obtindríem si es digitalitzés la seqüència original amb una freqüència inferior a l'establerta pel criteri de Nyquist. En aquest cas, suposarem que mostregem el senyal anterior a una freqüència de **8.000 Hz** i utilitzant l'esquema representat en la figura 41, amb el filtre antialiàsing. En aquest cas, com que la freqüència de mostreig és de 8.000 Hz, el filtre antialiàsing eliminarà tots els components de senyal que estiguin situats per sobre de 4.000 Hz. El senyal resultant és en el fitxer **Tones_8K_antialias.wav**.

1) Obriu el senyal amb l'Audacity, escolteu-lo, seleccioneu-ne un fragment d'aproximadament 1 segon i representeu-ne l'espectre (que hem representat en la figura 44). A partir d'aquest experiment és evident el següent:

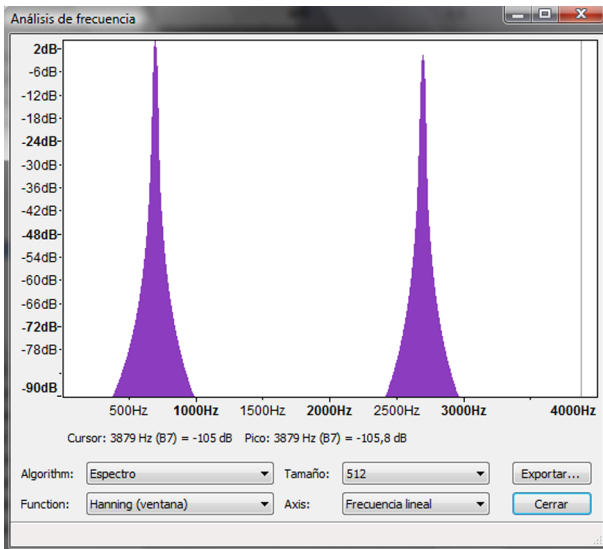
a) El senyal mostrejat a 8 kHz, amb filtre antialiàsing manca del component a 7.700 Hz, ja que ha estat filtrat.

Precaució

Podeu escoltar aquest senyal amb l'Audacity, però tingueu la precaució de reduir el volum dels auriculars fins a un nivell adequat, ja que el senyal és una superposició de tons purs modulats i pot resultar molt molesta a l'oïda.

b) En escoltar aquest senyal només sentim la superposició dels tons a 700 Hz i 2.700 Hz. Tots dos tons eren presents en el senyal original. Per tant, aquest senyal és la reproducció més fidel possible del senyal original, ja que conté tota la part de l'espectre que es pot reproduir correctament amb la freqüència de mostreig actual i manca dels components que no es poden reproduir correctament.

Figura 44. Espectre del senyal sintètic mostrejat a 8 kHz utilitzant un filtre antialiàsing



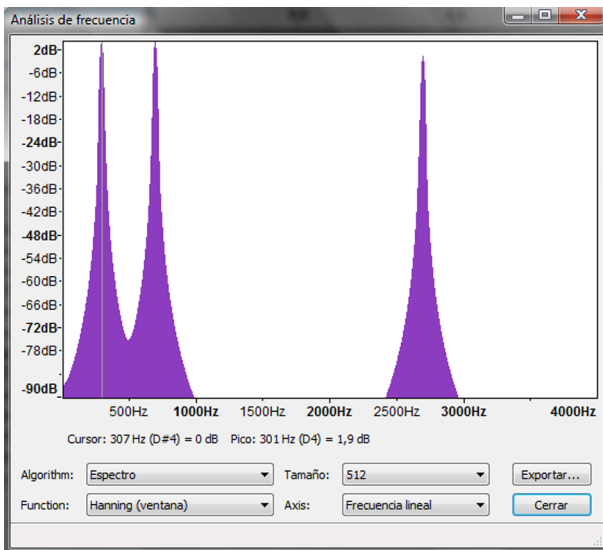
El component a 7.700 Hz ha estat eliminat pel filtre.

Finalment, analitzarem el resultat que s'obté quan el senyal original es mostra a 8 kHz **sense utilitzar el filtre antialiàsing**.

- 1) Carregueu el senyal **Tones_8K_NOAntialias.wav** en l'Audacity.
- 2) Seleccioneu un fragment d'aproximadament 1 segon i determineu l'espectre del senyal, el resultat del qual es representa en la figura 45.
- 3) Escolteu aquest senyal i compareu-lo amb l'original i amb el que s'obté utilitzant el filtre antialiàsing.

És clar que si no utilitzem filtre antialiàsing apareix un nou component espectral en el senyal que no era present en el senyal original. En el nostre cas, el nou component espectral apareix en la freqüència de 300 Hz. Aquest component és degut al mostreig incorrecte del component de 7.700 Hz. Fixeu-vos que la nova freqüència depèn tant de la freqüència del to original com de la freqüència de mostreig del senyal: $300 \text{ Hz} + 7.700 \text{ Hz} = 8.000 \text{ Hz}$. Des del punt de vista auditiu, el nou component produeix un so que no té cap relació amb l'original.

Figura 45. Espectre del senyal sintètic mostrejat a 8 kHz sense utilitzar un filtre antialiàsing



Apareix un component en la freqüència de 300 Hz que no era present en el senyal original.

Com a conseqüència directa d'aquesta experiència hauríem de concloure que sempre que es digitalitza un senyal a una freqüència inferior al doble de la seva amplada de banda (criteri de Nyquist) s'aprecia una **pèrdua de qualitat auditiva**.

No obstant això, aquesta pèrdua de qualitat és especialment greu si no s'utilitza un filtre antialiàsing, ja que poden aparèixer components freqüencials que no eren presents en el senyal original.

L'efecte auditiu de l'aliàsing en senyals sintètics és clar, però resulta força més complex d'apreciar en **senyals musicals**. Posarem alguns exemples d'aliàsing amb senyals musicals que es poden escoltar i analitzar utilitzant l'Audacity.

El primer exemple és amb el fragment d'àudio del *Panis angelicus* que ja hem utilitzat en aquesta activitat.

1) Carregueu els arxius **PA_48K.wav**, **PA_8K.wav** i **PA_8K_NOAntialias.wav**. El primer és el fragment musical mostrejat a 48.000 Hz, el que serveix com a referència de qualitat. Els altres dos fragments estan mostrejats a 8.000 Hz. El primer ha estat mostrejat utilitzant un filtre antialiàsing mentre que per al segon no s'ha utilitzat cap filtre.

2) Escolteu atentament els dos senyals. Apreciar l'aliàsing que es produeix en el segon senyal requereix força atenció. S'hauria d'apreciar una espècie de "soroll", de baixa freqüència que apareix en alguns fragments. Es tracta d'un soroll que no hi havia en el senyal original i el contingut espectral del qual varia en funció de les pròpies característiques espectrals del senyal original, per la qual

cosa en alguns fragments apareix de manera molt clara mentre que en d'altres no es pot apreciar. És important observar que el senyal mostrejat amb el filtre antialiàsing no té aquest soroll i s'escolta amb més claredat.

Evidentment, representa una pèrdua considerable respecte al senyal original perquè bona part de l'alta freqüència ha estat eliminada.

Es presenta un altre exemple en els fitxers `srv.wav`, `srv_noanti.wav`, `srv_anti.wav` que es correspon amb un fragment d'un **sol de guitarra elèctrica**. El primer senyal està mostrejat a 22.050 Hz i es pot considerar com la qualitat de referència. Els altres dos senyals es corresponen amb freqüències de mostreig de 4.400 Hz sense filtre antialiàsing i amb filtre antialiàsing. La qualitat d'aquests dos últims senyals és molt baixa a causa de la reducció dràstica de la freqüència de mostreig. No obstant això, després de repetides audicions, hauríeu de ser capaços d'escoltar un soroll de fons incoherent en el senyal obtingut sense l'ús del filtre antialiàsing i que no és present en el senyal en el qual sí que s'ha aplicat el filtre.

Exemple

En els arxius `gg_noanti.wav` i `gg_anti.wav` es proporcionen uns altres dos exemples de mostreig a freqüències reduïdes amb el filtre i sense el filtre.

2.3.10. L'aliàsing en vídeo i cinema

L'aliàsing es pot produir sempre que es faci el mostreig d'un senyal, per tant, no és exclusiu dels senyals d'àudio. En cinema i televisió, l'escena també es mostreja en el temps a raó de **24 fotogrames per segon** (cinema) o **25 fotogrames per segon** (televisió en sistema PAL). L'aliàsing hauria de ser, doncs, un fenomen a què estem força acostumats.

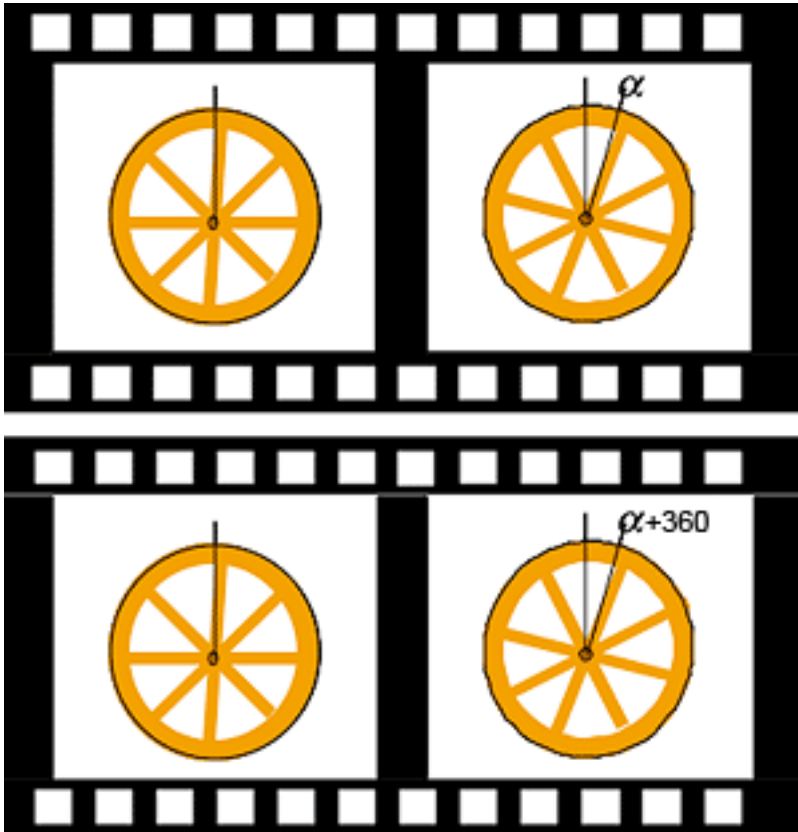
En efecte, un dels fenòmens que pot introduir el mostreig temporal de l'escena és l'aparició d'**efectes visuals** que falsegin o desvirtuïn el moviment real dels objectes. Aquests efectes són deguts al fet que la freqüència de mostreig escollida resulta insuficient per a una interpretació correcta del moviment. Així, tal com estableix el teorema del mostreig de senyals, si els moviments dels objectes són excessivament ràpids, hi pot haver components d'alta freqüència que, en mostrejar l'escena, s'encavalquin amb altres components i desvirtuïn el moviment real dels objectes.

Un dels exemples d'aquest fenomen apareix en moltes pel·lícules d'acció i de vegades fins i tot és provocat pel director per a produir en l'espectador la sensació de moviment. L'escena típica consisteix a centrar l'acció de la càmera en la roda d'un carro (en pel·lícules de l'Oest) o d'un automòbil. Quan el vehicle s'engega l'espectador observa i interpreta correctament el moviment de la roda, tanmateix, a mesura que la velocitat augmenta es produeix la sensació que els radis de la roda es paren o fins i tot giren en sentit contrari al de la marxa del vehicle.

Com s'il·lustra en la figura 46, aquest efecte és degut al fet que l'ull interpreta el moviment de la manera més simple possible a partir dels fotogrames que se li presenten. En la figura es representen dos fotogrames consecutius de la roda del carro que corresponen a dues situacions diferents però que, inevitablement, s'interpretaran de la mateixa manera, ja que les imatges que les defineixen són idèntiques. En la seqüència superior, la roda del carro experimenta

un gir d' α graus entre dos fotogrames consecutius, mentre que en la inferior el gir real de la roda és d' $\alpha + 360$ graus. En tots dos casos, el sistema visual interpretarà el mateix moviment aparent d' α graus, ja que la simetria de la roda no li permet distingir entre les dues seqüències. En aquest exemple, si la velocitat de gir de la roda fos tal que entre dos fotogrames consecutius es produís un gir de 360 graus, l'espectador no apreciaria cap moviment, ja que totes les imatges serien idèntiques. Es produeix un cas curiós quan l'angle de gir entre fotogrames és lleugerament inferior als 360 graus, ja que en aquest cas s'interpreta que el moviment de la roda és en sentit contrari al real. En aquest últim supòsit és evident que per al sistema visual resulta molt més factible interpretar que s'ha produït un petit moviment de la roda en sentit contrari al real que un moviment considerable en el sentit correcte.

Figura 46. Aliàs temporal de les imatges. En totes dues situacions, les seqüències d'imatges obtingudes són idèntiques, per la qual cosa el sistema visual interpretarà el mateix moviment.



Aquest efecte visual es coneix com l'**efecte estroboscòpic** i és provocat de vegades per a poder observar moviments de naturalesa periòdica d'alguns fenòmens físics la rapidesa dels quals no permet observar-los en condicions normals.

L'instrument utilitzat per a això es denomina **estroboscopi** i consisteix en una làmpada (flaix) que s'il·lumina a intervals regulars i la freqüència del qual i fase de disparament es pot controlar externament.

L'estroboscopi es pot utilitzar, per exemple, per veure detalladament la deformació que experimenta la membrana d'un altaveu quan és sotmesa a diferents excitacions periòdiques.

Per a fer l'experiment haurem d'alimentar l'altaveu amb un senyal sinusoidal a la freqüència a la qual volem fer l'anàlisi.

El sistema s'ha de situar en un entorn amb poca il·luminació, de manera que en disparar l'estroboscopi a la mateixa freqüència a què està vibrant l'altaveu podrem observar una imatge estacionària de la deformació de la membrana. Si modifiquem la fase dels disparaments, veurem les deformacions progressives que pren la membrana de l'altaveu en les diferents fases de la vibració. Fins i tot, si seleccionem una freqüència de disparament del flaix lleugerament diferent de la de la vibració de l'altaveu podrem alentir el moviment de la membrana a la velocitat desitjada.

Es proposa que el lector es plantegi quines possibles freqüències de l'estroboscopi utilitzaria per a poder observar el moviment d'una membrana que oscil·la a 300 Hz amb una velocitat aparent de 0,1 Hz. Tingueu en compte que, d'una banda, és convenient que la freqüència de l'estroboscopi permeti observar el moviment sense parpelleig i, d'una altra, que la freqüència de disparament no pot ser excessivament alta per la limitació pròpia del flaix.

Exemple

Es poden veure experiències semblants en diversos museus científics en els quals es mostren algunes aplicacions i efectes visuals obtinguts amb l'ajuda de l'estroboscopi. Una de les més espectaculars consisteix a alentir la caiguda lliure de les gotes d'aigua. En aquest cas, s'utilitza una cambra tancada en la qual es provoca un degoteig constant d'aigua. En il·luminar el flux de gotes d'aigua mitjançant un estroboscopi podem provocar que aquest s'il·lumini sempre que una gota d'aigua passi per una posició determinada, per la qual cosa veurem que la gota resta suspesa en l'aire. Una lleugera disminució o augment de la freqüència de tret produirà la sensació que la gota cau o puja lentament enmig. Es poden fer experiments similars per a visualitzar el vol d'insectes, analitzar les vibracions de màquines, etc.

2.3.11. Selecció de la freqüència de mostreig en sistemes d'àudio

El valor de la freqüència de mostreig en un sistema d'àudio o veu depèn de la qualitat final del senyal que es vulgui reproduir o emmagatzemar i de l'aplicació a què es destina el sistema. Quan es va definir el sistema *CD-digital audio* (CD d'àudio) es va prendre com a estàndard una freqüència de **44,1 kHz**, que permet treballar amb tot l'espectre audible (20 Hz a 20 kHz). El valor escollit verifica el criteri de Nyquist per a la banda d'àudio encara que, com ja hem comentat, sempre és necessari filtrar els senyals abans de mostrejar-los. Els filtres antialiàsing estan incorporats en els mateixos sistemes de digitalització. Així, qualsevol gravador multipista per a àudio professional incorpora uns filtres antialiàsing que compleixen amb els requisits marcats per l'estàndard d'àudio digital utilitzat.

44,1 kHz

La selecció de la freqüència de 44,1 kHz no és casual i es deu als primers sistemes de registre de senyals d'àudio en format digital que van aparèixer en la dècada dels seixanta. Aquests primers sistemes van ser desenvolupats per la BBC, utilitzant sistemes de vídeo modificats per a emmagatzemar la informació d'àudio en format digital. El sistema escollit emmagatzemava tres mostres d'àudio digital en l'espai de cinta que hauria d'haver correspost a una línia del senyal de televisió. Aquesta taxa de registre de dades dona lloc, per al sistema de vídeo utilitzat en aquella època a la Gran Bretanya, a una freqüència de mostreig de 44,1 kHz. Quan es va desenvolupar el sistema CD d'àudio es va prendre aquesta mateixa freqüència de mostreig com un tribut als primers sistemes d'àudio digital.

La freqüència de 44,1 kHz no és l'única freqüència de mostreig utilitzada en àudio digital. Quan es va definir el sistema DAT (*digital audio tape*) per al registre d'àudio en cintes magnètiques es van definir diverses freqüències de mostreig alternatives (32 kHz, 44,1 kHz i 48 kHz). L'objectiu original d'aquests canvis en la freqüència de mostreig era dificultar les còpies pirates de CD en formats de cintes magnètiques.

Els primers **minidiscos** que van sortir al mercat registraven la informació utilitzant freqüències de mostreig de **32 kHz**. Aquestes freqüències de mostreig exigeixen utilitzar un filtre antialiàsing que retalli la banda d'àudio útil entre 15 kHz i 16 kHz. Evidentment, la qualitat es redueix lleugerament, encara que només un nombre reduït de melòmans és capaç d'apreciar-ho.

Actualment, els sistemes portàtils **MP3** o **MP4** solen utilitzar una freqüència de mostreig de **44.100 Hz**, ja que els registres solen procedir d'enregistraments per a CD, encara que l'estàndard també permet utilitzar altres freqüències (32 kHz, 48 kHz, etc.).

El sistema **DVD de vídeo** utilitza la freqüència de **48 kHz** per al registre del senyal d'àudio. Això significa que, quan es produeix un DVD de vídeo partint d'una banda musical obtinguda en CD, s'ha de fer un canvi en la freqüència de mostreig.

El sistema **DVD d'àudio**, que no ha tingut gaire èxit comercial, pot utilitzar registres d'àudio amb freqüències de mostreig de **96 kHz** i **192 kHz**. Així, la banda d'àudio queda àmpliament coberta, amb la qual cosa s'espera resoldre per sempre la polèmica, promoguda per alguns melòmans amants dels vells discos de vinil, que les freqüències de mostreig del CD no permeten reproduir alguns matisos sonors que sí que es reproduïen amb els discos de vinil.

Els sistemes de **telefonía digital** utilitzen freqüències de mostreig més reduïdes, ja que la banda del senyal de veu és més reduïda que la musical i tampoc no es pretén obtenir qualitats de reproducció d'àudio digital. Recordem que en telefonía l'objectiu és la intel·ligibilitat del missatge i el reconeixement del locutor. No és necessària l'alta fidelitat. En aquest cas, la freqüència de mostreig utilitzada en la telefonía fixa per a la veu és de **8 kHz**.

Veurem que aquesta freqüència de mostreig exigeix un flux de dades considerable (64 kbps –transmetre 64.000 bits cada segon) que no és possible acomodar en els sistemes de **telefonía mòbil digital** ni en les comunicacions d'àudio per Internet. Per això s'han introduït diversos estàndards per a la **compresió/descompresió** de la informació d'àudio que intenten reduir el flux total de dades amb què es treballa sense afectar excessivament la qualitat del senyal. Les freqüències de mostreig de partida d'aquests sistemes se situen entre **6 kHz** i **20 kHz**.

2.3.12. Freqüència de reproducció del senyal d'àudio

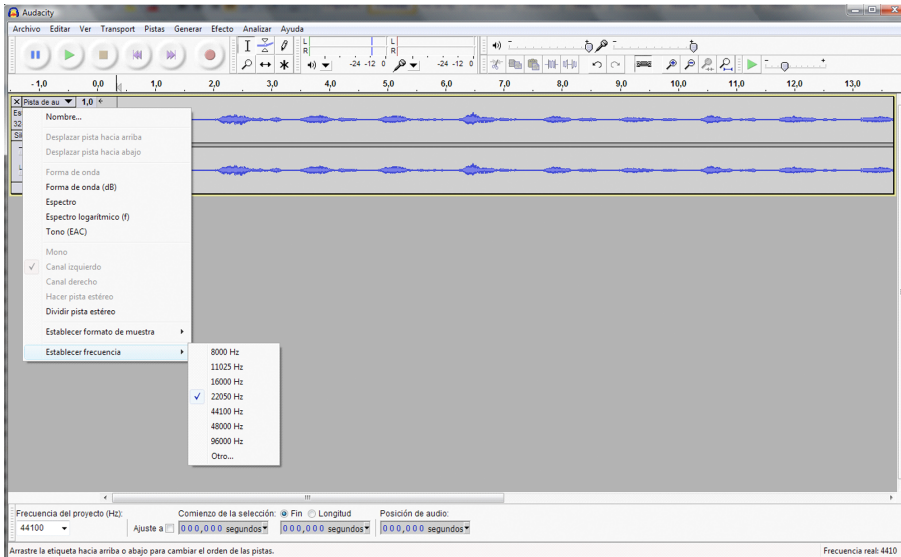
És molt important que la freqüència de mostreig utilitzada per a la reproducció del senyal d'àudio sigui **la mateixa** que la utilitzada per a la seva captura. Aquest fet és especialment important i ha de tenir-se en compte a causa de la multitud de freqüències de mostreig que es poden utilitzar en els sistemes d'àudio digital. Com ja hem comentat, és habitual disposar d'un senyal que s'ha adquirit a una freqüència de mostreig determinada i voler-lo reproduir en un sistema que no admet aquesta freqüència.

Exemple

Un exemple molt típic és la captura d'un senyal d'àudio procedent d'un CD d'àudio (44.100 Hz), editar-lo com a pista d'àudio d'un videoclip i reproduir-lo en un DVD de vídeo (sols admet reproduir àudio a 48.000 Hz). Podem trobar problemes similars per a passar de minidisc, DAT, MP3 a altres sistemes d'àudio.

- 1) Com a exemple per a veure els efectes de reproduir un senyal amb una freqüència de mostreig diferent de la que s'ha adquirit, es pot fer servir l'Audacity i carregar un dels senyals utilitzats en aquesta activitat (per exemple, **Panis Angelicus.wav**).
- 2) Podeu modificar la freqüència de mostreig de la reproducció obrint el desplegable d'**opcions** de la pista d'àudio i seleccionant la nova freqüència de mostreig (vegeu la figura 47).
- 3) Feu el canvi de freqüència de 44.100 Hz a 48.000 Hz i observeu els efectes (augment de la freqüència fonamental de la veu i disminució del temps total de reproducció). L'efecte produït és comparable al que tindríem en analògic en passar una cinta d'àudio a una velocitat superior a la que ha estat registrada.

Figura 47. Opcions per al canvi de la freqüència de mostreig del reproductor en l'Audacity



En tots aquests casos, el senyal s'ha de processar i remostrear a la freqüència de mostreig final del reproductor. Aquest processament es coneix amb el nom de **canvi de freqüència de mostreig** i és un dels aspectes que inclouen tots els sistemes d'edició d'àudio.

El canvi de la freqüència de mostreig es fa concatenant les operacions d'interpolació (augment de la freqüència de mostreig en un nombre enter) amb les operacions de delmació (disminució de la freqüència de mostreig en un nombre enter).

Així, per exemple, si volem passar d'una freqüència de mostreig de 44.100 Hz a 48.000 Hz hauríem d'interpol·lar el senyal en un factor 480 i delmar-la en un factor 441.

A la pràctica els valors d'interpolació i delmació de la freqüència de mostreig d'un senyal s'han de mantenir baixos perquè la complexitat computacional del processament d'àudio no sigui excessiva, per la qual cosa valors de 480 o 441 són absolutament impracticables. Així, doncs, a la pràctica, per a passar de 44.100 Hz a 48.000 Hz podríem aproximar-lo amb una interpolació en un factor 12 i una delmació en un factor 11. Amb aquests valors, si partim de 44.100 mostres per segon obtindrem un total de 48.109. Per tant, ens sobren 109 mostres cada segon, que haurem de descartar. La manera més habitual de descartar-les és eliminar aquestes 109 mostres i distribuir-les el més uniformement possible entre les 48.109 calculades i obtenir només 48.000 Hz cada segon.

2.4. Problemes/preguntes

1) Identifiqueu les freqüències de mostreig que s'utilitzen en els diferents sistemes d'àudio digital comercialitzats. Entre els sistemes d'àudio estudiats incloeu, com a mínim, els següents: CD d'àudio, DVD de vídeo, DAT, Minidisc, MP3, MP4, DAB, DVD d'àudio, SA-CD, DCC, NT, DAB+, DMB, MPEG-2 Layer 2, MPEG-2 LSF, NICAM 728, Dolby AC3, DTS, etc. Busqueu la documentació en bibliografia relacionada, pàgines de característiques tècniques de fabricants o documentació genèrica d'Internet.

2) Identifiqueu dos models de targetes de captura, processament i reproducció de so per a bus PCI o PCI Express per a PC. Determineu les freqüències de mostreig amb què poden treballar les targetes i el nombre màxim de canals d'àudio que es poden digitalitzar. Si és possible intenteu que els dos models siguin de diferents fabricants.

3) Identifiqueu dos models de targetes de captura, processament i reproducció d'àudio amb format ExpressCard. Compareu les prestacions d'aquests sistemes amb els identificats en la qüestió anterior.

4) Digitalitzeu un senyal musical procedent d'un sistema de reproducció amb casset compacte (*compact cassette*) analògic i mitjançant el programa Audacity estudieu l'espectre i amplada de banda del senyal obtingut. Proveu amb diferents freqüències de mostreig i determineu quin és el més adequat des del punt de vista auditiu.

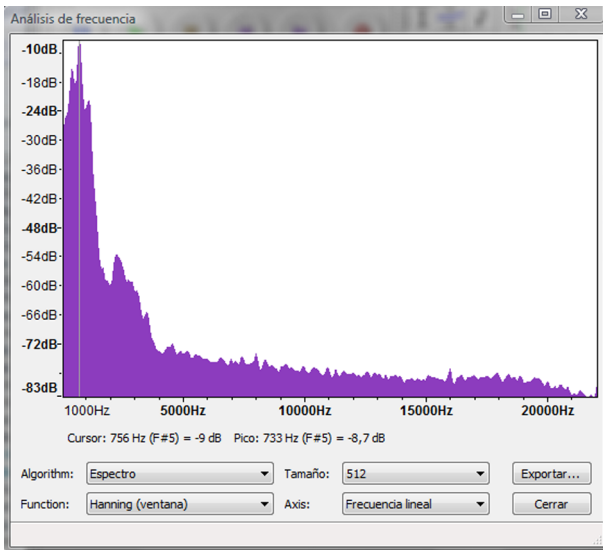
5) Una imatge digital es pot considerar com el mostreig en l'espai (files i columnes) d'una imatge analògica. Per tant, com que estem fent un mostreig és possible que aparegui el fenomen de l'aliàsing. Proporcioneu exemples en què pugui aparèixer el fenomen de l'aliàsing en imatges digitals. Indiqueu possibles maneres de fer/aplicar un filtre antialiàsing.

2.5. Avaluació

1) Considereu l'espectre del senyal d'àudio que es mostra en la figura 48 i que s'ha obtingut digitalitzant àudio amb l'Audacity. Supposeu que per especificar una amplada de banda prenem com a criteri la freqüència a partir de la qual el contingut espectral del senyal està 60 dBs per sota del màxim. Es demana el següent:

- a) Determineu la freqüència de mostreig amb què s'ha adquirit aquest senyal.
- b) Calculeu l'amplada de banda del senyal tenint en compte el criteri anterior.
- c) Determineu la freqüència de mostreig mínima a què considereu que es pot sentir aquest senyal sense pèrdues de qualitat apreciable.

Figura 48

**Solució**

a) El senyal s'ha mostrejat a 44.100 Hz, ja que la gràfica mostra fins a la meitat de la freqüència (22.050 Hz).

b) Tenint en compte el criteri dels -60 dBs respecte al màxim (-10 dBs,) haurem d'anar aproximadament als -70 dBs, que dona aproximadament 3.500 Hz d'amplada de banda.

c) Per a l'amplada de banda estimada en l'apartat anterior, una freqüència de mostreig de 7.000 Hz és suficient i proporciona la mateixa qualitat que el senyal original.

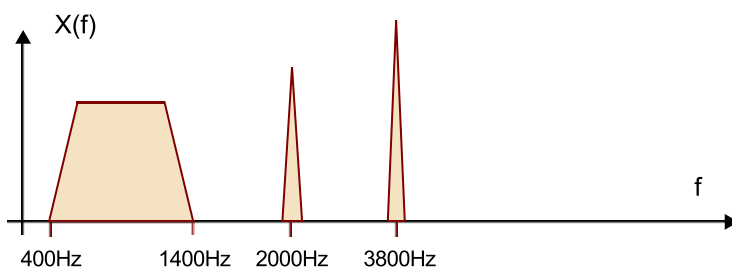
2) Considereu un senyal l'espectre del qual es pot aproximar pel de la figura 49.

a) Determineu l'espectre resultant si el senyal es mostreja a 10 kHz utilitzant un filtre antialiàsing.

b) Determineu l'espectre resultant si el senyal es mostreja 4,2 kHz utilitzant un filtre antialiàsing.

c) Determineu l'espectre resultant si el senyal es mostreja 4,2 kHz sense utilitzar un filtre antialiàsing.

Figura 49

**Solució**

a) Amb una freqüència de mostreig de 10 kHz l'espectre coincideix amb el de la figura.

b) El filtre antialiàsing haurà de filtrar 2,1 kHz, per tant, el to a 3.800 Hz desapareixerà.

c) Amb aquesta freqüència, si no utilitzem el filtre antialiàsing el to que originalment està en 3.800 Hz apareixerà en 400 Hz i se superposarà al senyal original.

3) Per canviar la freqüència de mostreig de 48.000 Hz a 44.100 Hz es proposa utilitzar una interpolació en un factor 11 i una delmació en factor 12.

a) Determineu la freqüència de mostreig final que obtindria.

b) Proposeu un procediment per a obtenir exactament 44.100 mostres per segon mantenint els factors d'interpolació i delmació de l'enunciat.

Solució

a) La freqüència de mostreig que obtindrem és de 44.000 Hz.

b) Haurem de crear 100 mostres més cada segon de les que hem calculat. Una possibilitat és repetir una de les mostres cada 10 ms. L'oïda no apreciarà la diferència i obtindrem un senyal real de 44.100 mostres per segon.

4) Imagineu un senyal que s'ha mostrejat a 44.000 Hz i que es vol reproduir a una freqüència de mostreig de 44.100 Hz. Es demana:

a) Si el senyal original té una durada de 50 minuts. Quant durarà el senyal reproduït a 44.100 Hz?

b) Supposeu que el senyal original conté un to de 800 Hz. Quina freqüència percebrem en reproduir-lo a 44.100 Hz?

Solució

a) El nombre total de mostres que tenim és de $44.000 \times 50 \times 60 = 132.000.000$. Si es reproduïxen a una freqüència de 44.100 mostres per segon, tindrem una durada de 2.993,1 segons, és a dir, 49 minuts i 53,1 segons.

b) El senyal de 800 Hz tindrà un període de $44.000/800$ mostres = 55 mostres. Quan aquestes mostres es reproduïxen a una velocitat de 44.100 obtenim un període d'1,24717 ms, que representa una freqüència de 801,81 Hz.

5) Es vol aprofitar el fenomen de l'aliàsing per a visualitzar la deformació de la membrana d'un altaveu quan s'excita a freqüències de 150 Hz i amb alta potència. Per a això, es vol alentir l'observació de la deformació de la membrana utilitzant un estroboscopi. Volem observar la membrana vibrant a 0,5 Hz (és a dir, una oscil·lació completa cada 2 segons). Determineu quina haurà de ser la freqüència de disparament del flaix de l'estroboscopi.

3. Activitat 3. Digitalització de senyals d'àudio: quantificació

3.1. Introducció

3.1.1. Contingut

Aquesta activitat se centra en l'estudi i anàlisi dels efectes de la codificació de les mostres en un nombre finit de bits. Per a això es proposa que l'estudiant faci diferents audicions d'un senyal digitalitzat amb diferents qualitats (nombre de bits) i n'avaluï l'efecte en la qualitat subjectiva del nombre de bits. Al seu torn, es proposa que s'analitzin els efectes de la quantificació tant en la forma d'ona del senyal, comprovant l'aproximació de nivell que es produeix, com en els efectes sobre l'espectre. Aquesta anàlisi es fa utilitzant senyals prèviament mostrejats i amb l'ajuda d'un programari genèric d'edició i anàlisi de senyals d'àudio.

Una vegada s'han dut a terme aquestes experiències bàsiques es formalitzen els conceptes que s'han introduït de manera intuïtiva.

- En primer lloc s'introdueix el concepte d'error de quantificació i s'interpreta com un soroll que s'afegeix al senyal. Per a valorar el nivell de soroll d'una manera objectiva i numèrica, s'introdueix el concepte de relació senyal-soroll i es proporcionen fórmules que permetin calcular-lo.

L'objectiu principal d'aquesta part és que l'estudiant sigui capaç d'associar nombres d'SNR en decibels, que habitualment es descriuen en les característiques tècniques dels equips d'adquisició, amb la qualitat perceptiva de l'àudio que s'ha escoltat en la part experimental.

- Tot seguit es defineix la codificació del senyal digital en format PCM considerant la codificació en binari natural i la codificació en complement a 2. Aquesta part de l'activitat es deixa com a opcional i vol justificar per què la major part dels sistemes d'adquisició utilitzen el complement a 2 com a format de representació de les mostres.
- Posteriorment s'introdueix el concepte de taxa de bit i es proporcionen diversos exemples per calcular-lo en funció dels paràmetres de digitalització del senyal. Es dóna especial importància a relacionar la taxa de bit amb la capacitat d'emmagatzemament i la capacitat de difusió de contin-

guts mitjançant xarxes de dades. Finalment, es proporcionen alguns exemples de sistemes comercials i es comenten els avantatges o inconvenients d'utilitzar un nombre de bits més alt o més baix. També, com a conclusió d'aquesta activitat es presenten els avantatges principals dels sistemes d'àudio digital davant els analògics.

3.1.2. Metodologia

En primer lloc es proposa que l'estudiant experimenti amb la qualitat auditiva d'un mateix senyal d'àudio codificat amb diferents nivells de quantificació. L'objectiu és que de manera natural s'estableixi una associació entre el nombre de bits de codificació i la qualitat del senyal, i es comprovin de manera auditiva aquests efectes i la seva importància. També es vol que es comprovi com les capacitats del sistema auditiu tenen un límit, i es justifiqui que no té sentit augmentar la resolució dels equips tecnològics més enllà d'aquest límit.

Paral·lelament i amb ajuda d'un programari genèric de visualització, edició i processament de senyals d'àudio es veuen els efectes de la quantificació en la forma d'ona i en l'espectre. Després d'aquests experiments l'estudiant haurà de tenir una idea clara del nombre de bits que necessita per a una audició correcta i de la caracterització en temps i en freqüència del soroll de quantificació.

En una segona part de l'activitat es formalitzen els conceptes de soroll de quantificació, relació senyal-soroll, codificació PCM, taxa de bits i avantatges dels sistemes digitals davant els analògics. Tots aquests conceptes s'exposen de manera autocontinguda en el text de l'activitat. A més, en la secció de problemes es proposen experiències alternatives que fomenten l'ús de bibliografia addicional o la cerca d'informació a Internet sobre alguns productes comercials o sistemes d'àudio. Al llarg del text s'aprofundeix en alguns conceptes específics, que sense ser fonamentals poden tenir cert interès per als estudiants més motivats des d'un punt de vista tecnològic. Aquests conceptes s'indiquen clarament com a opcionals encara que s'anima els estudiants a intentar seguir-los.

Finalment, es proposen diversos problemes semblants als que han aparegut com a exemples al llarg del text i s'inclouen les solucions.

3.1.3. Recursos

Els recursos principals utilitzats en aquesta activitat són:

- Arxius d'àudio predigitalitzats amb diferents resolucions (nombre de bits).
- Programari de captura i processament d'àudio Audacity (llicència lliure).
- Anàlisi de forma d'ona i anàlisi de freqüència dels senyals.

- Explicacions dels fonaments de la quantificació, el soroll, la relació senyal-soroll, la codificació PCM i la taxa de bits dels sistemes d'àudio digital.
- Proposta d'activitats de cerca d'informació en webs i recursos d'Internet per a aprofundir en els coneixements adquirits en aquesta activitat (webs de fabricants de targetes de captura, *wiki*, etc.).
- Problemes proposats per a verificar l'adquisició dels coneixements d'aquesta activitat.

3.2. Objectius

Els objectius principals que es persegueixen en aquesta activitat són:

- 1) Interpretar la funció dels dos elements clau en el procés de digitalització d'un senyal, la freqüència de mostreig i el pas de quantificació, aprofundint en aquest últim.
- 2) Relacionar, mitjançant experiments dirigits, la qualitat subjectiva d'un senyal d'àudio amb el pas de quantificació.
- 3) Relacionar el pas de quantificació amb el nombre de bits del quantificador.
- 4) Conèixer les característiques espectrals del soroll de quantificació.
- 5) Comprendre els efectes de la quantificació com una aproximació al senyal original que hi introdueix soroll.
- 6) Introduir el concepte de relació senyal-soroll i saber relacionar els valors obtinguts amb la qualitat subjectiva.
- 7) Comprendre el procediment de codificació d'un senyal en el format PCM (*pulse code modulation*).
- 8) Conèixer els avantatges de la codificació en complement a dos respecte a la codificació en binari convencional per a l'assignació de codis al senyal d'àudio.
- 9) Conèixer la definició de taxa de bits per a un sistema digital i en concret per als sistemes d'àudio digital.
- 10) Distingir entre els conceptes de quantificador, codificador i convertidor A/D.
- 11) Saber calcular la taxa de bits d'un sistema d'àudio digital en funció de les característiques dels paràmetres de mostreig.

12) Saber determinar l'amplada de banda i les capacitats d'emmagatzemament de sistemes per a la difusió i emmagatzemament d'àudio digital.

13) Conèixer les resolucions i nombre de bits que s'utilitzen en els sistemes d'àudio comercials més habituals.

14) Comprendre la necessitat de comprimir la informació PCM per a optimitzar la transmissió i l'emmagatzemament del senyal d'àudio digital.

15) Conèixer les tecnologies de captura i digitalització i els equips comercials que poden utilitzar-se en entorns d'edició tipus PC.

16) Conèixer els avantatges tecnològics principals dels sistemes digitals davant els analògics.

Aquests objectius estan relacionats amb les **competències de l'assignatura** següents:

- 1) Capacitat de capturar àudio i imatges de manera eficient i eficaç.
- 2) Capacitat de digitalitzar eficientment i eficaçment continguts d'àudio i imatge.
- 3) Capacitat d'interpretar i modificar informació auditiva de manera freqüencial.

I amb les **competències generals del grau** següents:

- 1) Capturar, emmagatzemar i modificar informació d'àudio, imatge i vídeo digitals aplicant principis i mètodes de realització i composició del llenguatge audiovisual.
- 2) Integrar i gestionar continguts digitals en aplicacions multimodals d'acord amb criteris estètics, tècnics i funcionals.
- 3) Atendre adequadament consultes sobre projectes, tecnologies i mercat de productes multimèdia avaluant de manera precisa l'entorn d'aplicació, els recursos i les alternatives tecnològiques disponibles.

3.3. Guia

Generalment els fenòmens físics naturals es manifesten com a **senyals analògics**.

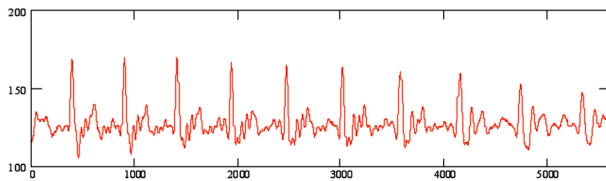
Podem esmentar un gran nombre d'exemples en què el fenomen físic es pot representar com una funció matemàtica, és a dir, com un senyal analògic.

Exemple

Entre aquests exemples: l'evolució de la temperatura en un punt d'observació, el cabal d'un riu, el creixement d'una planta, les variacions de pressió que formen les ones acústiques o la tensió recollida per un micròfon com a conseqüència d'aquestes ones acústiques. El senyal d'àudio és, doncs, en el seu origen, analògic.

En la figura 50 hem representat de manera esquemàtica el senyal de tensió recollit per un micròfon i que correspon a una vocal (la *o*).

Figura 50. Evolució en el temps de la tensió recollida per un micròfon en pronunciar la *o*



El senyal és analògic, és a dir, és continu en el temps i en amplitud.

Digitalitzar aquest senyal analògic és un procés d'aproximació que, com ja hem vist en l'apartat anterior, utilitza una **retícula de mostreig**. La qualitat d'aquesta aproximació depèn de la densitat de la retícula de mostreig tant en l'eix horitzontal (freqüència de mostreig) com en l'eix vertical (pas de quantificació).

En l'apartat anterior hem considerat els efectes de la **freqüència de mostreig** en el procés de digitalització del senyal. Per a valorar la incidència de la freqüència de mostreig s'ha suposat que el pas de quantificació és prou petit per a poder-ne menysprear els efectes en la pèrdua de qualitat del senyal.

Ara farem el procés invers: per analitzar els efectes del pas de quantificació suposarem que la freqüència de mostreig és prou alta i que el seu valor no condiciona la qualitat final del senyal en format digital.

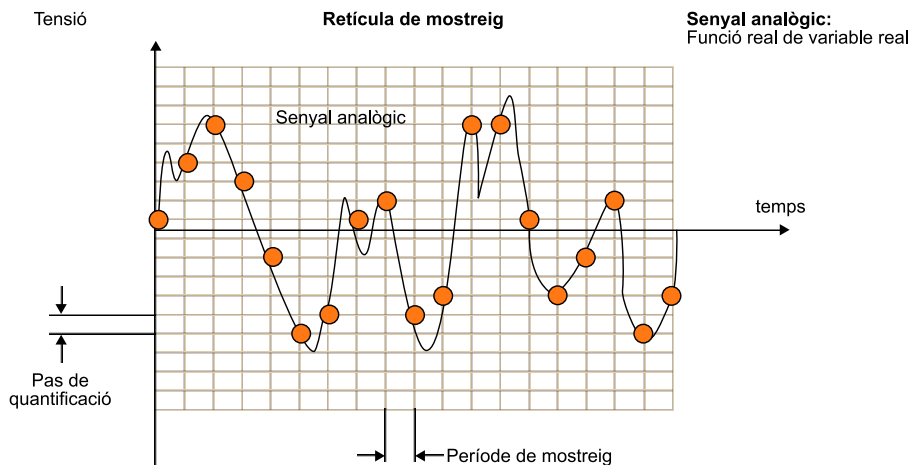
En aquesta activitat farem diferents proves auditives per valorar els efectes del pas de quantificació en el senyal digital. Amb això, volem determinar els criteris per a fixar aquest paràmetre en funció de l'aplicació i de la qualitat final que vulguem obtenir. Posteriorment descriurem el format d'àudio digital no comprimit PCM, analitzant-ne els avantatges i els inconvenients.

3.3.1. La retícula de mostreig i el pas de quantificació

En l'activitat anterior s'ha introduït el concepte de retícula de mostreig com una estratègia per a aproximar un senyal analògic mitjançant un conjunt numerable de punts. En la figura 34 es reproduïx aquesta retícula de mostreig en la qual definim la freqüència de mostreig i el pas de quantificació com la densitat de la retícula en els eixos horitzontal i vertical respectivament.

Es pot entendre intuïtivament que com més densa sigui aquesta retícula millor serà l'aproximació digital al senyal analògic. No obstant això, també és clar que augmentar la densitat de la retícula farà augmentar el volum d'informació associat al senyal digital, la qual cosa n'afectarà l'emmagatzemament i l'eficiència de la transmissió. Per tant, és molt important que la selecció d'aquests paràmetres es faci atenent els criteris de qualitat requerits per a una determinada aplicació. En efecte, no té sentit augmentar la freqüència de mostreig o disminuir el pas de quantificació més enllà del que pot percebre l'oïda humana.

Figura 51. Aproximació d'un senyal analògic mitjançant una retícula de mostreig



La retícula de mostreig es caracteritza pel pas horitzontal (període de mostreig) i el pas vertical (pas de quantificació).

En conseqüència, el sistema auditiu es converteix en el jutge últim que determina quina és la densitat adequada de la retícula de mostreig en cada eix. Tanmateix, com veurem, aquesta decisió no resulta trivial, ja que no tots els sistemes auditius són igual de crítics i, a més, és possible que el conjunt de proves que es fan per a fixar els paràmetres no tingui en compte alguns casos especials en què sí que es podrien apreciar diferències entre l'original analògic i l'aproximació digital.

Una de les característiques que hem estudiat en l'activitat anterior és el **teorema de Nyquist**, que estableix un límit teòric per a la freqüència de mostreig. Segons aquest teorema, si un senyal es mostra a una freqüència superior al doble de la seva amplada de banda, llavors el senyal analògic es pot recuperar exactament a partir de les seves mostres.

En altres paraules, no es produeixen pèrdues pel fet de mostrejar el senyal sempre que es compleixi el criteri de Nyquist.

Lamentablement, no passa el mateix en l'eix vertical. Quan fixem un nombre finit de possibles nivells, el valor original del senyal ja no es pot recuperar. En la figura 34 es veu que a partir de les mostres digitals (punts) només podrem obtenir una versió aproximada del senyal analògic original.

Per a analitzar l'efecte del pas de quantificació en la qualitat de senyal d'àudio és convenient fer proves auditives de qualitat. A la pràctica, el pas de quantificació es tria a partir de **proves subjectives** adreçades a múltiples usuaris.

Aquestes proves consisteixen en l'audició de diversos fragments musicals, quantificats amb diferents passos i que s'han de comparar amb el senyal analògic original.

El pas de quantificació escollit haurà de ser aquell per al qual una gran majoria dels usuaris (si és possible tots) no puguin distingir entre el senyal original i el senyal digitalitzat.

Evidentment, les proves s'han de fer amb una freqüència de mostreig prou alta perquè aquesta no afecti la qualitat del senyal.

En aquesta activitat reproduïrem de manera simplificada aquestes proves subjectives per veure l'efecte del pas de quantificació en la qualitat final del senyal.

Prendrem com a senyal de referència el senyal original d'un enregistrament per a CD d'àudio, que té una freqüència de mostreig de 44.100 Hz i un total de 65.536 passos de quantificació. En realitat es tracta d'un senyal digital, però si acceptem que les proves per a la definició del format CD d'àudio van ser correctes, llavors no hauria de ser possible apreciar diferències entre aquest senyal digital i l'original analògic⁹. L'arxiu original amb què treballarem és **NessunDorma.wav** i s'adjunta amb el material addicional d'aquesta activitat.

⁽⁹⁾Generalment, per abús del llenguatge, fem referència a un **senyal original analògic** quan en molts enregistraments el senyal original que pren en el *master* ja és digital, per la qual cosa no hi ha una versió analògica. No obstant això, en aquest text, entendrem per *senyal original analògic* el que escoltaríem en condicions ideals sense haver fet cap aproximació de digitalització.

1) Obriu aquest arxiu amb el programa Audacity i escolteu-lo. La qualitat és bona, ja que es correspon amb les mostres originals del CD a 44.100 Hz i utilitzant un total de 65.536 passos de quantificació¹⁰, és a dir, 16 bits.

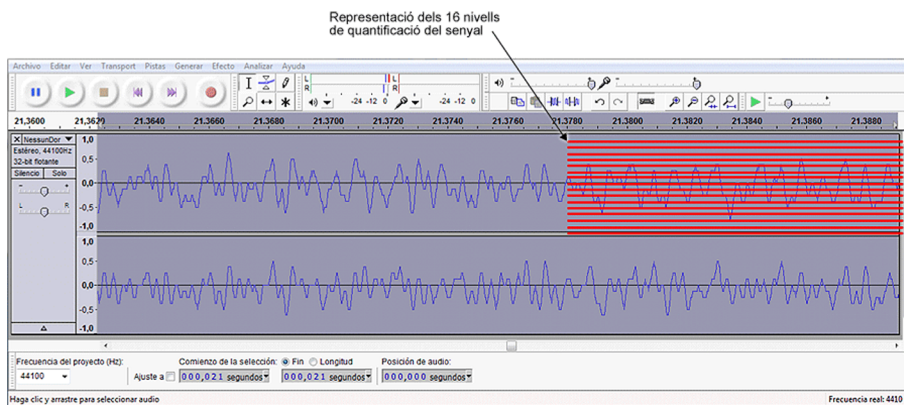
⁽¹⁰⁾El nombre de passos de quantificació sol ser sempre una potència de 2 perquè, com veurem més endavant, el nivell de quantificació es codifica mitjançant una paraula d'un nombre fix de bits. Així, amb 16 bits podem codificar fins a 65.532 nivells = 2^{16} nivells mentre que amb 15 bits podem codificar 2^{15} nivells = 32.768. Fixeu-vos que no té gaire sentit escollir un nombre de nivells de quantificació igual a 50.000, ja que per a codificar-los necessitarem utilitzar 16 bits per mostra i, per aquest mateix cost, en podem codificar fins a 65.532!

2) Obriu també l'arxiu **NessunDorma_4bit.wav**. Aquest arxiu s'ha obtingut a partir de l'original però modificant el nombre total de nivells de quantificació a 16 (4 bits – 2^4 nivells).

3) Escolteu atentament aquest arxiu i comproveu que l'efecte és que s'ha sumat un soroll considerable sobre el senyal original. Aquest soroll es denomina **soroll de quantificació** i es produeix perquè el pas de la retícula de mostreig en l'eix vertical és insuficient per a aproximar correctament el senyal. És indubtable que l'oïda humana és capaç de discriminar entre el senyal original i una vasta quantificació d'aquest que utilitza només 16 nivells. Veurem que a mesura que augmentem el nombre de nivells de quantificació aquest soroll disminuirà i arribarà un punt en què l'oïda serà incapaç de discernir entre el senyal original i l'aproximat.

4) Per veure amb més claredat l'efecte de la quantificació seleccioneu un fragment de curta durada del senyal de 4 bits i amplieu la finestra de zoom fins que pugueu veure els detalls de la forma d'ona. En la figura 52 es mostra un fragment del senyal en què s'han marcat els 16 possibles nivells de quantificació que s'estan utilitzant. En aquest gràfic es veu que l'efecte de la quantificació és una espècie d'aproximació en escala, per mitjà de la retícula de mostreig, al senyal original.

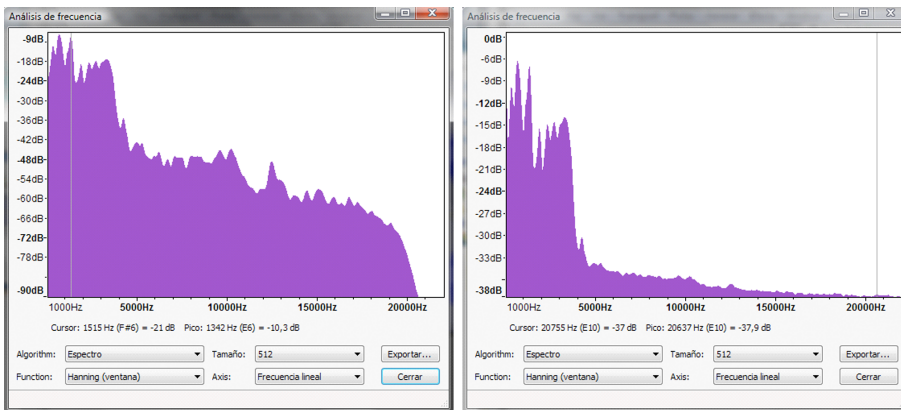
Figura 52. Representació de la forma d'ona d'un fragment del senyal NessunDorma quantificat a 4 bits



En vermell es mostren els 16 nivells de quantificació del senyal.

És molt interessant analitzar els senyals des del punt de vista espectral. Per a això, preneu un fragment de senyal relativament llarg (d'uns 10 o 20 segons) i compareu l'espectre del senyal original amb el del senyal quantificat amb 16 nivells. Analitzem un fragment relativament llarg perquè quedi la mitjana de diferents espectres i poder visualitzar el valor mitjà del soroll de quantificació. Aneu amb compte de seleccionar la mateixa finestra per als dos senyals. En la figura 52 mostrem els resultats que hem obtingut en el fragment que va des del segon 3 fins al 23.

Figura 53. Espectre d'un fragment del senyal NessunDorma original (16 bits –esquerra–) i el senyal quantificat (4 bits –dreta–)

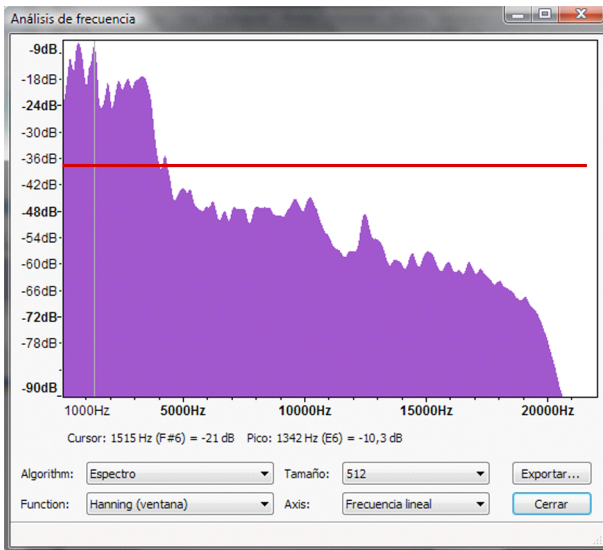


Els dos espectres s'han de comparar amb certa cura, observant que les escales dels eixos verticals són diferents i que l'aspecte general dels gràfics no ens hauria de confondre.

Essencialment, aquestes gràfiques ens mostren que els dos resultats són pràcticament idèntics en la regió de baixa freqüència, però que en la zona d'alta freqüència les diferències són significatives. Efectivament, fins als 4.000 Hz podem considerar, de manera aproximada, que els dos espectres tenen els mateixos components, en les mateixes freqüències i amb els mateixos nivells d'amplitud. A partir dels 4.500 Hz les diferències comencen a ser considerables, ja que per al senyal de 4 bits (16 nivells) predomina un component de soroll, d'espectre aproximadament pla i que se situa entorn de -38 dBs. Aquest component de soroll emmascara els components d'alta freqüència del senyal original que poden observar-se en la gràfica de l'esquerra.

En la figura 54 hem representat l'espectre del senyal amb 16 bits i una línia vermella que indica l'espectre del soroll. L'espectre del senyal de 4 bits pot considerar-se com la suma entre aquest soroll espectralment pla i l'espectre del senyal original. Observeu en la figura 53 que els nivells en la zona de baixa freqüència per a la gràfica de la dreta són una mica més grans que els de l'espectre original a causa de la suma del component de soroll.

Figura 54



Espectre del senyal quantificat a 16 bits (considerat original) amb la superposició del soroll (en vermell) d'espectre pla corresponent a la quantificació de 4 bits

Analitzem ara l'efecte sobre la qualitat del senyal i el seu espectre en augmentar o disminuir el nombre de passos de quantificació.

1) Obriu els arxius **NessunDorma_3bits.wav** i **NessunDorma_5bits.wav**, que corresponen a la quantificació del senyal amb 8 i 32 passos de quantificació respectivament, i compareu-los amb els resultats obtinguts amb la quantificació a 16 nivells.

2) Des del punt de vista auditiu resulta evident com el nivell de soroll disminueix progressivament a mesura que augmenta el nombre de bits del quantificador. En tots els casos la qualitat final és encara molt deficient, però la millora progressiva en augmentar els bits indica que amb alguns bits addicionals arribarem a obtenir qualitats acceptables.

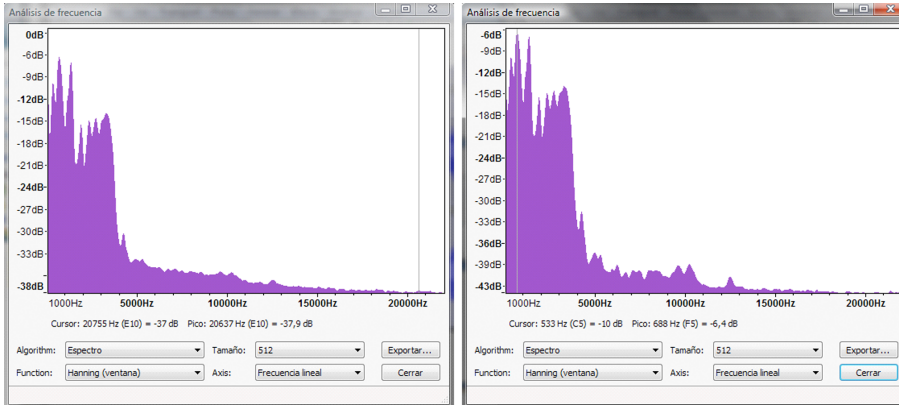
3) Respecte a l'anàlisi en freqüència dels senyals, és interessant observar com el nivell de soroll es redueix cada vegada que augmentem en 1 bit el quantificador.

Fixeu-vos que augmentar en 1 bit el nombre de passos de quantificació representa doblar el nombre total de nivells del quantificador. Així, en passar de 4 bits a 5 bits augmentem de 16 nivells a 32 nivells. Això significa que l'aproximació de la retícula de mostreig millora, ja que l'error de l'aproximació es reduirà a la meitat. La reducció de l'error en l'aproximació implica una reducció del soroll de quantificació que es fa evident en l'anàlisi espectral del senyal (i sens dubte, en l'anàlisi auditiu). En la figura 55 es compara l'espectre obtingut amb el senyal quantificat a 16 nivells (esquerra) i el quantificat a 32 nivells (dreta). Fixeu-vos en la reducció de nivell en l'espectre del soroll que va des d'uns -38 dBs en la gràfica de l'esquerra fins a una mica menys de -43 dBs

⁽¹¹⁾En aquestes gràfiques no resulta fàcil mesurar els nivells del soroll. Veurem més endavant que la diferència de nivell teòrica hauria de ser exactament de 6 dBs.

en la gràfica de la dreta¹¹. En l'apartat següent formalitzarem aquests resultats, proporcionant fórmules matemàtiques que ens relacionin directament la qualitat de l'audició amb el nombre de bits del quantificador.

Figura 55. Comparativa entre els espectres obtinguts amb un fragment de senyal quantificat amb 4 bits (esquerra) i el mateix fragment quantificat amb 5 bits (dreta)



Les parts de baixa freqüència són pràcticament iguals, però apareix una diferència de nivell de soroll entre els dos espectres d'aproximadament 6 dBs.

4) Analitzeu els espectres dels senyals quantificats a 6 bits, 7 bits i 8 bits i comproveu que el nivell de soroll es redueix progressivament en una quantitat constant. Recordeu que per a poder comparar els espectres és convenient que prengueu un fragment de senyal d'uns 10-20 segons i que sempre prengueu el mateix. Compareu també de manera auditiva la qualitat dels senyals obtinguts.

5) Escolteu els senyals quantificats amb 1 bit i 2 bits i analitzeu-ne els espectres. Tingueu en compte que en aquests dos senyals el soroll és molt important. Abaixeu el volum de l'altaveu o els auriculars per no fer malbé ni l'equip ni l'oïda. L'audició del senyal d'1 bit és autènticament espectacular. Fixeu-vos que la forma d'ona del senyal només pren dos possibles valors: positiu i negatiu. Tot i així, el sistema auditiu és capaç d'escoltar el passatge musical.

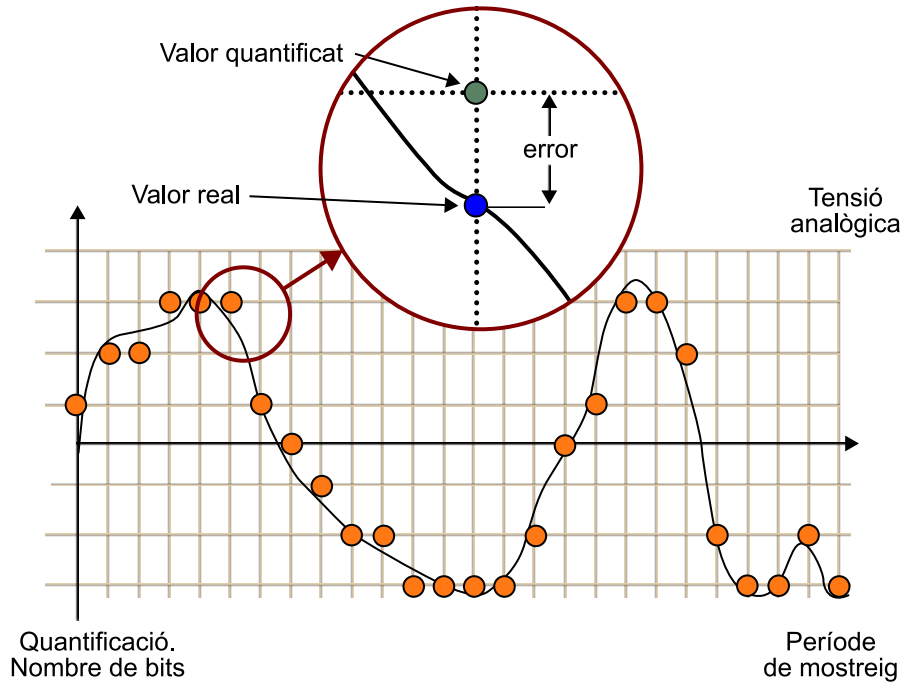
6) Augmenteu progressivament el nombre de bits: 9, 10, etc. i compareu auditivament aquests senyals amb el de 16 bits. Determineu el nombre mínim de bits amb els quals sou incapaços d'apreciar diferències amb el senyal del CD original (NessunDorma.wav).

3.3.2. Soroll de quantificació i relació senyal-soroll

En aproximar el senyal analògic mitjançant una retícula de mostreig s'introdueix un error en el nivell d'amplitud del senyal. En la figura 56 es mostra com el valor real del senyal en l'instant de mostreig ha de ser aproximat per un dels passos de quantificació existents en la retícula de mostreig. L'error que es produeix depèn del nivell del senyal. De vegades, és possible que l'error sigui molt petit perquè el senyal passa molt pròxim a un nivell de quantificació, mentre que en altres casos l'error pot ser més important. No obstant això, sí que està clar que l'error màxim mai no superarà la meitat de l'amplitud del

pas de quantificació. Efectivament, el nivell de quantificació més pròxim al senyal estarà sempre, com a molt lluny, a la meitat del pas de quantificació, ja que sempre podem aproximar el senyal pel nivell de quantificació més pròxim per dalt o per sota.

Figura 56. Representació gràfica de l'error que es produeix en aproximar el senyal per un nivell de quantificació.



Com que els nivells de senyal no es coneixen *a priori* podem assumir que l'error serà una variable aleatòria els valors de la qual estan compresos entre:

$$-\frac{\Delta}{2} \leq \text{error} \leq \frac{\Delta}{2}$$

en què Δ representa el pas de quantificació.

A més, també sembla raonable suposar que tots els valors d'error poden produir-se amb la mateixa probabilitat.

El pas de quantificació està relacionat amb el nombre de bits (nombre de nivells de quantificació) i els valors màxim i mínim entre els quals es defineix la retícula de mostreig.

En efecte, observem que en la figura 34 hem representat una retícula de mostreig que delimitava les possibles amplituds del senyal entre uns valors màxim i mínim. D'aquesta manera, el nombre de nivells de quantificació serà finit. D'altra banda, aquesta restricció no és important, ja que a la pràctica els nivells

del senyal sempre estaran delimitats pel mateix equip de mesura o captura (micròfon, amplificador), que segurament se saturarà si el fenomen físic (volum acústic) supera certs llindars.

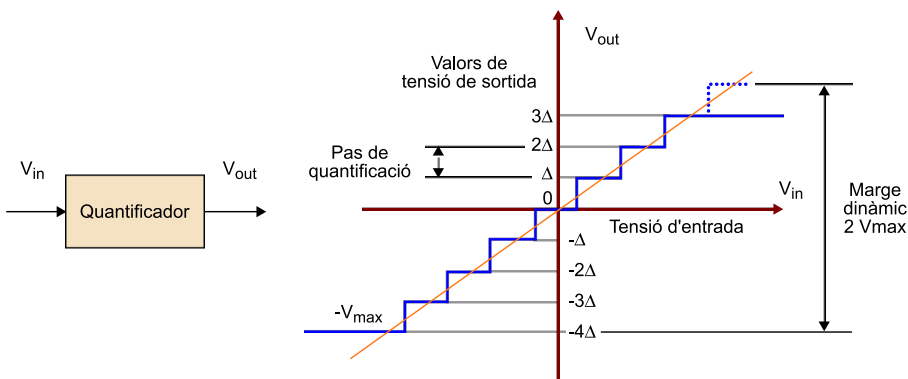
Una manera alternativa de representar la retícula de mostreig és definint els **quantificadors**.

Un quantificador és un sistema que aproxima el nivell del senyal d'entrada en un nombre finit de possibles nivells de sortida.

En la figura 58 es representa un quantificador de 8 nivells (3 bits). La tensió d'entrada es representa en l'eix horitzontal i pot prendre qualsevol valor. La tensió a la sortida¹² només pot prendre un dels 8 possibles nivells de quantificació indicats en la gràfica.

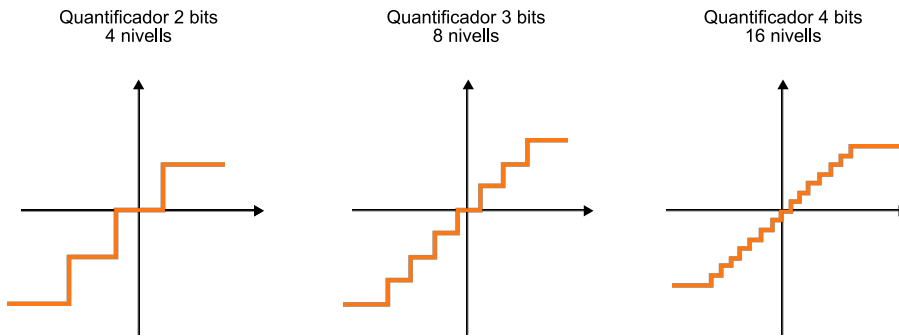
⁽¹²⁾En la majoria de sistemes pràctics els nivells de sortida no són simètrics respecte al 0. En la figura 58 es pot comprovar que el nivell mínim arriba fins a la tensió $-V_{ref}$ però que el valor màxim no arriba a assolir $+V_{ref}$. Descriurem amb més profunditat aquesta asimetria més tard, encara que de moment val la pena comentar que quan el nombre de bits és prou elevat els nivells màxim i mínim es poden considerar aproximadament simètrics.

Figura 58. Relació entrada-sortida d'un quantificador de 8 nivells (3 bits)



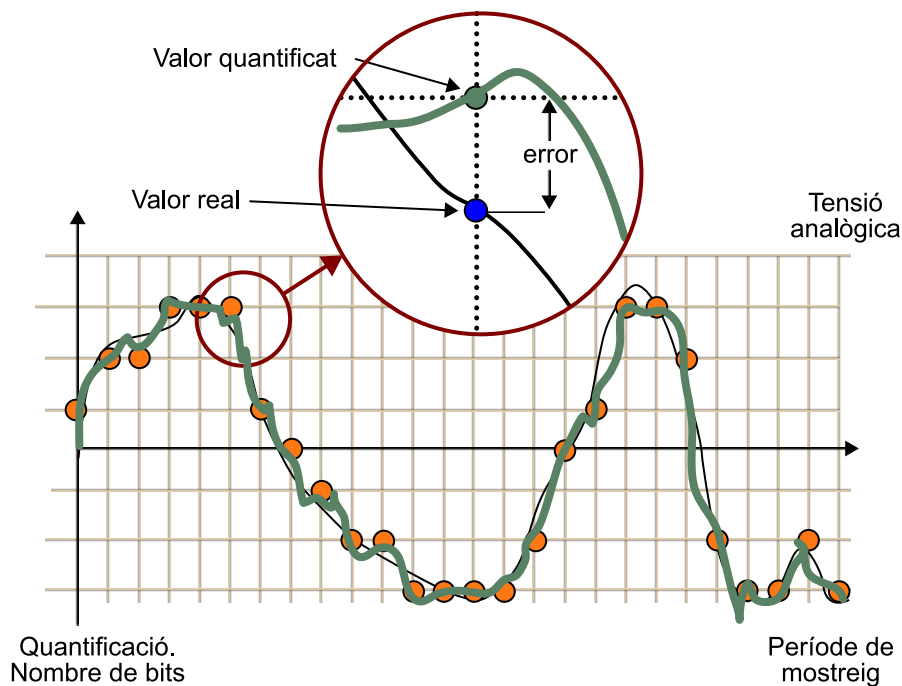
En la figura 59 es mostren alguns exemples de quantificadors per a diferents quantitats de bits. S'observa que en augmentar el nombre de bits tindrem més nivells de quantificació i, per tant, l'aproximació entre el senyal d'entrada i el de sortida serà millor.

Figura 59. Corbes d'entrada-sortida per a quantificadors de 2 bits, 3 bits i 4 bits



Hem comprovat que l'error de quantificació és equivalent a un soroll i que és per això que moltes vegades es coneix amb el nom de **soroll de quantificació**. En la figura 60 s'intenta representar aquesta equivalència entre l'error i el soroll. En efecte, el procés de quantificació pot interpretar-se com que s'ha sumat un soroll al senyal i que aquest soroll és el que ha provocat que la mostra prengui un valor de tensió que no era l'original (sinó l'aproximació de la nostra retícula de mostreig).

Figura 60. Interpretació de l'error de quantificació com un soroll que se suma al senyal (soroll de quantificació).



3.3.3. La relació senyal-soroll

Per a valorar des d'un punt de vista quantitatiu la qualitat d'un sistema d'àudio se sol utilitzar la relació senyal-soroll (SNR: *signal to noise ratio*), que es defineix com el quocient entre la potència del senyal i la potència de soroll i sol expressar-se en decibels mitjançant una relació logarítmica.

Així tenim:

$$SNR = 10 \cdot \log_{10} \frac{P_{signal}}{P_{noise}}$$

Com podem veure en aquesta fórmula, la relació senyal-soroll representa el quocient entre la potència del senyal i la potència del soroll. La qualitat perceptiva de l'àudio està relacionada amb aquest quocient. Evidentment, com menor sigui la potència de soroll millor serà la qualitat perceptiva. Al seu torn, si la potència de soroll disminueix, el resultat del quocient augmenta, proporcionant uns valors numèrics d'SNR més alts.

La conclusió és que com és alts són els valors numèrics d'SNR millor és la qualitat perceptiva del senyal d'àudio.

D'altra banda, pot semblar una mica estrany que es prengui el logaritme del quocient de potències en comptes del quocient de potències pròpiament dit. Això es deu, en part, al comportament logarítmic de l'oïda humana quant a l'apreciació de potències de soroll. Cada vegada que la potència de soroll es redueix a la meitat apreciem una reducció progressiva d'aquest, si la tornem a reduir a la meitat tornem a apreciar la mateixa reducció, etc. En altres paraules, en prendre el logaritme del quocient de potències, els nombres d'SNR que obtenim expressen en una escala aproximadament lineal la qualitat perceptiva del senyal.

Quan es pren el logaritme del quocient de potències i es multiplica per 10, les unitats obtingudes reben el nom de **decibels**.

En àudio digital la relació senyal-soroll només es pot calcular mitjançant aproximacions genèriques, ja que la potència del senyal depèn del mateix senyal i possiblement canviarà al llarg d'un enregistrament en funció del contingut musical. En efecte, en passatges de baix volum o en *pianissimo*, la potència de senyal serà molt més baixa que en els *fortissimos*. Per tant, a la pràctica, el que se sol prendre com a potència de senyal en l'expressió de l'SNR és un valor mitjà, vàlid per a una àmplia gamma de senyals musicals.

En canvi, la potència de soroll és més fàcil de calcular, ja que coneixem les característiques d'error en l'aproximació del senyal i sabem que la seva amplitud depèn del pas de quantificació, que al seu torn depèn del nombre de bits del quantificador.

A partir de la figura 58 podem establir la relació següent entre el pas de quantificació, els límits de la retícula (tensió màxima) i el nombre de bits:

$$\Delta = \frac{2 \cdot V_{\max}}{2^N}$$

En què V_{\max} representa les tensions límit del quantificador (retícula) i N el nombre de bits del quantificador. L'expressió s'obté dividint l'excursió màxima dels nivells d'entrada $V_{\max} - (-V_{\max}) = 2V_{\max}$ pel nombre total de nivells de quantificació 2^N .

A partir d'aquesta última expressió queda clar que en augmentar el nombre de bits del quantificador es redueix el pas de quantificació Δ i, per tant, es redueix l'amplitud màxima i la potència del soroll.

En concret, l'amplitud del soroll es reduirà en la meitat cada vegada que augmentem el quantificador en 1 bit.

Aquest resultat és molt important i s'ha de recordar. De fet, és un resultat molt intuïtiu al qual podem arribar sense necessitat de fórmules. En efecte, quan el quantificador s'augmenta en 1 bit, el nombre de nivells es dobla, per la qual cosa l'amplitud màxima de l'error de quantificació es redueix a la meitat.

Si l'amplitud del soroll es redueix a la meitat la potència (que es calcula com la mitjana de les amplituds al quadrat) es reduirà en una quarta part. Podem expressar aquest resultat matemàticament com a:

$$\begin{aligned} SNR_{N+1bit} &= 10 \cdot \log_{10} \frac{P_{\text{signal}}}{P_{\text{Noise } N+1}} = 10 \cdot \log_{10} \frac{P_{\text{signal}}}{(P_{\text{Noise } N}/4)} = \\ &= 10 \cdot \log_{10} \frac{P_{\text{signal}}}{P_{\text{Noise } N}} + 10 \cdot \log_{10} 4 = SNR_{Nbit} + 6,02dB \end{aligned}$$

En definitiva, en augmentar en 1 bit el quantificador, la relació senyal-soroll millora en 6 dBs.

A la pràctica es pot arribar a obtenir una expressió genèrica per a la relació senyal-soroll com una fórmula simple que només depèn del nombre de bits del quantificador:

$$SNR = A + 6.02 \cdot N$$

En què el valor de la constant A depèn de la potència del senyal.

Generalment, es pren un valor d' A **igual a 1,76**, que es correspon amb la potència màxima del component de senyal. D'aquesta manera, l'equació anterior es transforma en el següent:

$$SNR = 1.76 + 6.02 \cdot N$$

Problemes

Vegem alguns exemples utilitzant aquesta fórmula.

1) Determineu la relació senyal-soroll del sistema CD d'àudio i calculeu la potència del soroll de quantificació d'aquest sistema com una fracció de la potència de senyal.

Solució

Aprofitant que el nombre de bits del sistema CD d'àudio és de 16 i utilitzant l'equació anterior obtenim:

$$SNR = 1.76 + 6.02 \cdot 16 = 98 \text{ dBs}$$

Amb 98 dBs la relació entre la potència de senyal i la potència de soroll de quantificació pot expressar-se mitjançant:

$$98 = 10 \cdot \log_{10} \frac{P_{\text{signal}}}{P_{\text{Noise}}} \Rightarrow P_{\text{Noise}} = \frac{P_{\text{Signal}}}{10^{9.8}} = 1,58 \cdot 10^{-10} P_{\text{signal}}$$

2) Determineu la relació senyal-soroll d'un sistema de telefonia digital que utilitza 8 bits per mostra.

Solució

Particularitzant la fórmula anterior per a 8 bits obtenim que la relació senyal-soroll d'un sistema de telefonia digital és de 49,7 dBs.

3.3.4. Codificació PCM (*pulse code modulation*) del senyal d'àudio

Hem vist que digitalitzar un senyal analògic consisteix, essencialment, a aproximar-lo mitjançant una retícula de mostreig i hem analitzat les limitacions que introdueixen la freqüència de mostreig i el pas de quantificació en la qualitat d'aquesta aproximació. No obstant això, encara ens queda per definir com es construeix el senyal digital, és a dir, la seqüència de bits que representa la informació del senyal.

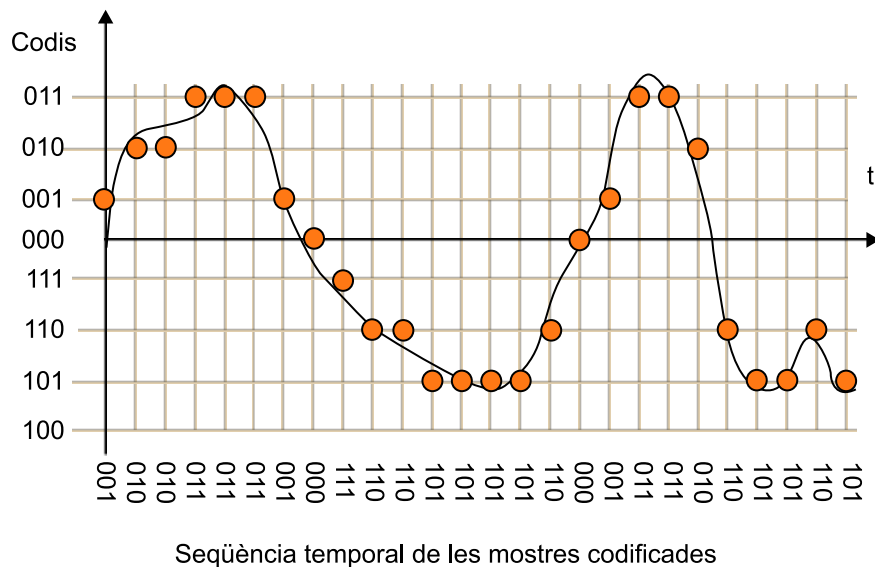
El procés més directe per a convertir el senyal en una seqüència de bits és assignar un codi binari a cada nivell de quantificació. D'aquesta manera, per a indicar el valor de cada mostra n'hi haurà prou de proporcionar la paraula binària associada.

Aquest procediment s'il·lustra de manera esquemàtica en la figura 61, en la qual es mostra un quantificador de 8 nivells en què cada nivell té assignada una paraula codi de 3 bits diferent. Per a transmetre o emmagatzemar la informació associada al senyal n'hi haurà prou a enviar o emmagatzemar de manera seqüencial les paraules codi de cada mostra. En la figura 61 es mostra una seqüència d'exemple –001 010 010 011 011 011 001 000 111 110 110 101 101 101...– que es correspondria amb l'aproximació a la forma d'ona representada.

PCM

Aquest tipus de codificació en què associem una paraula binària de longitud fixa a cada mostra es coneix amb el nom de **PCM** (*pulse code modulation*) i va ser patentat el 1937 per Alec H. Reeves.

Figura 61. Exemple de la codificació PCM d'un senyal analògic com una seqüència de bits



001 010 010 011 011 011 001 000 111 110 110 101 101 101 ...

Malgrat la seva simplicitat, aquesta codificació és, sens dubte, la més utilitzada. Entre altres raons, perquè és fàcilment realitzable en maquinari, tant en aplicacions d'àudio com en aplicacions d'imatge o vídeo. No requereix cap mena de processament especial.

En imatges digitals els formats *bitmap* i RAW són diferents versions de la codificació PCM. En vídeo, els formats 4:2:2 i 4:2:0 estandarditzats en la recomanació ITU-601 també són una codificació PCM. Però el format PCM també es troba present en molts equips que utilitzen com a base altres codificadors per emmagatzemar o transmetre la informació. Així, una gravadora o reproductor portàtil en MP3 captura la informació d'àudio en PCM i posteriorment la processa, per comprimir-la i emmagatzemar-la de manera més eficient en un format comprimit com MP3, MP4 o WMA. Al seu torn, durant la reproducció, la informació codificada en format comprimit es processa i es converteix a PCM abans d'aplicar-la als circuits electrònics que s'encarregaran de convertir-la a analògica i d'ampliar-la.

El maquinari du a terme la codificació PCM mitjançant **convertidors A/D** (analògic/digital). Un convertidor A/D és un circuit integrat que s'encarrega de la digitalització dels senyals i que sol incloure el subsistema de mostreig i quantificació del senyal.

Exemple

Podem trobar exemples d'aquest tipus de codificació en el format WAV d'àudio, el *disc* compacte d'àudio, el DAT (*digital audio tape*) i l'enregistrament d'àudio PCM que hi ha en molts DVD de vídeo.

En essència, de manera conceptual i molt simplificada un convertidor A/D seria un circuit que té com a entrades el senyal analògic i la freqüència de mostreig i que a la seva sortida proporciona la seqüència de bits binària associada a la codificació PCM del senyal.

Un element que generalment està associat al convertidor A/D és el convertidor D/A, que fa la funció inversa, és a dir, converteix la seqüència de bits PCM en els valors de tensió del quantificador. Generalment el convertidor D/A també incorpora un filtre passabaix per a passar al senyal analògic que s'aplicarà als amplificadors.

És important distingir entre els tres conceptes, molt pròxims, que han aparegut en aquest apartat: el **quantificador**, el **codificador** i el **convertidor A/D**.

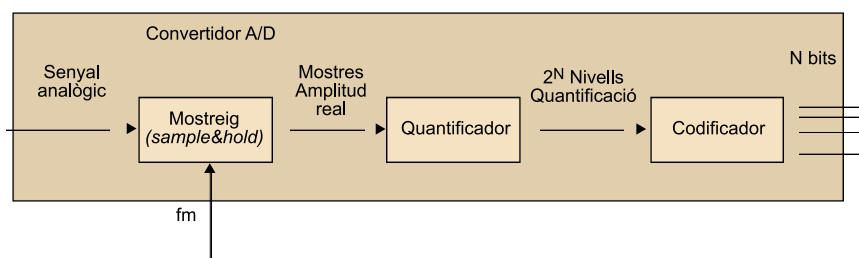
- El **quantificador** aproxima la tensió del senyal d'entrada per un nivell escollit entre un nombre finit de possibilitats. Un quantificador de 16 nivells només admet 16 possibles valors de tensió a la sortida.
- El **codificador** (PCM) associa una paraula binària a cada un d'aquests nivells finits del quantificador. En cas que el quantificador sigui de 16 nivells s'assignarà una paraula de 4 bits a cada un dels nivells¹³.
- El **convertidor A/D** és una realització pràctica d'un sistema de digitalització que converteix el senyal analògic en una seqüència de bits codificada en PCM. Internament el convertidor aproxima el senyal d'entrada a un dels nivells de quantificació, però a la sortida ens proporciona directament el codi PCM associat. Si connectem un convertidor D/A¹⁴ directament a la sortida d'un convertidor A/D obtindrem els nivells de quantificació als quals s'ha aproximat el senyal d'entrada.

⁽¹³⁾ Generalment el nombre de nivells del quantificador és una potència de 2, ja que el codificador utilitza paraules codi amb un nombre de bits fix. En efecte, es poden associar fins a 2^N nivells de quantificació a una paraula d' N bits.

⁽¹⁴⁾ Molts circuits integrats D/A inclouen un filtre analògic passabaix per reconstruir el senyal analògic, per tant, en aquests casos no és possible obtenir els nivells de quantificació originals.

La figura 62 intenta relacionar de manera esquemàtica aquests elements. El subsistema de mostreig s'identifica amb els circuits de mostreig i manteniment que s'encarreguen de congelar el nivell de tensió del senyal d'entrada en l'instant de mostreig.

Figura 62



Representació gràfica de la relació entre el quantificador, el codificador i el convertidor A/D

A la pràctica, la seqüència de bits associada a un senyal PCM requereix algun tipus de senyalització addicional que ajudi a interpretar quan comença o acaba un paquet de mostres. Aquesta senyalització depèn de l'aplicació i pot variar considerablement en funció d'aquesta.

- Un exemple seria el format CD d'àudio en què les mostres PCM s'estructuren en paquets de 33 bytes. Aquests 33 bytes contenen un total de 6 mostres PCM de cada un dels dos canals (6 mostres x 2 canals x 2 bytes/mostra = 24 bytes) més 9 bytes addicionals que s'utilitzen per a sincronismes, subcodis de control i de *display* i bits de redundància per a la correcció d'eventuals errors.
- Un altre exemple són els formats WAV, AUD, AIFF, etc., que especifiquen les capçaleres i les agrupacions de les mostres PCM en fitxers d'àudio.

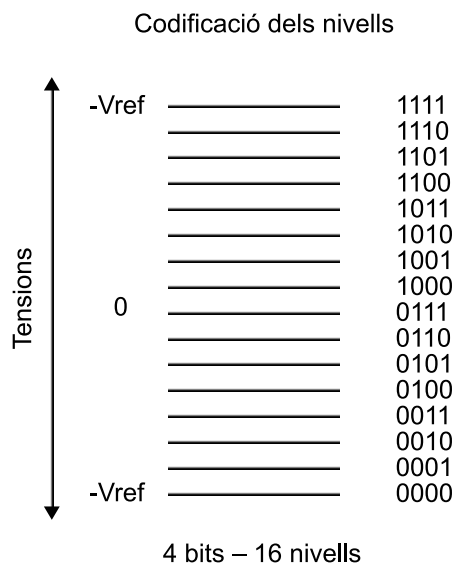
3.3.5. La selecció de les paraules codi en PCM (opcional)

Hem vist que un codificador PCM assigna una paraula codi d' N bits a cada un dels 2^N nivells del quantificador. En principi, la paraula codi assignada a cada nivell pot ser arbitrària sempre que cada nivell tingui una i només una paraula codi associada o, el que és el mateix, es pugui establir una **relació biunívoca** entre paraules codi i nivells de quantificació.

La pregunta que ens formulem en aquest apartat és si hi ha un criteri per a l'assignació de les paraules codi que resulti més convenient que d'altres.

Considerarem com a possible assignació de paraules codi la que es mostra en la figura 63 i que consisteix a numerar consecutivament cada un dels nivells, des del més baix (paraula només amb zeros) fins al més alt (paraula només amb uns). Aquesta sembla una assignació de paraules codi bastant natural i que no manca de sentit pràctic, ja que pot establir-se una fórmula senzilla que relacioni el nivell de quantificació amb el nombre binari.

Figura 63



Assignació lineal dels codis binaris als nivells de quantificació.
Exemple amb 16 nivells/4 bits

En efecte, la fórmula que relaciona el nivell de quantificació amb la paraula binària depèn del següent:

$$V_{in} \cong 2V_{REF} (b_1 2^{-1} + b_2 2^{-2} + b_3 2^{-3} + b_4 2^{-4}) - V_{REF}$$

En què resulta fàcil comprovar que la paraula només amb zeros quedarà assignada amb la tensió $-V_{REF}$.

La paraula només amb uns quedarà assignada al **nivell de quantificació màxim**, que en aquest cas és $(7/8)V_{REF}$, com es pot comprovar si se substitueixen tots els bits per 1 en l'equació anterior.

Encara que aquesta assignació sembla força simple, directa i fins i tot pràctica pateix d'un problema fonamental, i és que no es poden aplicar operacions aritmètiques directament sobre els valors codificats. Què volem dir amb això? Y, per què és tan important?

Quan digitalitzem un senyal és habitual que hi vulguem fer operacions digitals. Aquestes operacions poden ser un filtratge, una equalització, efectes sonors, etc. En qualsevol cas, les mostres s'han de multiplicar, sumar, etc., per la qual cosa seria desitjable que els resultats d'aplicar les operacions als nivells de quantificació i a les paraules codi fossin les mateixes.

Posem un exemple perquè quedi més clar:

Amb el sistema de codificació lineal anterior, si tenim la mostra 0111 i la mostra 1000, el resultat de la suma dels codis binaris és 1111. Fins aquí tot sembla correcte, però ara qüestionem-nos quins nivells de quantificació representen els codis anteriors i veurem que hem comès un error important. En efecte, el codi 0111 es correspon amb el nivell $-V_{REF}/8$ i el codi 1000 amb el nivell 0 V. Evidentment, en sumar tots dos nivells esperem obtenir com a resultat $-V_{REF}/8$, és a dir, el codi 0111, i tanmateix hem obtingut el codi 1111, que es correspon amb el nivell de $7V_{REF}/8$.

Sembla que les propietats del nostre codi no són tan bones com ens esperàvem. No podem aplicar directament les operacions aritmètiques en les paraules codi, ja que els seus resultats no es corresponen amb les operacions aritmètiques aplicades en els nivells de tensió. Penseu que realment la informació important és en els nivells de quantificació, en les tensions associades a les mostres, ja que al final de la cadena sempre hem de tornar a convertir el senyal en analògic per a poder-lo escoltar. **Com podem solucionar aquest problema?** Bé, hi ha dues alternatives, una de molt més complexa que l'altra:

1) **Mantenir el codi d'assignació lineal** però fer les operacions mitjançant taules. Això significa que sabem que les operacions sobre les paraules binàries no conserven les propietats additives i multiplicadores i que ens construïm nostres pròpies taules. Per exemple, en la taula de suma tindrem la regla següent: $1000 + 0111 = 0111$ (per coherència amb els resultats obtinguts en el paràgraf anterior). Amb aquesta solució sumar (i multiplicar) es redueix a buscar resultats en taules. Hem trobat una solució al problema però és una solució complexa i poc eficient, ja que buscar els resultats en una taula requereix

reix molt més temps que aplicar les regles de l'aritmètica, per a la qual cosa hi ha processadors especialment dissenyats. A més, necessitarem una memòria addicional per a emmagatzemar les taules de resultats.

2) La segona solució consisteix a **trobar un sistema de codificació** que conservi les propietats de la suma i el producte i sobre el qual es puguin aplicar les operacions directament en les paraules codi. Aquesta codificació existeix, és la que més s'utilitza a la pràctica i es coneix amb el nom de *complement a 2*.

3.3.6. PCM en complement a 2 (CA2) (opcional)

La codificació en complement a 2 (CA2) consisteix a numerar consecutivament en binari natural els nivells de quantificació positius.

Així, el nivell de quantificació de 0 volts rep la paraula codi només amb zeros. El primer nivell de quantificació positiu rep la paraula 0001 (estem suposant un quantificador de 4 bits) i així successivament. Els nivells de quantificació positius queden, doncs, codificats en binari natural.

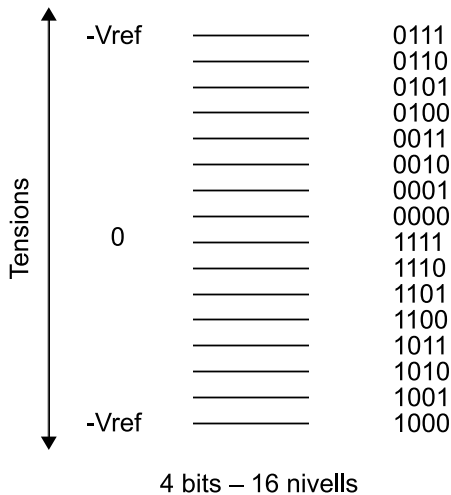
Per a obtenir la paraula codi associada a un nivell de quantificació negatiu hem de procedir d'acord amb l'algorisme següent (posem un exemple per a codificar el nivell -5):

- 1) Escriure la paraula codi del nivell positiu associat ($+5 \rightarrow 0101$)
- 2) Complementar a 1 la paraula anterior 1010 (canviar els zeros per uns i viceversa)
- 3) Sumar la unitat a la paraula obtinguda en l'etapa anterior 1011 (CA2)

La notació en CA2 és molt eficient, ja que simplement requereix complementar a 1 i sumar la unitat als nombres negatius per a obtenir el codi que els representa. A més, podem identificar fàcilment tots els nombres negatius, que són els que començaran per 1. A partir del codi en CA2 podem obtenir el valor absolut del nombre aplicant el mateix algorisme, és a dir, tornant a complementar el nombre a 2.

En la figura 64 es mostra el resultat d'assignar un codi en CA2 a un quantificador de 4 bits. Fixeu-vos que tots els nivells positius comencen per 0 mentre que els negatius comencen per 1.

Figura 64



Assignació dels codis mitjançant complement a 2 a un quantificador de 4 bits

Per a comprovar que el CA2 manté les operacions aritmètiques podem posar un exemple amb els nivells 6 (0110) i -3 (1101). En aquest cas la suma hauria de ser 3 que es correspon amb $0110 + 1101 = 0011$.

3.3.7. La taxa de bits i l'emmagatzemament de la codificació PCM

La digitalització del senyal i la seva codificació en PCM representen una manera alternativa de transportar o emmagatzemar el senyal d'àudio. En sistemes analògics els mecanismes de transport més utilitzats han estat les modulacions de ràdio (AM o FM) o cable (telefonía), mentre que els sistemes d'emmagatzematge s'han basat fonamentalment en la cinta magnètica o discos de vinil. En tots els casos, s'envia o emmagatzema un senyal analògic que representa la informació acústica que volem reproduir a distància o en el futur. En digital els mecanismes per al transport o l'emmagatzemament continuen essent, en molts casos, els mateixos (ràdio, cable, cinta magnètica, etc.), encara que en algunes aplicacions han aparegut noves tecnologies de suport (discos magnètics, suports òptics, etc.).

El canvi fonamental de la transició al món digital és la naturalesa del senyal mateix, constituït ara per bits que suporten la informació dels nivells de cada mostra.

En aquest apartat calcularem un dels paràmetres més importants dels sistemes d'àudio digital: la **taxa de bits associada**. La taxa de bits ens permet calcular, entre altres coses, quant espai d'emmagatzemament necessitarem per a emmagatzemar una cançó, quanta amplada de banda es requereix per a enviar un canal de ràdio o si és possible enviar un material audiovisual perquè es rebí en temps real per Internet.

La taxa de bits és el nombre de bits per segon que produeix el procés de digitalització d'un senyal.

En efecte, quan digitalitzem un senyal en PCM, en prenem mostres de manera periòdica i cada una la codifiquem amb un nombre de bits fix. Per tant, a mesura que digitalitzem el senyal es produeix un flux constant de bits que el representen i que s'hauran d'emmagatzemar o enviar a un hipotètic receptor. Per tant, la taxa de bits depèn dels paràmetres bàsics de la digitalització: la freqüència de mostreig i el nombre de nivells de quantificació. Posarem alguns exemples que ens ajudin a entendre aquest concepte i a relacionar-lo amb la capacitat d'emmagatzemament i la velocitat de transmissió.

Per a calcular la taxa de bits per segon d'un sistema digital determinat s'ha de calcular el producte entre el nombre de bits per mostra i el nombre de mostres per segon. En funció del sistema es poden incloure diversos canals d'àudio (estèreo, multicanal, etc.) que també s'han de tenir en compte. Així, en general, tenim:

$$R_T = M_{\text{canals}} \times f_{\text{mostreig}} \times N_{\text{bits}}$$

Exemple 1: Disc compacte d'àudio

Hem vist que el sistema de disc compacte d'àudio utilitza una freqüència de mostreig de 44.100 Hz i que cada mostra es codifica amb 16 bits. En tractar-se d'un sistema estereofònic s'han de codificar dos canals d'àudio. Podem calcular la taxa de bits multiplicant aquests tres paràmetres:

$$R_{CD} = 2 \text{ canals} \times 44.100 \frac{\text{mostres}}{\text{segon}} \times 16 \frac{\text{bits}}{\text{mostra}} = 1.411.200 \text{ bps} = 1.411,2 \text{ kbps}$$

És a dir, el senyal d'àudio digital d'un CD produeix un total d'1,4 Mbps, la qual cosa representa una taxa certament considerable que en pot limitar l'ús en algunes aplicacions. En efecte, en un reproductor d'àudio portàtil amb una capacitat de 4 GB es poden emmagatzemar un total de 70 hores de música codificada en MP3. En canvi, si emmagatzemem la música directament en el format de CD d'àudio, tenim la capacitat següent:

$$\frac{4 \text{ GB}}{1,41 \text{ Mb/s}} = \frac{4.000 \times 8 \text{ Mb}}{1,41 \text{ Mb/s}} = 22.695 \text{ s} = 6,3 \text{ hores}$$

Per tant hi ha una relació d'aproximadament un factor 10 respecte a la taxa de bits del format CD al MP3. Aquesta reducció es fa augmentant considerablement el processament que ha de fer tant el codificador com el descodificador d'MP3. A més, la compressió és amb pèrdues, la qual cosa significa que el senyal que reproduïm solament és una aproximació al senyal original del CD. No obstant això, els algorismes de codificació utilitzats actualment utilitzen models avançats del sistema auditiu humà per aconseguir que aquest no pugui diferenciar entre el senyal original en CD i el senyal comprimit.

Exemple 2. Veu digital per a telefonia

En telefonia digital es tria una freqüència de mostreig de 8 kHz i s'utilitza un quantificador de 8 bits. Amb això s'obté una qualitat acceptable que permet la intel·ligibilitat de la veu i el reconeixement del locutor. Examinarem amb més detall aquests paràmetres en una activitat posterior. La taxa de bit obtinguda per a aquests paràmetres de digitalització és:

$$R_T = 8 \frac{\text{bits}}{\text{mostra}} \times 8.000 \frac{\text{mostres}}{\text{s}} = 64.000 \text{ bits/s} = 64 \text{ kbps}$$

Històricament, aquest valor de 64 kbps té certa importància, ja que amb els primers sistemes digitals de telefonia, que es van denominar **xarxa digital de serveis integrats** (XDSI), es proporcionaven canals de 64 kbps a l'usuari, és a dir, quan un abonat contractava una línia XDSI se li proporcionava una amplada de banda de 64 kbps. El valor concret de 64 kbps es devia al fet que en la línia digital s'havia de poder acomodar un senyal de veu digitalitzada amb la qualitat de telefonia. Compareu aquest valor de 64 kbps amb les amplades de banda que s'ofereixen amb les línies ADSL (12 Mbps - 20 Mbps), la qual implica un increment en un factor d'aproximadament 300.

3.3.8. Nivells de quantificació i sistemes d'àudio

El nombre de nivells de quantificació utilitzat en un sistema d'àudio depèn de les seves característiques i de la qualitat final que vulguem obtenir. Quan es va definir el sistema **CD d'àudio** l'objectiu era obtenir un sistema d'alta fidelitat la qualitat del qual fos superior a tots els sistemes d'àudio domèstics que hi havia fins aleshores. Es van fer extensives proves subjectives per a determinar el nombre de bits òptim utilitzant experts en àudio. Aquestes proves s'han repetit de manera simplificada en aquesta activitat.

Els resultats obtinguts van revelar que la gran majoria d'oients estaven bastant satisfets amb 13 bits per mostra i que no era possible distingir entre el senyal original i el digital si s'utilitzaven 14 bits per mostra.

Durant l'especificació del format, feta conjuntament per Sony i Philips, es va discutir llargament si s'especificaven 14 bits (Philips defensava aquesta opció perquè tenia tecnologia de mostreig de 14 bits en aquell moment) o s'ampliava a 16 bits (opció defensada per Sony amb l'argument que totes les memòries estaven estructurades al byte, per la qual cosa la tecnologia, a llarg termini, seria més econòmica).

Finalment, es va escollir com a **estàndard** utilitzar 16 bits per mostra, és a dir, un total de **65.536 nivells** de quantificació que, en principi, són més enllà de la capacitat de discernir entre senyals analògics i digitals de l'oïda humana.

14 bits

El sistema CD d'àudio es defineix al final dels anys setanta i el primer reproductor apareix el 1982 quan la tecnologia de digitalització encara és incipient. Per això, la major part

dels enregistraments fets en aquesta primera època es fan només amb 14 bits deixant els 2 bits menys significatius a zero.

Televisió NICAM 728

El sistema d'àudio digital per a televisió NICAM 728 utilitza una freqüència de mostreig de 32 kHz i transmet les mostres amb un total de 10 bits. No obstant això, aquests 10 bits s'obtenen a partir de la compressió dinàmica sobre un senyal que originalment s'ha mostrejat amb 14 bits per mostra.

La compressió dinàmica utilitzada en el NICAM permet obtenir una qualitat molt semblant als 14 bits utilitzant únicament 10 bits. Com s'aconsegueix? El procediment és complex i requereix cert processament del senyal. Essencialment es tracta d'agrupar paquets de 32 mostres i determinar la que té més valor absolut. Els 10 bits es trien tenint en compte el valor d'aquesta mostra de manera que només s'enviaran a partir del primer bit que és diferent de 0. Així, per exemple, si el valor absolut màxim del paquet de 32 mostres és 0010 0110 1101 11, no enviarem els dos primers bits, ja que prenen el valor 0. Com que el sistema només envia 10 bits, tampoc no s'enviaran els dos últims (que es consideraran poc significatius).

Resumint, si l'amplitud supera un determinat llindar els 10 bits que es transmeten són els 10 bits més significatius, descartant els 4 bits menys significatius de totes les mostres del paquet. En canvi, si el valor de la mostra és molt petit només es transmeten els 10 bits menys significatius. En general, sempre es transmeten 10 bits que se seleccionen en funció del nivell de la mostra de més valor absolut. Paral·lelament, juntament amb la informació de comprovació d'errors, s'envia informació al receptor sobre els bits que s'han escollit, perquè pugui reconstruir el senyal de manera aproximada. La qualitat final que s'obté amb el sistema NICAM és acceptable i, encara que en alguns passatges musicals es poden apreciar diferències significatives respecte al senyal de CD d'àudio, el resultat final és considerablement superior al que obtindríem mostrejant el senyal amb només 10 bits per mostra.

Un altre sistema especialment interessant d'analitzar és el **DVD d'àudio**, en el qual es defineixen diferents modes i qualitats d'àudio, admetent quantificadors de 16 bits, 20 bits i 24 bits. La pregunta que sorgeix immediatament és si resulta necessari fer un mostreig de 20 bits o 24 bits quan hem comentat (i comprovat experimentalment en aquesta activitat) que el sistema auditiu humà no és capaç d'apreciar diferències entre el senyal analògic original i un senyal quantificat amb 14 bits. S'ha de tenir en compte que 24 bits representa una relació senyal-soroll molt important i que, fins i tot, amb l'estat actual de la tecnologia hi ha pocs fabricants que ofereixin sistemes d'adquisició fiables de 24 bits. **Quin és doncs l'interès de mostrejar un senyal d'àudio a 20 bits o 24 bits?** Doncs, bé, el motiu és doble.

1) D'una banda, els sistemes d'àudio domèstic poden fer un processament del senyal cada vegada més avançat. Avui en dia, és fàcil trobar equips que condicionen l'equalització del senyal d'àudio en funció de les característiques de la sala o que afegeixen efectes de presència, d'èmfasi, de reverberació, etc. al senyal d'àudio per tal que la seva audició resulti més confortable i adequada a l'entorn. Tots aquests efectes es duen a terme aplicant operacions a les mostres d'entrada.

Aquestes operacions poden constituir un nombre important de multiplicacions i sumes amb la qual cosa es produeixen arrodoniments dels resultats parcials que s'obtenen durant el tractament del senyal. Tots aquests arrodoniments afecten el senyal d'àudio i seran més importants com més petit sigui el nombre de bits originals amb què s'ha mostrejat. El resultat final és que, després d'un processament relativament complex, la qualitat del senyal pot quedar

Bits

En aquest punt cal comentar que la major part dels equips professionals són capaços de digitalitzar els senyals amb 20 bits, 22 bits o 24 bits.

afectada a causa dels efectes d'arrodoniment de les mostres i que el que el que inicialment pensàvem que tenia una qualitat comparable als 16 bits hagi estat arrodonit en excés i la relació senyal-soroll sigui molt menor que l'esperat. Per a evitar aquests efectes és convenient dur a terme les operacions de processament amb la màxima precisió possible, per la qual cosa és recomanable mostrejar el senyal amb un nombre de bits superior a 16.

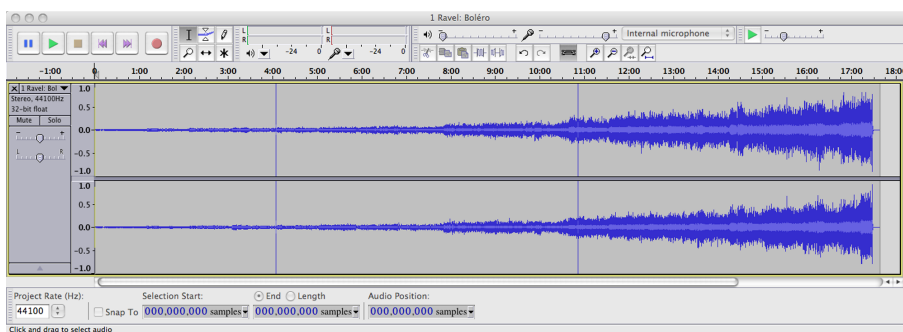
En aplicacions professionals és molt important que el senyal original tingui el màxim nombre de bits possible, ja que durant les etapes d'edició i postproducció es du a terme un gran nombre d'operacions per al condicionament dels senyals que en podria afectar la qualitat. Una vegada s'ha editat el senyal, és possible reconvertir-lo a 16 bits per a registrar-lo en un CD d'àudio.

2) Un altre dels aspectes amb certa importància per a considerar l'augment del nombre de bits és la gran dinàmica que poden tenir algunes composicions musicals, especialment de música clàssica.

Per posar un exemple molt clar, considerem el *Bolero* de Ravel, una peça clàssica, molt popular, que és una lliçó d'orquestració de Maurice Ravel. En aquesta peça Ravel utilitza una frase musical que es va repetint, entrant cada vegada nous instruments de l'orquestra. La composició s'inicia amb la percussió, amb un volum molt baix, i la potència acústica augmenta a mesura que avança. En una audició domèstica, amb un equip d'alta fidelitat, és molt probable que posem el volum inicial de l'equip a un nivell excessiu i que hàgim de corregir el volum (abaixant-lo) a mesura que el tema avança, per iniciativa pròpia o per protestes del veí. En aquest cas es tracta d'un exemple en què el senyal que s'ha registrat al començament del tema és un senyal d'un nivell molt baix i que, per tant, només utilitza un nombre molt reduït dels 65.536 nivells de què disposa el convertidor.

Aquesta situació s'intenta il·lustrar de manera gràfica en la figura 65, en què es mostra una captura de la composició completa del *Bolero* de Ravel (la durada de la peça musical és d'uns 18 minuts).

Figura 65. Forma d'ona del *Bolero* de Ravel



Fixeu-vos que el nivell del senyal durant els primers minuts és molt reduït, per la qual cosa no s'utilitzen tots els nivells de quantificació.

És important observar en aquesta gràfica com el nivell de volum augmenta progressivament, aprofitant tota la dinàmica del convertidor en la part final de la composició. No obstant això, queda clar que durant els primers minuts l'amplitud del senyal és molt petita i només s'utilitza una petita part de tots els nivells de quantificació.

Des d'un punt de vista pràctic això significa que durant aquesta primera fase de la composició la relació senyal-soroll que tenim és considerablement més baixa que la que tenim en la part final.

Relació senyal-soroll

Fixeu-vos que la relació senyal-soroll disminueix perquè la potència del senyal baixa, mantenint la del soroll. No obstant això, si aquest efecte és molt important és possible que el soroll de quantificació arribi a ser audible.

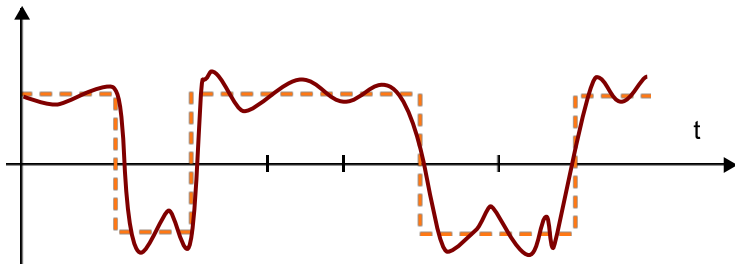
La solució a aquest problema consisteix a reduir també el soroll de quantificació, de manera que la relació senyal-soroll es mantingui en uns nivells acceptables per a l'oïda. La reducció de la potència de soroll de quantificació significa reduir-ne l'amplitud augmentant el nombre de bits del quantificador. És en aquestes circumstàncies i en aquest tipus de passatges i composicions musicals on pot tenir certa importància l'augment del nombre de bits a 20 o 24.

3.3.9. Avantatges i inconvenients de la digitalització del senyal d'àudio

Són bastants els avantatges que ofereixen els sistemes d'àudio digital respecte als sistemes analògics convencionals. L'avenç de la tecnologia ha facilitat enormement el desenvolupament dels sistemes digitals, fent que en general siguin, ara com ara, menys costosos que els analògics i amb uns avantatges molt significatius quant a emmagatzemament i possibilitats de difusió. Entre altres, s'haurien de destacar els avantatges següents:

- **La qualitat i la capacitat de reproduir el so en fidelitat** molt alta que ofereixen els sistemes digitals no té comparació amb el que es podria obtenir amb sistemes analògics. Hem vist que si escollim el nombre de bits del quantificador prou elevat no és possible distingir entre un senyal original i la seva aproximació digital. En aquestes condicions, la virtut principal dels sistemes digitals és que el senyal no es degrada a causa del seu emmagatzemament o la seva difusió. En efecte, una de les característiques principals dels sistemes digitals és la capacitat de regeneració, és a dir, que si rebem el senyal amb soroll, distorsions, interferències ratllades del mitjà, etc., gairebé sempre és possible determinar el valor original dels bits (vegeu la figura 79). Això significa que la qualitat es manté pràcticament inalterada encara que les condicions de transmissió o d'emmagatzemament no siguin les ideals. Així, per exemple, un CD admet un cert grau de brutícia i ratllades sense que la qualitat de la reproducció quedi afectada. Amb el mateix nivell de brutícia i ratllades en un disc de vinil, la qualitat quedaria greument afectada.

Figura 79. Comparativa entre un senyal digital (línia discontinua) i el mateix senyal deteriorat per sorolls, distorsions, etc.



Si el grau de soroll no és excessiu sembla raonable que un processador pugui recompondre la informació original.

- **Inserció de metadades.** El fet de codificar els senyals mitjançant una seqüència de bits repercuteix en el fet que es pugui incorporar un altre tipus d'informació diversa a més de l'àudio. Aquesta informació pot contenir els títols dels temes musicals, els intèrprets, el gènere musical, etc.
- **Protecció dels continguts.** La informació en format digital es pot encriptar de manera molt eficient i dificultar o impedir l'accés a la informació als usuaris que no estiguin autoritzats. Els continguts es poden protegir per a preservar la confidencialitat, autenticar l'autoria o proporcionar serveis de pagament mitjançant accés condicional.
- **Processament digital de la informació.** En tractar-se d'informació digital, es pot transformar directament mitjançant processadors digitals avançats, que poden executar un nombre elevat d'operacions per segon per a adaptar els senyals a les característiques de l'auditori o generar els efectes desitjats. A més, aquests algorismes són molt flexibles en el sentit que es poden modificar amb facilitat, simplement canviant el programari o el microprogramari (*firmware*) del processador.
- **Capacitat d'inserir codis correctors.** Ja hem comentat que els sistemes digitals són més robustos davant els efectes no desitjats que sofreix el senyal en ser transmès pel canal. No obstant això, si el nivell de soroll és prou important pot provocar errors en el procés de regeneració dels bits. Aquests errors es poden corregir si s'introdueixen **codis redundants**, que protegeixin el senyal davant els errors. La major part de sistemes digitals inclouen codis redundants per a la protecció d'errors que els confereix una gran robustesa.
- **Compressió de les dades.** Encara que la codificació PCM requereix una taxa de bits molt important (1,4 Mbps per al cas del CD d'àudio), la informació digital es pot processar i comprimir en uns factors molt importants sense que l'oïda humana pugui apreciar les diferències entre la informació original i la comprimida. Els factors de compressió que s'aconsegueixen amb els codificadors MP3 i MP4 són de l'ordre 10 (la mida del fitxer és 10 vegades més petita).

Compressors de dades

Aquests compressors permeten emmagatzemar senyals d'alta qualitat en dispositius de memòria d'estat sòlid de mida molt reduïda, la qual cosa permet aconseguir equips portàtils molt sofisticats que de cap manera no s'haurien pogut obtenir utilitzant sistemes analògics.

3.4. Problemes/preguntes

1) Busqueu informació i feu un breu resum sobre les característiques bàsiques del format WAV d'àudio. Intenteu esbrinar si hi ha diferents versions, quina mena de capçaleres s'inclouen, i quina informació porten les capçaleres sobre alguns paràmetres bàsics del PCM com la freqüència de mostreig, el nombre de bits per mostra i el nombre de canals. Identifiqueu com s'empaqueten les mostres d'àudio.

2) Busqueu informació sobre l'empaquetament de la trama de mostres en el sistema NICAM 728¹⁵. Elaboreu un breu informe en què com a mínim es responguin les qüestions següents:

- Quant dura una trama de NICAM 728?
- Quantes mostres s'envien?
- Quants canals s'envien?
- Quins tipus es poden identificar?
- Com se seleccionen els 10 bits per mostra?
- Com se senyalitzen, mitjançant la paritat, els 10 bits que s'han escollit?
- Per què rep el nom de NICAM 728?

3) Busqueu un parell d'exemples musicals, diferents del *Bolero* de Ravel, en què la dinàmica de la composició recomani augmentar més enllà de 16 el nombre de bits del quantificador.

4) Construïu una taula resum en la qual s'inclouin diferents sistemes d'àudio domèstic indicant la freqüència de mostreig, el nombre de bits de quantificació i si utilitzen codificació PCM o fan algun tipus de compressió de les dades. Si voleu, podeu ampliar aquesta taula a sistemes d'àudio professionals.

5) Identifiqueu productes comercials d'adquisició de senyals d'àudio que puguin treballar amb resolucions de 16 bits, 20 bits i 24 bits en un entorn d'edició basat en ordinadors personals de tipus PC o MAC. Esbrineu en les característiques tècniques d'aquests productes si les relacions senyal-soroll que es proporcionen coincideixen amb les que esperem obtenir de productes que siguin capaços de capturar a les resolucions esmentades.

⁽¹⁵⁾No és necessari que inclogui informació sobre la modulació utilitzada per a la transmissió dels bits juntament amb el senyal de televisió. En aquesta experiència només estem interessats en la part de la informació digital d'àudio.

6) Compareu les característiques d'un sistema d'adquisició d'àudio integrat en un ordinador portàtil (escolliu una marca/model determinat) amb les característiques que es poden obtenir de mòduls de mostreig que es connecten en bus PCI Express. Compareu les relacions senyal-soroll, el nombre de bits, el nombre de canals que es poden mostrejar, etc.

7) Obriu l'arxiu **NessunDorma.wav** amb el programa Audacity i determineu-ne la mida a partir del nombre total de mostres, el nombre de canals i sabent que cada mostra està codificada amb 16 bits. Anoteu el nombre de bytes que obteniu. Paral·lelament, amb l'explorador del Windows o el Finder, determineu la mida de l'arxiu *.wav que representa aquest senyal (per a saber el nombre exacte de bits heu d'entrar en les propietats del fitxer). Coincideixen totes dues mides? Si no coincideixen, podeu explicar a què és degut? Quin percentatge de bits en un fitxer wav correspon a informació de capçaleres/sincronització?

8) Digitalitzeu un senyal musical procedent d'un sistema de reproducció de casset compacte analògic i mitjançant el programa Audacity utilitzant una freqüència de mostreig de 44.100 Hz i 16 bits per mostra. Avalueu el soroll amb què s'ha capturat el senyal. Possiblement, si no es disposa d'un sistema de captura professional el nivell de soroll en el senyal serà molt superior al soroll de quantificació. Intenteu mesurar experimentalment aquest soroll a partir de la forma d'ona del senyal o del seu espectre. A partir dels resultats obtinguts determineu el nombre de bits mínim del quantificador que manté el soroll de quantificació per sota del soroll de mesura.

9) Una imatge digital es pot considerar com el mostreig en l'espai (files i columnes) d'una imatge analògica. Intenteu generar una imatge en blanc i negre amb diversos nivells de gris (per exemple, amb el Photoshop). En la mesura que sigui possible modifiqueu el nombre de nivells de gris que es presenten per pantalla i determineu quants nivells de gris és capaç de discriminar el sistema visual humà. Compareu el nombre de nivells de gris que pot discriminar el sistema visual humà amb el nombre de nivells de quantificació que es necessiten en el senyal d'àudio.

3.5. Avaluació

1) Determineu la relació senyal-soroll d'un sistema de DVD d'àudio que utilitza 24 bits per mostra. Calculeu la fracció de la potència de senyal que representa el soroll de quantificació en aquest sistema.

2) Supposeu un quantificador de 5 bits on s'utilitza una codificació en CA2 i que pot treballar amb unes tensions màxima i mínima de +/-5 volts respectivament. Determineu:

a) El pas de quantificació.

Solució

La relació senyal-soroll en un sistema de 24 bits estarà determinada pel següent:

$$SNR = 1.76 \cdot 6 \cdot 24 = 145,76 \text{ dB}$$

A partir d'aquest valor es pot calcular la relació entre la potència de senyal i la potència de soroll, que estarà determinada pel següent:

$$10 \log_{10} \frac{P_{\text{senyal}}}{P_{\text{soroll}}} = 145,76 \Rightarrow \frac{P_{\text{senyal}}}{P_{\text{soroll}}} = 10^{14,576} \Rightarrow P_{\text{senyal}} = \frac{P_{\text{soroll}}}{3,77 \cdot 10^{16}}$$

b) L'error màxim que es cometrà en aproximar una tensió analògica per un nivell de quantificació.

c) Calculeu l'SNR del convertidor i indiqueu, a partir del que heu vist en aquesta activitat, si creieu que l'oïda humana serà capaç d'apreciar el soroll de quantificació.

d) Determineu els codis associats als nivells +13 i -7.

e) Comproveu que la suma dels codis en CA2 de l'apartat anterior coincideix amb la suma dels nivells associats.

Solució

a) El pas de quantificació es pot calcular tenint en compte que tot el marge de valors d'entrada (10 volts) ha de ser repartit de manera uniforme entre els nivells de quantificació.

$$\Delta = \frac{5V - (-5V)}{2^5} = \frac{10}{32} = 0,3125 \text{ Volts}$$

b) Hem vist que l'error màxim sempre és la meitat del pas de quantificació, per tant, tindrem un error màxim de 0,15625 volts.

c) L'SNR serà la corresponent a un convertidor de 5 bits. Per tant, aplicant la fórmula obtenim la relació senyal-soroll següent:

$$SNR = 1,76 + 6,02 \cdot 5 = 31,86 \text{ dB}$$

D'acord amb els exercicis fets en aquesta activitat, el soroll de quantificació¹⁶ que s'obté amb 5 bits és audible.

d) El codi associat al nombre 13 és directament el seu codi binari natural, és a dir, 01101. Per a determinar el codi associat al -7 hem d'expressar el nombre 7 en binari natural, complementar-lo a 1 i sumar-hi la unitat. Obtenim el següent:

- Codi binari natural 7: 00111
- Complement a 1 (-7): 11000
- Complement a 2 (-7): 11001

e) Si sumem en binari natural els dos nombres anteriors obtenim $01101 + 11001 = 0011^{17}0$, que coincideix amb el codi en CA2 del nombre 6, resultat de sumar 13 i -7.

⁽¹⁶⁾S'ha de tenir en compte que l'SNR no depèn dels nivells d'entrada del convertidor, que en aquest cas són de +/-5 volts. Això és així perquè suposem que el senyal està condicionat per a cobrir tot el marge de valors d'entrada que admet el convertidor. El pas de quantificació també és proporcional al marge d'entrada, de manera que si s'augmenta el marge de nivells d'entrada també augmenta el nivell de soroll i la relació senyal-soroll es manté constant. Aquesta és una de les peculiaritats de la fórmula que estem utilitzant: la qualitat (SNR) d'un convertidor només depèn del nombre de bits.

⁽¹⁷⁾Fixeu-vos també en aquest punt que quan se suma $01101 + 11001$ obtenim un ròssec (*carry*) igual a la unitat (és a dir, el resultat de la suma és de 6 bits 100110). Aquest ròssec s'ignora en totes les operacions fetes en CA2, ja que estem treballant amb 5 bits. Aquesta circumstància s'ha de tenir en compte quan es defineixen els algorismes de processament, ja que es poden produir desbordaments en funció dels valors dels nombres amb què s'opera. No obstant això, aquest problema no és específic de la notació en CA2 sinó que també es produeix en binari natural o en CA1.

3) Determineu la taxa de bits d'un sistema d'àudio digital en què es mostregen 5 canals d'àudio convencionals i 1 canal de *subwoofer* (àudio multicanal). Els 5 canals d'àudio es mostregen a 96 kHz i prenen un total de 20 bits per mostra. El canal de *subwoofer* es mostreja a 300 Hz i pren 16 bits per mostra.

Solució

Per a determinar el nombre total de bits per segon n'hi ha prou de considerar:

- Nombre de bits per segon canal convencional:

$$96.000 \text{ mostres/segon} \times 20 \text{ bits/mostra} = 1.920.000,0 \text{ bps}$$

- Nombre de bits per segon canal *subwoofer*:

$$300 \text{ mostres/segon} \times 16 \text{ bits/mostra} = 4.800,0 \text{ bps}$$

Així, el total de bits per segon és el resultat de sumar el canal de subwoofer pels 5 canals convencionals:

$$5 \times 1.920.000 + 4.800 = 9.604.800 \text{ bps}$$

4) Un estudi d'enregistrament vol enviar urgentment un total de 100 cintes d'1 hora de durada en format DAT (*digital audio tape*). Les cintes són estereofòniques (2 canals) i estan registrades a una freqüència de mostreig de 48 kHz a 16 bits per mostra. Ens plantegem enviar aquesta informació per mitjà d'una línia dedicada que pot transmetre 2 Mbps o enviar-la com un paquet per mitjà d'un repartidor o *courier*, amb el qual tenim garantit el lliurament en 24 hores. Quin dels dos sistemes resulta més ràpid? Quin és el nombre crític de cintes a partir del qual un sistema resulta més ràpid que l'altre?

Solució

En primer lloc determinarem el nombre total de bits que tenim en les 100 cintes de DAT.

- El nombre de bits per segon del DAT serà:
2 canals x 48.000 mostres/s x 16 bits/mostra = 1.536.000 bps
- El nombre de bits contingut en una cinta (1 hora = 3.600 s) serà:
1.536.000 bps x 3.600 s = 5,5296 10⁹ bits
- Per la qual cosa el nombre de bits que s'han de transmetre per a enviar les 100 cintes són:

Nombre de bits (100 cintes DAT); 5,5296 10¹¹ bits

Si la línia per la qual fem la transmissió té una capacitat de 2 Mbps, necessitarem un total de 5,5296 10¹¹/2 10⁶ = 276.480 segons, que corresponen a 4.608 minuts o 76,8 hores. En definitiva, resulta més ràpid enviar les cintes per mitjà d'un repartidor per servei de 24 hores.

Per a determinar el nombre crític de cintes a partir del qual un servei resulta més rendible que l'altre hem de calcular el nombre total de bits que es poden transmetre en 24 hores mitjançant una línia de 2 Mbps:

$$2 \text{ 10}^6 \text{ bits/s} \times 24 \text{ h} \times 3.600 \text{ s/h} = 1,728 \text{ 10}^{11} \text{ bits}$$

Com que cada cinta té un total de 5,5296 10⁹ bits tenim que el quocient entre tots dos nombres ens dona un total de 31,25 cintes. És a dir, si el nombre de cintes és 31 surt rendible enviar-les a través de la nostra línia de 2 Mbps.

5) Repetiu el problema anterior però ara suposant que volem transmetre 100 cintes de DV (*digital video*) d'1 hora de durada. Supposeu en aquest cas que la taxa de bits del sistema DV és de 25 Mbps.

Solució

El problema és pràcticament idèntic a l'anterior, però ara la capacitat de bits d'una cinta de DV és francament superior. En efecte, passem dels 1,536 Mbps que estan registrats en un DAT als 25 Mbps que estan registrats a la cinta DV. La relació entre tots dos factors és que les cintes DV tenen una capacitat 16,276 vegades superior a una cinta de DAT.

Segons aquesta relació podem extrapolar els resultats:

a) Es necessiten 1.250 hores per enviar les 100 cintes de DV per mitjà d'una línia de 2 Mbp.

b) El límit perquè resulti rendible enviar les cintes per mitjà de la línia de 2 Mbps¹⁸ és d'1,92 cintes.

⁽¹⁸⁾Fixeu-vos que en aquest problema la taxa de bits que conté la cinta DV és considerablement superior a la velocitat de transmissió que estem considerant. En efecte, suposem que es transmeten bits a una velocitat de 2 Mbps i que la cinta conté un total de 25 Mbps. Això significa que per a transmetre un fragment de vídeo corresponent a un segon necessitem un total de 12,5 segons. Per tant, la transmissió no es pot dur a terme en temps real. Per la mateixa regla de tres, la transmissió d'una cinta d'1 hora ens ocuparà un total de 12,5 hores. Fixeu-vos que el cas és considerablement diferent del cas del DAT, en què la transmissió es podia fer en temps real, ja que la capacitat del canal era superior a la capacitat de la cinta.

4. Activitat 4. Filtratge de senyals d'àudio

4.1. Introducció

4.1.1. Contingut

En aquesta activitat es presenten els conceptes bàsics sobre el filtratge i l'equalització de senyals d'àudio, parts essencials en el processament digital de senyals d'àudio. Per la seva importància, aquesta activitat s'ha dividit en dues parts amb l'objectiu de facilitar la consolidació dels coneixements.

En aquesta primera part ens centrarem en els conceptes de **filtratge** i en la concatenació de diversos filtres per a obtenir dels efectes desitjats.

La segona part es dedicarà principalment als **equalitzadors paramètrics i gràfics**. A tota hora s'intenta que els conceptes que es presenten estiguin acompanyats d'exercicis pràctics que permetin mostrar de manera auditiva els efectes d'un determinat tipus de filtratge.

L'activitat s'inicia amb una presentació dels diferents tipus de filtres que s'usen a la pràctica per al condicionament de senyals d'àudio i s'intenta distingir entre els denominats **filtres** i els denominats **equalitzadors**. A continuació es presenten els filtres passabaix i passaalt, que al llarg de l'activitat es consideraran filtres bàsics i a partir dels quals es pot construir qualsevol altre tipus de filtre. S'experimenta amb diferents programes per a veure els efectes d'aquests filtres en els senyals d'àudio.

Posteriorment es presenten els principis de **connexió de filtres en sèrie i en paral·lel**, la qual cosa permet introduir els filtres de pas de banda com una connexió en sèrie d'un filtre passabaix i un passaalt. També s'experimentarà amb diferents senyals i programes per a veure els efectes i aplicacions d'aquests filtres.

Finalment, s'introdueixen els **filtres de shelving**, que se solen considerar com una primera versió d'equalitzadors. Els equalitzadors més utilitzats, com el paramètric i el gràfic, s'estudiaran amb detall en la segona part de l'activitat.

4.1.2. Metodologia

S'utilitza la mateixa metodologia que s'ha estat utilitzant en activitats anteriors i que cerca que l'alumne pugui anar comprovant els conceptes exposats mitjançant l'audició dels efectes del filtratge dels senyals. En aquesta activitat

es continua utilitzant el programa Audacity per a la captura, edició i processament dels senyals d'àudio, però també s'introdueix una **miniaplicació de Java** (*Java applet*) que implementa diferents tipus de filtratge sobre el senyal d'àudio.

L'avantatge principal de la miniaplicació de Java és que permet modificar els paràmetres dels filtres en temps real, i veure'n l'efecte de manera immediata.

En tot cas, a mesura que s'introdueixen els conceptes es proposen una sèrie d'exercicis amb aquests programes. Els exercicis són guiats i no haurien de ser gaire difícils de fer. El més important és que l'alumne se centri sobre els resultats i els conceptes implícits en els exercicis.

Al final de l'activitat es proposen alguns exercicis addicionals que s'espera que l'estudiant resolgui amb un grau superior d'autonomia i que n'interpreti correctament els resultats. Finalment, es proposen diversos problemes semblants als que han aparegut com a exemples al llarg del text i s'inclouen les solucions.

4.1.3. Recursos

Els recursos principals utilitzats en aquesta activitat són:

- Programari de captura i processament d'àudio Audacity (llicència lliure)
- Programari de demostració de filtres i equalitzadors basat en miniaplicacions de Java
- Anàlisi de forma d'ona i anàlisi de freqüència dels senyals mitjançant el programa Audacity
- Captura de senyals de veu per a la seva equalització
- Arxius d'àudio predigitalitzats amb diferents característiques
- Explicacions dels fonaments de l'equalització i les diferències amb el filtratge
- Explicacions de les aplicacions principals dels equalitzadors
- Proposta d'activitats de cerca d'informació en webs i recursos d'Internet per a aprofundir en els coneixements adquirits en aquesta activitat (aplicacions de programari d'equalització, fabricants d'equalitzadors, etc.)

- Activitats proposades per a verificar l'adquisició dels coneixements d'aquesta activitat

4.2. Objectius

Els objectius principals que es persegueixen en aquesta activitat són:

- 1) Introduir els conceptes clau per al processament de senyals d'àudio mitjançant filtres.
- 2) Introduir els conceptes bàsics de filtres passabaix i passaalt.
- 3) Relacionar, mitjançant experiments dirigits, els efectes del filtratge sobre la qualitat i percepció sonora d'un senyal d'àudio.
- 4) Relacionar els efectes del filtratge en la representació en freqüència del senyal d'àudio.
- 5) Comprovar que les realitzacions dels filtres que utilitzem a la pràctica s'allunyen del comportament ideal, permetent el pas de residus de senyal que resulten audibles.
- 6) Conèixer els principis bàsics d'associació en sèrie i en paral·lel de filtres.
- 7) Introduir els conceptes de filtratge de pas de banda i filtratge de banda eliminada mitjançant la concatenació en sèrie/paral·lel de filtres bàsics.
- 8) Introducció als equalitzadors i controls de to basats en els filtres de *shelving*.
- 9) Aprofundir en el maneig de diverses pistes d'àudio, la generació d'efectes de filtratge i la barreja d'aquests.

Aquests objectius estan relacionats amb les **competències de l'assignatura** següents:

- 1) Capacitat de capturar àudio i imatges de manera eficient i eficaç
- 2) Capacitat de digitalitzar eficientment i eficaçment continguts d'àudio i imatge
- 3) Capacitat d'operar de manera digital amb senyals digitals d'àudio i imatge
- 4) Capacitat d'interpretar i modificar informació auditiva de manera freqüencial

I amb les **competències generals del grau** següents:

- 1) Ser capaç d'analitzar un problema en el nivell d'abstracció adequat a cada situació i aplicar habilitats i coneixements adquirits per a abordar-lo i resoldre'l.
- 2) Capturar, emmagatzemar i modificar informació d'àudio, imatge i vídeo digitals aplicant principis i mètodes de realització i composició del llenguatge audiovisual.
- 3) Integrar i gestionar continguts digitals en aplicacions multimodals d'acord amb criteris estètics, tècnics i funcionals.
- 4) Atendre adequadament consultes sobre projectes, tecnologies i mercat de productes multimèdia avaluant de manera precisa l'entorn d'aplicació, els recursos i les alternatives tecnològiques disponibles.

4.3. Guia

El filtratge i l'equalització d'un senyal és un dels procediments utilitzats amb més freqüència en la captura, edició, producció i reproducció de senyals d'àudio. En aquesta activitat definirem els conceptes de filtratge i equalització i veurem diferents alternatives i sistemes per a dur-los a terme. L'activitat està dividida en dues parts. En aquesta primera ens centrem en la introducció dels filtres i l'experimentació amb ells, deixant la part d'equalització i les seves aplicacions per a la segona part.

Per a comprendre el procés de filtratge i equalització és important tenir en compte la representació del senyal d'àudio en el domini de la freqüència, que ja s'ha tractat amb cert detall en l'activitat 1.

El filtratge i l'equalització consisteixen a emfatitzar o atenuar algunes regions freqüencials respecte a d'altres, per a aconseguir un determinat efecte.

És possible modificar el balanç entre baixes i altes freqüències d'un instrument musical o d'una veu humana ajustant els guanys que proporciona l'equalitzador a diferents zones de l'espectre.

No és fàcil distingir entre els conceptes de filtratge i d'equalització, ja que es confonen i s'intercanvien sovint. De fet, l'equalitzador sol estar constituït per un o diversos filtres, la qual cosa augmenta la confusió.

En alguns textos, es pren com a definició de filtre la d'un sistema l'objectiu del qual és deixar passar una determinada banda de freqüències i eliminar-ne d'altres.

Els filtres passabaix, de pas de banda i passaalt que estudiarem en aquesta activitat corresponen a aquesta definició de filtre.

En canvi, en un equalitzador l'objectiu és **emfatitzar/atenuar** unes bandes de freqüències respecte d'altres, però sense necessitat d'eliminar-les completament.

S'entén que un equalitzador és un sistema la funció del qual pot ser controlada per l'usuari d'una manera directa i intuïtiva, sense necessitat de conèixer a fons els fonaments del filtratge de senyals.

A més, hi ha uns tipus de filtres, denominats *de shelving* (prestatgeries), que es poden considerar com a mig camí entre un filtre convencional i un equalitzador.

En la figura 66 es compara de manera esquemàtica, un filtre convencional, un filtre de *shelving* i un equalitzador. La representació del filtre correspon a la gràfica de l'esquerra. En aquest exemple, es tracta d'un filtre passabaix, és a dir, un filtre que permet el pas dels components de baixa freqüència del senyal (fins a 800 Hz, en el nostre exemple) i que elimina els components de freqüència més alta. Per comprendre l'efecte d'un filtre passabaix en un senyal d'àudio pensem que, si s'aplica aquest filtre, la part de baixa freqüència del senyal se sentirà igual, mentre que els components que superen la freqüència de 800 Hz es deixaran de sentir.

Per a qualsevol filtre es defineixen tres zones que es denominen:

- la banda de pas,
- la banda atenuada i
- la banda de transició.

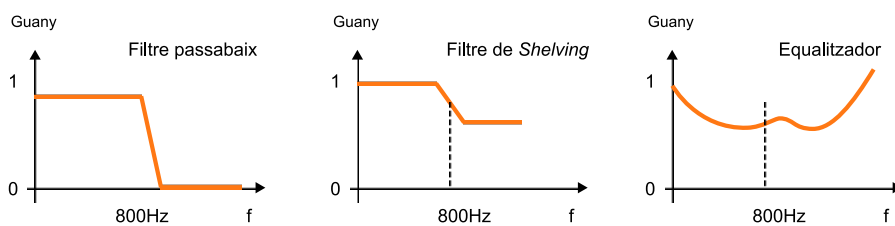
En l'exemple de filtre passabaix de la figura 66, la **banda de pas** és la zona que té guany unitat, és a dir, tots els components freqüencials que passaran inalterats a través del filtre. La **banda atenuada** és la banda que té un guany zero, és a dir, aquests components de freqüència seran eliminats, no apareixeran a la sortida. La **banda de transició** és la que s'utilitza per a passar de la banda de pas a la banda atenuada. En ella, el guany varia de manera progressiva entre un valor unitat (no modificar el senyal) i un valor zero (eliminar-lo).

Filtres *shelving*

Aquests filtres permeten que l'usuari configuri guanys diferents per a les regions de freqüència baixa i alta.

En els filtres ideals la banda de transició no està definida, i es passa directament de la banda de pas a l'atenuada. En els filtres reals tampoc no és possible tenir una banda de pas amb un valor de guany exactament igual a la unitat, ni una banda atenuada amb un valor exactament igual a zero. Entre la banda de pas i la banda de rebuig hi ha una banda de transició en la qual el guany del filtre passa del valor màxim al mínim. De manera esquemàtica, el filtre es representa assignant un valor de guany unitat (senyal sense modificar) en les freqüències que permetem el pas del senyal, i un valor de guany zero en les freqüències que no es deixen passar, encara que a la pràctica els guanys no seran exactament iguals a la unitat i zero (és a dir, ni s'anul·laran completament ni passaran exactament igual).

Figura 66. Representació comparativa entre un filtre convencional, un filtre de *shelving* i un equalitzador



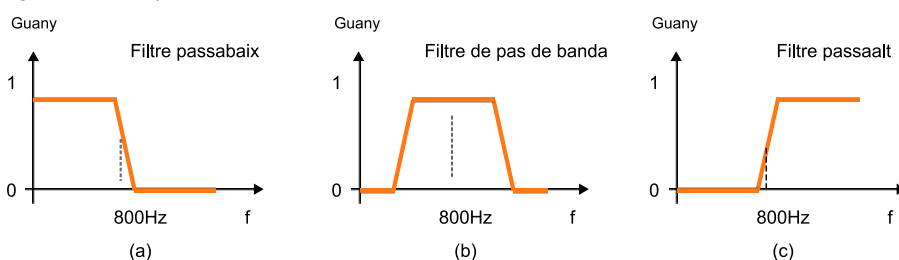
Tipus de filtre

Hi pot haver diversos tipus de filtres en funció de les freqüències que permeten passar o eliminar. Així, per exemple, un filtre passaalt deixa passar les altes freqüències i elimina les baixes, mentre que un filtre de pas de banda deixa passar les freqüències intermèdies, eliminant els components de freqüència alta i baixa.

En la figura 67 es representen esquemàticament un filtre passabaix, un filtre de pas de banda i un filtre passaalt (d'esquerra a dreta respectivament). El concepte d'aquests filtres convencionals sempre està lligat al fet de deixar passar uns determinats components freqüencials i eliminar-ne la resta.

En canvi, el **filtre de shelving**, representat en la gràfica central de la 66 es caracteritza per tenir dos guanys diferents, un per a la zona de freqüències baixes i un altre per a la de les freqüències altes. Generalment l'usuari pot controlar la freqüència a què es produeix la transició i els valors de guany per a cada una de les regions. En el nostre exemple, la freqüència de transició està situada en els 800 Hz i els guanys són de $G = 1$ per als components de freqüència inferior als 800 Hz, i de $G = 0.6$ per a les freqüències superiors.

Figura 67. Exemples de filtres convencionals



a) Passabaix, b) Pas de banda, c) Passaalt

L'equalitzador es representa en la part de la dreta de la figura 66. La característica d'aquests filtres és que l'usuari disposa d'una gran flexibilitat per a configurar la resposta en freqüència del sistema.

De fet, el nom d'**equalitzador** procedeix del fet que és un sistema amb el qual tractarem de compensar alguna deficiència o filtratge previ que ha sofert el so, i intentarem igualar o restaurar amb aquest sistema les condicions originals.

Veurem que hi ha diferents mecanismes per a permetre que l'usuari pugui configurar de manera fàcil i intuïtiva la forma de la corba de resposta en freqüència de l'equalitzador. En el cas concret representat en la figura 66, l'equalitzador augmentaria el volum dels components de baixa i d'alta freqüència

4.3.1. Exemples amb filtratge passabaix

Farem alguns exercicis pràctics per analitzar i escoltar els efectes d'un filtre passabaix sobre un senyal d'àudio. Per a això utilitzarem una aplicació Java desenvolupada en la Universitat de Helmut-Schmidt a Hamburg (Alemanya), concretament pel professor Udo Zölzer, i que es pot descarregar de la pàgina web <http://ant.hsu-hh.de/dasp/>.

Aquesta pàgina web inclou altres exemples de tractament digital de senyals d'àudio que acompanyen el llibre següent: Udo Zölzer (2008, agost). *Digital Audio Signal Processing*. (2a. ed.). John Wiley. Encara que aquest text és molt més avançat que els continguts que presentem en aquest curs, si voleu podeu descarregar-vos altres exemples de quantificació i mostreig o efectes de senyals d'àudio. La pàgina web inclou aplicacions escrites en Java i en MATLAB. En la versió Java es permet descarregar la versió executable o, si es vol, el codi font¹⁹. La versió en MATLAB requereix tenir instal·lat el MATLAB a l'ordinador per a la seva execució.

⁽¹⁹⁾La descàrrega del codi font només es recomana per als alumnes que estiguin interessats a programar els algorismes digitals que filtren els senyals. Aquests algorismes són força complexos i no són l'objectiu principal d'aquest curs.

Alternativament, es pot descarregar l'aplicació Java de la pàgina de materials addicionals de l'assignatura. Es tracta d'un fitxer comprimit (**filtros.zip**) que en descomprimir-lo generarà un directori **Filters** en el qual s'inclouen diversos fitxers. Descomprimiu el fitxer en el directori que vulgueu, situeu-vos en aquest directori i carregueu el fitxer eqApplet.html en un navegador (Internet Explorer, Firefox, Safari). Haureu de tenir instal·lat el Java²⁰ perquè l'aplicació s'executi correctament.

⁽²⁰⁾Si no teniu instal·lat el Java, el podeu descarregar des d'<http://www.java.com/es>. De tota manera, normalment, si no està instal·lat, el mateix navegador anirà a la pàgina d'instal·lació del Java.

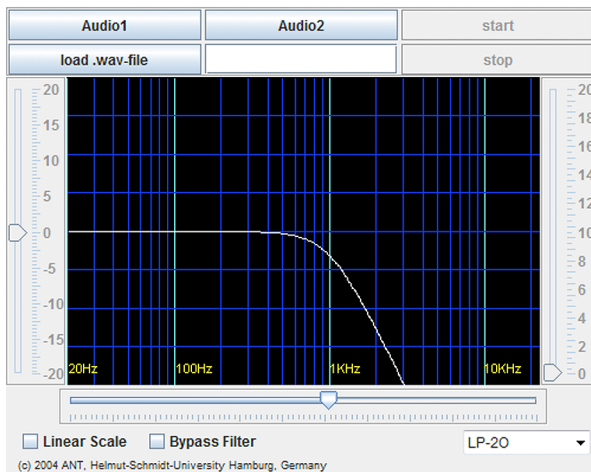
A més haureu de permetre que el navegador executi els *scripts* d'aquesta pàgina, opció que generalment estarà bloquejada en la configuració de seguretat per defecte del navegador.

Una vegada permeteu l'execució de *scripts* i del programa, apareix l'aplicació Java de filtratge de senyals que es veu en la 68 i que utilitzarem per a diverses experiències que proposa aquesta activitat. Els botons *Audio 1* i *Audio 2* permeten carregar dos arxius d'àudio per defecte, amb els quals es mostren els efectes dels diferents filtres. També apareix el botó ***Load.wave file***, que permet utilitzar el fitxer que vol l'usuari i que usarem més endavant per a veure els efectes del filtratge respecte a diferents captures d'àudio.

La casella de selecció ***Linear Scale*** s'utilitza per a triar si l'escala de l'eix horitzontal de la gràfica ha de ser lineal o logarítmica. Com que el comportament de l'oïda en la discriminació de freqüències és logarítmic, per a la major part de les proves, convé no tenir activada aquesta casella de selecció. En la representació en escala lineal (casella activada) les freqüències es representen en una escala lineal entre 0 Hz i 20 kHz. En aquesta representació, els 10 kHz, una freqüència extremadament alta des del punt de vista subjectiu, queda en la meitat de la gràfica. Quan la casella no està activada, des del punt de vista auditiu, la relació entre l'escala horitzontal i l'efecte auditiu resulta més natural.

L'altra casella de selecció es denomina ***Bypass Filter*** i resulta molt útil per a avaluar l'efecte del filtre. Si la casella no està activada, l'aplicació Java processa i modifica el senyal d'acord amb la resposta del filtre que es mostra en la gràfica. Quan s'activa la casella podem sentir el senyal d'àudio original, sense ser processat pel filtre. Resulta útil commutar aquesta casella de selecció per a comprovar l'efecte del filtratge en el senyal. La barra de desplaçament situada per sota de l'eix horitzontal permet modificar la freqüència de tall del filtre. La resposta en freqüència del filtre es mostra en la gràfica de manera dinàmica en moure la barra de desplaçament. Fixeu-vos que, tal com hem definit anteriorment, el filtre intenta proporcionar un guany unitat a la banda de pas i un guany nul a la banda eliminada. Fixeu-vos també que l'escala en l'eix horitzontal és logarítmica, la qual cosa significa que el guany unitat correspon a 0 dBs [$10 \log_{10}(1) = 0$] i que el guany nul de la banda atenuada correspondrà a $-\infty$ dBs. En la casella de selecció situada en la part inferior esquerra de l'aplicació podem seleccionar diferents tipus de filtres. Per a aquest apartat és convenient tenir seleccionat el filtre LP-2n. (*low pass 2nd order* o passabaix de segon ordre).

Figura 68. Aplicació Java eqApplet.html que mostra l'efecte d'un filtratge passabaix en un senyal d'àudio.



1) Activeu els senyals d'àudio *Audio 1* i *Audio 2* i apliqueu el filtre passabaix amb diferents freqüències de tall. Moveu la barra de desplaçament associada a la freqüència de tall i observeu l'efecte que té sobre el so. Activeu i desactiveu la casella *Bypass* per comparar el senyal filtrat amb l'original. Comproveu com alguns instruments de percussió i els més aguts desapareixen en disminuir la freqüència de tall.

2) Activeu l'escala lineal i modifiqueu la freqüència de tall del filtre passabaix mitjançant la barra de desplaçament. Anoteu per a quin valor de la freqüència de tall comenceu a notar canvis significatius entre el senyal original i el senyal filtrat. Podeu donar una explicació a aquest fenomen? Quina?

3) Desactiveu l'escala lineal (la podeu mantenir activada si voleu). Observeu com la potència del senyal disminueix en reduir la freqüència de tall del filtre. Podeu explicar per què?

Ara utilitzarem el programa Audacity²¹ per a capturar senyal de veu i aplicar-hi un filtratge passabaix. El programa Audacity permet que l'usuari configuri les característiques d'un filtre passabaix, de manera que podrem analitzar el senyal en el domini de la freqüència i veure quins són els efectes del filtratge.

⁽²¹⁾ Seguiu el procediment descrit en l'activitat 1 d'àudio per a activar el programa Audacity.

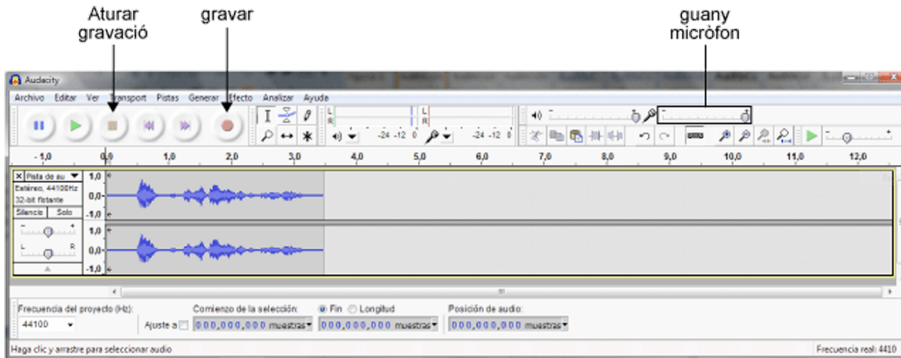
1) Obriu l'aplicació Audacity i captureu un fragment de senyal de veu, parlant directament davant de l'ordinador. Un fragment d'uns 3 o 4 segons hauria de ser suficient per a dur a terme l'experiència correctament, però, si es vol, es pot prendre un fragment més llarg. En tot cas, intenteu evitar silencis llargs i parleu de manera constant durant la captura.

En la figura 69 es mostra una captura de la pantalla d'Audacity amb l'adquisició d'un senyal d'àudio usant el micròfon integrat²² en un ordinador. En la 69 també s'indica quina és la barra de desplaçament per a controlar el guany del micròfon, el botó d'enregistrament del senyal i el botó de parada

⁽²²⁾ Si no es disposa de micròfon integrat es pot fer l'enregistrament amb un micròfon extern o, si ho preferiu, fer l'enregistrament d'àudio amb un dispositiu extern i capturar-lo posteriorment com a senyal de línia.

d'enregistrament. Si es vol, es poden seleccionar els fragments previs i posteriors al registre i tallar-los (eliminar-los) utilitzant el menú **Editar + Cortar** (Ctrl + X).

Figura 69. Captura d'una pantalla d'Audacity amb la captura del senyal d'àudio procedent d'un micròfon



2) Escolteu el senyal d'àudio capturat pel micròfon diverses vegades. Després seleccioneu tot el senyal i calculeu-ne la representació en el domini de la freqüència. Recordeu que per a calcular la representació en el domini de la freqüència heu d'executar els menús **Analizar + Análisis de Espectro**. El resultat per al nostre exemple es mostra en la figura 70. Resulta interessant notar que el senyal mostra molt poc contingut en freqüències altes, la qual cosa és propi del senyal de veu, que té una energia molt baixa més enllà dels 3.000 Hz.

Figura 70. Anàlisi en freqüència del senyal de veu digitalitzat amb un micròfon integrat

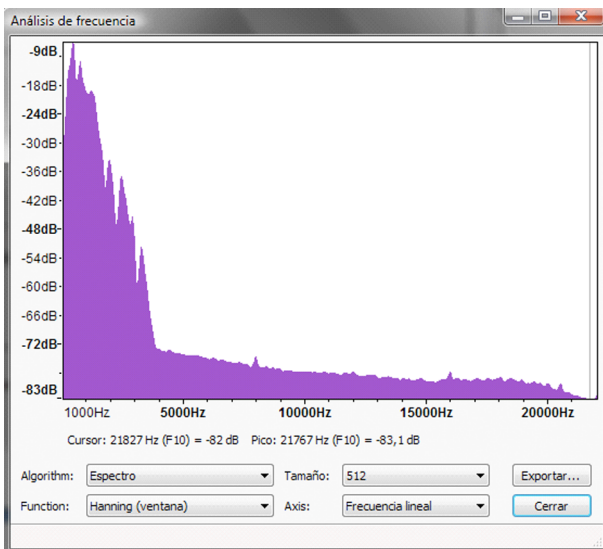
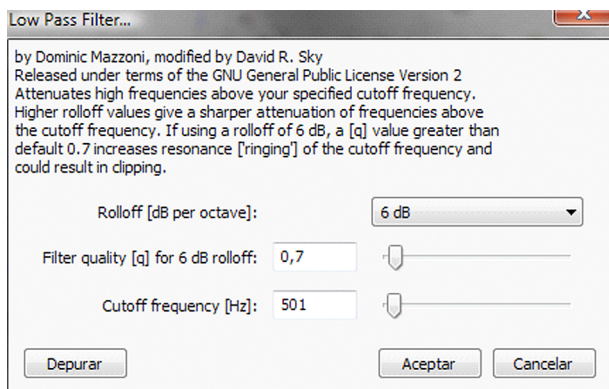


Figura 71. Selecció de paràmetres per al filtre passabaix



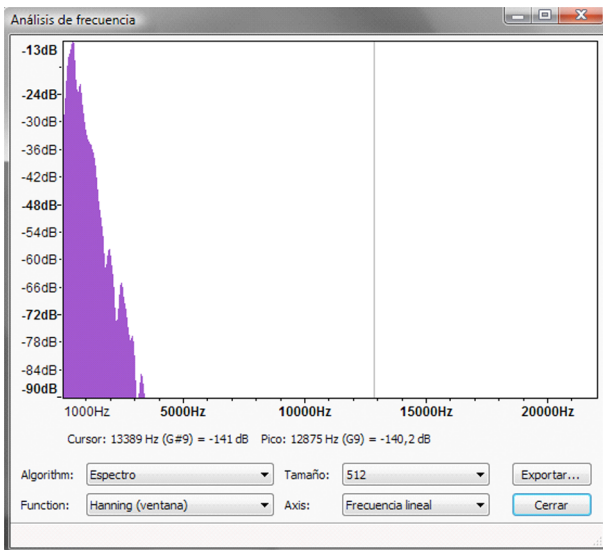
3) Guardeu el senyal d'àudio anterior en un **fitxer WAV**. Per a això, en el menú **Archivo** del programa Audacity seleccioneu l'opció **Exportar** i seleccioneu com a format de sortida WAV (Microsoft) *signed* 16 bit PCM.

4) Filtreu el senyal d'àudio anterior mitjançant un filtre passabaix amb una freqüència de tall de 500 Hz. Per a això, en el programa Audacity seleccioneu l'opció **Efecto\Filter\Low Pass** i modifiqueu els paràmetres del filtre perquè coincideixin amb els que es mostren en la figura 71. Apliqueu el filtre i escolteu el senyal resultant. Per a comparar el senyal filtrat amb l'original podeu anar al menú **Editar** i utilitzar les opcions **Deshacer Filtrado** i **Rehacer Filtrado**.

5) Podeu comparar el resultat del filtre generat amb l'Audacity i el filtre generat amb l'aplicació **eqApplet**. Per a això, reactiveu la finestra de l'aplicació Java amb els filtres d'àudio, i amb la barra de desplaçament seleccioneu una freqüència de tall del filtre passabaix d'uns 500 Hz. Carregueu el fitxer d'àudio que heu desat en el fitxer anterior mitjançant l'opció **Load .wav file** i comproveu que els dos filtres generats tenen el mateix efecte des del punt de vista auditiu.

6) Finalment, seleccioneu el senyal d'àudio filtrat i analitzeu-ne l'espectre, seguint els mateixos passos utilitzats en el punt 2). El resultat de l'anàlisi espectral es mostra en la figura 72 en què pot comprovar-se com l'efecte del filtre ha estat el de reduir els components d'alta freqüència del senyal de veu.

Figura 72. Resultat de l'anàlisi en freqüència del senyal de veu processat mitjançant un filtre passa baix

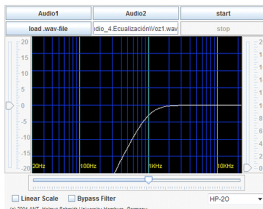


4.3.2. Exemples amb filtratge passaalt

En aquest apartat repetirem la mateixa experiència anterior però utilitzant un filtre passaalt. Per tant, en aquest cas, processarem el senyal permetent el pas dels components d'alta freqüència i eliminant les baixes freqüències. Des del punt de vista auditiu, eliminarem els sons greus i mantindrem els aguts.

Per a començar amb l'experiència auditiva se suggereix carregar l'aplicació Java de filtres d'àudio i seleccionar el filtre HP-HP-2O (passaalt - 2n. ordre; *high pass - second order*). En la figura 73 es mostra la finestra de l'aplicació, amb la gràfica del filtre amb una freqüència de tall seleccionada d'1 kHz. Els botons, caselles i barres de desplaçament tenen la mateixa interpretació que en l'apartat anterior.

Figura 73. Exemple de l'aplicació Java per al filtratge d'àudio utilitzant un filtre passaalt de segon ordre



La freqüència de tall està seleccionada en 1 kHz.

Així, doncs, repetirem les proves que hem fet en l'apartat anterior, però ara amb un filtre passaalt. En l'exercici següent es resumeixen les proves que es recomanen fer amb aquest filtre.

1) Carregar els senyals *Audio 1* i *Audio 2* que es proporcionen en la miniaplicació de Java i modificar la freqüència de tall del filtre passaalt amb la barra de desplaçament. Activar i desactivar la casella *Bypass Filter* per a veure l'efecte del filtre amb més claredat. Observar com desapareixen els sons més greus del senyal mentre es mantenen els més aguts (platerets de les percussions).

2) Indiqueu les freqüències de tall en què comenceu a observar canvis significatius entre el senyal original i el filtrat.

3) Observeu el fenomen de disminució de la potència del senyal quan es modifica la freqüència del filtre.

4) Utilitzeu el mateix senyal d'àudio que heu registrat en el punt 2) de l'apartat anterior (filtre passabaix) i representeu l'anàlisi en freqüència del senyal original i del senyal filtrat amb el programa Audacity. Observeu els efectes del filtratge sobre l'espectre. Se suggereix prendre una freqüència de tall de 1.000 Hz o superior, perquè els efectes del filtratge resultin més clars i evidents. Amb aquest tipus de filtres, en desaparèixer els components de baixa freqüència, que són els de més energia en el senyal d'àudio, l'efecte del filtratge sobre l'espectre hauria de ser molt més clar. Fixeu-vos que, generalment, obtindreu dues gràfiques amb escales verticals diferents, ja que l'espectre del senyal de veu filtrat té una energia molt més baixa. En la figura 74 es mostra el resultat que hem obtingut amb el filtratge passaalt d'una veu humana amb una freqüència de tall de 1.000 Hz i un filtre de 6 dB per octava (segon ordre, equivalent als filtres passaalt de la miniaplicació de Java). Per veure amb més claredat l'efecte del filtre sobre l'espectre pot ser convenient fer la presentació en escala logarítmica. Aquest resultat es mostra en la figura 75.

Figura 74. Representació en freqüència dels espectres d'un senyal de veu original (esquerra) i filtratge passaalt amb una freqüència de tall de 1.000 Hz i un filtre de segon ordre

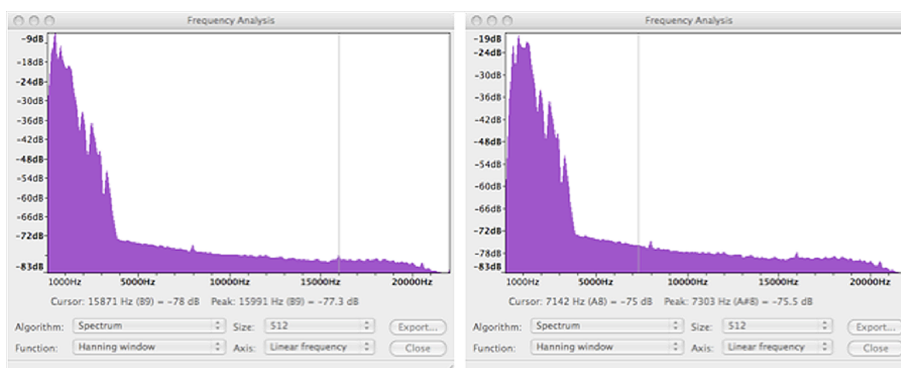
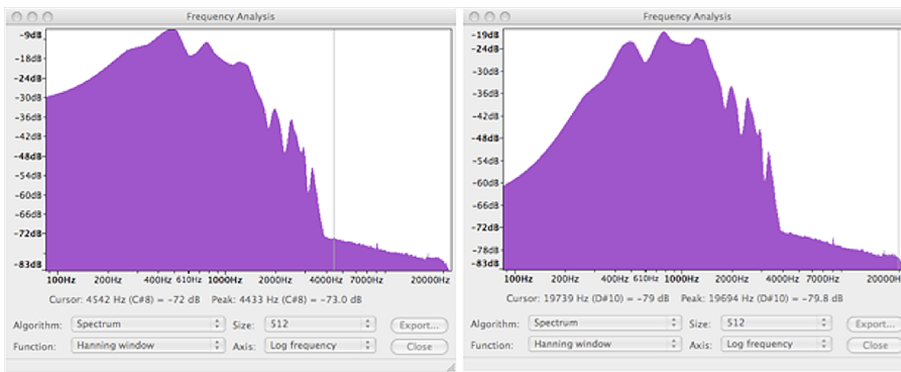


Figura 75. Comparativa entre els mateixos espectres que en la figura anterior però ara representats utilitzant l'eix horitzontal en escala logarítmica



La representació logarítmica facilita la visualització de l'espectre en les regions de baixa freqüència.

5) Filtrar amb la miniaplicació de Java el mateix senyal del punt anterior, amb la mateixa freqüència de tall del filtre i comprovar que l'efecte audible és el mateix.

4.3.3. Combinacions de filtres en sèrie i en paral·lel

Els filtres passa baix i passa alt estudiats en els apartats anteriors es poden considerar com uns mòduls bàsics per a construir un altre tipus de filtres més complexos mitjançant l'associació en sèrie o en paral·lel de diversos mòduls de filtratge. En aquest apartat veurem alguns exemples senzills i en apartats posteriors estendrem l'associació en sèrie i en paral·lel per a filtres més complexos.

En la figura 76 es mostra l'esquema bàsic de les connexions en sèrie i en paral·lel de dos filtres qualssevol. Atès que l'efecte final és que s'obté un senyal de sortida per a una entrada, podem intentar buscar un **únic filtre equivalent**, que farà la mateixa funció que la connexió en sèrie o en paral·lel dels dos primers.

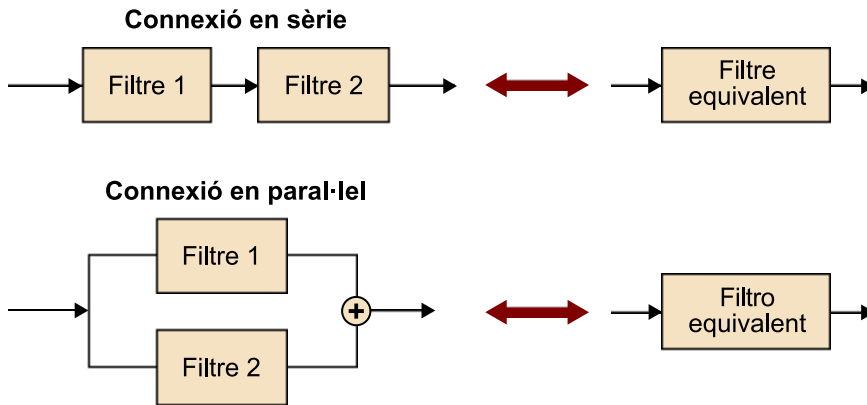
En una connexió en sèrie, els filtres s'apliquen l'un darrere l'altre de manera consecutiva. Així, el senyal d'entrada passa primer a través del filtre 1 i la seva sortida s'utilitza com a entrada del segon filtre.

Per a determinar el sistema equivalent, haurem de tenir en compte que el segon sistema processa els senyals que ja ha processat el primer.

Així, si un component en freqüència ha estat atenuat pel primer sistema en un factor 0,9 i el segon sistema li proporciona una atenuació de 0,5, llavors l'efecte global sobre aquest component és la d'una atenuació de $0,9 \times 0,5 = 0,45$. Per tant, sembla raonable que l'efecte final sobre un component sigui el resultat del producte de cada un dels efectes dels filtres. A més, d'acord amb aquest raonament és indiferent que un component sigui processat primer per un filtre i després per un altre o al revés, podent-se intercanviar l'ordre dels

filtres. D'altra banda, quan els sistemes estan connectats en paral·lel cada un du a terme el processament directament sobre el senyal d'entrada i se sumen les sortides. En aquest cas, el sistema equivalent serà la suma dels dos filtres.

Figura 76. Associació en sèrie i en paral·lel de filtres i filtres equivalents



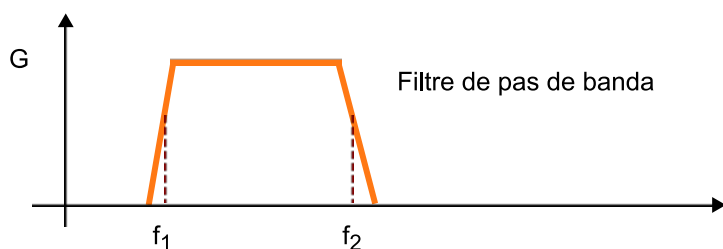
4.3.4. Filtres de pas de banda

Un filtre de pas de banda és un filtre que només deixa passar les freqüències situades dins d'una banda compresa entre una freqüència de tall inferior f_1 i una freqüència de tall superior f_2 , tal com es representen en la figura 77. Les freqüències baixes, amb una freqüència més baixa que f_1 són eliminades pel filtre. Anàlogament, les freqüències altes, amb una freqüència superior a f_2 tampoc no passen a través del filtre.

Un filtre de pas de banda es pot aconseguir com la connexió en cascada entre un filtre passabaix i un filtre pas alt, tal com es mostra en la figura 78. En aquest cas, la freqüència de tall del filtre passabaix és f_2 , eliminant les freqüències superiors a aquest límit. La freqüència de tall del filtre passaalt és f_1 , per tant, s'eliminen totes les freqüències inferiors a aquest valor.

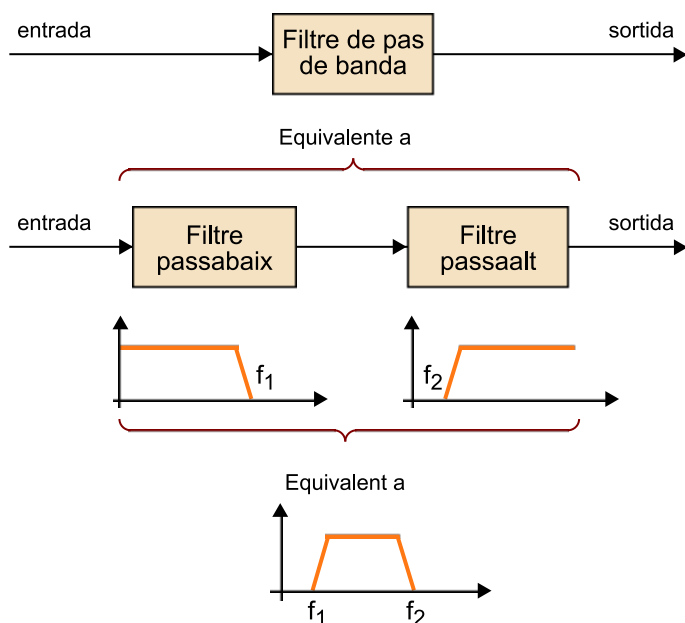
En resum, el filtre ha eliminat les freqüències superiors a f_2 i les inferiors a f_1 per la qual cosa a la sortida només tindrem els components que siguin entre aquestes dues freqüències. Els filtres de pas de banda tenen diverses aplicacions, sobretot en el context de comunicacions. En àudio, es poden utilitzar per a separar la veu humana de sorolls o instruments musicals. En efecte, els components freqüencials de la veu humana estan situats entre uns 300 Hz i uns 3.400 Hz. D'aquesta manera, si posem un filtre de pas de banda amb aquestes freqüències de tall podrem eliminar els sorolls que tinguin freqüències fora d'aquests límits.

Figura 77. Exemple de resposta d'un filtre de pas de banda, caracteritzat per dues freqüències de tall



Només els components que estiguin compresos entre aquestes dues freqüències podran passar a través del filtre.

Figura 78. Exemple d'un filtre de pas de banda obtingut mitjançant la connexió en sèrie d'un filtre passabaix i un filtre de pas de banda.



Vegem l'efecte d'un filtre de pas de banda sobre un tema musical amb veu cantada. Si voleu, podeu capturar amb l'Audacity un fragment qualsevol de la vostra música favorita. Nosaltres farem l'exercici amb un fragment del tema *El marit de la perruquera*, de Pedro Guerra. Podeu trobar aquest arxiu en format WAV amb el material addicional associat a aquesta activitat.

Per començar obrirem l'arxiu **PedroGuerra_1.wav** amb l'Audacity i l'escoltarem en la versió original. En la figura 79 es mostra la forma d'ona del fitxer que es representa en la pantalla de l'aplicació. Després procedirem al filtratge. Aconseguir un bon efecte de filtratge de pas de banda per a aïllar la veu humana és un procés que requereix força paciència i que a més solament funcionarà de manera aproximada, ja que, d'una banda, els filtres no són ideals, per la qual cosa no són capaços d'eliminar completament les freqüències indesitjades.

A més, i no és menys important, gran part dels instruments musicals que s'utilitzen en qualsevol enregistrament també tenen components freqüencials en la regió espectral de la veu humana. Per això, serà totalment impossible eliminar aquests components musicals del senyal.

Com a primera aproximació a fer es recomana utilitzar un filtre passabaix amb una freqüència de tall de 700 Hz i 12 dB per octava (això significa que la transició entre la banda de pas i la banda atenuada és més ràpida que la que hem utilitzat fins ara). Per al filtre passaalt hem pres una freqüència de tall de 300 Hz i també 12 dB per octava. Els paràmetres que s'han seleccionat per a cada filtre es mostren en la figura 80. No obstant això, aquests valors només són una primera aproximació. Es recomana que proveu altres freqüències de tall per veure la seva influència des del punt de vista auditiu. Recordeu que podeu trobar els filtres en l'opció *Efectos\Filtrado* i que els podeu aplicar en l'ordre que vulgueu (són intercanviables). Recordeu també que podeu cancel·lar els filtres aplicats al senyal mitjançant el menú *Edició\Deshacer*.

Figura 79. Representació de la forma d'ona del senyal PedroGuerra1.wav

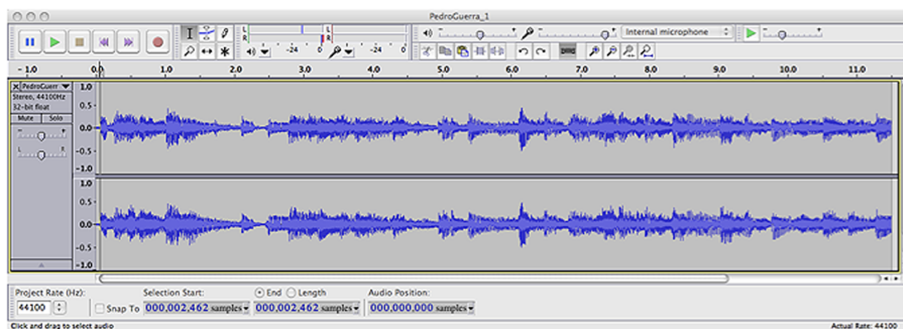
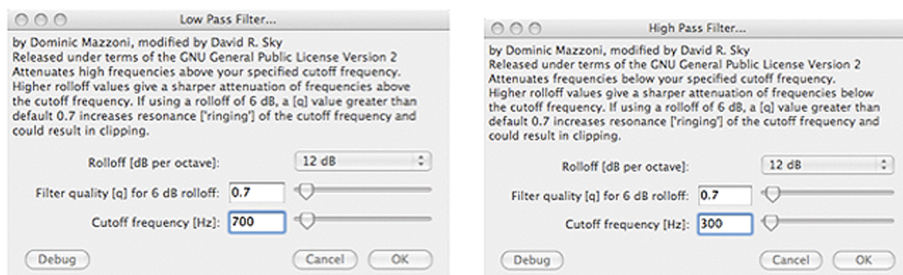


Figura 80. Exemple dels paràmetres seleccionats per als filtres passaalt i passabaix

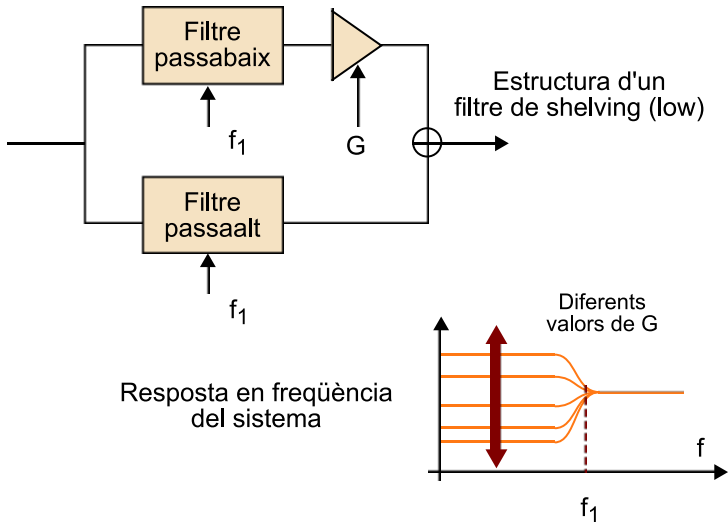


4.3.5. Filtres de *shelving*

Els filtres de *shelving* (prestatergia) són uns filtres que es poden aconseguir mitjançant la connexió en paral·lel d'un filtre passabaix amb un filtre passaalt, tots dos amb la mateixa freqüència de tall, que pot ser controlada per l'usuari. Hi ha dues configuracions típiques. En la figura 81 es mostra una configuració denominada *low shelving filter* (**filtre de *shelving* de baixa freqüència**) en la qual s'aplica un guany variable a la sortida del filtre passabaix, mentre que la sortida del filtre passaalt es manté constant. Aquesta configuració pot produir diferents respostes en funció del valor del guany tal com es mostra en la figura

81. És interessant notar que el nom de *shelving* (prestatgeria) procedeix dels diferents nivells que es poden obtenir en la banda de freqüències baixes en variar el valor del guany i que recorda una prestatgeria.

Figura 81. Representació de l'estructura d'un filtre de *shelving* com una configuració en paral·lel d'un filtre passabaix i un filtre passaalt



La configuració alternativa és el filtre de *shelving* d'alta freqüència (*high shelving filter*) que s'obté mantenint fix el guany de la branca de baixa freqüència i modificant la d'alta freqüència.

Aplicació dels filtres de *shelving*

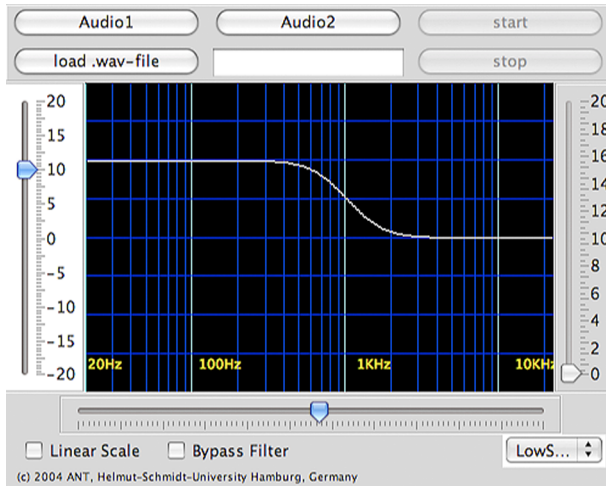
Els filtres de *shelving* s'utilitzen en moltes aplicacions com, per exemple, per al control de to mitjançant tècniques digitals. El filtre de *low shelving* s'usa com a control de to per a baixos (*bass*) mentre el de *high shelving* s'usa per al control de tons aguts (*treble*). També s'utilitzen per a filtrar el soroll d'un micròfon en la captura de la veu humana. En efecte, el filtre de *low shelving* és útil per a eliminar els *pops* (explosions que es produeixen en parlar prop d'un micròfon) o per a eliminar el *rumble* (soroll aleatori produït per alguna font de so elèctrica) o els xiulets (*hiss*).

En aquest exercici farem algunes proves amb els filtres de *shelving* utilitzant la miniaplicació de Java que ja hem usat per als filtres passabaix i passaalt. En aquest cas, hem d'activar l'aplicació `eqApplet.html`, i en la casella de selecció seleccionar els filtres `LowShelvingFilter` i `HighShelvingFilter`. El funcionament de la miniaplicació és semblant al que ja hem explicat per als filtres passabaix i passaalt. En aquest cas es disposa de dos controls de desplaçament. Un és en l'eix horitzontal de la gràfica i s'utilitza per a ajustar la freqüència de tall dels filtres. L'altre és en l'eix vertical i s'utilitza per a controlar el valor de guany del filtre passabaix (en el cas del `LowShelvingFilter`) o del filtre passaalt per a ell (`HighShelvingFilter`). La figura 82 mostra l'aspecte de l'aplicació `eqFilter` quan el `LowShelvingFilter` està seleccionat amb un valor més gran que la unitat per al guany del filtre passabaix.

Es proposa que de manera més o menys autònoma es facin diverses proves de filtratge de senyals amb aquest tipus de filtres utilitzant les mostres d'àudio que hem estat usant en aquesta activitat. Els filtres de *shelving* se solen considerar com un primer tipus de filtres equalitzadors, molt simples i dels quals només

podem controlar dos paràmetres, que són la freqüència de tall i el guany d'un dels filtres. En l'activitat següent veurem altres tipus d'equalitzadors més complexos.

Figura 82. Filtre de *shelving* passabaix fet amb l'aplicació eqFilter en què es pot comprovar que disposem de control sobre la freqüència de tall i el guany del bloc de passabaix.



4.4. Problemes/preguntes

1) En l'exercici 1 ("Exemples amb filtratge passabaix") hem capturat un senyal de veu al qual hem aplicat diversos tipus de filtres passabaix i hem vist els efectes del filtre en la representació freqüencial. No obstant això, hem vist que el senyal de veu té molt poc contingut en alta freqüència i, per tant, no és necessari mostrejar-la a 44.100 Hz. En aquest primer exercici es proposa que es repeteixin les proves fetes en aquest exercici utilitzant una freqüència de mostreig de 8.000 Hz (haureu de modificar la freqüència de mostreig del projecte en la casella inferior esquerra del programa Audacity). Representeu i compareu els espectres que s'obtenen amb el senyal filtrat i sense filtrar.

2) Genereu un senyal compost per tres tons amb les freqüències de 400 Hz, 1.000 Hz i 1.700 Hz utilitzant el programa Audacity. Recordeu el procediment descrit en l'apartat 1 per a generar diferents pistes i combinar-les en un únic senyal. Determineu les amplituds dels tres components perquè siguin aproximadament iguals però no se saturi el senyal. Es demana que feu les proves següents amb el senyal obtingut:

- a) Escolteu el senyal original amb la superposició dels tres tons.
- b) Representeu l'espectre del senyal original i observeu l'aparició dels tres tons.
- c) Feu un filtratge passabaix del senyal amb una freqüència de tall de 900 Hz i un pendent de 6 dB.

d) Feu el mateix filtratge anterior amb altres pendents (3 dB, 12 dB, etc.). Observeu els efectes audibles en modificar el pendent del filtre.

e) Determineu l'espectre del senyal obtingut en l'apartat c i comproveu la reducció en la potència dels components en 1.000 Hz i 1.700 Hz.

f) Feu un filtre de pas del senyal amb una freqüència de tall de 1.500 Hz i un pendent de 6 dB.

g) Representeu l'espectre del senyal obtingut en l'apartat f.

h) Genereu una pista d'àudio amb el contingut del filtratge de l'apartat c i una altra amb el contingut de l'apartat f. Combineu aquestes dues pistes d'àudio en una nova pista i escolteu el senyal resultant. Quins tons sentiu?

i) Representeu l'espectre obtingut en l'apartat h i comproveu que s'ha eliminat el to central.

3) Els filtres de banda eliminada són els complementaris dels filtres de pas de banda. També estan definits per dues freqüències de tall f_1 i f_2 però, en aquest cas, tots els components que hi ha entre aquestes dues freqüències s'eliminen. Així com els filtres de pas de banda s'obtenen com la connexió en sèrie d'un filtre passaalt i un filtre passabaix, els filtres de banda eliminada es poden obtenir mitjançant la connexió en paral·lel d'aquests dos tipus de filtres. En l'apartat h del problema anterior hem obtingut un filtre de banda eliminada. En aquest problema volem filtrar un senyal musical amb veu humana, reduint la presència del cantant mitjançant l'ús d'un filtre de banda eliminada. Per això es proposa que escolliu un fragment musical amb veu cantada (o utilitzeu l'arxiu **PedroGuerra1.wav** que es proporciona amb la pràctica). Feu les operacions següents per fer un filtre de banda eliminada:

a) Carregueu en l'Audacity el fitxer que voleu processar.

b) Genereu dues pistes noves i copieu l'arxiu original en cada una.

c) Apliqueu un filtre passabaix amb una freqüència de tall d'entre 100 Hz i 120 Hz aproximadament i 12 dB per octava a la primera de les pistes noves que heu generat.

d) Apliqueu un filtre passaalt a l'altra pista, amb una freqüència de tall d'uns 4.000 Hz i també amb 12 dBs per octava.

e) Poseu en *mute* la pista original i escolteu la superposició de les altres dues.

f) Determineu l'espectre de la superposició d'aquests dos senyals i compareu-lo amb l'espectre del senyal original.

g) Compareu el senyal original amb l'obtingut en l'apartat *i*.

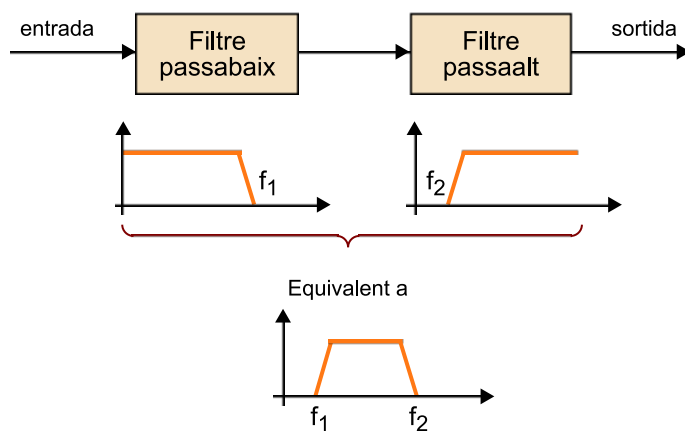
4.5. Avaluació

1) Determineu i representeu de manera esquemàtica el filtre que obteniu quan es connecten en sèrie un filtre passabaix amb freqüència de tall $f_1 = 3.000$ Hz amb un filtre passaalt amb freqüència de tall $f_2 = 500$ Hz.

Solució

S'obté un filtre de pas de banda que permet el pas dels components freqüencials que estan situats entre 500 Hz i 3.000 Hz. El resultat es mostra de manera esquemàtica en la figura 83.

Figura 83. Exemple de la concatenació d'un filtre passabaix i un filtre passaalt per a l'obtenció d'un filtre de pas de banda (problema 1)



2) Determineu i representeu de manera esquemàtica el filtre que obteniu quan es connecten en sèrie un filtre passabaix amb freqüència de tall $f_1 = 300$ Hz amb un filtre passaalt amb freqüència de tall $f_2 = 1.500$ Hz.

Solució

En principi, en aquest cas, si els filtres fossin ideals no s'obtindria cap senyal en la sortida. En efecte, el primer filtre elimina tots els components que estiguin per sobre de 300 Hz, per la qual cosa l'entrada al segon filtre no conté cap component superior a 1.500 Hz, que són els que hi podrien passar. A la pràctica, com que els filtres no són ideals obtindrem un petit residu de senyal.

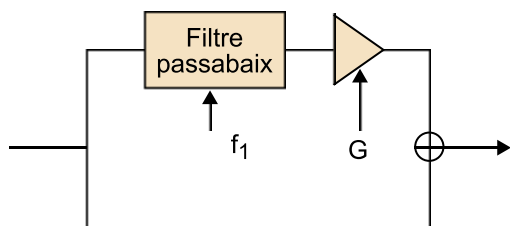
3) Determineu i representeu de manera esquemàtica el filtre que obteniu quan es connecten en sèrie un filtre passabaix amb freqüència de tall $f_1 = 800$ Hz amb un filtre passaalt amb freqüència de tall $f_2 = 500$ Hz.

Solució

En aquest cas obtenim un filtre passabaix amb una freqüència de tall de 500 Hz, la més petita de les dues. En efecte, els components amb una freqüència inferior a 500 Hz poden passar a través dels dos filtres, mentre que els components superiors són eliminats per algun dels filtres o tots dos. Resumint, en connectar en sèrie dos filtres passabaix obtenim un nou filtre passabaix amb una freqüència de tall igual a la mínima dels dos filtres. Anàlogament podem raonar que en connectar en sèrie dos filtres passaalt obtenim un filtre passaalt amb una freqüència de tall igual al màxim de les dues.

4) Raonar quin tipus de filtre podem obtenir utilitzant l'estructura representada en la figura adjunta quan els valors de guany es poden modificar entre -1 i 9 .

Figura 84



Solució

S'obté un filtre de *shelving* de baixa freqüència. Aquesta és una estructura alternativa per a obtenir filtres de *shelving*. En efecte, quan el guany és -1 , els components de baixa freqüència que provenen de la branca superior es cancel·laran amb els components que vénen de la branca inferior, i s'obtindrà un senyal nul en la regió de baixa freqüència (situat per sota de la freqüència de tall). En canvi, quan el guany de la branca superior sigui de 9 , la suma amb la branca inferior donarà un guany total de 10 en la zona de baixa freqüència. El filtre de *shelving* tindrà, doncs, una corba de variació de guany de la part de baixa freqüència que pot oscil·lar entre 0 i 10 .

5. Activitat 5. Equalització de senyals d'àudio

5.1. Introducció

5.1.1. Contingut

Aquesta activitat s'ha de considerar com una segona part de l'activitat de filtratge de senyals en la qual s'aprofundeix en el concepte d'equalització i es presenten els dos tipus d'equalitzadors més utilitzats: els **equalitzadors gràfics** i els **paramètrics**. Els dos tipus d'equalitzadors són introduïts com a models evolucionats de filtres, que permeten que l'usuari controli la forma de la resposta en freqüència i que poden ser interpretats com a associacions en sèrie o en paral·lel de filtres bàsics. En tots dos casos, l'estudiant es familiaritza amb el funcionament d'aquests sistemes utilitzant eines de programari, la major part de les quals ja s'ha utilitzat en altres activitats.

Per a comprendre els filtres paramètrics és necessari introduir els conceptes de freqüència central, guany i amplada de banda com a paràmetres de control d'un filtre de pas de banda. La dependència de l'amplada de banda amb la freqüència central s'explica mitjançant la definició del factor de qualitat i s'experimenta mitjançant el filtratge de senyals de mostra. També es proporcionen diversos exemples d'aplicació en els quals s'utilitzen els equalitzadors per a millorar la qualitat del senyal d'àudio i s'introdueixen models comercials d'equalitzadors paramètrics.

Els equalitzadors gràfics són més intuïtius, ja que generalment s'incorporen a molts equips d'àudio domèstic o aplicacions informàtiques de reproducció musical. En aquest cas es presenten els conceptes bàsics de descomposició de l'equalitzador com un banc de filtres i es determinen les freqüències centrals de cada filtre en funció del nombre de bandes de l'equalitzador.

5.1.2. Metodologia

S'utilitza la mateixa metodologia que s'ha estat utilitzant en activitats anteriors i que cerca que l'estudiant pugui anar comprovant els conceptes exposats mitjançant l'audició dels efectes del filtratge dels senyals. En aquesta activitat es continua utilitzant la mateixa miniaplicació de Java que ja s'ha utilitzat en l'activitat anterior però ara amb l'opció de **PeakFilter** que simula un equalitzador paramètric. També s'utilitza una nova miniaplicació de Java que simula un equalitzador gràfic d'octaves i es continua utilitzant el programa Audacity

per a la captura, edició i processament dels senyals d'àudio. Al llarg de la pràctica es proposen diversos exercicis que s'han de fer utilitzant els fitxers d'àudio que es proporcionen.

Comentem que la major part d'aquests exercicis no té solucions tancades, ja que es tracta d'ajustar els equalitzadors perquè la percepció del senyal d'àudio millori i això, en definitiva, és un aspecte subjectiu. Amb aquests exercicis només s'espera que l'estudiant s'acostumi al maneig de diferents equalitzadors sense pretendre convertir-se en un expert en condicionament de so.

Al final de l'activitat es proposen alguns exercicis addicionals perquè l'estudiant els resolgui amb més autonomia i dels quals s'espera que interpreti correctament els resultats obtinguts.

Finalment, es proposen diversos problemes semblants als que han aparegut com a exemples al llarg del text i s'inclouen les solucions.

5.1.3. Recursos

Els recursos principals utilitzats en aquesta activitat són:

- Programari de captura i processament d'àudio Audacity (licència lliure).
- Programari de demostració de filtres i equalitzadors basat en miniaplicacions de Java.
- Anàlisi de forma d'ona i anàlisi de freqüència dels senyals mitjançant el programa Audacity.
- Captura de senyals de veu per a la seva equalització.
- Arxius d'àudio predigitalitzats amb diferents característiques.
- Explicacions dels fonaments de l'equalització i les diferències amb el filtratge.
- Explicacions de les aplicacions principals dels equalitzadors.
- Proposta d'activitats de cerca d'informació en webs i recursos d'Internet per a aprofundir en els coneixements adquirits en aquesta activitat (aplicacions de programari d'equalització, fabricants d'equalitzadors, etc.).
- Activitats proposades per a verificar l'adquisició dels coneixements d'aquesta activitat.

5.2. Objectius

Els objectius principals que es persegueixen en aquesta activitat són:

- 1) Introduir els conceptes clau utilitzats per a l'equalització de senyals d'àudio.
- 2) Definir el concepte de filtre paramètric i familiaritzar l'estudiant amb el control dels paràmetres de guany, freqüència central i amplada de banda.
- 3) Relacionar, mitjançant experiments dirigits, els efectes d'un equalitzador paramètric en la qualitat i la percepció sonora d'un senyal d'àudio.
- 4) Conèixer les aplicacions dels equalitzadors en el context de la captura, reproducció i tractament del senyal d'àudio.
- 5) Introduir els diagrames de blocs d'un equalitzador paramètric en funció de filtres de pas de banda.
- 6) Conèixer productes i equips d'equalització comercials.
- 7) Introduir el concepte dels filtres *notch*, com una solució per a atenuar els efectes de l'acoblament de senyals d'àudio.
- 8) Definir el factor de qualitat d'un filtre i comprendre el concepte de filtres amb factor de qualitat constant.
- 9) Introduir el diagrama de blocs d'un equalitzador gràfic com un banc de filtres de pas de banda.
- 10) Proporcionar equacions necessàries per a calcular les freqüències centrals d'un equalitzador gràfic per a qualsevol nombre de bandes.

Aquests objectius estan relacionats amb les **competències de l'assignatura** següents:

- 1) Capacitat de capturar àudio i imatges de manera eficient i eficaç.
- 2) Capacitat de digitalitzar eficientment i eficaçment continguts d'àudio i imatge.
- 3) Capacitat d'operar digitalment amb senyals digitals d'àudio i imatge.
- 4) Capacitat d'interpretar i modificar informació auditiva de manera freqüencial.

I amb les **competències generals del grau** següents:

1) Ser capaç d'analitzar un problema en el nivell d'abstracció adequat a cada situació i aplicar habilitats i coneixements adquirits per a abordar-lo i resoldre'l.

2) Capturar, emmagatzemar i modificar informació d'àudio, imatge i vídeo digitals aplicant principis i mètodes d'elaboració i composició del llenguatge audiovisual.

3) Integrar i gestionar continguts digitals en aplicacions multimodals d'acord amb criteris estètics, tècnics i funcionals.

4) Atendre adequadament consultes sobre projectes, tecnologies i mercat de productes multimèdia avaluant de manera precisa l'entorn d'aplicació, els recursos i les alternatives tecnològiques disponibles.

5.3. Guia

En l'activitat anterior hem estudiat diversos tipus de filtres. Els filtres constitueixen un dels mètodes més importants per a **processar i condicionar els senyals d'àudio**. Hem vist que hi ha dos tipus de filtres bàsics a partir dels quals podem aconseguir altres tipus de filtres més complexos mitjançant les associacions en sèrie i en paral·lel.

Aquests filtres bàsics són el passabaix i el passaalt. Tots dos estan caracteritzats per un únic paràmetre: la **freqüència de tall**. El filtre passabaix permet que tots els components del senyal d'àudio situats per sota d'aquesta freqüència de tall passin a través del filtre, mentre que els components de freqüència superior són eliminats. El filtre passaalt fa l'efecte invers, deixant passar els components d'alta freqüència i eliminant les baixes.

Aquests dos filtres bàsics tenen la freqüència de tall com a únic paràmetre de control. No obstant això, quan associem aquests filtres en sèrie o en paral·lel²³ podem obtenir més paràmetres de control. Així, hem vist que quan connectem en sèrie un filtre passaalt i un passabaix podem obtenir un **filtre de pas de banda**. El nou filtre de pas de banda té dos paràmetres de control: la **freqüència de tall inferior i la superior**. També hem vist que els filtres de *shelving*, que s'obtenien com l'associació en paral·lel d'un filtre passabaix i un passaalt, tenien dos paràmetres de control: la freqüència de tall i el guany d'una de les bandes (que anomenàvem *la prestatgeria*). Recordem que els controls de to de baixos i aguts (*bass i treble*) que es proporcionen en la majoria d'equips d'àudio es poden veure com dos filtres de *shelving* en els quals l'usuari té el control sobre el guany d'una de les bandes i la pot augmentar o atenuar.

⁽²³⁾ Les associacions de filtres en sèrie i en paral·lel s'han explicat en l'activitat 4 d'àudio.

En aquesta activitat ens centrarem en uns tipus de filtres una mica més complexos, que proporcionen més flexibilitat per al condicionament dels senyals d'àudio gràcies a un nombre superior de paràmetres que l'usuari pot controlar. En general, aquests filtres reben el nom d'**equalitzadors**. Tal com ja hem comentat, tots els equalitzadors es poden obtenir associant filtres bàsics.

Els dos tipus d'equalitzadors més utilitzats són els **paramètrics** i els **gràfics**, i n'analitzarem els principis de funcionament, hi experimentarem i veurem algunes de les seves aplicacions en els apartats següents. Els filtres de *shelving* i els filtres de *notch* també es denominen, en alguns textos, *equalitzadors*. Els primers ja s'han estudiat en l'activitat anterior. En aquesta activitat veurem que els filtres de *notch* es poden considerar com un cas particular d'un filtre paramètric.

5.3.1. Equalitzadors paramètrics

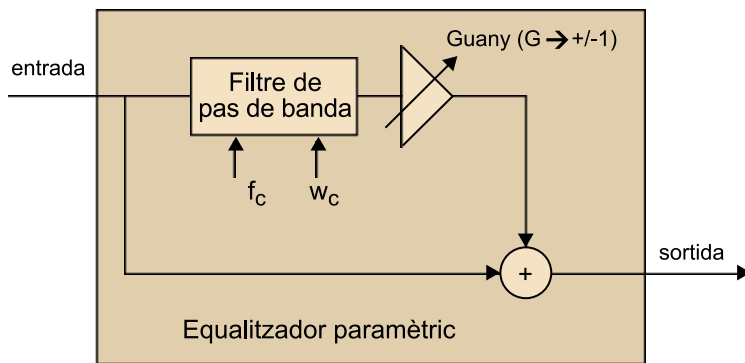
Un filtre paramètric és un filtre que es pot obtenir a partir d'un filtre de pas de banda utilitzant una estructura com la indicada en la figura 85.

La particularitat del filtre és que permet que l'usuari controli, de manera independent, els paràmetres següents del filtre:

- a) El guany (G)
- b) La freqüència central del filtre de pas de banda (f_c)
- c) L'amplada de banda del filtre de pas de banda (W)

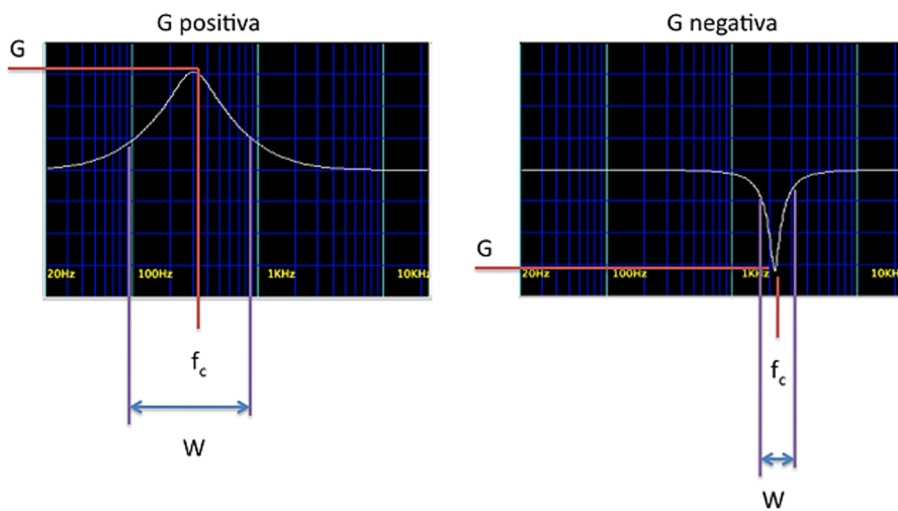
En la figura 86 es mostren alguns exemples de les corbes de resposta d'un filtre paramètric per a diferents valors dels paràmetres de disseny. En general, les corbes que obtenim tenen una forma de campana, que pot ser cap amunt o cap avall en funció del valor del paràmetre G . Quan G és positiu, la campana és cap amunt, mentre que quan G pren valors negatius la campana és cap avall. Fixeu-vos que quan el valor de G és igual a 0 la branca del filtre de pas de banda en la configuració de la figura 85 no té cap efecte i, per tant, el filtre tindrà una resposta totalment plana (no es modifica el senyal). El control de la freqüència central f_c modifica la posició de la campana en l'eix horitzontal, mentre que el control de l'amplada de banda W controla l'amplada de la finestra. En la figura 86 es mostren dues possibles respostes en freqüència obtingudes per a diferents valors dels paràmetres de control G , f_c i W .

Figura 85. Esquema d'un filtre equalitzador paramètric representat en funció d'un filtre de pas de banda



Els paràmetres que l'usuari pot controlar són el guany (G), la freqüència central (f_c) i l'amplada de banda (W)

Figura 86. Exemple de les respostes en freqüència dels filtres paramètrics per a diferents valors dels paràmetres



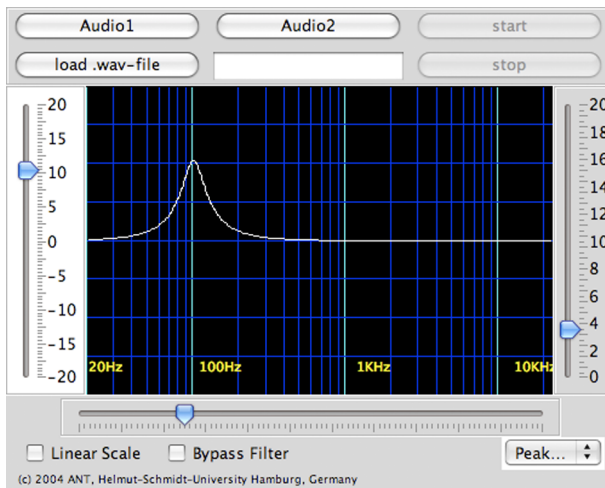
En l'exemple de l'esquerra tenim un valor de G positiu, que dona lloc a un filtre amb forma de campana no invertida. La gràfica de la dreta correspon a una G negativa i dona lloc a una forma de campana invertida.

En tot cas, possiblement la millor manera de comprendre els filtres paramètrics és utilitzant-los i escoltant els efectes que produeixen en diferents senyals d'àudio.

Per a veure l'efecte d'aquests filtres sobre exemples d'àudio usarem la mateixa miniaplicació²⁴ de Java que hem utilitzat en l'activitat anterior per a il·lustrar els efectes dels filtres passabaix, passaalt i *shelving*. En aquest cas, executem de nou la miniaplicació **eqApplet.html** i seleccionem el filtre PeakFilter. Veurem que apareix una pantalla com la que es mostra en la figura 87 i que proporciona tres controls independents a l'usuari. Els controls són en les barres verticals esquerres (G , guany), dreta (W , amplada de banda) i en la barra de desplaçament horitzontal (f_c , freqüència central).

⁽²⁴⁾Aquesta miniaplicació ha estat desenvolupada pel grup de Processament de Senyal de la Universitat Helmut-Schmidt d'Hamburg sota la direcció del professor Udo Zölzer.

Figura 87. Miniaplicació de Java amb la implementació d'un filtre paramètric (PeakFilter)



Per a familiaritzar-se amb els filtres paramètrics se suggereix que s'actui sobre els diferents controls mentre s'està reproduint algun dels senyals d'àudio que se subministren per defecte (botons *Audio 1* i *Audio 2*), amb això es pot tenir una primera aproximació de com cada un dels paràmetres afecta el senyal d'àudio. Mentre es fan aquestes proves es poden veure diversos aspectes que poden ajudar a comprendre algunes característiques dels filtres paramètrics i que es resumeixen a continuació:

a) Quan el guany del filtre està situat en un valor 0 la resposta del filtre és completament plana. En efecte, quan posem aquest guany a 0 es desactiva la branca superior del filtre de pas de banda representat en la figura 85, de manera que el sistema deixa passar el senyal d'entrada sense fer-hi cap modificació. Fixeu-vos que encara que es modifiquin els valors de la freqüència central (barra de desplaçament inferior) o de l'amplada de banda (barra de desplaçament vertical de la dreta) no obtenim cap efecte.

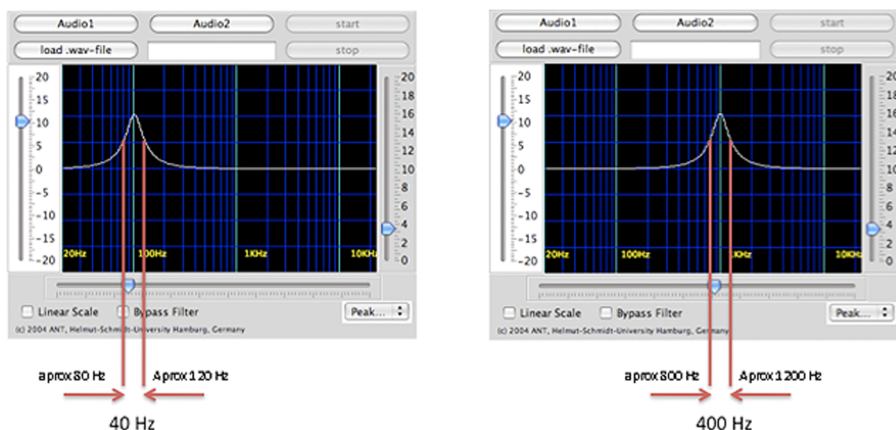
b) Quan el guany del filtre és positiu, la resposta del filtre és sempre de campana. En canvi, quan el guany és negatiu obtenim una campana invertida.

c) L'amplitud de la campana es controla mitjançant la barra de desplaçament vertical dreta. Els valors numèrics que apareixen al costat de la barra vertical representen el **factor de qualitat** del filtre, denominat Q . El factor de qualitat és un paràmetre directament relacionat amb l'amplada de banda del filtre, que definirem més endavant. De moment, n'hi ha prou d'observar que en augmentar el factor de qualitat disminueix l'amplada de banda del filtre, i s'obté un filtre molt més estret.

d) Quan mantenim fixos el guany (G) i l'amplada de banda (factor de qualitat) del filtre però desplaçem la freqüència central (f_c) observem que, d'acord amb l'esperat, la campana del filtre es desplaça en freqüència. No obstant això, hauríem d'observar que l'eix de freqüències està representat en escala logarítmica, per la qual cosa si l'amplada de la campana se'ns manté constant

significa que l'amplada de banda real del filtre augmenta amb la freqüència central. Veurem amb més detall aquest efecte quan analitzem la definició del factor Q . Per ara, fixeu-vos que en la figura 88 es representen dos diagrames del filtre en els quals el guany i el factor Q són constants i en els quals només hem variat la freqüència central del filtre. Tal com es veu en aquesta ultima figura, l'amplada de banda del filtre quan la freqüència està centrada en 100 Hz és aproximadament de 40 Hz, mentre que quan està centrada en 1.000 Hz és de 400 Hz (aquests valors s'han de prendre només com a aproximacions calculades sobre les gràfiques).

Figura 88. Comparativa entre les amplades de banda del filtre paramètric mantenint el factor Q constant i usant diferents freqüències centrals.



e) Quan posem el guany al valor mínim (barra de desplaçament vertical esquerra en l'extrem inferior) i el factor de qualitat al màxim (barra de desplaçament vertical dreta en l'extrem inferior), obtenim el que es coneix com un **filtre notch**, és a dir, un filtre que elimina una freqüència concreta. Podem modificar el valor d'aquesta freqüència actuant sobre la freqüència central del filtre.

1) En aquest exercici utilitzarem un equalitzador paramètric per a millorar la qualitat d'un senyal de veu. Per a això es proposa:

a) Digitalitzeu un fragment de la vostra pròpia veu utilitzant el programa Audacity. Si és possible, utilitzeu directament els micròfons integrats a l'ordinador (en aquest exercici, no es tracta d'obtenir un registre excel·lent sinó de comprovar que, si el registre és de mala qualitat, amb una equalització correcta el resultat es pot millorar una mica). El temps de registre hauria de ser aproximadament d'1 minut, per a disposar de temps suficient per a ajustar els paràmetres de l'equalitzador i avaluar la qualitat del so.

b) Exporteu el senyal registrat en el format WAV (Microsoft) *signed 16 bits*.

c) Activeu la miniaplicació de Java **eqApplet.html** i seleccioneu el filtre **PeakFilter**. Seleccioneu el botó **load .wav file** i seleccioneu l'arxiu d'àudio que acabeu d'exportar.

d) Escolteu unes quantes vegades el fragment d'àudio i actueu amb els diferents paràmetres del filtre fins a obtenir una equalització que us resulti satisfactòria.

Nota

A causa de la gran variabilitat dels equips de captura i la subjectivitat de la mateixa equalització, no podem proporcionar una solució a aquest exercici. És una solució vàlida qual-sevol equalitzador paramètric que introdueixi una millora subjectiva en la percepció del senyal de veu que ha registrat.

2) Ara prendrem un fragment musical de guitarra clàssica i intentarem equalitzar-lo. En aquest cas, l'arxiu original ha estat preprocessat per a intentar simular un problema que tenen alguns instruments de corda, i és que la caixa de ressonància pot amplificar unes freqüències més que altres, de manera que s'obté un instrument amb un so que no està correctament equilibrat. Per a compensar aquest efecte se solen usar filtres paramètrics.

L'objectiu del filtre paramètric és compensar l'efecte de la caixa de ressonància, atenuant les freqüències en què es produeix la ressonància de l'instrument.

3) Carregueu l'arxiu **Alhambra1.wav** que es proporciona juntament amb el material d'aquesta activitat en la miniaplicació de Java i intenteu, després d'uniques audicions, determinar els paràmetres d'un filtre equalitzador que condicione de manera correcta el senyal d'àudio. Igual que amb l'exercici anterior, la solució final pot no resultar evident per als que no siguin experts musicals. No us preocupeu, l'objectiu és que us familiaritzeu amb els efectes dels paràmetres d'aquest tipus de filtres i que vegeu com afecten la percepció del so. No pretenem que amb aquests simples exercicis els estudiants es converteixin en prestigiosos enginyers de so. Per a donar una pista sobre quin podria ser el resultat final, es proporciona l'arxiu d'àudio **Alhambra2**, que realment es correspon amb el senyal original de *Records de l'Alhambra* de Francisco Tárrega enregistrat pel músic Pepe Romero.

4) Ara canviarem la guitarra clàssica per l'elèctrica i veurem com es poden utilitzar els filtres paramètrics per a generar efectes d'àudio. Carregueu el fitxer **SRV_Pride&Joy1.wav** en la miniaplicació de Java. El fitxer correspon a un fragment introductori del tema *Pride and Joy* interpretat per Stevie Ray Vaughan. Els efectes de filtratge sobre senyals d'àudio s'utilitzen en rock i en música electrònica. Essencialment consisteixen a modificar algun paràmetre d'un filtre en el temps, fent-ho oscil·lar entre dos valors amb la mateixa cadència que la música. En aquest exemple, pot ser particularment interessant seleccionar un filtre amb un guany positiu, pròxim al valor màxim i un factor de qualitat de valor aproximadament 4. Amb aquests paràmetres fixos desplaçeu el cursor de la freqüència central del filtre al ritme de la música i vegeu quin

és l'efecte de filtratge que es produeix. Provoqueu un altre tipus d'efectes de modulació per filtratge actuant sobre els paràmetres del filtre i modificant-ne un en el temps.

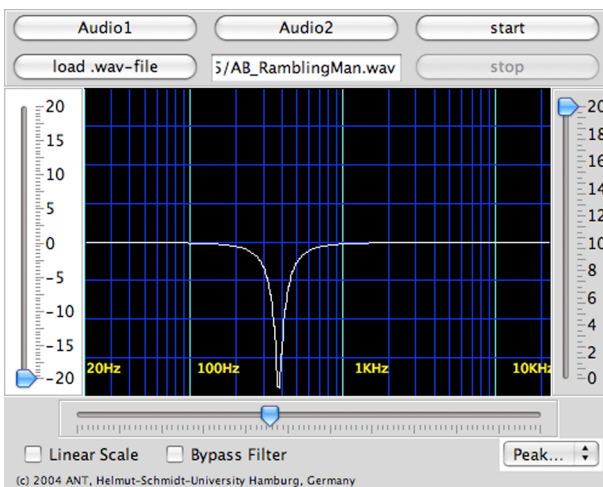
5) Finalment, veurem una de les aplicacions principals dels filtres de *notch*, que és la reducció de l'efecte Larssen. Aquest fenomen va ser descrit per primera vegada per l'enginyer danès Søren Larssen, que va explicar les causes que el provoquen.

Es tracta d'un efecte característic que es produeix per un problema de realimentació entre la font d'emissió i la font de reproducció del so.

És habitual que es produeixi en locals amplificats quan el micròfon captura el so emès pels altaveus i, al seu torn, aquest so s'amplifica i es torna a enviar als altaveus. Informalment, es coneix amb el nom d'*acoblament* (el micròfon s'acobla!).

El fenomen també pot produir-se amb el *pickup* (sensor de vibració) d'una guitarra elèctrica, que en aquest cas substitueix al micròfon. En tot cas, el resultat és l'aparició d'un to molest que és habitual en concerts de rock, sales de festes o conferències. La manera més habitual de reduir aquest efecte és utilitzant un filtre *notch*. Els filtres *notch* són filtres que tenen una resposta pràcticament plana en totes les freqüències excepte en una freqüència concreta, que és la que volem eliminar. Podem aconseguir un filtre *notch* amb miniaplicació de Java `eqApplet.html` posant el guany del filtre al mínim (màxim valor negatiu) i l'amplada de banda del filtre també al mínim. La freqüència que volem eliminar es pot seleccionar mitjançant el paràmetre de selecció de la freqüència central. En la figura 89 es mostra la resposta d'un filtre *notch* obtinguda amb la miniaplicació de Java.

Figura 89. Exemple d'un filtre *notch*



En aquest exercici carregarem el fitxer **AB_RamblingMan.wav**, que conté un registre del tema musical *Rambling Man* interpretat pels Allman Brothers. Aquest registre ha estat manipulat i s'hi ha inserit un to molest l'amplitud del qual hauria de ser reduïda pel filtre *notch*. Se suggereix que s'escolti el fitxer diverses vegades, activant i desactivant el filtre fins a trobar la freqüència que redueix millor l'amplitud del to.

És important notar que en el nostre exemple el to no desapareix completament. Això és perquè els filtres no són mai ideals i no eliminen completament els senyals que volem. No obstant això, a la pràctica i en **directe**, aquest tipus de filtres solen ser suficients per a eliminar completament l'efecte Larssen. Penseu que quan utilitzem aquests filtres en un directe, atenuem el valor de la freqüència capturada pel micròfon, de manera que l'efecte de realimentació a què fèiem referència desapareix per complet.

5.3.2. El factor de qualitat en els filtres paramètrics

Hem comentat que els filtres paramètrics permeten controlar tres paràmetres, que són la **freqüència central**, el **guany** i l'**amplada de banda de la resposta**.

L'amplada de banda se sol controlar utilitzant un paràmetre que es denomina **factor de qualitat** i que generalment està representat amb la lletra **Q**. Així, doncs, a la pràctica, l'usuari especifica un factor de qualitat per a cada filtre. El factor de qualitat relaciona la freqüència central de la campana amb la seva amplada de banda d'acord amb l'equació següent:

$$Q = \frac{f_c}{W}$$

En què f_c representa la freqüència central del filtre, i W l'amplada de banda, tal com s'han representat en la

Per a mesurar l'amplada de banda W s'utilitzen les freqüències en les quals la campana del filtre atenua la potència del senyal a la meitat respecte al màxim.

Posarem alguns exemples per aclarir la interpretació del factor de qualitat. Suposem que l'usuari selecciona un factor de qualitat de valor unitat. En aquest cas, quan la freqüència central del filtre sigui de 100 Hz l'amplada de banda del filtre també serà de 100 Hz. Al seu torn, quan la freqüència central del filtre sigui de 1.000 Hz, la seva amplada de banda també serà de 1.000 Hz. Si el factor de qualitat fos 4, l'amplada de banda del filtre sempre seria una quarta part de la freqüència central. Així, doncs:

a) Quan el factor de qualitat és alt, l'amplada de banda del filtre es redueix.

b) Per a un valor de Q fix, l'amplada de banda del filtre és proporcional a la freqüència central del filtre. Si augmentem la freqüència central, l'amplada de banda del filtre augmenta de manera proporcional.

Aquesta relació de proporcionalitat entre la freqüència central i l'amplada de banda del filtre està especialment adaptada a les característiques del sistema auditiu. En efecte, el sistema auditiu també té un comportament basat en aquest mateix principi, ja que la diferència perceptual de canvi de freqüència entre dos senyals de 200 Hz i 250 Hz és la mateixa que es produeix entre dos tons de 2.000 Hz i 2.500 Hz. Fixeu-vos que en el primer cas les freqüències estan separades per 50 Hz mentre que en el segon ho estan per 500 Hz. Així, quan augmentem la freqüència central dels tons també hem d'augmentar la seva separació (amplada de banda) de manera proporcional. Si utilitzem dos tons de 2.000 Hz i 2.050 Hz (ara estan separats per 50 Hz), l'oïda pràcticament no apreciarà diferències significatives entre ells. La proporcionalitat entre l'amplada de banda del filtre i la freqüència central quan Q es manté constant justifica que, quan es representa el filtre en una escala logarítmica de freqüències, l'amplada de la campana es mantingui constant.

Finalment, comentem que el nom de *factor de qualitat* no ha de confondre el lector, ja que solament és una mesura de l'amplada de banda del filtre. En definitiva, una bona equalització és la que corregeix els problemes d'un sistema de registre o de reproducció d'àudio. És possible que la correcció adequada s'hagi de fer amb filtre amb un Q elevat o amb un Q baix, depenent en gran part de les causes que poden deteriorar el senyal.

5.3.3. Aplicacions dels equalitzadors paramètrics

Els filtres paramètrics es poden utilitzar per a condicionar els senyals d'àudio en diferents circumstàncies. El mateix nom d'*equalitzador* ja ens indica quines són les seves aplicacions principals.

Equalitzar significa 'igualar', i essencialment consisteix a modificar les característiques freqüencials d'un so per a tornar-hi l'equilibri original.

Així, generalment intentarem equalitzar un so quan sospitem que els equips de captura o reproducció n'han deteriorat d'alguna manera les característiques originals. En els paràgrafs següents pretenem fer un breu resum dels usos principals dels equalitzadors paramètrics.

1) **Compensar resposta en freqüència de micròfons.** Generalment els micròfons direccionals utilitzats per a la captura de veu augmenten la seva resposta a baixes freqüències a mesura que s'aproximen a la font de so. Aquest problema es coneix amb el nom d'*efecte de proximitat* i resulta en un augment desproporcionat de la quantitat de so greu a mesura que ens apropem al mi-

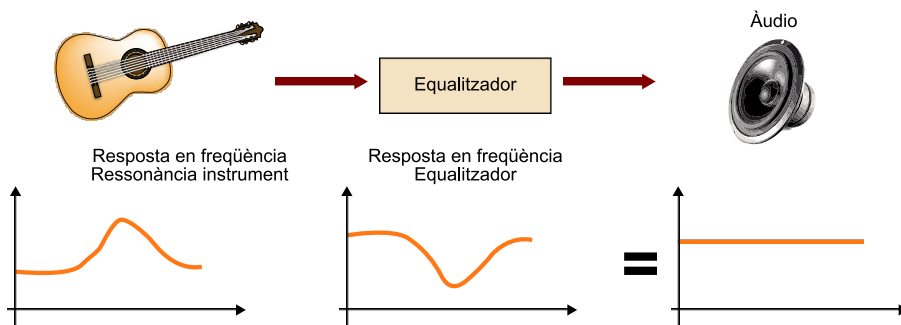
cròfon. Alguns vocalistes aprofiten l'efecte de proximitat de manera deliberada per millorar les característiques de la seva veu. Per a reduir l'efecte de proximitat es pot utilitzar un filtre passaalt (que elimini les altes freqüències) o un filtre paramètric, amb la campana invertida i que elimini part d'aquestes baixes freqüències. Encara que molts micròfons ja incorporen internament un equalitzador, no sol ser suficient i per a obtenir una qualitat màxima se sol haver d'incorporar un equalitzador de la taula de barreges.

2) Compensar les característiques direccionals dels micròfons. La majoria dels micròfons presenten una resposta direccional que depèn de la freqüència. En general, en augmentar la freqüència els micròfons es tornen més directius, i capturen principalment el so que procedeix de la direcció en què estan apuntant. Aquestes característiques no són importants quan el micròfon està situat a prop de la font de so (*close-miking*), però poden resultar molt significatives quan ens allunyem de la font (registre de música clàssica). En efecte, la pèrdua de direccionalitat en les baixes freqüències fa que es capturi millor el camp reverberant, és a dir, el so que no procedeix directament de l'instrument sinó de les reflexions de l'ona a les parets del recinte. Com a conseqüència, les altes freqüències queden perjudicades i són capturades amb menys intensitat. La solució al problema sol ser un filtre paramètric amb una campana significativa en la zona d'alta freqüència.

3) Correcció de sorolls durant la captura de senyal. Els equalitzadors paramètrics també s'utilitzen per a reduir diversos sorolls desagradables que es produeixen durant la captura de senyals. Entre aquests sorolls hem de destacar les explosions d'aire que apareixen en parlar prop d'un micròfon (*pops*), la captura de soroll procedent d'aire o vent en un micròfon, el soroll de la pua de la guitarra en tocar les cordes, el lliscament dels dits de la mà esquerra en instruments de corda, el pedal del piano, etc. Evidentment, tots aquests sorolls tenen contribucions en un marge de freqüències molt ampli, per la qual cosa no es poden eliminar completament mitjançant el filtratge, encara que sí que podem reduir-los i fer que resultin menys molestos.

4) Compensar efectes de ressonància d'instruments musicals. Una de les aplicacions principals dels equalitzadors paramètrics consisteix a compensar les possibles deficiències de ressonància en alguns instruments musicals. En efecte, en determinats instruments és possible que les caixes acústiques emfatitzin o atenuïn de manera excessiva un marge de freqüències de l'instrument respecte a un altre. Aquests problemes es poden compensar aplicant un equalitzador paramètric en el so capturat. En la figura 90 es mostra l'essència d'aquest principi. La resposta en freqüència de l'instrument representa la seva naturalesa tímbrica que en l'exemple que ens ocupa té una certa tendència al predomini d'aguts. L'equalitzador pot compensar aquesta resposta de l'instrument per a donar una resposta tímbrica més plana. En aquest sentit, l'equalitzador iguala o equilibra els diferents components freqüencials de l'instrument.

Figura 90. Exemple de l'ús d'un equalitzador paramètric per a ajustar la ressonància d'un instrument musical o modificar el timbre



Evidentment, l'equalitzador també es pot utilitzar per a modificar o ajustar el timbre d'un determinat instrument als nostres propis gustos. Modificar el timbre és modificar la contribució relativa de cada un dels components en freqüència i això es pot aconseguir utilitzant un equalitzador. Els mòduls equalitzadors paramètrics són d'ús habitual en guitarra elèctrica, i molts guitarristes els utilitzen per a adaptar les característiques tímbriques de l'instrument en funció de les peces musicals. En la figura 91 es mostra una fotografia d'un d'aquests mòduls d'equalització per a guitarra elèctrica.

Figura 91. Mòdul de filtre paramètric per a guitarra. Observeu els tres controls de guany, amplada de banda i freqüència.



El pedal s'utilitza per a activar i desactivar el filtre.

5) Reducció d'efectes de realimentació d'àudio. Ja hem comentat en el punt 4) de l'exercici 1 que una de les aplicacions més importants dels filtres de *notch* és eliminar els efectes de realimentació en sistemes d'amplificació electrònica d'àudio (efecte Larssen). En aquest cas, el filtre s'ajusta per a atenuar la freqüència a què es produeix la interferència tonal, de manera que es redueix la realimentació i s'evita l'aparició de l'efecte.

6) Alteració del balanç en la barreja d'instruments. Aquesta és una de les aplicacions més importants i justifica que les taules de barreges incorporin un nombre considerable de filtres per a poder fer l'equalització independent de cada una de les pistes d'àudio. Per a il·lustrar el concepte bàsic utilitzarem la en la qual se suposa que estem intentant fer la barreja entre un instrument solista i la resta de l'orquestra. La gràfica de l'esquerra mostra la superposició dels dos espectres i es veu que l'espectre del solista estarà emmascarat per l'orquestra. Això significa que tindrem un so apagat per la mateixa orquestra, amb un solista que sembla situar-se per darrere d'aquesta. Utilitzant filtres paramètrics podem atenuar lleugerament una banda de l'orquestra (amb un filtre paramètric de campana invertida) i augmentar aquesta mateixa banda per al solista (amb un filtre paramètric de campana no invertida), tal com mostrem en la

Ara, el solista ha guanyat presència respecte a l'orquestra, per la qual cosa el so estarà més equilibrat. En el cas concret de vocalistes se solen utilitzar filtres centrats en la freqüència de 1.500 Hz, amb uns factors de qualitat entre 5 i 7.

7) Equalitzadors paramètrics multibanda. La importància dels equalitzadors paramètrics en el condicionament dels senyals d'àudio fa que en molts casos no en tinguem prou amb un únic equalitzador paramètric i n'hàgim de connectar uns quants en cascada. A la pràctica hi ha filtres paramètrics multibanda que combinen diferents equalitzadors en un únic equip. Així, un equalitzador paramètric de cinc bandes no és més que una connexió en sèrie de cinc equalitzadors paramètrics. Si només en volem utilitzar un, posarem el guany de tots els altres a 0, i obtindrem un guany pla. Quan tots els equalitzadors actuen, el resultat del conjunt és el producte de les respostes en freqüència de cada un. Si cada equalitzador es programa amb una freqüència central diferent i les amplades de banda no se superposen, els diferents filtres es poden programar de manera bastant independent. En la figura 93 es mostra un esquema que interconnecta tres equalitzadors paramètrics en sèrie, i en la figura 94 es mostra un equip professional que disposa de cinc equalitzadors paramètrics i un filtre *notch*. Molts equalitzadors multibanda incorporen també filtres passa-baix i passa-alt, com el que es mostra en la figura 95.

Figura 92. Exemple d'una equalització per a alterar el balanç d'una barreja musical

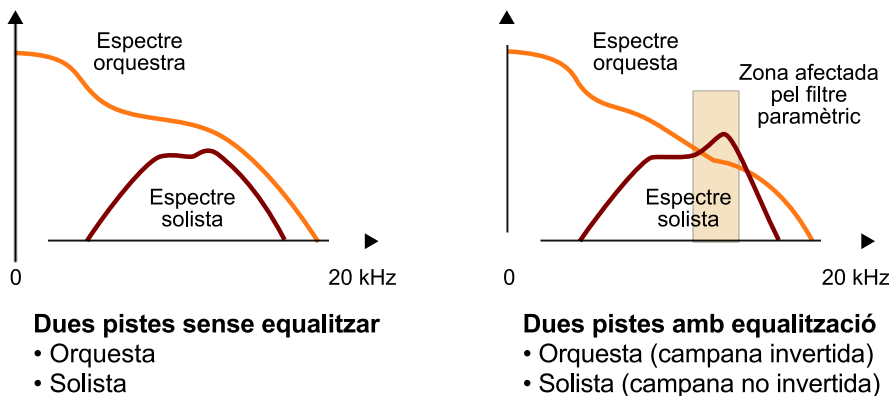


Figura 93. Exemple d'un equalitzador paramètric multibanda (tres bandes)

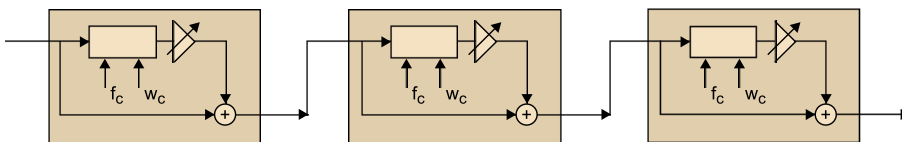


Figura 94. Equalitzador paramètric professional de cinc bandes més un filtre de notch



Figura 95. Equalitzador paramètric multibanda que incorpora un filtre passaalt i un de passabaix.



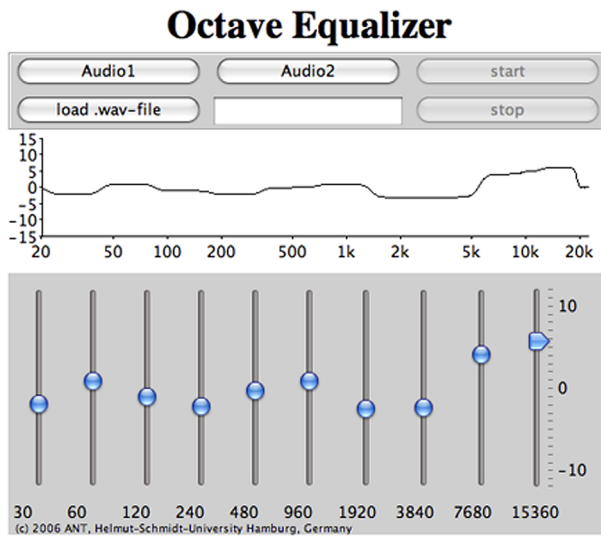
5.3.4. Equalitzadors gràfics

Els equalitzadors gràfics són una alternativa als filtres paramètrics, i també proporcionen una gran flexibilitat a l'usuari i resulten considerablement més intuïtius d'utilitzar. De fet, l'equalitzador gràfic és un element habitual en moltes cadenes d'àudio i en molts programes de reproducció musical per a ordinador. Des d'un punt de vista intuïtiu, un equalitzador gràfic ens proporciona una sèrie de controls de desplaçament que ens permeten configurar la corba de la resposta en freqüència del filtre equalitzador. Possiblement, la manera més simple de comprendre'n el principi de funcionament és utilitzant-ne un i comprovant l'efecte que té en els senyals d'àudio.

En aquest exercici utilitzarem una miniaplicació de Java per a familiaritzar-nos amb l'ús d'un equalitzador gràfic. En concret, utilitzarem una miniaplicació desenvolupada pel mateix centre²⁵ que ha desenvolupat l'`eqApplet.html`, que ja hem utilitzat per a il·lustrar els filtres i els equalitzadors paramètrics. Per a iniciar l'execució de la miniaplicació s'ha de descomprimir el contingut del fitxer `octaveEq.zip` en una carpeta de l'ordinador i executar l'arxiu `octaveEqApplet.html`. Haurem de donar permís al navegador per a executar el codi de la miniaplicació i apareixerà una pantalla com la que es mostra en la figura 96

⁽²⁵⁾Recordem que es tracta del Departament de Processament de Senyal de la Universitat Helmut-Schmidt d'Hamburg, dirigit pel professor Udo Zölzer.

Figura 96. Equalitzador gràfic d'octaves proporcionat per la miniaplicació de Java octaveEqApplet.html



El maneig d'un equalitzador gràfic és molt intuïtiu. En el nostre cas disposem de deu controls lliscants que cobreixen tota la banda de freqüències d'àudio. La miniaplicació ens permet veure com es modifica la resposta en freqüència de l'equalitzador a mesura que actuem sobre aquests controls. Se suggereix que es practiqui amb diferents corbes d'equalització, observant els efectes que es produeixen en els dos temes musicals d'exemple. Fixeu-vos que es tracta d'un filtre bastant efectiu, ja que cada control pot augmentar el guany en 10 dBs (potència 10 vegades superior) o atenuar-lo en 10 dBs (potència 10 vegades inferior). Per tant, cada una de les bandes admet actuar sobre la potència en una relació 1 a 100. Evidentment, quan tots els controls estan en la posició de 0 dBs tenim una resposta del filtre totalment plana, equivalent a no fer cap filtratge.

El filtre anterior es denomina **filtre d'octava** perquè les freqüències centrals de cada control tenen una relació d'octaves. En efecte, la primera banda està centrada en la freqüència de 30 Hz, la segona la trobem una octava²⁶ per sobre, és a dir, a 60 Hz, i així consecutivament, doblant cada vegada la freqüència.

Cada un dels controls actua sobre una freqüència central i sobre la seva banda de freqüències adjacents. És important notar que aquestes bandes de freqüència són més amples a mesura que augmenta la freqüència central. Com en el cas dels filtres paramètrics, es tracta de filtres en què la relació entre la freqüència central i l'amplada de banda es manté constant.

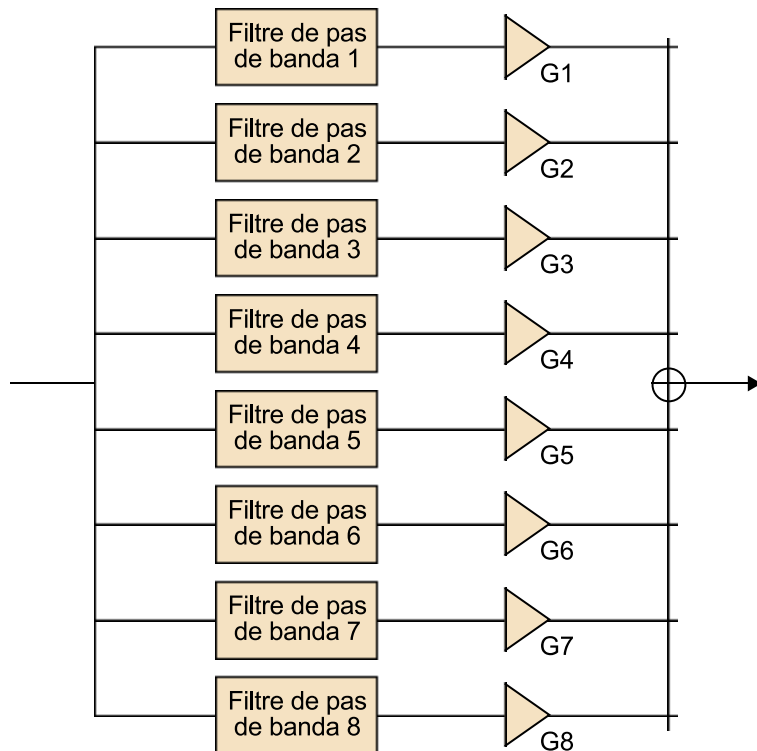
⁽²⁶⁾En terminologia musical una octava equival a doblar la freqüència. El nom es deu al fet que hi ha vuit notes musicals entre una nota determinada i la següent amb freqüència doble (do, re, mi, fa, sol, la, si, do).

5.3.5. Estructura d'un equalitzador gràfic

En la figura 97 es mostra l'estructura d'un equalitzador gràfic que s'ha descompost com una associació en paral·lel de diversos filtres de pas de banda. El senyal d'entrada es descompon en un banc de filtres de pas de banda que es controlen de manera independent mitjançant els controls de guany. El nombre de filtres que s'utilitzen es denomina *les bandes de l'equalitzador*, que per al

nostre exemple seran un total de vuit bandes. Comercialment es poden trobar equalitzadors gràfics amb una gran varietat de bandes, des de tres o cinc fins a vint o més. El nombre total de bandes determina la flexibilitat del filtre per a configurar la corba de resposta en freqüència.

Figura 97. Estructura d'un equalitzador gràfic de vuit bandes



En la figura 98 s'indica com les freqüències centrals estan distribuïdes en la freqüència d'acord amb una regla logarítmica, és a dir, la freqüència central de la banda següent s'obté multiplicant la freqüència de la banda anterior per una constant. Aquesta constant depèn del nombre de bandes, en concret, normalment, si hem de cobrir tota la banda des de 20 Hz fins a 20 kHz, pren el valor $C = 10^{3/B}$, en què B representa el nombre de bandes⁽²⁷⁾. La freqüència central del primer filtre també es pot calcular en funció d'aquest valor de C utilitzant la fórmula següent:

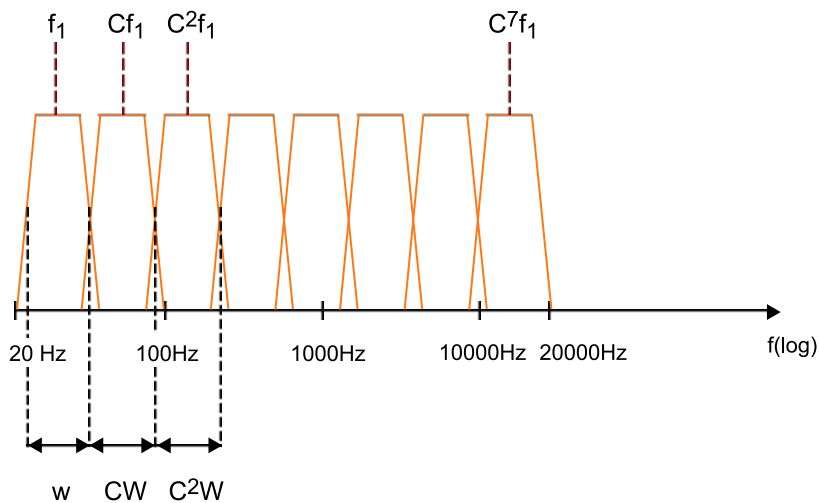
$$f_1 = 20 \cdot \sqrt{C}$$

La majoria de vegades, si apliquem directament aquesta fórmula obtenim freqüències amb molts decimals, per la qual cosa el resultat se sol aproximar a un nombre enter proper⁽²⁸⁾.

(27) Fixeu-vos que per a un equalitzador de deu bandes tindrem $C = 2$, és a dir, la freqüència central de la banda següent s'obté duplicant la freqüència central de la banda actual. Per tant, un equalitzador de deu bandes serà sempre un equalitzador d'octaves.

(28) En el cas concret de l'equalitzador d'octava del filtre de la miniaplicació, si apliquem directament la fórmula anterior obtenim una freqüència de 28,28427 Hz, que ha estat aproximada per 30 Hz

Figura 98. Representació dels filtres de pas de banda que componen un equalitzador gràfic.



En la mateixa figura 98 també podem veure com l'amplada de banda de cada filtre augmenta d'acord amb la constant multiplicadora C . La relació entre la freqüència central del filtre i la seva amplada de banda és, per tant, una constant. Per això, quan els filtres es representen utilitzant un eix de freqüències logarítmic, tots tenen la mateixa amplada.

En aquest exercici es proposa que s'experimenteu utilitzant l'equalitzador gràfic d'octaves proporcionat en miniaplicació de Java que heu utilitzat en l'exercici anterior utilitzant els diferents exemples d'àudio que s'han processat mitjançant un equalitzador paramètric en exercicis anteriors. Per a això, es proposa que dugueu a terme les experiències següents:

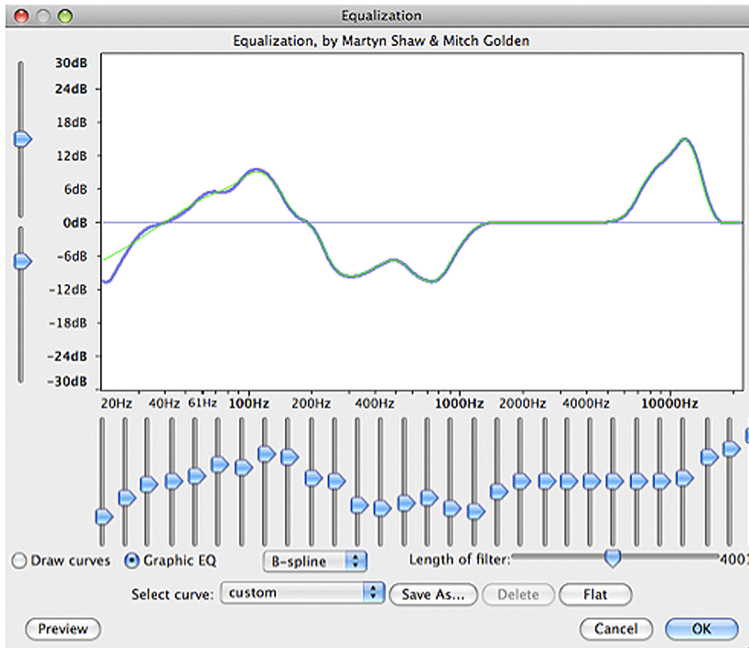
- 1) Escolteu unes quantes vegades el senyal de veu que heu enregistrat en l'exercici 1 i ajusteu els paràmetres de l'equalitzador gràfic per obtenir un so satisfactori. Compareu les corbes de resposta en freqüència que heu obtingut amb les que va obtenir en l'exercici 1.
- 2) Equalitzeu el senyal **Alhambra1.wav** fins a obtenir el timbre desitjat per a l'instrument musical.
- 3) Intenteu reduir la interferència tonal que apareix en l'enregistrament **AB_RamblingMan.wav** mitjançant aquest equalitzador gràfic.

El programa Audacity també disposa d'un equalitzador gràfic excel·lent al qual es pot accedir des del menú *Efectos\Filtros\Ecualizador*. En aquest cas, tal com mostrem en la es proporcionen un total de trenta bandes per a ajustar la corba de resposta, la qual cosa proporciona una gran flexibilitat a l'usuari. També es disposa d'un mode que permet que l'usuari dibuixi directament la corba que vol (activant la casella *Draw Curves*). Aquesta opció que el tècnic de so dibuixi directament la corba que vol obtenir sobre un terminal gràfic és cada vegada més utilitzada en les modernes taules de barreja, basades en gran part en aplicacions de programari genèriques que controlen sistemes de filtres

programables. Es proposa que en aquest exercici s'experimenti amb diversos arxius d'àudio de què disposeu per a veure la funcionalitat de l'equalitzador que es proporciona amb l'Audacity.

L'equalitzador de l'Audacity incorpora *presets* d'equalització també diferents que es poden obtenir a partir del desplegable *Select Curve* i permet que l'usuari gravi les seves pròpies corbes d'equalització per a recuperar-les en altres sessions.

Figura 99. Equalitzador gràfic de vint-i-nou bandes inclòs en l'aplicació Audacity



5.4. Problemes/preguntes

1) En els exercicis 1 i 5 hem capturat un senyal de veu al qual hem aplicat filtres d'equalització paramètrics i gràfics. No obstant això, hem vist que el senyal de veu té molt poc contingut en alta freqüència i, per tant, no és necessari mostrejar-lo a 44.100 Hz. En aquest primer exercici es proposa que es repeteixin les proves fetes en els exercicis anteriors utilitzant una freqüència de mostreig de 8.000 Hz (haureu de modificar la freqüència de mostreig del projecte en la casella inferior esquerra del programa Audacity). Representeu i compareu els espectres que s'obtenen amb el senyal equalitzat i sense equalitzar.

2) Genereu un senyal compost per tres tons amb les freqüències de 400 Hz, 1.000 Hz i 1.700 Hz utilitzant el programa Audacity (aquest senyal ja es va generar en el problema 2 de l'apartat 4, apartat "Problemes/preguntes"). Es demana que feu les proves següents amb el senyal obtingut:

a) Escolteu el senyal original amb la superposició dels tres tons.

b) Mitjançant els equalitzadors paramètrics de la miniaplicació de Java **eqApplet.html**, genereu un filtre *notch* que atenuï el senyal de 1.700 Hz i escolteu el resultat.

c) Mitjançant els equalitzadors paramètrics de la miniaplicació de Java **eqApplet.html**, genereu un filtre amb una campana de guany màxim centrat en la freqüència de 1.000 Hz i escolteu-ne els efectes.

d) Mitjançant l'equalitzador gràfic de l'aplicació **octaveEqApplet.html**, genereu un filtre que elimini els components de 400 Hz i 1.700 Hz i escolteu el resultat obtingut.

e) Compareu el resultat anterior quan feu la mateixa equalització amb el filtre equalitzador gràfic incorporat a l'Audacity.

3) Trobeu diferents fabricants, models, preus i característiques d'equalitzadors comercials per a les aplicacions següents:

a) Pedals de filtres paramètrics per a guitarra.

b) *Racks* de filtres paramètrics multibanda. Determineu-ne les característiques.

c) Mòduls d'equalitzadors gràfics per a equips domèstics d'alta fidelitat.

d) Mòduls d'equalitzadors gràfics que es puguin usar per a l'equalització de sales.

e) Taules de barreges per a discjòqueis. Identifiqueu els equalitzadors que es proporcionen per a cada canal i el tipus.

f) Taules de barreges professionals. Identifiqueu els equalitzadors/filtres que inclouen en cada canal i el seu tipus.

4) Seleccioneu una aplicació de programari de reproducció de senyals d'àudio (Winamp, iTunes, etc.) i esbrineu si permet que l'usuari configuri el seu propi equalitzador gràfic. De quantes bandes consta? Comproveu si les freqüències centrals coincideixen amb les fórmules proporcionades en la guia d'aquesta activitat. Si el programari de reproducció d'àudio conté *presets* d'equalització, comproveu-ne les característiques i escolteu l'efecte que tenen en diferents temes musicals.

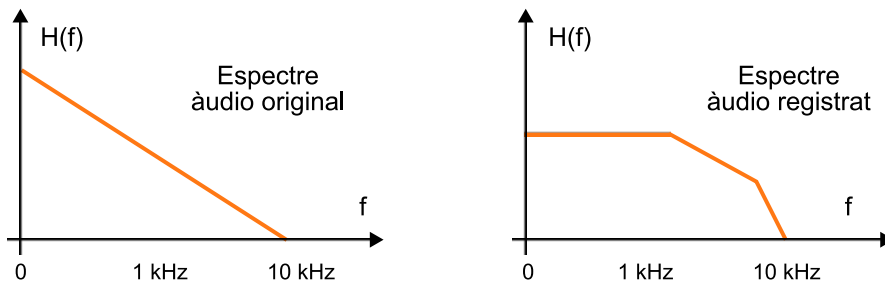
5) Si disposeu d'un reproductor d'àudio portàtil comproveu si permet que l'usuari configuri un equalitzador gràfic. De quantes bandes consta? Quines són les freqüències centrals? Conté *presets*? Les corbes les *presets* coincideixen amb les obtingudes en 4)?

6) Busqueu informació a Internet per determinar a què fan referència les diferents corbes de *preset* que es proporcionen en l'equalitzador de l'Audacity (AES, Columbia 78, RCA Victor 1938, RCA Victor 1947, Decca FFRR LP, etc.)

5.5. Avaluació

1) Suposeu que estem enregistrant una font d'àudio l'espectre de la qual té una forma triangular com la que es mostra en la figura adjunta (esquerra). Una vegada enregistrat el senyal n'analitzem l'espectre i veiem que té la forma representada en la gràfica de la dreta.

Figura 100

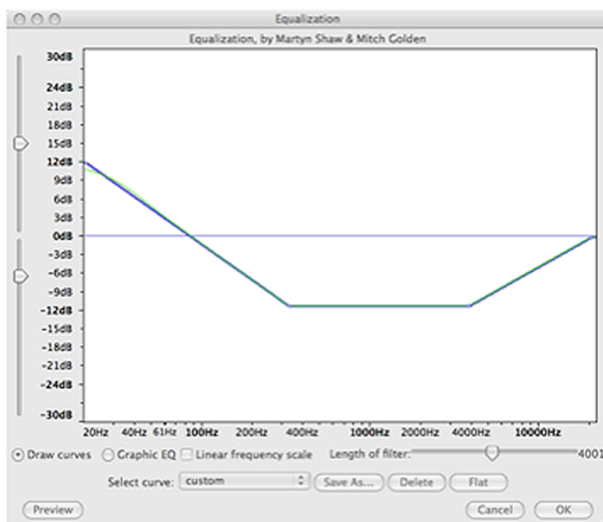


Determineu de manera aproximada quina haurà de ser la corba d'equalització del senyal.

Solució

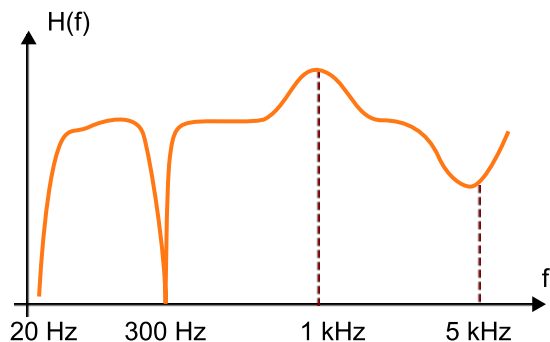
El filtre d'equalització haurà de restaurar l'espectre a la forma original, per tant, haurà de tenir una forma com la representada en la figura:

Figura 101



2) Considereu la corba d'equalització següent, que sabem que s'ha obtingut mitjançant un sistema de filtres multibanda. Indiqueu quants filtres hi han intervingut i quines característiques tenen.

Figura 102

**Solució**

El sistema està compost per quatre filtres actius:

- a) Un filtre passabaix
- b) Un filtre *notch* centrat en 300 Hz
- c) Un filtre paramètric amb guany positiu i centrat en 1 kHz
- d) Un filtre paramètric amb guany negatiu i centrat en 5 kHz

3) Determineu les freqüències centrals d'un equalitzador gràfic de cinc bandes.

Solució

D'acord amb les fórmules proporcionades $C = 10^{3/5} = 4$. Per tant, la primera freqüència prendrà el valor $f_1 = 20 \times 2 = 40$ Hz. La resta de freqüències seran de 160 Hz, 640 Hz, 2.560 Hz i 10.240 Hz.

4) Un filtre paramètric té un $Q = 6$. Determineu l'amplada de banda del filtre quan la freqüència central és de 150 Hz i quan és de 3.000 Hz.

Solució

Utilitzant la definició de $Q = f/W$ obtenim que l'amplada de banda estarà determinada pel quocient entre la freqüència central i el factor de qualitat. Així, en aquest cas tindrem una amplada de banda $150 \text{ Hz}/6 = 25$ Hz i $3.000/6 = 500$ Hz en cada un dels dos casos.

Bibliografia

Bibliografia bàsica recomanada (espanyol)

Ken C. Pohlman. *Principios del Audio Digital*. McGraw-Hill Interamericana de España.

Bibliografia bàsica recomanada (anglès)

David M. Howard; Jamie Angus. *Acoustics and Psychoacoustics* (3a. ed.). Focal Press.

Bibliografia complementària

Udo Zölzer (2008, agost). *Digital Àudio Signal Processing* (2a. ed.). John Wiley.

John Watkinson. *An Introduction to Digital Audio*. Focal Press.

José Chilitos Valenzuela. *Audio digital: Conceptos Básicos y Aplicaciones*. Backbeat Books.

