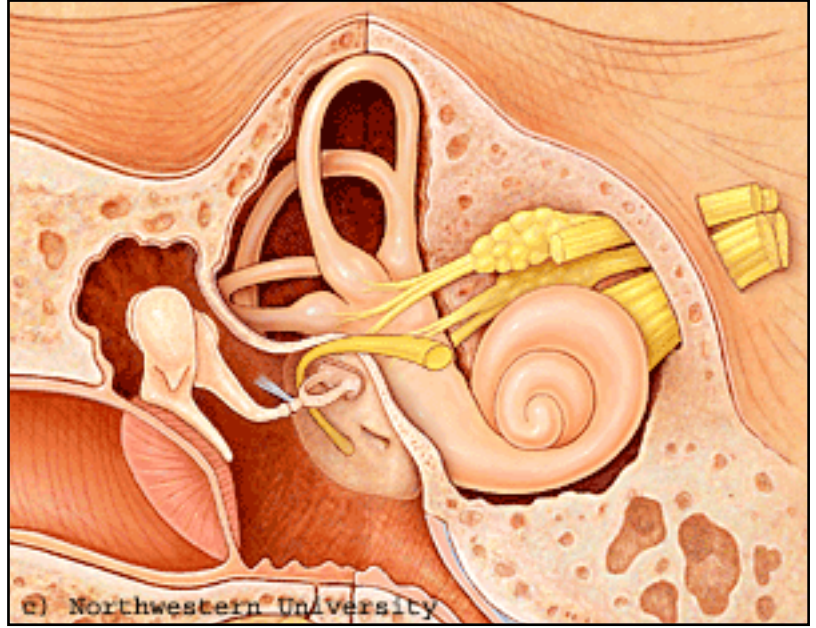


**İÜDK Müzikoloji Bölümü,
Geleneksel ve Modal Müzikler Anabilim Dalı,
“Osmanlı Dönemi Karşılaştırmalı Müzik” Programı
ODKM4089 Ses Teknolojileri I
DERS NOTLARI**

Prof. Dr. Ozan Yarman

Havada, karada, suda, 20 ila 20,000 Hertz (Hz) arasındaki titreşimler kulağımıza ses olarak gelir. Hertz, saniyedeki “atım” (pulse) sayısıdır ($f = n/s$). Atım, düzenli olarak kendini tekrar eden “ortam bozulmasıdır”.

Bin haneli rakamlar KiloHertz (kHz) olarak da ifade edilebilir ($20,000 Hz = 20 kHz$). Duyma bandının altı “sub-sonic” veya “infra-sound”, üstü ise “ultra-sonic” veya “ultra-sound” bölge olarak adlandırılır.



Kulağın işleyişini ve beyne giden sinyallerin sinirsel mekanizmasını ele alan anatomi ile fizyoloji daha çok tıbbın ve felsefenin konusu olduğundan, “bir şekilde sesi duyarız” diyelim.

Ses dalgaları ortamda madde yoksa (yani vakumda/uzay boşluğunda) ilerlemez. Dalga olması için dolu ortam gerekir. Ortamda ne kadar yoğun madde varsa ve madde ne kadar sıcaksa, o kadar hızlı ilerler ses.

Havada, 21 santigrad derecede (oda sıcaklığında), saniyede 344 metre ilerler ses dalgaları.

Suda saniyede 1.5 km hızla yayılırlar.

Katı maddelerde çok daha çabuk giderler. Örneğin çelik içinde saniyede 5 km hızla yol alırlar.

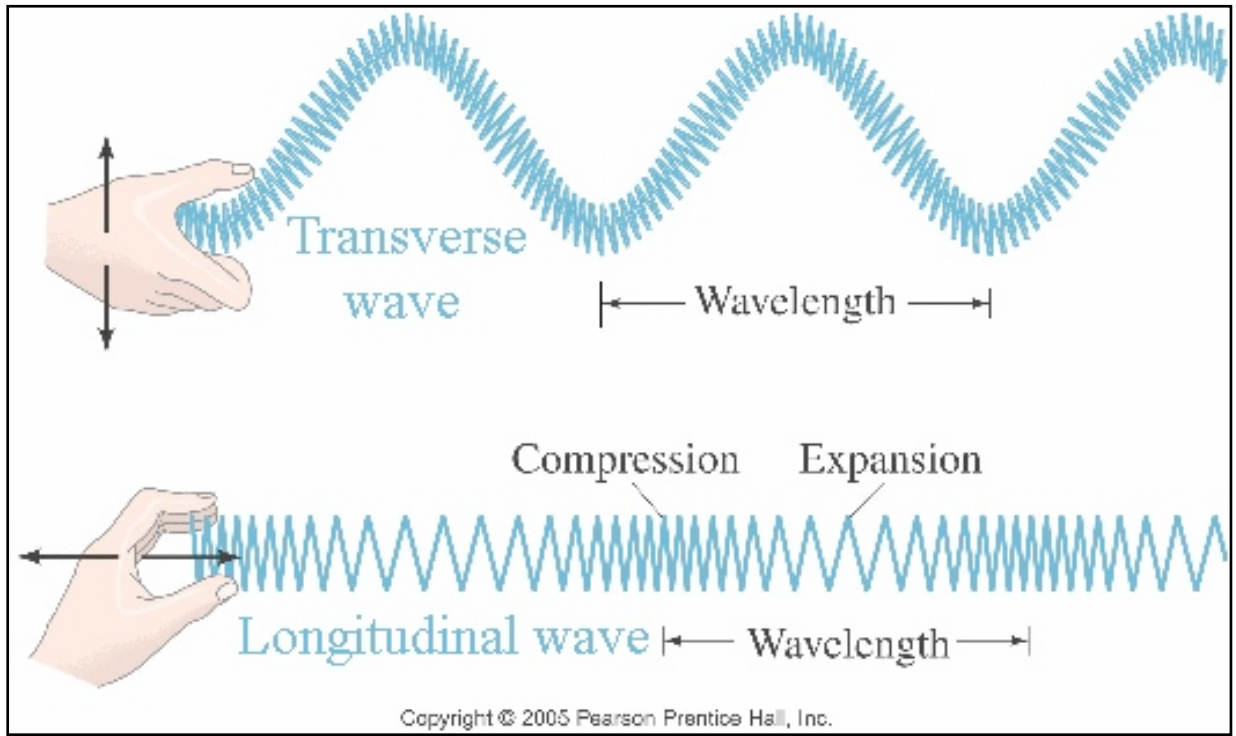
Gaz, katı veya sıvı ortamların sıcaklığı arttıkça, sesin bunların içinde yayılma hızı da artar.

Sesin havadaki hızını (c veya v olarak temsil edilir) hesaplamak için formülümüz:

[20.05 diye bir *adiyabatik indis ile gaz sabiti sayısı* ÇARPI karekök içinde (273+ *Santigrad*)]

Yukarıdaki 273 sayısı, *Santigrad (Celsius)* dereceden *Kelvin* dereceye dönüştürüm içindir.

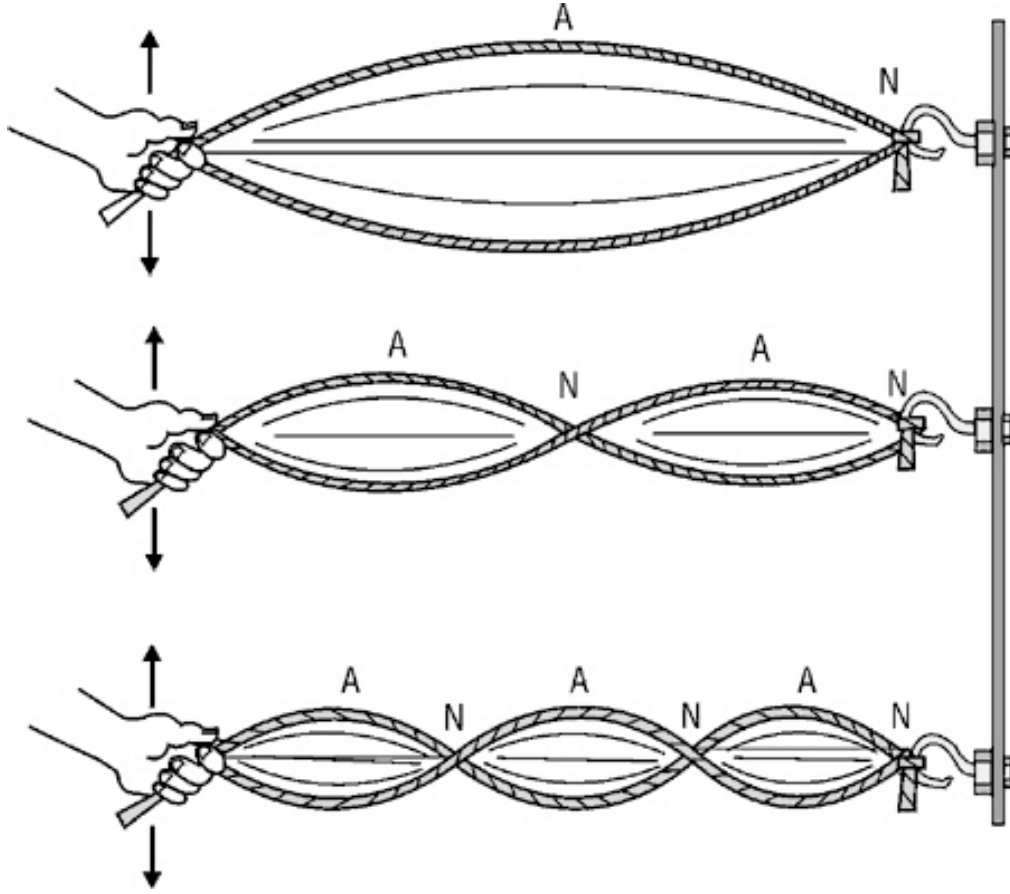
Ses dalgaları gerili tel üzerinde, tıpkı deniz yüzeyindeki su dalgaları gibi, “*transverse*” denilen modda ilerler: Türkçesi “enlemesine kalkıp-inerek”. Boru veya kamış içinde ise “*longitudinal*” denilen modda ilerler: Bunun da Türkçesi “boylamasına itiş-kakışlı”.



Enlemesine giden (yani transverse) ses dalgası, iki uçtan gergin bir telin bittiği noktadan geri dönüş yaparak arkadan gelen dalgayla üst-üste geçeceğinden, sonuç, İngilizce’de bilinen adıyla, “*standing (stationary) wave*”, yani “duran (kararlı) dalga” olur.

Bu üst-üste geçme olayı, aslında, İngilizce “*interference*” denilen “*girişim*”dir ve ileride ayrıntılandırılacaktır.

İki ucundan kısırılmış ipte, duran dalganın oluşumuna dair örnek şekil aşağıdadır:

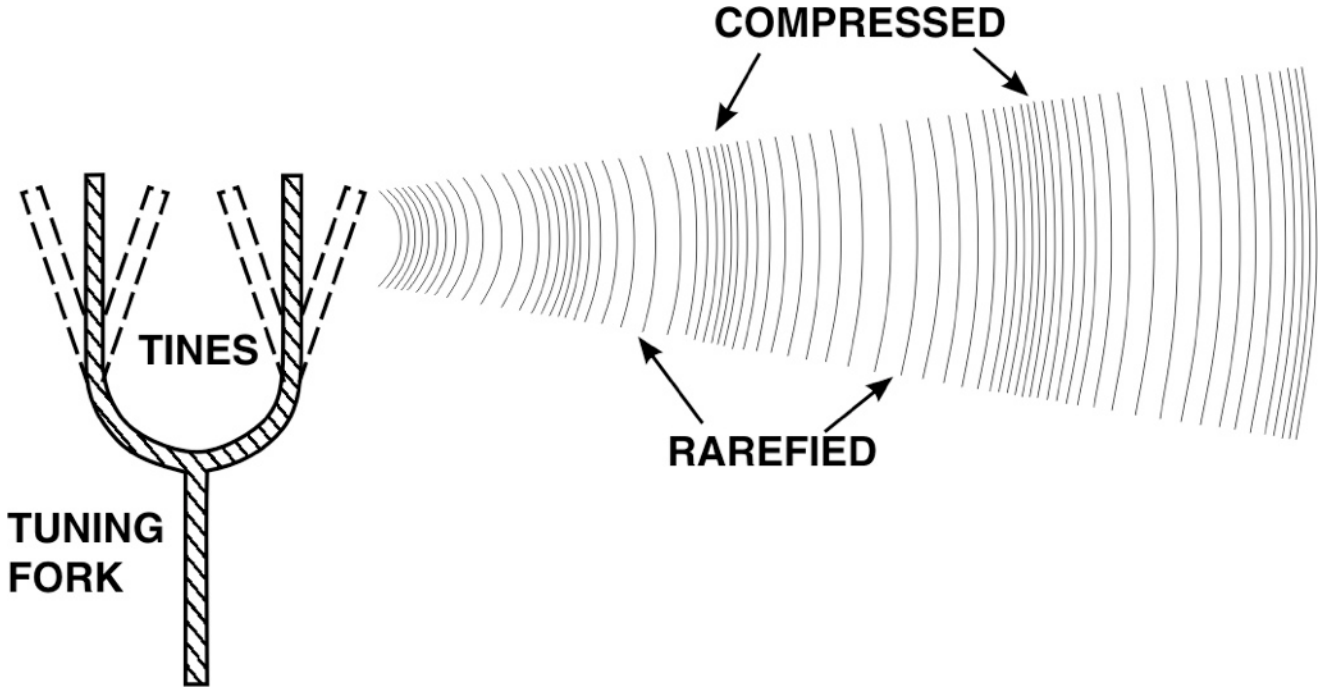


Nitekim, Pişano, Keman, Gitar gibi algılarda tellere yakın odaklı kamera ekimi yapılırsa, buna benzer grntler izlenebilir. Őekilde iřaretli boėumların (N: *nodes*) ve gbeklerin (A: *anti-nodes*) “armonikler”, dolayısıyla “algı tınısı” ile iliřkisi ileride aıklanacaktır.

Boru veya kamıř iinde de, anlatıldıėı gibi, hava stunu dahilinde ilerleyen dalganın ulardan tepmesi sz konusu olacaktır; ki o noktadan itibaren geri dnř yaparak, gelmeye devam eden dalga ile st-ste geme meydana geleceėinden, sonu, yine İngilizce adıyla, “*standing (stationary) wave*”, yani “duran (kararlı) dalga” olur.

Her iki halde de (*transverse* vs. *longitudinal*) aslında madde gitmez; gelgelelim, atomlar/ molekller kendilerine ait konumda bir miktar salınarak, hareketi komřu atomlara/ molekllere aktarır. Bu aktarımın belirli bir kesitine yoėunlařtıėımızda, kendini dzenli tekrar eden atımları (*pulsations*) grebiliriz.

Havada ses dalgası iřte bu Őekilde, ama kresel itiř-kakıřlı (*longitudinal*) bir tarzda, kaynaktan dıřarı doėru  boyutlu yayılır. Buna İngilizce’de *propagation of sound waves*, yani “ses dalgalarının ilerleyiři” denilir.



Yukarıdaki şekilde görüldüğü üzere, sıkışan kısımları müteakiben seyrekleşen kısımlar geliyor ve bunlar sürekli yer değiştiriyor. Yani, alabildiğince sıkışan kısımlar bir müddet sonra alabildiğince açılıyor ve aynı anda alabildiğince açılan kısımlar da alabildiğince daralmaya başlıyor ve bu değişimler düzenli olarak tekrar ediyor. Basıncın bazı kısımlarda artmasına İngilizce “*compression*” (sıkışma), diğer kısımlarda azalmasına “*rarefaction*” (seyrekleşme) deniyor. Gerili bir tel titreştiğinde, havadaki atomları/molekülleri işte böyle harekete geçiriyor.

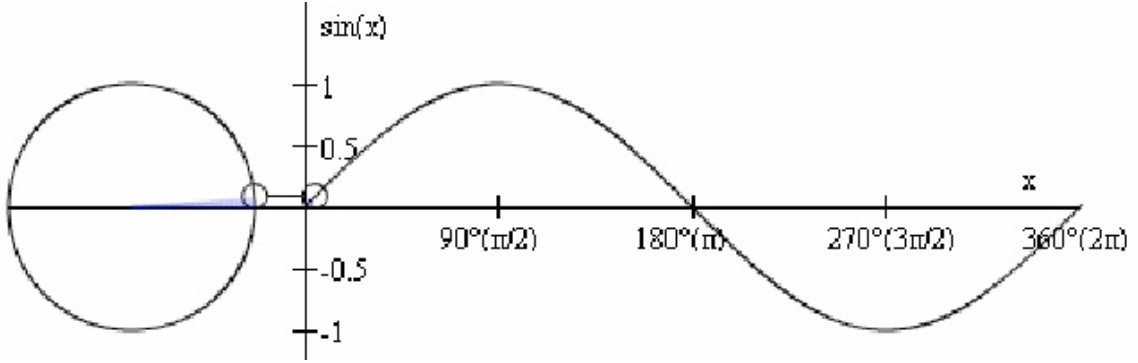
Havada ilerleyen ses dalgası bir yüzeye (mesela duvara) çarparsa,

1. ya tamamen yansıyabilir (*reflection*); ki bu da yankılanmaya yol açar,
2. ya enerji kaybına uğramaklı - yani frekans kaymasıyla - kırılarak yansır (*diffusion*),
3. ya da enerjisi soğurulur/yutulur (*absorption*).
4. O arada engelin öte tarafına ulaşabilir (*transmission*).

Enerji kayıplı engeli aşarsa, buna İngilizce “*transmission loss*” (aktarım kaybı) denir. Duysal olarak, kapıların kapalı tutulduğu bir odada yapılan müziği koridordan dinlemek ile o odanın içinde dinlemek arasındaki fark buna örnek olarak verilebilir.

Hertz “frekans” birimidir ve bu da saniyedeki titreşim adedidir. Müzikte her bir nota bir frekans (veya frekans öbeği) ile eşleştirilir. Yüksek frekans tiz sese, düşük frekans ise pest sese karşılıktır. Sözelimi, 500 *Hz* saniyede 500 kere titreşime (yahut atıma) denktir.

Mevzubahis frekansı tel üstünde “enlemesine kalkıp-inen” (*transverse*) dalga biçiminde görüldüğüne benzer olarak resmetmek istersek, bu bir “sinüs eğrisinin” (*sinüsoidin*) her saniye 500 kere ardı-ardınca dizilmesi olarak düşünülebilir. Sinüs eğrisi, 360 derecelik bir açıyı takiben ortaya çıkan dairesel hareketin zaman (x) eksenine izdüşümüdür ve çizilişi şöyledir:



Ses saniyede 500 Hz titreşiyor demekle, saniyenin 500’de 1’inde bir defa titreşiyor demek aynı şeydir. Demek ki, 500 Hz ’de bir sinüs eğrisinin çiziminin tamamlanması $1/500 = 0.002$ saniyede gerçekleşiyormuş, ki bu da 2 milisaniye eder. Yani:

$1/\text{frekans} = \text{period}$ (diğer bir deyişle, saniye veya milisaniye cinsinden “çevrim”).

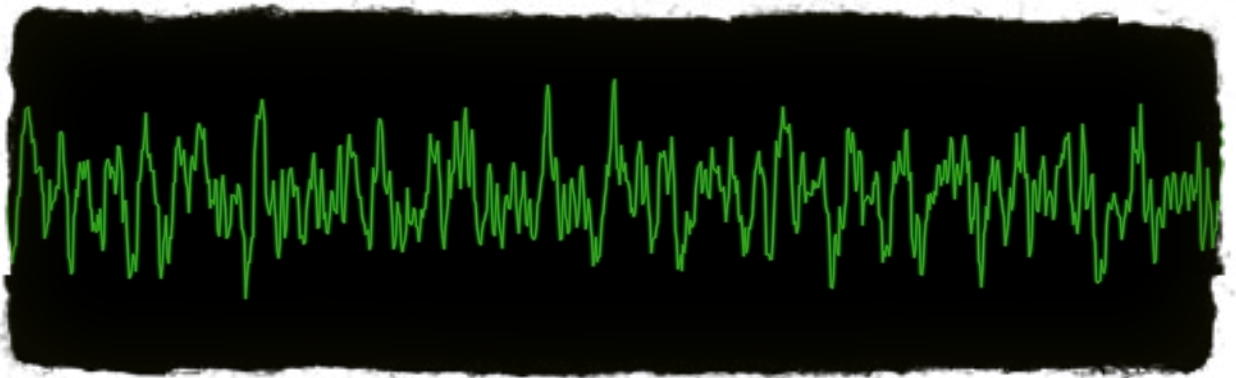
Çevrimin tamamlanması için katedilen mesafeye sesin dalgaboyu (*wavelength*) denir. Formülü:

$$\text{Lambda (yani dalgaboyu)} = (c/f) \text{ veya } (v/f)$$

Oda sıcaklığında 500 Hz için dalgaboyu $344/500 = 0.7$ metre yapar.

Sinüs eğrisini çizdirmekte kullanılan dairenin yarıçapı ne kadar büyük olursa, eğrinin x -ekseninden tepe noktasına yüksekliği o kadar fazla olacaktır. Bu yüksekliğe dalganın genliği (*amplitude*) denir. Daha fazla genlik, daha volümlü, yani gür ses demektir. Gürlük arttıkça, aynı ses, engellere rağmen işitilebilir. Bu da tıpkı daha yüksek deniz dalgasının kıyıdaki seti aşıp yola taşabilecek kadar enerjili olabilmesi örneğindeki gibi zihinde canlandırılabilir...

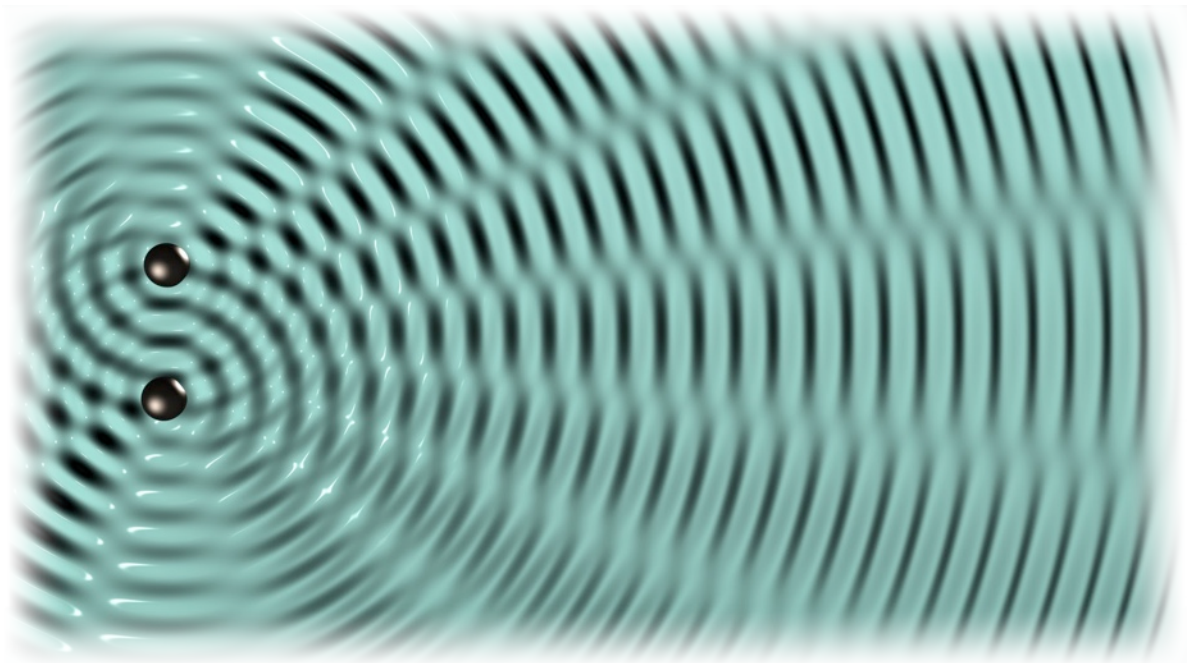
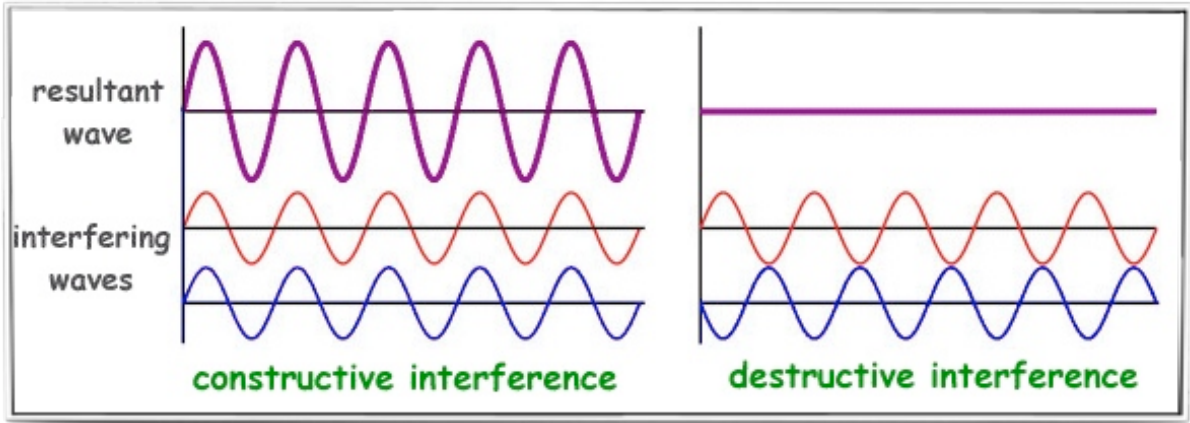
Elektronik ortamda ses dalgasının bu tarzda iki boyutlu resmedilmesi mümkündür ve çok yaygındır. Literatürde, “*waveform*” (dalgaformu), “*audio signal*” veya basitçe “*signal*” olarak da adlandırılır ve örnek şekli aşağıdadır:



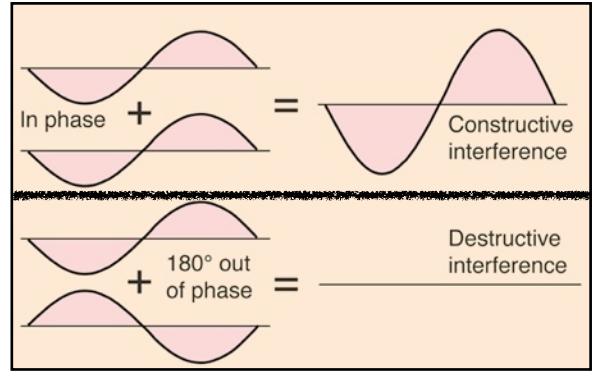
Tek başına sinüs eğrisi, armoniklerden yoksun olduğundan, tınısız (*timbreless*) çalgı sesi olarak da adlandırılabilir. Akort çatalı, ıslık, kristal bardağa ıslak parmak çalınması, vs... bu eğrinin hopperlörlerden üretilmesine yakın bir duyum verir.

Sinüs eğrisini çizdirmeye x -ekseninin 0 derecesinden başlamayıp, başka bir açı derecesinden de başlayabiliriz. Bu durumda, referans sinüs eğrisine kıyasla “kayık” bir sinüs eğrisi çizdirmemiz söz konusu olur. Veya ilk sinüs eğrisine kıyasla birkaç milisaniye gecikmeli aynı sinüs eğrisini çizdirebiliriz. İlkinin göre ikincinin açısal değeri “faz” (*phase*) olarak adlandırılır.

“Faz farkına” (*phase difference*) sahip iki aynı frekansta sinüs eğrisi aynı zamanda ve genlikte üst-üste binerse, sonuç bunların toplamı olur; yani sesin frekansı değişmediği halde gürlüğü değişir. Farklı zamanlarda veya genliklerde de sinüs eğrileri üstüste bindirilebilir ve bileşke genlik yer yer veya hepten sönümlenebilir. Öyle ya da böyle, faz farkı yüzünden İNTERFERANS (*interference* = girişim) meydana gelir. Örnek grafikler aşağıdadır:



Sözelimi, 180 derece faz farkına sahip (*out of phase*) iki sinüsoid üst-üste gelirse, birbirlerini yokederler ve ses duyulmaz. Buna İngilizce’de *destructive interference* (yıkıcı girişim) denir. Aynı fazda (*in phase*) üst-üste gelirlerse, bu kez *constructive interference* denilen yapıcı girişim olur ve sesin volümü artar.



Faz farkı, bir çalgıdan çıkan sesin dalgası üzerine, duvardan yansıyan öncül dalgaının

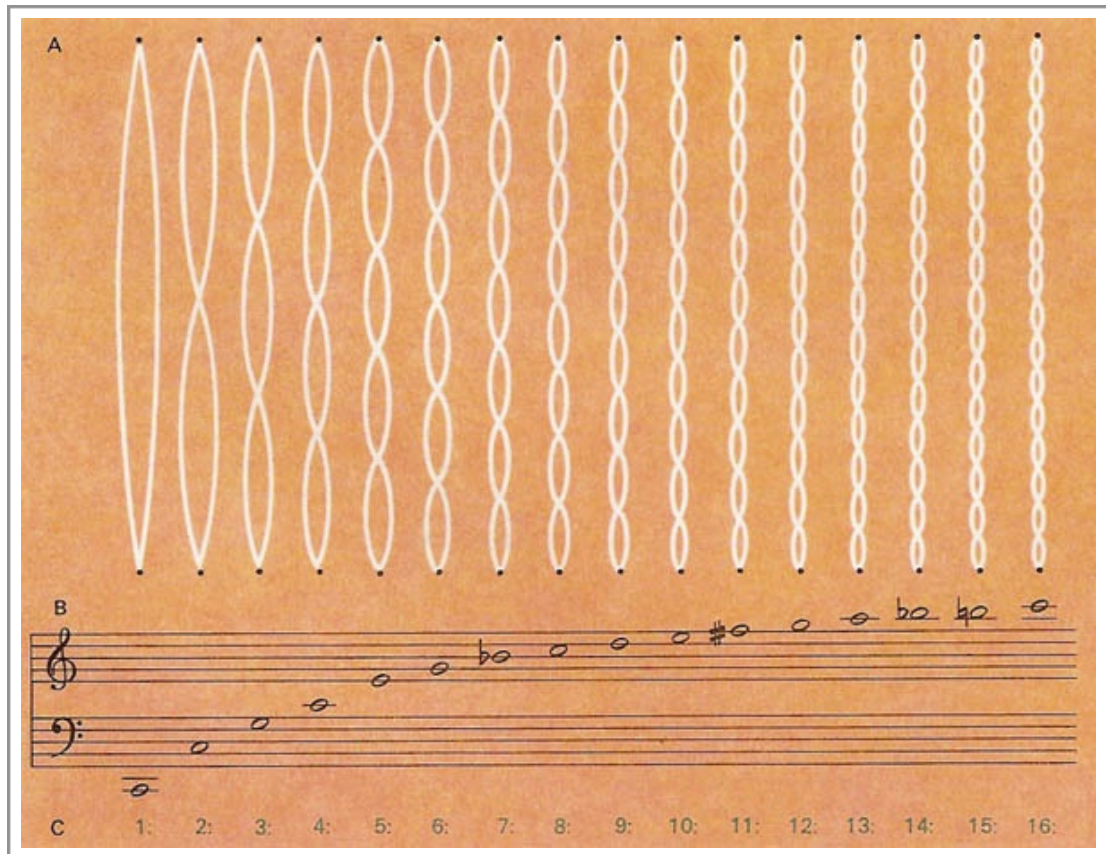
farklı anlarda üst-üste binmesiyle oluşabileceği gibi, birbirinden belli bir mesafeyle ayrı iki mikrofona yapılan kaydın aynı dijital miksaj kanalına aktarılmasıyla da oluşur.

[Faz: $0.00x$ saniye olarak ikinci sinüsoidin giriş anı farkı x frekans x 360 derece]

Diğer taraftan:

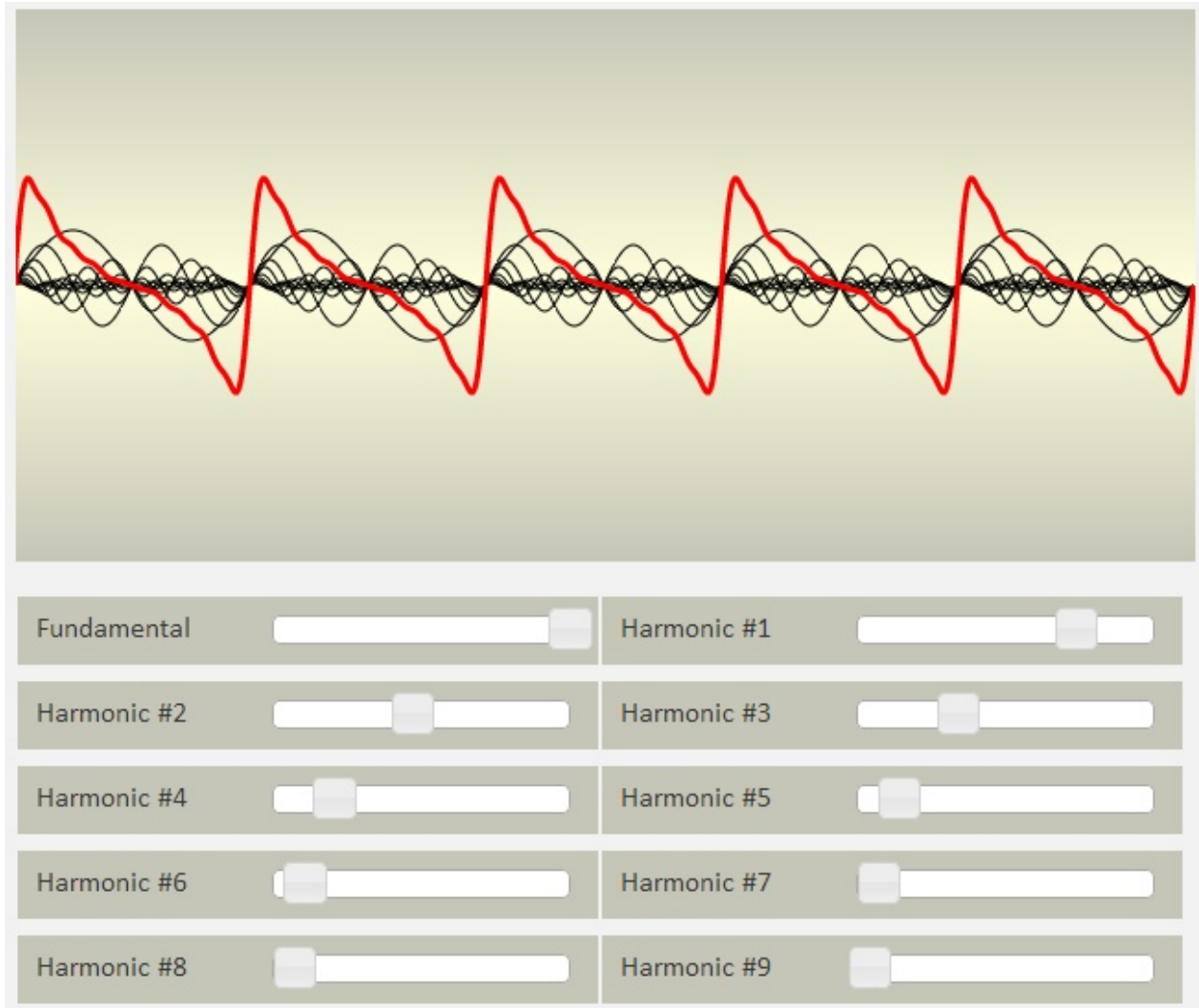
[giriş anı farkı = $\text{sesin ilk çıktığı veya kayda alındığı noktadan ikinci noktaya uzaklık} / c$]

Piyano, Gitar, Ney, Kemançe, vs... gibi çalgıların tınısını (ses rengini = *timbre*) meydana getiren “armonikler”dir; “selenler” yahut “doğuşkanlar” diye de adlandırılırlar. İngilizce karşılıklarını sırasıyla şöyle saptayabiliriz: “*harmonics*”, “(upper) *partials*”, “*overtones*”.



Çalgılarda, temel ses (*fundamental*) ve 1. armonik olan ilk sinüs eğrili duran dalga üstüne, iki kat, üç kat, dört kat... (hatta yarım kat, vs...: *lower partial/sub-harmonic*) sıkışık sinüs eğrili duran dalgalar aynı zamanda biner. Diğer armonikler, yukarıdaki şekilde görüldüğü gibi, birbirlerine göre eşit uzaklıklarda konuşlu boğumlardan (*nodes*) ve göbeklerden (*anti-nodes*) oluşur. Bunlar farklı genliklerde ve milisaniye farkıyla giriş anlarında üstüste binebilir. Genel kural şudur: Armonik ne kadar yüksek bir sayı ise, genliği de o kadar düşüktür. Sözgelimi, 5. armoniğin çalgı tınısına katkısı 1. armoniğe göre yüzde 4-5 kadardır.

Armoniklerin katışımı sonucu, İngilizce “*complex tone*” (tınılı ses) denilen ve portede notalarla ifade edilen çalgı sesi oluşur. Buna dair örnek bir dalgaformu aşağıdadır:



Piyano örneğinde, armonikler tam katlar yerine küsurlu katlarla tırmanırlar ve bu da oktavların muhakkak 2/1 oranından (ileride göreceğimiz üzere, 1200 *sentlik* aralıktan) geniş (örneğin birinci oktavda 1204 *sente*, diğer oktavlarda daha da artan değerlere göre) akortlanmasını gerektirir. İngilizce’de “*octave stretch*” de denilen bu durumun fiziksel tanımı yapılırken *inharmonicity* = *dersizlik* kavramı kullanılır. Dersizlik, boğumların ve göbeklerin konuşlanmasındaki minik eşitsizlikler ve tutarsızlıklardır. Piyanoların bu yüzden en pest ve en tiz tuşları arasında çeyrek-sese varan, ancak kulağı rahatsız etmeyen farklar oluşur.

Piyanoda armonikler “sempatetik rezonans” (*sympathetic resonance/vibrations*) sayesinde birebir gösterilebilir. Buna göre, 65 Hz’lik DO sesinin içinde 131 Hz’lik oktav üstteki Do, 197 Hz’lik oktav+beşli üstteki Sol, 263 Hz’lik iki oktav üstteki do, vs... sesleri de vardır. Şu da var ki, Piyanoya özgü 12-sesli temperamanda, bazı armoniklerden sapmalar olması kaçınılmazdır.

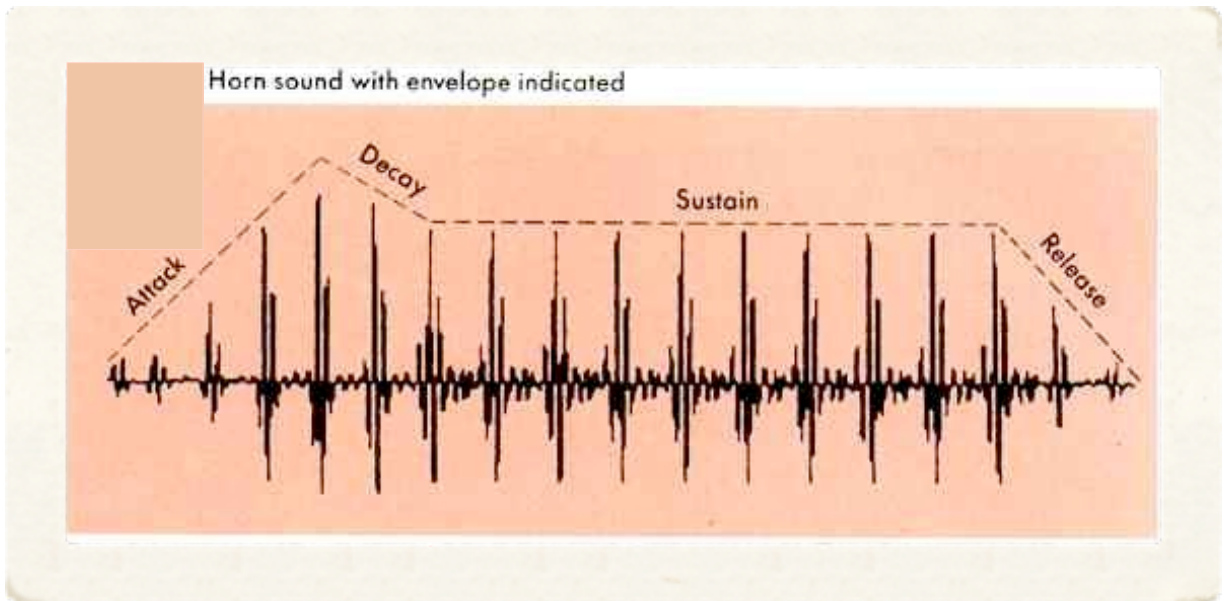
Tek başına armoniklerinden yoksun 65 Hz olsaydı, saf sinüs dalgası olurdu. Tam veya küsurlu sayıda kat kat sinüs eğrileri de içiçe olduğundan, susturucusu sessizce kaldırılan Piyano tuşlarına denk gelen teller, “sempatetik rezonans” sayesinde çınlatılabilmektedir. İşte, çalgı tınısı böyle armoniklerden oluşmaktadır.

Çalgı tınısını oluşturan diğer bir olgu ise ZARF’tır (*envelope*). Buna göre, ses dalgasının genliğinde yol boyu (zaman içinde) değişim sözkonusudur. Genliğin kısa sürede maksimuma çıktığı başlangıç anını müteakiben, genliğin iyi-kötü değişmeden kaldığı bir süreklilik durumu ve nihayetinde bırakım olacaktır. İngilizce’de, genliğin başlangıçtaki *zıplayışına* ATTACK, *sürüşüne* SUSTAIN, *bırakımına* DECAY adı verilir.

Elektronikte ise, bırakım genellikle ikiye ayrılır ve SUSTAIN’den önce ve sonra olacak şekilde konumlanır. Buna göre, sondaki kısım RELEASE (*sönümlenme*) diye adlandırılır:

Attack
Decay
Sustain
Release

Kısaltması: ADSR.



Gerçekte, herhangi bir akustik çalgıda farklı armoniklerin ayrı ayrı zarfları olduğu söylenebilir.

Armoniklerin temel sesin katları olan frekanslar olduğu görüldükten sonra, bunların birbirleriyle oluşturduğu müzikal aralıkların hesaplanması sorunu ortaya çıkar. Sözelimi, 3. armoniğin 2. armoniğe oranı 3/2 olan tam beşlidir ve 4. armoniğin 3. armoniğe oranı 4/3 olan tam dördüldür. Ancak bunların açılımı olan 1.5 ve 1.3333... sayılarından yola çıkarsak, tam beşli ile tam dördü farkının büyük ikili olması gerektiğini ve bunun oranının 9/8 ile ifade edilen 1.125 sayısı olduğunu çarpma-bölme işlemleri yapmadan bulamayız.

İnsan zihni ise, tam dördüye karşılık olan DÖRT ADIMDA 5'li => 4'lü => 3'lü => 2'li bulunması örneğindeki müzikal aralık adlarından görüleceği üzere, müzik teorisinde toplama-çıkarma yapmaya daha yatkındır. Müzikal oranları anlamlandırmanın güçlüğü karşısında, *logaritmik* bir aralık birimi olan *sent* devreye girer ve bu birim Ses Teknolojilerinde sıklıkla kullanılır.

İki frekansın birbirine olan oranı (bağıl frekans = *relative frequency*) hangi formülle sente dönüştürülür? Bir oranı sente dönüştürmeye yarayan ve herbiri yekdiğeriyle eşdeğer formüller, şimdilik hesap makinesi kolaylığına başvurursak, şöyledir:

1. $(\log_{10} \text{oran}) * 1200 / (\log_{10} 2) = \text{oranın sent adedi } n$
2. $(\log_2 \text{oran}) * 1200 = \text{oranın sent adedi } n$
3. $(\ln \text{oran}) * 1200 / (\ln 2) = \text{oranın sent adedi } n$

Peki, sent adedi bağıl frekansa nasıl dönüştürülür? Yani eldeki orana gerisin geri nasıl döneriz? Bunun için formüller ise şöyledir:

- A) $2^{(n \{ \text{yani sent adedi} \} / 1200)}$
- B) $(1200 / n) \sqrt[2]{2}$
- C) $\exp [(n / 1200) * (\ln 2)]$

Tabii, bunların hepsi oktavın tam 1200 sent (2/1 kesiri) alındığı varsayımına dayalıdır. Oktav farklı bir oran veya sent adedi olacaksa (mesela 2.029091 bağıl frekansının karşılığı olan 1225 sent alınacaksa), ilgili rakamlar (yani 2'nin ve 1200'ün geçtiği yerler) değiştirilir.

Şu hatırlatmayı da yapmanın yeridir: Herhangi bir frekanstan bir aralık kesiri miktarınca yukarıda olan diğer frekansı bulmak için, sözkonusu frekansı o oranla çarparsız; tersi olarak ise, aşağısındaki frekansı bulmak için o orana bölersiz.

Örneğin, 1 sentin bağıl frekansı = $\exp [(1/1200) * (\ln 2)] = 1.000577789506555$ olduğuna göre, 440 Hz'ten 1 sent yükselmek için: $440 \times 1.00058 = 440.25423 \text{ Hz}$ işlemi yapılır. Kademe kademe, sent sent yükselmek için ise aşağıdaki formüllere başvurulur:

$$f * 2^{(n\{\text{yani } 1,2,3,\dots 1200\} / 1200)} \quad \text{veya}$$

$$f * (1200\sqrt[2]{2})^{n\{\text{yani } 1,2,3,\dots 1200\}} \quad \text{yahut}$$

$$f * \exp \left[(n\{\text{yani } 1,2,3,\dots 1200\} / 1200) * (\ln 2) \right]$$

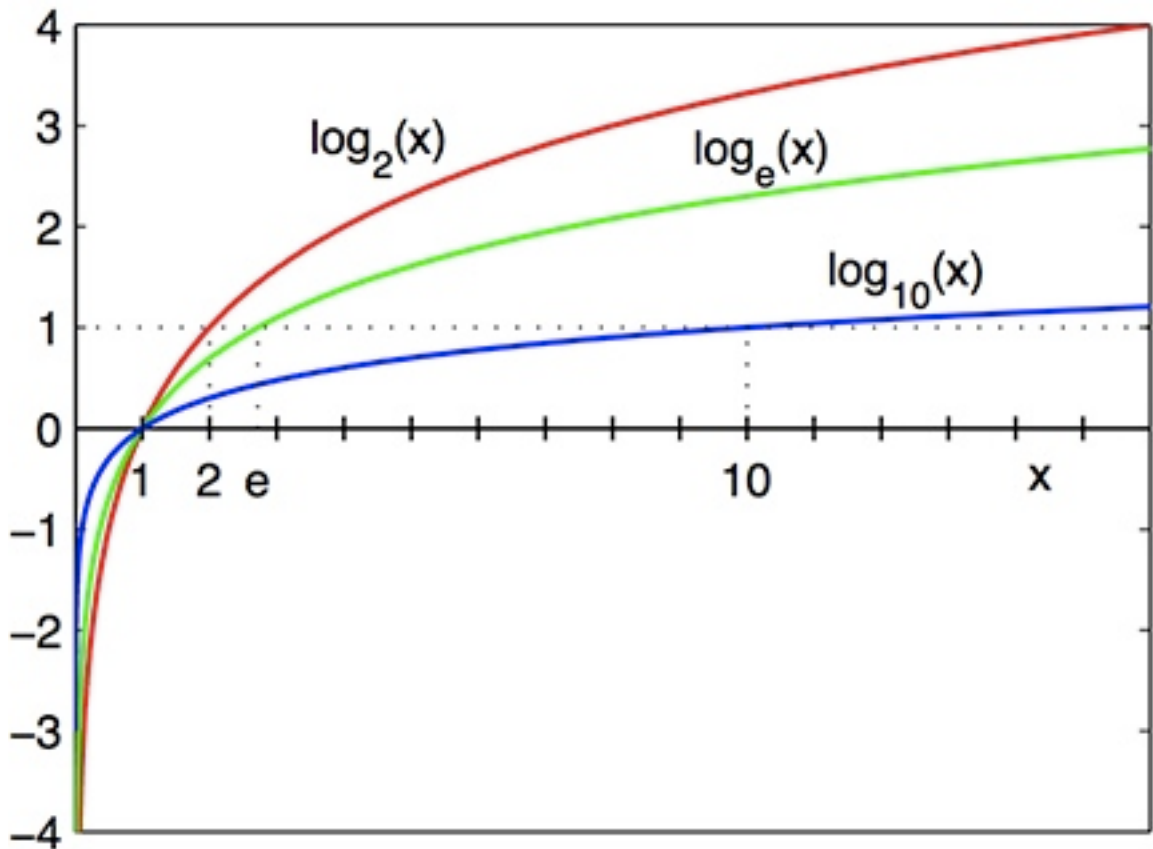
“Üssel işlem” demek olan ve aşağıdaki grafikte eğrilerle örneklendirilen *logaritma* konusunu basitçe şöyle anlatabiliriz:

Bir sayının 10 tabanlı logaritması demek, kaç tane 10 rakamını birbiriyle çarparsam o sayıyı elde ederim demektir.

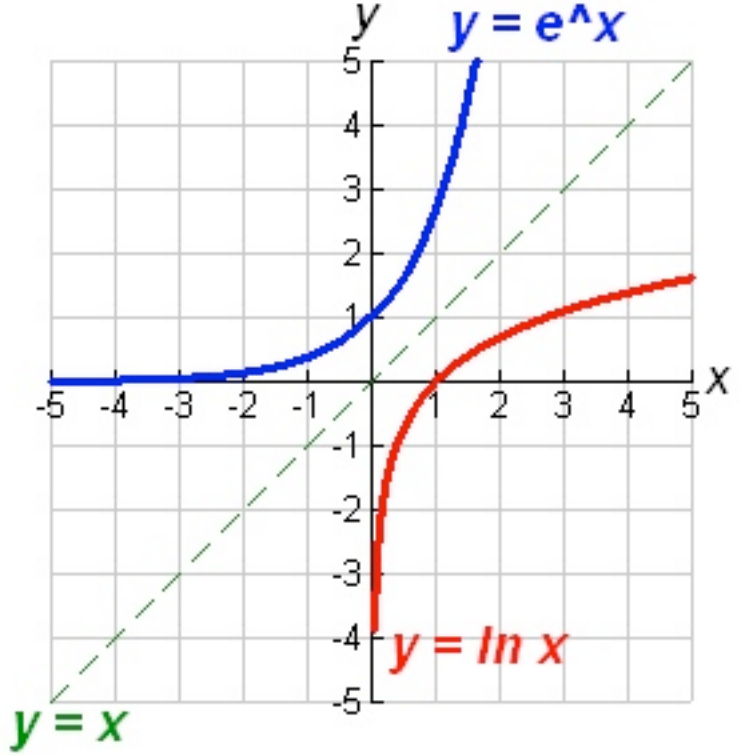
$$\log_{10} 1000 (= 10 \times 10 \times 10) \\ = 3$$

Bir sayının iki tabanlı logaritması demek, kaç tane 2 rakamını birbiriyle çarparsam o sayıyı elde ederim demektir.

$$\log_2 16 (= 2 \times 2 \times 2 \times 2) \\ = 4$$



Yukarıda “Napierien logaritma” ($\ln$ veya $\log_e$) da geçti; ki 17. Yüzyılda yaşamış matematikçi *John Napier*’e izafe edilmektedir. Buna “doğal logaritma” da deniyor ve tabanı (*Euler* sayısı denilen) 2.718281828459045... rakamına dayalıdır. Bu rakam ise, *eksponansiyel* (e veya $\exp$) denilen fonksiyon ile ilişkilidir. **Eksponansiyel eğrisi** yandaki şekilde göstermiştir:



Doğal logaritma eğrisi
eksponansiyel fonksiyonunun tersidir.

Doğal logaritma “ $\log_e$ ” veya “ $\ln$ ” diye yazılır. Bir sayının e tabanlı logaritması demek, kaç tane 2.718281828459045... rakamı birbiriyle çarpılırsa o sayı elde edilir demektir.

$\ln 148.41316...$ ($= 2.7183... \times 2.7183... \times 2.7183... \times 2.7183... \times 2.7183...$)
 $= 5$

Ses Teknolojileri kapsamında “logaritmanın” önemi ortadadır; zira, beyinde sesin volümünü ve müzikal aralıkları algılayışımız logaritmiktir. (Volüm farklılıklarını bu şekilde regülarize eden *desibel* birimini daha sonra göreceğiz.) Yani, logaritma dendiğinde aslında anlaşılması gereken, büyük sayısal atlayışlarla kendini gösteren fiziksel değer değişimlerini eşit gıdımlara ayarlanmış bir ızgaraya yerleştirme operasyonudur.

Örneklendirmek üzere şu seriye bakalım:

1, 10, 100, 1000, 10000, 100000 ...

Burada her yeni sayı bir öncekinden 10 kat büyük. Yani, sağa doğru gittikçe her yeni sayının fazladan bir sıfırı var. Şu şekilde de ifade edilebilirler (^ üssü demektir):

$10^0 = 1 = 1$
 $10^1 = 10 = 10$
 $10^2 = 100 = 10 * 10$
 $10^3 = 1000 = 10 * 10 * 10$
 $10^4 = 10000 = 10 * 10 * 10 * 10$
 .
 .

Görülebileceği gibi, her yeni sayı ile ilk sayı arasında büyük atlayışlar söz konusu olmaktadır. 1'e kıyasla yüz tam olarak 100 kat yukarıda. 1'e kıyasla onbin ise tamı tamına 10000 kat yukarıdadır. Dolayısıyla doğrusal (lineer) olmayan (yani eşit gıdımlı olmayan) bir büyüme söz konusudur. Halbuki, logaritma sayesinde, yukarıdaki seriyi eşit gıdımlı bir seriye izdüştürebilmekteyiz:

$$1=0, 10=1, 100=2, 1000=3, \text{ vs...}$$

zira, 0-1-2-3-4-5 serisindeki her rakamın önü arkası 1 gıdımdır. Buna göre:

$$\log_{10}(100) = \text{"yüz rakamı içinde kaç tane on rakamı var?"} = 2$$

Benzer olarak:

$$\log_2(8) = \text{"sekiz rakamı içinde kaç tane 2 var?"} = 3$$

zira,

$$\begin{aligned} 2^0 &= 1 &= 1 \\ 2^1 &= 2 &= 2 \\ 2^2 &= 4 &= 2 * 2 \\ 2^3 &= 8 &= 2 * 2 * 2 \\ 2^4 &= 16 &= 2 * 2 * 2 * 2 \\ & \cdot \\ & \cdot \end{aligned}$$

Diğer bir deyişle, felanca tabanlı logaritma demek suretiyle, o taban sayısının üslerini alarak ilerleyen herhangi bir rakam serisi linearize edilebilir (eşit gıdımlı doğrusal bir seriye dönüştürülebilir). 10 tabanlı logaritma dendiğinde 1, 10, 100, 1000... şeklinde yükselen bir seri 0,1,2,3... serisine oturtulur. 2 tabanlı logaritma dendiğinde ise 1, 2, 4, 8, 16 diye yükselen bir seri 0,1,2,3... serisine oturtulur. İşte, *müzikte aralıkları da bu şekilde bir linearizasyon sayesinde algılamaktayız*. Bu yüzden aralık hesabında *sent* birimi işleri kolaylar.

Daha da ileri gitmeden önce şunu bilmek gerekir ki, temel alacağımız bir sesin bağıl (rölatif) frekansı o sesin frekansının kendisine bölünmesidir. Yani, temel sesi DO notası olarak ele alırsak, sözgelimi $250 \text{ Hz} / 250 \text{ Hz} = 1/1$ olur. Çalgı tınısını oluşturan üst armonikleriyle birlikte düşünersek, Do'nun oktavının kendisine oranı $500 \text{ Hz} / 250 \text{ Hz} = 2/1$, oktav +beşlisi ise $750 \text{ Hz} / 250 \text{ Hz} = 3/1$ olarak hesaplanır. Böylece ilk üç armoniği ele aldık.

Üçüncü armoniğin oktav ötesine taşmasından ötürü, oktav sınırlaması içinde çalışmak istersek, bundan oktav aralığını çıkarmamız (onu 2/1'e bölmemiz) gerekir:

$$\frac{3}{1} \times \frac{1}{2} = \frac{3}{2}$$

Yani, artık oktav+beşliyi değil, sırf beşli aralığının bağıl frekansını ($750 \text{ Hz} / 500 \text{ Hz}$ veya $375 \text{ Hz} / 250 \text{ Hz} = 3/2$ oranını) alarak devam edebiliriz. Dolayısıyla, muhayyel düz bir çizgi üzerinde 1 (temel ses) ile 2 (temel sesin oktavı) farazi noktaları arasında $3/2 = 1.5$ konumundaki noktayı kastetmiş oluruz. Çalgı teli üstünde gösterilirse, açık telin üçte birine dokunulmak suretiyle üçte ikisi titreştirilir ve açık telin DO'suna kıyasla beşli aralığı yukarıdaki SOL sesi işitilir. Gelgelelim, buradaki tutarsızlık hemen göze çarpmış olacak, zira 1 ve 2 arasındaki 1.5 muhayyel noktasının temel ses ile oktavına eşit uzaklıkta (0.5'er birim) olmasından farklı olarak, tel üzerinde parmakla dokunulan SOL noktasının eşikten olan uzaklığı, temel sesin oktavına olan uzaklığının iki katıdır, aynısı değildir.

Şöyle bir durum daha var ki, bu bir oktavlık ses sahasını araları bir puan olan iki farazi noktanın işaretlendiği muhayyel düz bir çizgi olarak düşünmemiz halinde, bunun tam duysal ortası (yani yarı konumu) aslında SOL değil "FA diyez" olmalıdır (zira, *oktavın yarısı artık dörtlü veya eksilmiş beşli olarak adlandırılan aralığı verir*). Halbuki, tel üstünde temel ses ile oktavının görsel yarısı olan konum, açık telin dörtte üçünün titreştirilmesiyle oluşan FA sesidir; çünkü tam dörtlü aralığının bağıl frekansı olan $4/3$ oranı, telin aritmetik dört eşit parçaya bölünmesinden sonra, o dört parçanın ilki dışında kalan tel uzunluğunun titreşimini ifade eder. Diğer yandan, açık teli simgeleyen 1 ile oktavı simgeleyen 2 muhayyel noktaları ayarınca, FA sesinin bağıl frekansı tam $1.33333...$ noktasını işaret etmektedir ve bu da söz konusu iki muhayyel noktanın arasının ($0.33333...$ 'erlik birimden) üç eşit parçaya bölünmesinden sonra ilk kısmına isabet eder. Oysa yine bilinmektedir ki, oktavın duysal üçte biri FA değil Mİ olmalıdır.

Diğer bir deyişle, aritmetik eşit bölünmeler, gerek tel üzerindeki, gerek mevzubahis muhayyel hat üzerindeki konumlandırmalar itibariyle, müzikal aralık dünyasında duyumsadığımız eşit bölünmelerle örtüşme sergilememektedir.

Bunları Piyanoda eşit yarım sesler üzerinden sınayabiliriz. Yani, tam beşli aralığı aslında uçları temel ses ile temel sesin oktavı olan muhayyel bir hat üzerinde ortanın biraz daha yukarısında, tam dörtlü aralığı ise çok değil biraz aşağısında bulunmalıdır.

Şimdi *logaritma* konusuna geri dönebiliriz. Logaritma fonksiyonu sayesinde bağıl frekansları (müzikal oranları) duyumumuzla eşdeğer bir ıskalaya oturtabilmekteyiz.

Önce bahsi geçen aralıkları linearize edelim (0 ile 0.30103 noktalarının uçları olduğu eşit gıdımlı muhayyel bir çizgiye yayalım):

$$\begin{aligned}\log_{10} (1/1) &= 0 \\ \log_{10} (4/3) &= 0.1249387366083 \\ \log_{10} (3/2) &= 0.176091259055681 \\ \log_{10} (2/1) &= 0.301029995663981\end{aligned}$$

Ancak, 0.301 küsuratlı bir nokta olduğundan, şimdi yukarıdaki sayıları topluca, lastik ipe dizili boncuklar misali, 0 noktası ile 1 noktası arasına çekştirelim (*normalization*):

$$\begin{aligned}\log_{10} (1/1) / \log_{10} (2/1) &= 0 \\ \log_{10} (4/3) / \log_{10} (2/1) &= 0.415037499278844 \\ \log_{10} (3/2) / \log_{10} (2/1) &= 0.584962500721156 \\ \log_{10} (2/1) / \log_{10} (2/1) &= 1\end{aligned}$$

Görüldüğü gibi, 0 ila 1 noktalarının işaretlendiği eşit gıdımli muhayyel bir hat üstünde, gerçekten de yarı noktanın (0.5'in) biraz altına tam dörtlü aralığı yerleşirken, biraz üstüne de tam beşli aralığı yerleşti. Duyum ile kağıt üstündeki sayısal temsil örtüşmüş oldu. Bundan sonra, yukarıdaki sonuçları *sent* değerine dönüştürmek işten değil, zira oktavda 1200 adet olduğu kabul edilen sent birimine dönüştürmek için sonuçları 1200 ile çarpmak yeterli:

$$\begin{aligned}[\log_{10} (1/1) / \log_{10} (2/1)] \times 1200 &= 0 \text{ sent (başlangıç sesi)} \\ [\log_{10} (4/3) / \log_{10} (2/1)] \times 1200 &= 498.0449991346128 \text{ sent (dörtlü aralığı)} \\ [\log_{10} (3/2) / \log_{10} (2/1)] \times 1200 &= 701.9550008653872 \text{ sent (beşli aralığı)} \\ [\log_{10} (2/1) / \log_{10} (2/1)] \times 1200 &= 1200 \text{ sent (oktav)}\end{aligned}$$

Bir frekans oranının sent adedini bulmak için 2 tabanlı logaritma kullanılması işleri kolaylar, zira 10 tabanlı logaritma kullanıldığında, işlemi tamamlamak için oktavın (yani 2/1'in) 10 tabanlı logaritmasına bölme yapmak gerekir.

Öte yandan, oktav tam 2 değil de, Piyanodaki “*inharmonisite*” (dersizlik) nedeniyle küsuratlı olacaksa (mesela 2.001 alınacaksa), 2 tabanlı logaritma işe yaramaz. Diğer iki hazır seçenek 10 tabanlı ve “*e*” tabanlı logaritmalardır.

ln kullandığımızda da, 10 tabanlı logaritmada olduğu gibi, oktav olarak seçilecek aralığın sözkonusu tabandaki logaritmasını almak ve işlemde bunu bölen kılmak gerekecektir.

2 yerine, daha önce belirtildiği gibi, 2'ye yakın küsuratlı bir sayı girilebilir. Özellikle dijital ortamda oktav ile oynamak kolaydır. Sözelimi, *Dynamic Tonality* özellikli Viking Sentessayzerde, armonikleri kimıldatılmış çalgı sesleri sentezlenebilmektedir. Logic Pro gibi stüdyo kayıt programlarında ise, TUNING penceresinde “oktav çekştirme” (*octave stretch*) özelliği bulunmaktadır.

Ses şiddetini ölçmekte kullanılan “*Desibel*” (Bel’in 10’da 1’i) birimi de logaritmiktir. Ses Teknolojilerinde, basınç, güç ve voltaj değerlerindeki görece büyük değişiklikler, algılanan seste görece küçük değişimlere sebep olduğundan, logaritmik bir ıskalanın tercih edilmesi uygundur. Bu sayede, büyük atlayışlarla büyüdüğü halde “etkisi minimal” olan değerleri kısa atlayışlarla (az rakamsal farklarla) kağıt üstünde ifade edebiliriz. Desibel’in formülü şöyledir:

$$10 \times \log_{10} (WATT / \text{referans } WATT) = \text{DESİBEL } (dB)$$

Watt burada (farklı parlaklıkta ışık veren ampullerde olduğu gibi) hoparlörün harcadığı güçtür. Bu da NEWTON birimince kuvvetin *metre/saniye* cinsinden hızla çarpımıdır. Diğer bir deyişle, “güç, birim zamanda harcanan enerjidir”. Yahut VOLT x AMPER olarak da ifade edilebilir. Volt elektriksel yük birikimi (gerilim) olup, Amper de teldeki akımdır. Gerilimden akış olursa güç ürer. Sadece genel-geçer bilgi olarak verildiğini anımsayalım. Formül için bize lazım olan bilgi, sesi üretmek için ne kadar elektriksel güç harcandığıdır.

Üretilen ya da üretilecek desibelin ne kadar yüksek bir elektriksel veya mekanik güce karşılık olduğunu ise şu dönüştürücü formülle buluruz:

$$WATT = \text{referans } WATT \times 10^{(dB / 10)}$$

Ayrıntılı matematiksel hesapları ve örnekleri, SES KAYIT VE MÜZİK TEKNOLOJİLERİ (Ufuk Önen - Çitlenbik Yayınları) başlıklı kitapta bulabilirsiniz.

Şimdilik şu kadarının bilinmesi kafidir ki, güçteki her 2 kat artış, desibeldeki kabaca 3 puan artışa denktir. Aslında 3 puan değil de, tam olarak “[10 x log₁₀ 2] = 3.010299956639812” kadardır. Bununla birlikte, ilk ondalık hanesinden sonraki küsurat duyum açısından önemsizdir.

Yine bilinmesi kafidir ki, güçteki her 10 kat artış, desibeldeki tamı tamına 10 puan artışa denktir: “[10 x log₁₀ 10] = 10”.

Bin haneli WATT değerlerini ifade ederken, sağdaki üç sıfırı atıp kW (kiloWATT) da diyebiliriz. Bu örnek itibariyle, *referans WATT* 1 ise, 1 kW (1000 W) çıkış 30 desibel şiddetinde ses üretir.

VOLTAJ olarak ele aldığımızda ise, gerilimdeki her 2 kat artış desibeldeki 6 puan artışa, 2 kat azalma da 6 puan azalmaya denktir ([20 x log₁₀ 2] = 6.020599913279624).

Voltaj 3 kat artar yahut düşerse, kabaca artı/eksi 10 dB, 10 kat artar yahut düşerse, artı/eksi tam 20 dB fark oluşur. Volt referanslı dB standartlarına bir sonraki dönemde bakılacaktır.

Ses dalgaları bir yüzeye çarptığında o yüzeyi ittirirler. Müzik çalan bir hoparlörün membranının önüne kağıt konduğunda, membran tarafından iteklenen hava kağıdı çalkalar. Ses her yüzey üzerinde bu türden basınç meydana getirir. Düzenli ve sürekli ses dalgaları düzenli ve sürekli basınç üretirler. Böylesi bir basınç üzerinde yapılacak ölçüm SOUND PRESSURE LEVEL (SPL) olarak adlandırılır. SPL ölçümünde esas alınan fiziksel birim PASCAL olup, Pascal'ın açılımı metrekare başında NEWTON kuvvettir:

$$1 \text{ Pa} = 1 \text{ N/m}^2 = 10 \text{ dyne / cm}^2$$

Duyuma eşiği (*hearing threshold*) 0.00002 Pa olarak kabul edilir.

Bu referans itibariyle, 0.02 Pa “konuşma seviyesindeki” ses basıncına karşılıktır. Kulağın ağrı eşiği (*pain threshold*) ise 200 Pascal seviyesinde ve üstündedir. Uzun süre böylesi bir basınca maruz kalındığı takdirde, kulak zarının yırtılması kaçınılmaz olur.

Ses basıncı (SPL) seviyelerini ölçeklendirmek için yine logaritmaya başvurulur. Formülümüz şudur:

$$\text{dB SPL} = 20 \times [\log_{10} (\text{PASCAL} / \text{referans PASCAL})]$$

Konuşma seviyesindeki ses basıncına karşılık olan 0.02 Pa, bu hesapça, 60 dB SPL değerindedir. Yani, fısıltı seviyesinden konuşma seviyesine Pascal'daki 1000 kat artışı, sadece 60 puan desibel artışı olarak yazabilmekteyiz.

Sergilenen desibelin ne kadar yüksek bir basınca tekabül ettiğini ise şu dönüştürücü formülle buluruz:

$$\text{Pa} = 10^{(\text{dB SPL} / 20)} \times \text{referans PASCAL}$$

Kaynaktan yayılan ses ne kadar uzağa giderse, şiddeti o kadar azalır. Açık havada ses kaynağı ile ölçüm yapılan nokta arasındaki mesafeyi 2 katına taşıdığımızda ve ölçümü tekrarladığımızda, ses şiddetinin $([20 \times \log_{10} 2] = 6.020599913279624)$ yani 6 dB SPL eksildiğini göreceğizdir.

Bir örnek vermek gerekirse, 3 metre mesafeden dersin hocasının konuyu anlatırkenki sesinin basıncını 70 dB ölçmüşseniz, 6 metre mesafeden yapacağınız ölçüm sonucu sesi $70 - 6 = 64$ dB SPL şiddetinde bulursunuz.

Buna dair formülümüz şudur:

ikinci ölçüm dB SPL = birinci ölçüm dB SPL - $[20 \times \log_{10} \{\text{ikinci uzaklık} / \text{birinci uzaklık}\}]$

Örneğimiz itibariyle:

$$\text{SPL}_2 = 70 \text{ dB SPL} - [20 \times \log_{10} 6 \text{ metre} / 3 \text{ metre}]$$

$$\text{SPL}_2 = 70 \text{ dB SPL} - [20 \times \log_{10} 2]$$

$$\text{SPL}_2 = 70 \text{ dB SPL} - [20 \times 0.30103]$$

$$\text{SPL}_2 = 70 \text{ dB SPL} - 6.0206$$

$$\text{SPL}_2 = 63.9794001 = 64 \text{ dB SPL}$$

Eğer birden çok kaynaktan aynı anda ses geliyorsa, bunların bileşke ses basınç seviyesi için formülümüz:

$$\text{toplam ses basınç seviyesi} = 10 \times \log_{10} \{10^{(\text{SPL}_1/10)} + 10^{(\text{SPL}_2/10)} + \dots\}$$

Demek ki, hoca ders anlatırken bir arkadaşımız faraza gür şarkı söylüyorsa (80 desibel) ve bir başka arkadaşımız kemençede hafif taksim yapıyorsa (60 desibel), hocanın konuşma seviyesinin yanısıra (70 desibel) aşağıdaki altı çizili değerler devreye girecektir ve bileşke ses basıncını yukarıdaki formüle göre hesaplayabileceğizdir:

$$\text{SPLT} = 10 \times \log_{10} \{10^{(\underline{70} \text{ dB SPL} / 10)} + 10^{(\underline{80} \text{ dB SPL} / 10)} + 10^{(\underline{60} \text{ dB SPL} / 10)}\}$$

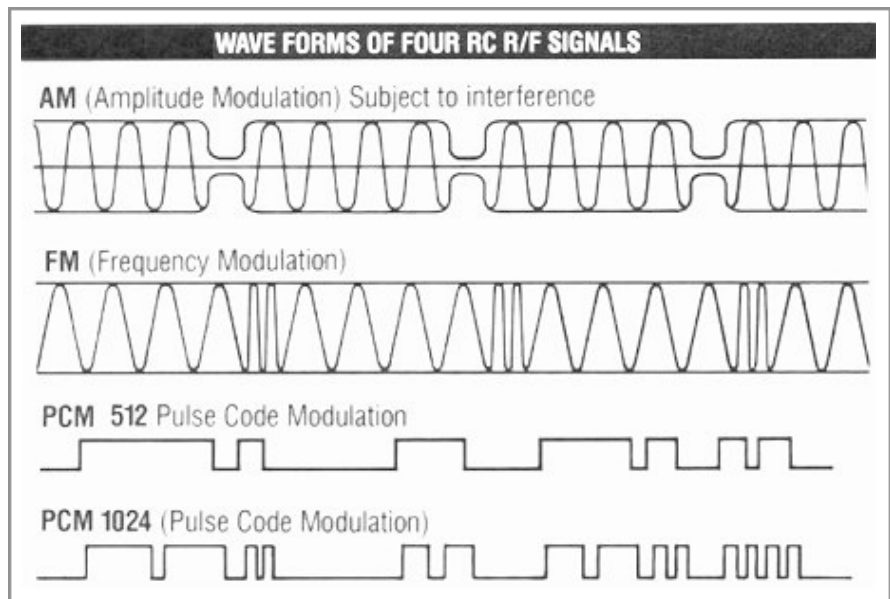
$$\text{SPLT} = 10 \times \log_{10} \{10,000,000 + 100,000,000 + 1,000,000\}$$

$$\text{SPLT} = 10 \times \log_{10} \{111,000,000\}$$

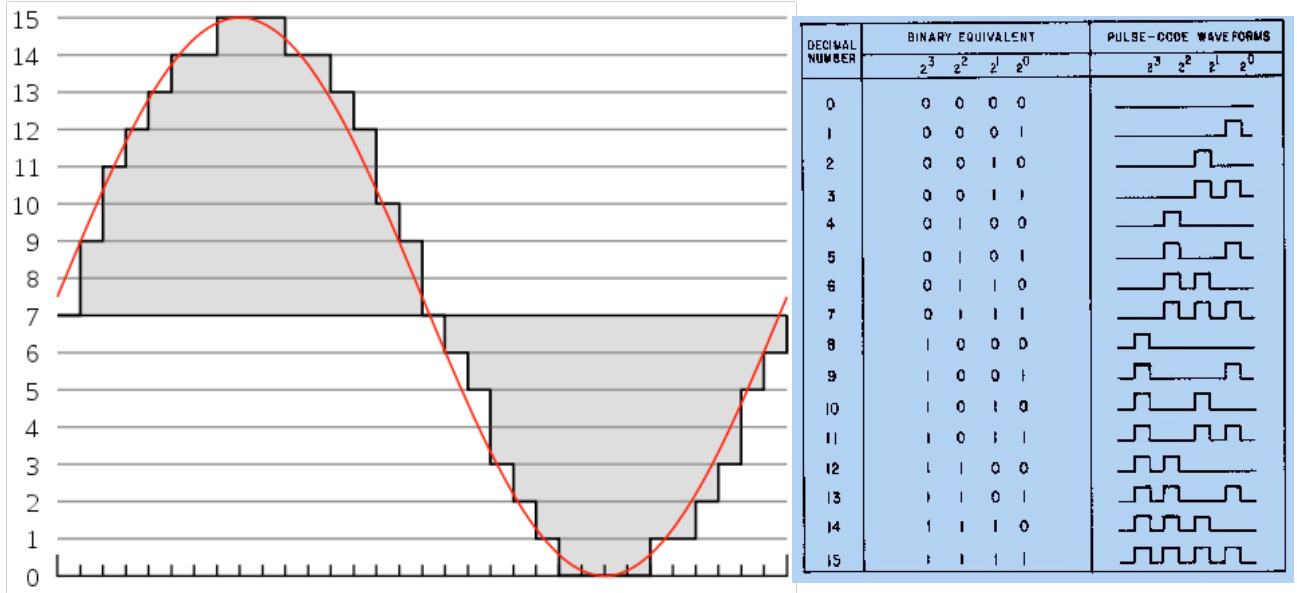
$$\text{SPLT} = 10 \times 8.045322978786658$$

$$\text{SPLT} = 80.45 \text{ dB SPL}$$

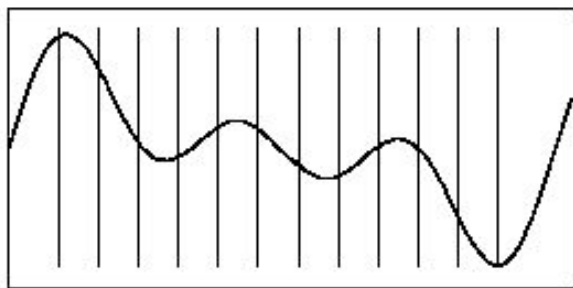
Dijital ses teknolojisinin temelleri, 1937'de Alec Reeves tarafından, Fransa'da "International Telephone & Telegraph" şirketinde çalıştığı esnada, *Pulse Code Modulation* (PCM) olayını keşfettiğinde atılmış olmaktadır. Bugün bile radyolarda geçerli olan AM ve FM yöntemleri ile bir karşılaştırma, yandaki grafikte görülmektedir.



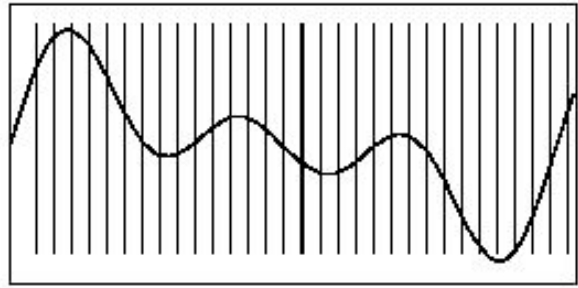
PCM, Windows .WAV ses dosyalarında da görmüş olabileceğiniz harflerdir. Peki ne demek bu PCM? En basitinden anlatmak gerekirse, *analog* (kesintisiz ve sürekli fiziki değişkenlerle ifade olunan) sinyalin belirli zaman aralıklarında nicelendirilerek (kuvantize edilerek) ikili sayı sisteminde (yani *dijital* olarak) temsil edilmesidir. Bir *sinüsoidin* bu şekilde temsili aşağıda gösterilmiştir:



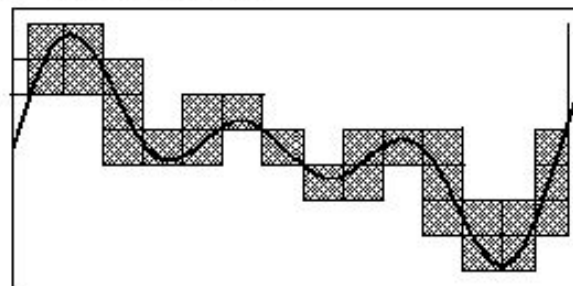
Böylece, herhangi sinüsoidal veya farklı şekle sahip bir dalgaformu, dikeylemesine eşit dilimlere doğranıp, her bir dilimi 0 ve 1 cinsinden “binary sayılarla” anlatılarak, “*periodic sampling*” (düzenli aralıklarda örnekleme) elde edilmiş olur. Buna dair grafik aşağıdadır:



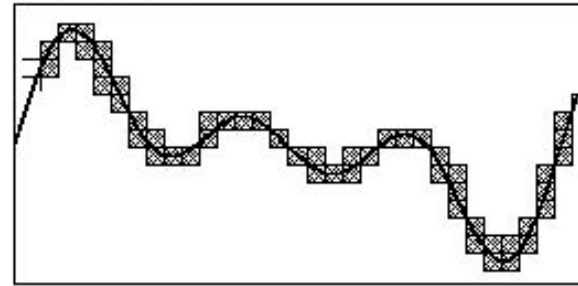
lower sample rates take fewer snapshots of the waveform



faster sample rates take more snapshots.....

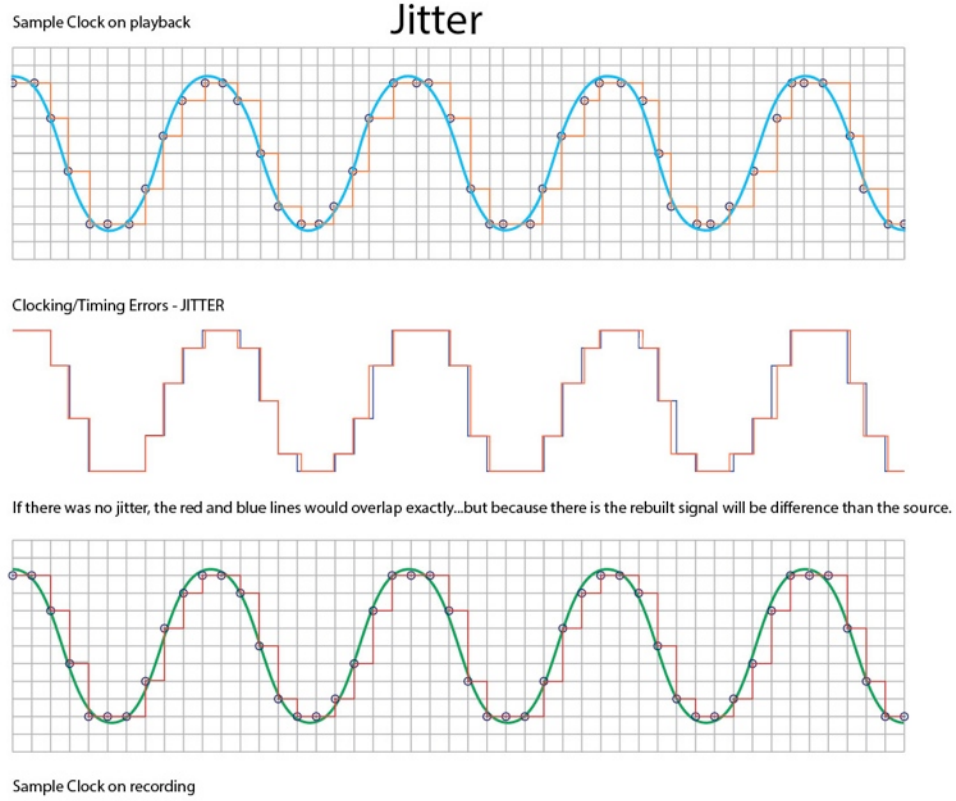


resulting in a rough recreation of the waveform.



resulting in a smoother and more detailed recreation of the waveform.

Periyodik örnekleme işleminde, referans saatteki teklemelerden dolayı, kayma, attırma veya sapmalar (*Jitter*) da meydana gelebilir tabii. Bu ise dijitaldeki temsilin orijinaline göre (algılanamasa bile) bozulması demektir ve buna dair örnek bir şekil aşağıda verilmiştir:



PCM sayesinde gerçekleşen dijital devrimin bugünkü popülaritesine ulaşması 1990'ları bulmuştur. Dijital teknolojilerin de etkisiyle, 1980'lerden itibaren "plak devri" kapanmıştır. Bu geçiş dönemi boyunca, analog ve dijital yanyana yürümüştür. Nitekim, büyük ve orta çaplı prodüksüyon stüdyolarında, analogtan dijitale geçiş, anlatıldığı kadarıyla, epey sürüncemeli olmuştur.

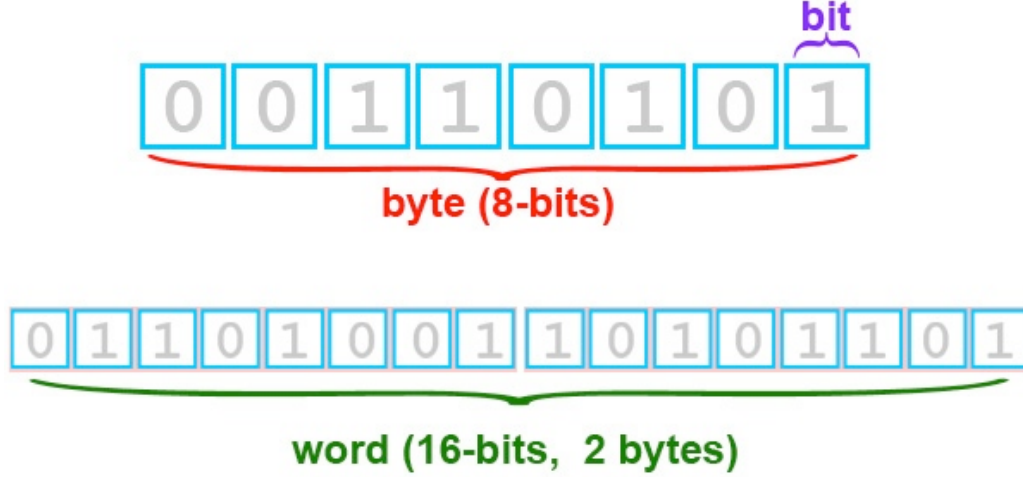
Fakat, editing ve miksing yönünden hayli üstün imkanlar sunduğu için, dijital büyük oranda galip gelmiştir. Günümüzde dijital her alanda vazgeçilmez bir teknolojidir.

Analog teyplere miks yapıp editlemeye kıyasla, dijital:

1. daha ucuz olması,
2. aynı derecede "sıcak" sesler sunabilmesi,
3. analog converter (dönüştürücü) teknolojisindeki kalite ve performans gelişimi,
4. analog kayıtlardaki gibi kayıplar yaşanmaması,
5. editing ve otomasyon gibi üstün yönleri

itibariyle pek revaç bulmuştur.

Analog sinyali, anlatıldığı gibi, ikili (binary) sisteme çevirme işlemine “*digitizing*” (dijitize etmek) denmektedir. Bu sayede, sürekli bir eğrinin 0 ve 1 cinsinden sayılarla basamaklı temsili sağlanır. Dijitizasyonda, en küçük birim olarak 0 veya 1 alınabilmektedir. Bunlar BINARY DIGIT’in kısaltılmışı olan “*bit*” olarak adlandırılır. 8 adet *bit* yanyana gelirse, buna da BYTE denmektedir. Örnek bir temsil aşağıda sunulmuştur:



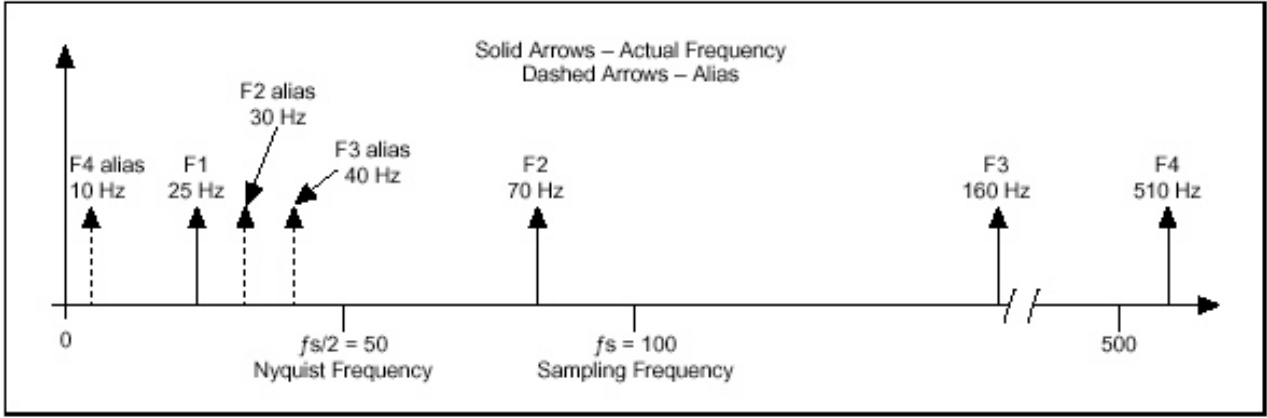
Hardisklerde MegaByte (MB), GigaByte (GB), TeraByte (TB), işte, bu 8-bit’lik lokmanın milyon, milyar, trilyon kez yanyana konabileceği kadar geniş bir depolama alanını işaret eder.

Dijital ortamda analog sinyalin bu şekilde bit’lerle temsil edilmesinden ötürü, ses seviyeleri dBFS (*Desibel Full Scale*) ıskalasında değerlendirilir. Buna göre, eksi rakamlardan başlanarak 0 rakamına doğru geliriz. dBFS’de 0, teorik nominal nokta kabul edilmektedir. Yani, sinyaldeki aşırılık yüzünden seste distorsiyon (*distortion* = bozulma, cırtlama, kırtlama) meydana gelmeden önceki farazi son basamak... Ancak, kısık sinyalleri yükseltebilmek için, genellikle 0’ın üzerine +10 hatta +20 puan hareket alanı daha eklenebilmektedir.

Analog sinyalden saniyede kaç örnek (SAMPLE) alınırsa, bu bize SAMPLING RATE (örnekleme oranı) sayısını verir. Endüstri standardı olarak bugün 44,100 *Hz* ve 48,000 *Hz* kabul edilmektedir. 44.1 *kHz* saniyede 44100 örnek, 48 *kHz* saniyede 48000 örnek alındığını bildirir. 96 *kHz* düzeyinde çalışan sistemler dahi vardır, ki o zaman saniyede 96000 örnek alındığı anlaşılmaktadır.

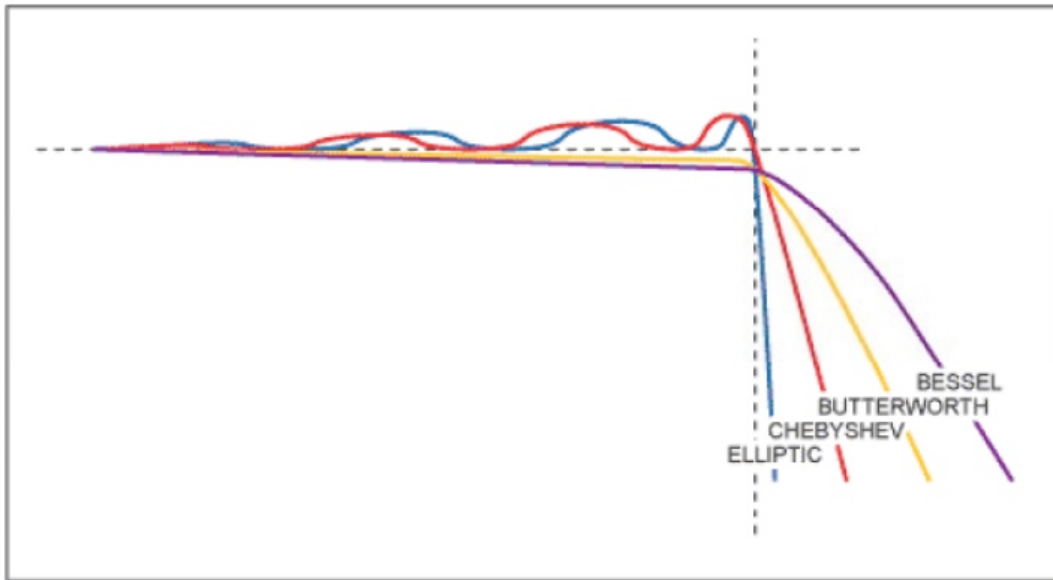
Sampling rate, yani “örnekleme oranı”, aynı zamanda, müsait olan frekans aralığını belirler. Ne kadar yüksek bir oran sözkonusu ise, o kadar geniş bir frekans bandı müsaittir. Bunun prensibini 1920’lerde BELL LABS’de çalışırken *Harry Nyquist* bulmuş ve kuramsallaştırmış.

Nyquist kuramınca, dijital ortamda kaydedilebilecek veya üretilebilecek en yüksek frekans, SAMPLING RATE’in yüzde 45’i ila 50’si kadar olabilmektedir. Buna göre, 20 *kHz* frekansa sahip bir sesi dijitize edebilmek için, 40 *kHz* kadar örnekleme çözünürlüğü gerekir. 20 *kHz*’in üzerinde ise “*alias* frekanslar” denen tuhafliklar ortaya çıkar. “*Alias*”, sahte, kabalaşmış anlamına gelmektedir ve bunların oluşumuna dair örnek bir şekil aşağıda sunulmuştur:



Alias frekanslar, *anti-alias filtrelerden* geçirilerek imha edilmelidir. Duyulabilecek en tiz frekans olan 20 kHz itibariyle düşünürsek, buradan yukarısını bıçak gibi kesecek filtre olmadığı bir yana, filtrelerdeki muhtemel arızalar ve aksamalar da hesaba katılarak, tize doğru (duyulamayacak bile olsa) daha fazla ses alanı alınmalıdır. Bu yüzden, 44.1 kHz SAMPLING RATE kalitede alt sınır kabul edilir. Bugün, müzik CD'leri bu esasa göre hazırlanmaktadır. Daha büyük prodüksüyonlar ise 96 kHz 'ten aşağısına tamah etmemektedir. Halbuki, teorik olarak CD kalitesi gereğinden fazla iş görmektedir. Bazı filtreleri gösteren şekil aşağıdadır:

Most common filter types



Eğer, sözelimi, 10 kHz 'lik bir örnekleme oranı ile çalışırken $7,000\text{ Hz}$ 'lik sinyalleri kaydetmeye kalkarsak, bunlar Nyquist teorisinde bulgulandığı üzere, kaydedilmeyip, orijinal sinyale ait olmayan, $10 - 7 = 3\text{ kHz}$ 'lik ALIAS (sahte) frekanslar üretilecektir. Kaydedilmesi olanaksız frekans ne kadar tiz olursa, üretilen sahte frekans da, duyurulmak istenen frekansın en yakın SAMPLING RATE katsayısından farkı kadar olacaktır.

Yapılması gereken, sistemin kaydetmesine olanak bulunmayan müteceviz frekansları elimine etmektir, ki bunun için anlatıldığı üzere dijital filtreler kullanılır ve sonuçlar tatmin edicidir.

Buna rağmen, mükemmeli ararken, yüksek dijital çözünürlükten ödün verilmemesi gerektiğini savunanlar az değildir. Nitekim, 48 kHz ve 96 kHz örnekleme oranı ile çalışan makinalar işi garantiye almış sayılmaktadır.

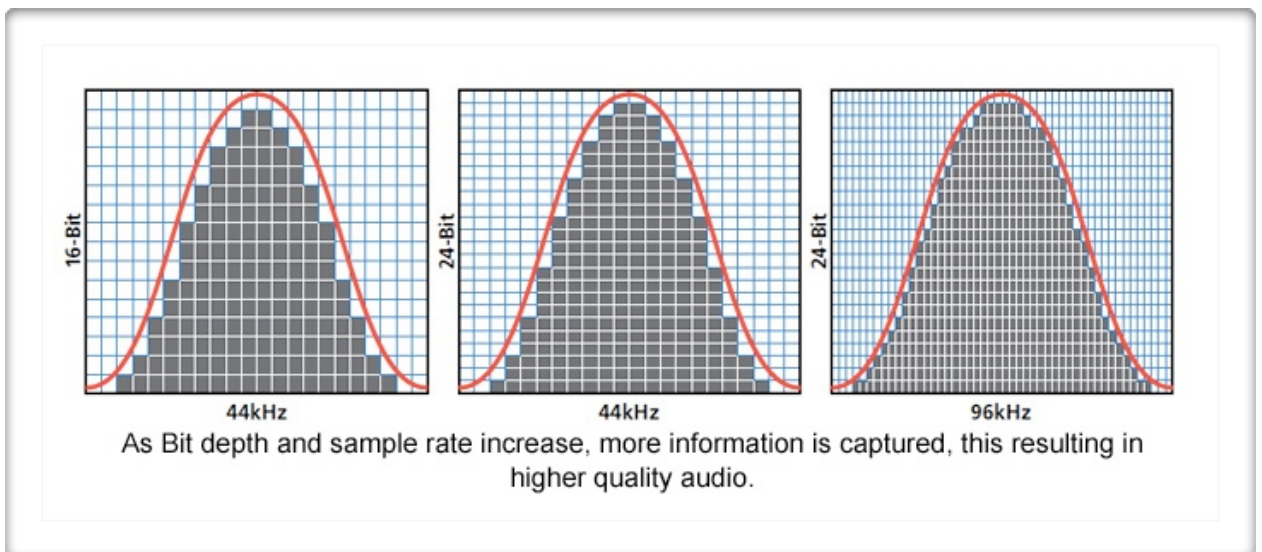
Bir analog sinyali dijitalde çevirirken, *oversampling* (sample rate x 128), sonra filtre ve nihayet *downsampling* (bölü 128) işlemleri sayesinde, gürültü geniş bir frekans aralığına yayılarak ve insan kulağının duymayacağı bölgelere dağıtılarak, yokedilmeye çalışılmaktadır.

Downsampling yapan devreye DECIMATOR adı verilir. Dijitalden analoga geçişte ise çarpım sayısı daha düşük olmaktadır ve *anti-alias filtre* ne kadar işe yarıyor, daha ne kadar geliştirilebilir, bunu anlamaya hizmet eder; yani, daha ziyade test maksatlı düşünülebilir.

Analog dalgaformunun dijitalde temsili için ne kadar bit kullanılırsa, o kadar iyi taklit söz konusu olur. Dijital teknoloji kuvantizasyon (*quantization*) demek olduğundan, ayrıntı neyse, o imkan nisbetinde sayılara yuvarlama söz konusudur: Yani, irrasyonel sayıların virgülden sonraki *n*'inci haneden sonra en yakın mümkün sayıya yuvarlanması... 8-bit çözünürlükte bir ses kaydı bu yüzden çok kaba-saba olacaktır (ilkel telefon teknolojisiyle yapılan konuşmalarda olduğu gibi).

16-bit CD kalitesi için geçerlidir. 24-bit ve 32-bit çözünürlükler tercih edilen en yüksek seviyeler olarak günümüzde kullanılmaktadır. Yüksek bit rakamı (*bit-depth*) sayesinde, sesin dalgaformunun detayı ve kalitesi artmaktadır.

16-bit 65536 (2^{16}) farklı seçenek sunar. 24-bit 16,777,216 (2^{24}) seçenek sunar. Sözelimi, bilgisayar ekranlarında 24-bit renk seçimi yaptığımızda, işte bu kadar farklı renk (göremesek bile) üretilebiliyor demektir. Dijital ses çözünürlüklerine dair örnek grafik ise aşağıdadır:

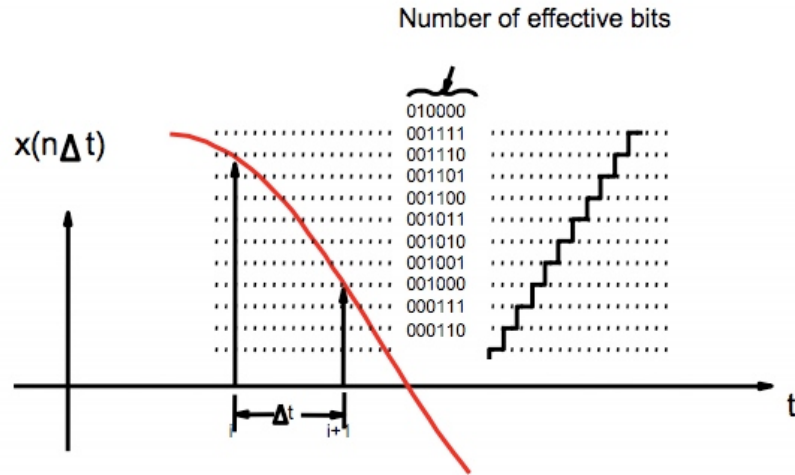


8-bit kayıtlarda “dinamik alan aralığı” (teoride en sönük ile en gür üretilebilecek seslerin arası) 48 dB’dir. 16-bit için 96 dB’dir. 24 bit için 144 dB’dir. Buna dair grafik aşağıdadır:

Dynamic range and bits

$$\text{dB} = 20 \log 2^N \text{ (number of ADC bits)}$$

12 bits	72 dB
14 bits	84 dB
16 bits	96 dB
24 bits	144 dB



16-bit çözünürlükte bir kayıttan yola çıktığımızda, 24-bit veya 32-bit çözünürlükte çalışan sinyal işlemcilerle iş görüp sonra tekrar 16-bit'e dönersek, veri kaybı olur; zira burada da yukarıda bahsedilen sayısal yuvarlama geçerlidir. Yüksek bit sayılı bir seviyeden düşük bit sayılı bir seviyeye gelindiğinde, *bit truncation* denilen ayrıntı kaybı yaşanır.

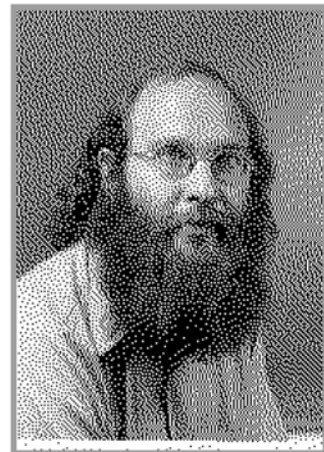
Bu kaybı telafi etmek için, gelişigüzel *white noise* (beyaz gürültü) sinyale uygulanabilir. Bu uygulama DITHER diye isimlendirilir. Grafik dünyasıyla paralellik kurabileceğimiz örnek bir şekil aşağıdadır:



Original
(8 bits)

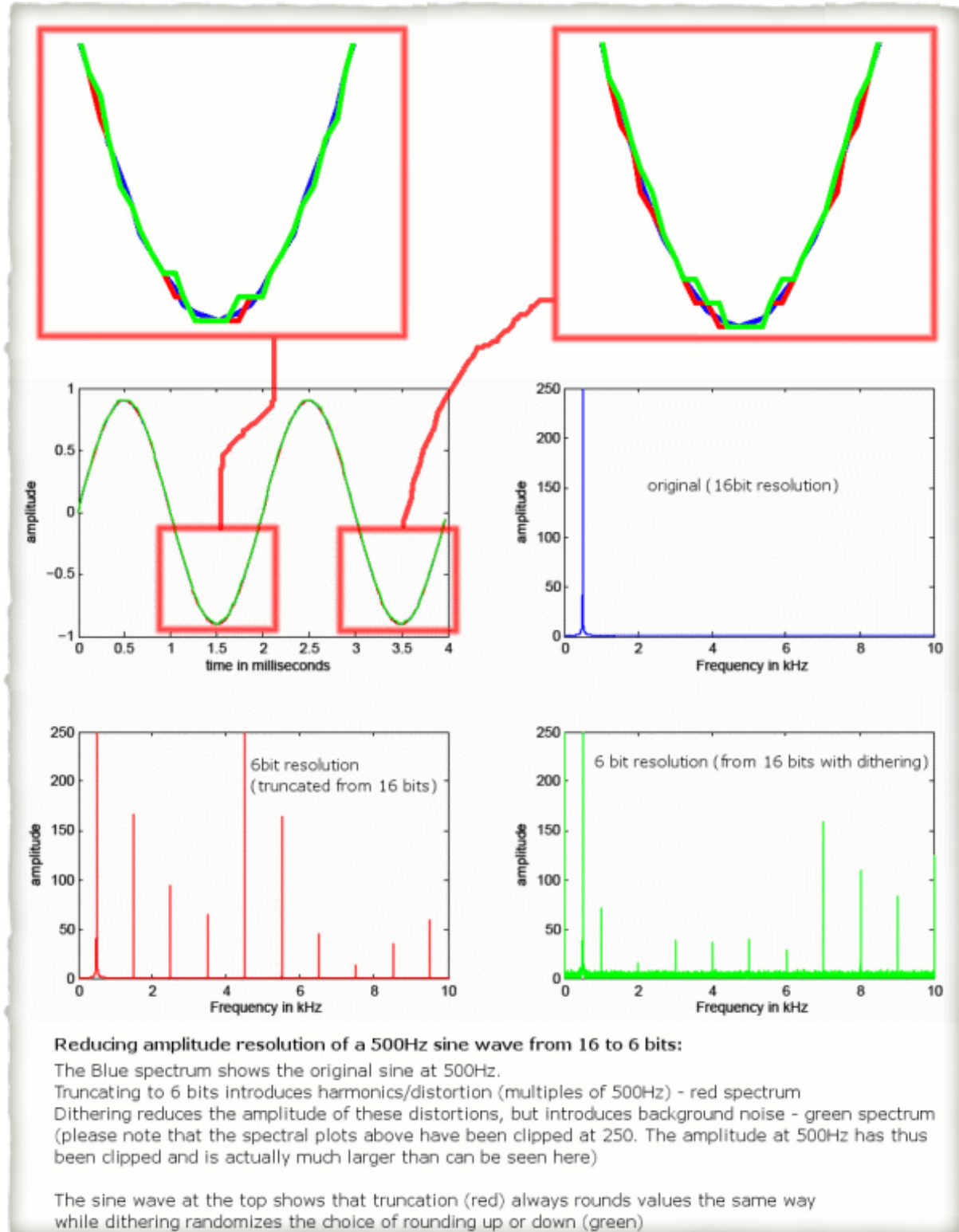


Uniform
Quantization
(1 bit)



Floyd-Steinberg
Dither
(1 bit)

Dither'in dinamik aralık açısından önemi büyüktür. 24-bit'ten 16-bit'e inildiğinde kaybedilen 48 dB'lik dinamik alan aralığı, Dither sayesinde geri kazanılabilir. Böylece, -96 dBFS şiddetindeki bölgeler sinyale geri katılabilir (yani, *noise floor* denilen "dip gürültü eşiği" duyulamayacak kadar sönük seviyelere geriletilir). Böylece, kulağa 115 dB dinamik aralığa sahipmiş gibi kendini hissettiren 16-bit çözünürlüğünde müzik kaydı elde etmek mümkün hale gelmektedir. Dalgaformuna *dither* katılmasına dair örnek bir şekil aşağıdadır:



Beyaz gürültü, insan kulağının hassas olduğu ve olmadığı bölgeler dikkate alınarak uygulanırsa, gürültünün en fazla algılandığı yerlerde gürültüyü alçak, diğer yerlerde normal seviyede tutmak suretiyle, NOISE-SHAPED DITHER gerçekleştirilir. Bu sayede, 16-bit'lik kayıtlarda 20-bit performansı yakalanabileceği ortaya çıkmaktadır.

Dinamik alan anlatılırken atlanılan şu husus şimdi tahkim edilecektir: “*Signal-to-Error ratio*” veya “*Signal-to-Noise ratio*” denilen büyüklük, analogdan dijitale dönüştürüm yapıldığında, orijinal dalgaformuna kıyasla, dijitalizasyon yüzünden sinyalin ne kadar kabalaştığını, yani ne kadar gürültü eklendiğini (*desibel* birimince) belirtmektedir. Oran ne kadar büyük bir sayı verirse, dönüştürülen sinyal o kadar temiz dijitale aktarılabilmiş demektir.

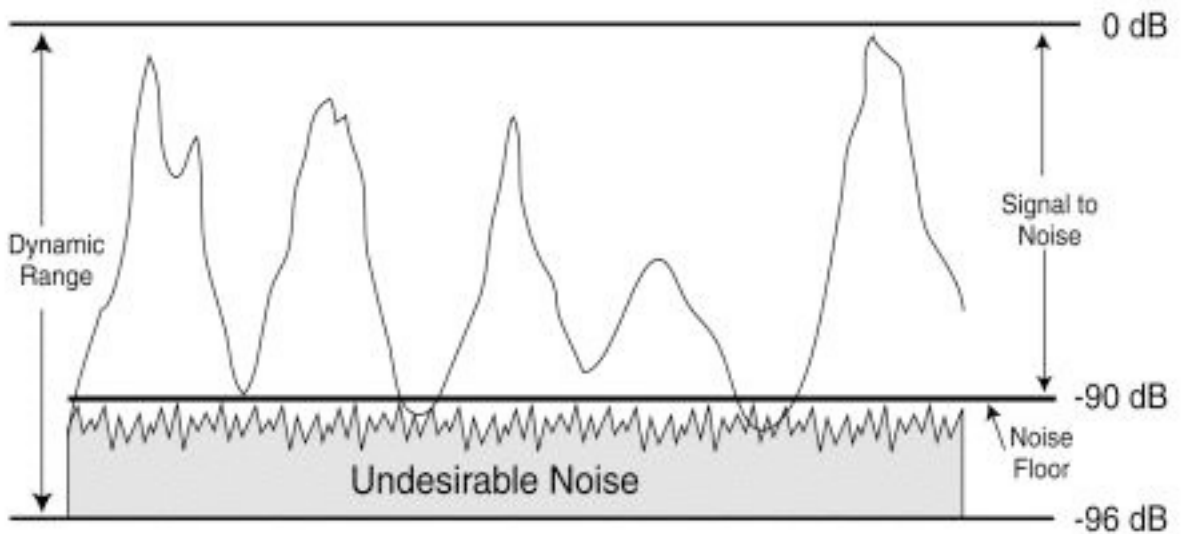
Sözkonusu oranı bulmanın formülü, dinamik alan için verilen örnek şekilde de görüldüğü üzere, şöyledir:

$$\begin{aligned} S/E &= 20 \log_{10} \{2^n\} \\ &= 20 \times n \times \log_{10} 2 \\ &= 6.020599913279624 \times n \text{ kadar dB} \end{aligned}$$

n sayısı kaç bit çözünürlükte çalıştığımızı belirteceğimiz kısımdır. Mesela, 16-bit ayrıntıda iş görölüyor ise, Sinyalin hataya oranı

$$6 \times 16 = 96.33 \text{ dB olur.}$$

Yani, eklenen her *bit* çözünürlük, hata payında (gürültüde) 6 *desibel* azalmaya denktir. S/E, aynı zamanda, idealdeki “dinamik alan aralığını” (pratikte en sönük ses tabanı ile en gür ses tavanı arasını) da bize bildiriyor. Dinamik alana karşılık S/E aşağıda gösterilmiştir:



Gelelim MIDI konusuna. *Musical Instrument Digital Interface*, 1982 yılında NAMM tarafından temelleri atılan, elektronik çalgılar ve cihazlar arasında müzik performans bilgilerinin paslaşılabilirdiği bir protokoldür. NAMM ise *International Association of Music Merchants* (Uluslararası Müzik Tüccarları Cemiyeti) başlığının kısaltmasıdır. İlk MIDI aygıtları ROLAND ve SEQUENTIAL CIRCUITS firmaları tarafından 1983 yılında piyasaya sürülmüştür.

MIDI sayesinde, tek kişilik stüdyo kurmak ve müzisyenlerle uğraşmaksızın demo, aranjman, beste yapabilmek mümkün hale gelmiştir. MIDI, akustik çalgılardan elde edilen örneklenmiş sesler (SAMPLE SOUNDS, SAMPLE LIBRARY, SOUNDSET, vs...) üzerinden notaları tetikleyebileceği gibi, dijital sentezlenmiş sesleri de tetikleyebilir. Bu sayede, canlı performans öncesi bir parçanın nasıl duyulacağına dair sağlıklı fikir edinilebilir.

Günümüzde, MIDI ile çalışan sistemler akustik çalgıları çok iyi taklit edebilen bir düzeye ulaşmıştır. Bu o kadar böyledir ki, büyük prodüksüyonlarda ve film müziklerinde destek çalgılar olarak MIDI ile yapılan seslendirmelere güvenilebilmektedir. Teknolojinin ilerlemesiyle birlikte, asılları ile farkını ayırt etmek artık neredeyse imkansızdır.

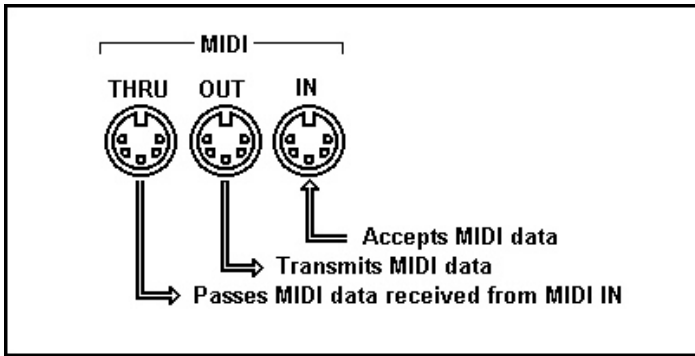
MIDI “*real-time*” çalışır. Yani o anda ne yapıyorsanız onu duyarsınız. Tabii, MIDI aygıtlarında kullanılan bağlantılar ne kadar uzun kablolarla yapılırsa, veya dijital ortamda ne kadar çok giriş-çıkışlı bir bağlantı şematiği uygulanıyorsa, milisaniyeler cinsinden o kadar gecikme yaşanabilir. MIDI performans bilgilerinin önden okunarak yol alınması gereken haller de olabilmektedir: Mesela, *Sibelius 7* ile çalışırken *Wallender Note Performer* ses paketini (SOUNDSET) kullandığımızda, gerçeğe yakın sesler elde edebilmek için, yazılı notaların bir adım önden okunup 1-2 saniye gecikmeli seslendirme yapılması kaçınılmaz olmaktadır. Bu gecikmelere LATENCY denir. (Bu bağlamda, birçok MIDI çalgı ünitesinin arasında eşzamanlılığı sağlamak üzere, MIDI’nin referans aldığı saat [*system clock*] üzerinden “MIDI SYNC” yapılabilmesi sözkonusudur.) Herşeye rağmen, bu tür MIDI uygulamaları hep “*real-time*” kategorisinde değerlendiriliyor.

MIDI işlemlerinin gerçekleştirildiği yazılım veya donanım platformlarına *ardışımlayıcı* (*sequencer*) denir. Ardışımlayıcı ile tüm MIDI verilerini kaydedebilir, editleyebilir, birçok aşamadan geçirerek dönüştürebilirsiniz. İlk başlarda MIDI ardışımlayıcılar donanım şeklinde piyasaya çıkmış: Ya tek başlarına özel “*stand-alone sequencer*” ünitelerin içinde yahut *sentessayzer klavyelerin* içinde bulunuyorlarmış. Halen pek çok sentessayzerin ardışımlayıcı özellikleri mevcuttur. Ancak bilgisayar teknolojisinin hız kazanmasıyla, yazılım biçiminde ardışımlayıcılar piyasaya sürülmüştür. En meşhurları *Pro Tools*, *Logic*, *Nuendo*, *Cubase* gibi markalardır. En popüler ilk ardışımlayıcı CUBASE imiş. Ben tercihen *Logic Pro 8* kullanıyorum mesela.

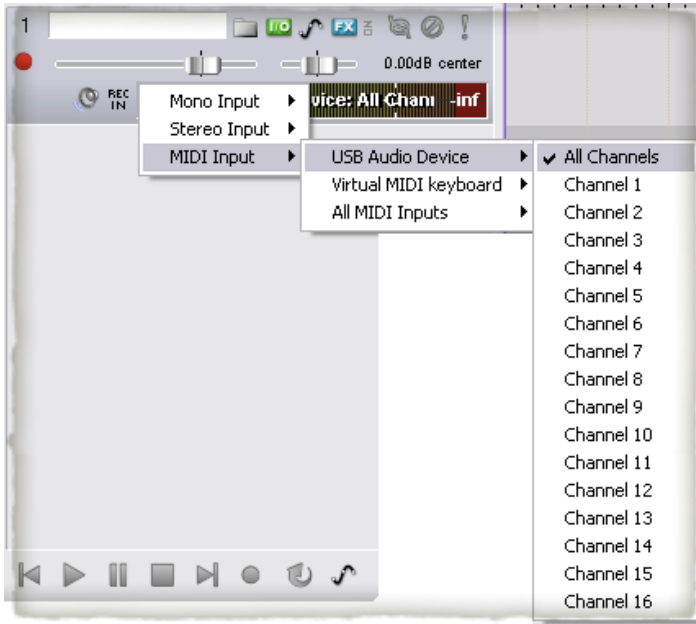
Ardışımlayıcılar artık MIDI’nin ötesinde AUDIO sinyallerin her tür kayıt ve manipölasyonuna imkan verecek şekilde tasarlanıyorlar. Bu itibarla, piyasada sözkonusu platformlara *sequencer* yerine “*DAW: Digital Audio Workstation*” veya “*Host*” (evsahibi) adı veriliyor.

Bilgisayar ile iş yapmanın rahatlığı düşünüldüğünde, yazılım tarzı ardışımlayıcıların niye bu kadar yaygınlaştığı kolaylıkla anlaşılabilir. Günümüz DAW yazılımları bize, ancak kullandığımız bilgisayarın kapasitesi ile sınırlanacak sayıda kanal, kontrol özellikleri, yedekleme imkanları ve dosya paylaşım ayrıcalıkları sunuyor.

MIDI kablolar ve donanım bağlantı şematiklerine aşina olduğumuzu varsayıyoruz. Asıl önemli kısım şudur: Her MIDI Girişi (*port*) 16 tane MIDI kanalına (*channel*) müsaade etmektedir. MIDI kanal bilgileri 4-bit çözünürlükte kodlanır. Mesela, 7. kanal 0110 şeklinde binary (0 ve 1 sayılarından ibaret) kodlanır. Bunlara dair şekiller aşağıdadır:

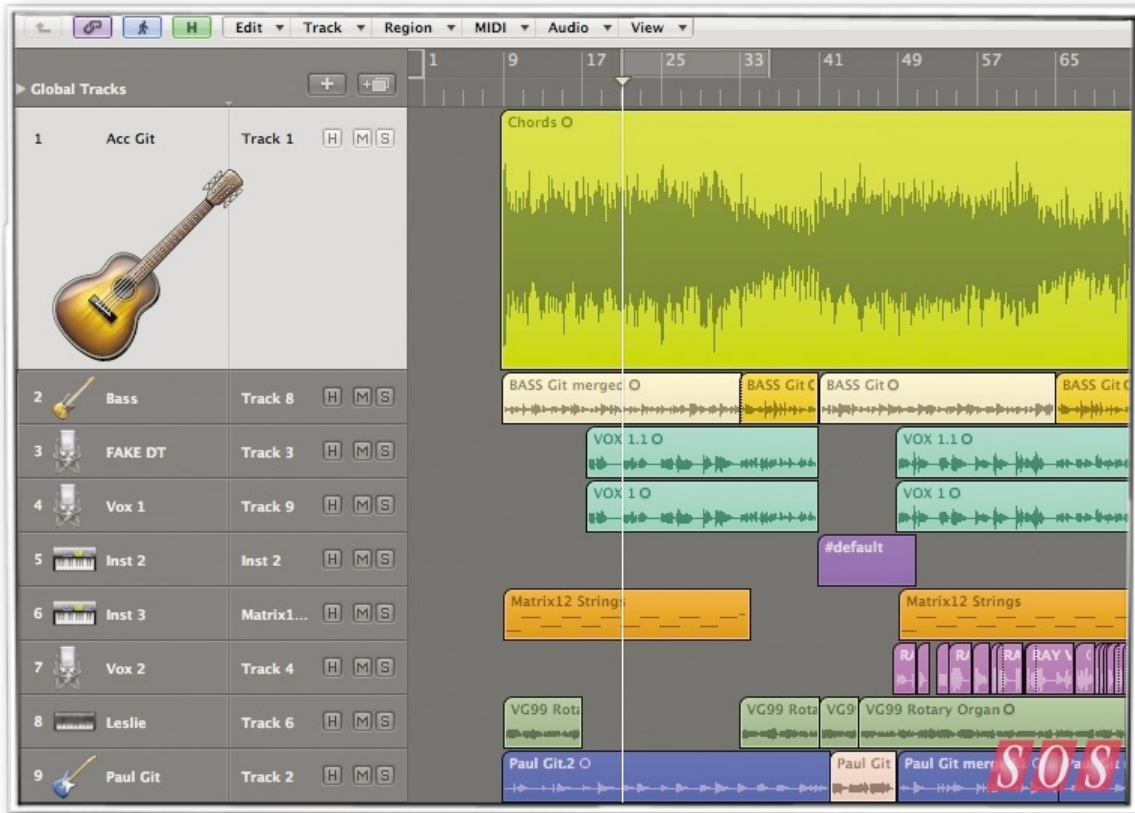


Hexadecimal	Binary	Decimal	MIDI Channel
00	0000	0	Channel 1
01	0001	1	Channel 2
02	0010	2	Channel 3
03	0011	3	Channel 4
04	0100	4	Channel 5
05	0101	5	Channel 6
06	0110	6	Channel 7
07	0111	7	Channel 8
08	1000	8	Channel 9
09	1001	9	Channel 10
0A	1010	10	Channel 11
0B	1011	11	Channel 12
0C	1100	12	Channel 13
0D	1101	13	Channel 14
0E	1110	14	Channel 15
0F	1111	15	Channel 16



Bir MIDI cihazı hangi kanaldan veri alacak şekilde ayarlanmışsa, sadece o kanalın aktardığı (*transmit*) veriyi alır. Çoğu MIDI enstrümanı, birden çok kanaldan aynı anda veri kabul edecek şekilde (*receive*) ayarlanabilir. Bu özelliğe sahip MIDI aletleri *multitimbral* diye isimlendirilir.

PORT sayısı ne kadar çok olursa, mümkün kanal sayısı 16'şar artar (16, 32, 48, 64, vs...). Ardışımlayıcılar AUDIO ve MIDI bilgilerini bu şekilde, her bir Kanal bir TRACK'e karşılık gelecek şekilde ayarlar. TRACK, ardışımlayıcı içinde sinyal veya MIDI veri kaydının nakledileceği, ekranda görünen patıkadır. Unutulmamalıdır ki, MIDI ses taşımaz! MIDI sadece hazırdaki sesleri tetikleyen bilgiyi taşır. Aşağıdaki resimde, *Logic Pro*'nun çalışma penceresinde AUDIO ile MIDI'nin birlikte olduğu örnek bir oturum sunulmuştur:

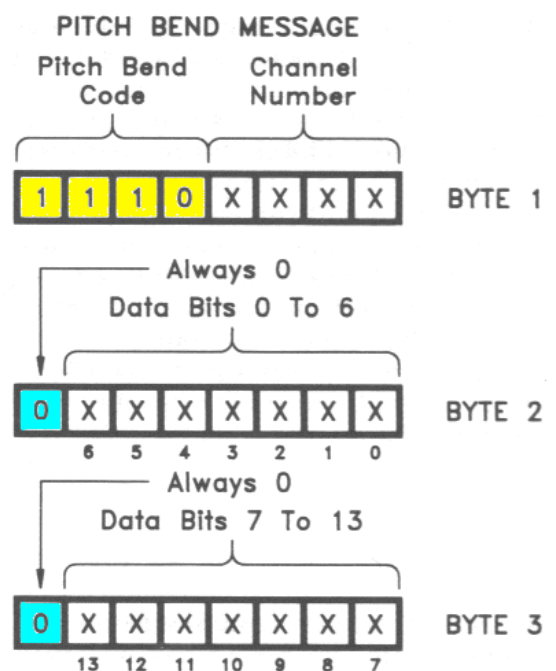


MIDI mesajları iki cinstir: STATUS message (1 byte) ve iki adet DATA message (1+1 = 2 byte). İlki *durum bilgisi* taşır, diğeri *veri bilgisi* taşır. Durum mesajları, iletilen bilginin türüne ve kanal numarasına dairdir. Veri mesajları ise, bunların değerlerini belirtir. STATUS+DATA mesajlarının işleyişine dair örnek bir şekil aşağıda verilmiştir:

Sözgelimi; STATUS: 2. kanalda bir notaya basıldı, DATA: basılan nota C4'tür, şiddeti felan değerindedir, perde bükme belirtildiği kadardır...

1 BYTE = 8 bit. Bu önemli. Bunda dikkat edilmesi gereken şu: Bu 8 bitlik paketin ilk (en sol sıradaki) biti MSB (*Most Significant Bit*), yani “en mühim bit” diye isimlendirilir. STATUS mesajlarında MSB’nin aldığı değer her zaman 1, DATA mesajlarında her zaman 0’dır. Yani, MIDI mesajlarının hangi cins olduğu bu şekilde ayırt edilir. Geriye de zaten 7-bit kalıyor. İki adet 7-bit’lik DATA mesajı toplamda 14-bit çözünürlük demektir.

14-bitlik bir MIDI DATA mesajına dair örnek vermek istersek:



STATUS mesajı dese ki BEN PERDE BÜKME MESAJIYIM ve 6. KANALDAYIM...
ilk 7-bitlik DATA mesajı da dese ki, 0-127 arasında 45. değerim,
ikinci 7-bitlik DATA mesajı da dese ki, 0-127 arasındaki 69. değerim;

işte burada ilk 7-bit MOST SIGNIFICANT BYTE, ikinci 7-bit ise LEAST SIGNIFICANT BYTE olarak tanımlanır. Bu iki adet 7-bit sayesinde, 0 ila 127 arasında tam 128'er değer belirtilebilir. MSB ve LSB kısaltmaları görüldüğünde asıl kastedilen bunlar olacaktır. Birinci veri mesajı ile, sözgelimi, 100 sent olarak saptanmış olan eşit yarım ses aralığı sahasında (-100 sent _ 0 sent _ +100 sent) aşağı yukarı -64 ve + 63 adım gidilebilecek olduğuna göre (zira 0 dendiğinde perde bükme yok demektir), 69. değer belirtilmek suretiyle 0 ila +100 sent olan kısım 64'e bölünüp $(69 - [\{127+1\}/2]) = 5$ ile çarpılarak *COARSE AMOUNT* denilen kaba ayar miktarı $([100/64] \times 5 = 7.8125 \text{ sent})$ bulunacaktır.

LSB de işin içine girdiğinde, MSB'deki 69. (7.8125 sent) ile 70. (9.375) değerler arası (ki 1.5625 senttir) 128'e bölünüp 45 ile çarpılacak ve 69. MSB değerine *FINE TUNING* katılacak demektir:

$$\begin{aligned} &7.8125 \text{ sent} + \{(9.375 - 7.8125 \text{ sent}) \times (45/128)\} \\ &= 7.8125 \text{ sent} + (1.5625 \times 0.3515625) \\ &= 8.36181640625 \text{ sent} \end{aligned}$$

Elde ettiğimiz bu son sayıyı $8192 / 100 = 81.92$ ile çarparsak, 14-bit çözünürlükte, yani +/- 8192 ünite sahasında (0 ila 16383 bandında) nerede olduğumuzu buluruz: 8,877.

Diğer bir deyişle, 8192 değerindeki 0. pozisyonundan ("perde bükme yok" konumundan) yukarı $8877 - 8192 = 685$ değer gidilecektir, ki bu da üstte belirttiğim 8.362 sente karşılıktır.

Sent adedinden 14-bit PERDE BÜKME (PBU) değerine dönmenin yolu ise şöyledir:

$$(8192 / \text{referans perde bükme aralığı}) \times \text{sent adedi}$$

Farklı bir örnek vermek gerekirse, MIDI ortamında bir notayı 256/243 oranına sahip bakiye aralığı kadar, 200 sent olarak belirlenmiş eşit tam ses bükme sahası dahilinde ve tam 2/1 seçilmiş oktava göre bükmek istesek, girmemiz gereken PBU (*pitch bend unit*):

$$\begin{aligned} &(8192 / 200) \times \{[\log_{10} 256/243] / [\log_{10} \text{oktav aralığı}]\} \times 1200 \\ &= 40.96 \times (0.022633691713537 / 0.301029995663981) \times 1200 \\ &= 40.96 \times 0.075187496394218 \times 1200 \\ &= 3,695.615822768597974 \text{'tür;} \end{aligned}$$

yani 3696'ya yuvarlanır.

Bu sayıyı $8192 / 200 = 40.96$ 'ya bölersek, gerçekten de yaklaşık 90 sentlik bakiye aralığına karşılık geldiğini görürüz:

$$3696 / 40.96 = 90.234375 \text{ sent}$$

Tam doğrusu ise aslında:

$$\log_2 (256/243) \times 1200 = 90.224995673063064 \text{ senttir.}$$

Fark 0.009379326936936 sent olup, 0.01 sent hata ile bakiye oranı tutturulmuş olur.

Tabii, ortalama insan kulağının, olağanda işitebileceği frekans tayfının (20-20,000 Hz) orta bölgelerinde 1 sentlik farkı bile güç-bela duyabileceği akıldan çıkartılmamalıdır. Diğer bir deyişle, 14-bit MIDI perde bükme çözünürlüğü beklentilerin de ötesinde akort ayrıcalığı sağlamaktadır.

Mus2 gibi nota programlarından MIDI Export (ihraç) yaparken, eğer partitürde 12-ton Eşit Temperaman dışında perde bükme bilgileri varsa, şu hususlara dikkat edilmelidir:

1. Perde bükme sahası kaynak programda nasıl belirlenmiş ise, hedef programda da birebir aynı olmalıdır. Sözgelimi, kaynak programda her nota aşağı yukarı maksimum 100 sent kımıldatılabiliyorsa, hedef programdaki MIDI çalgılarının perde bükme sahası da 100 sente ayarlanmalıdır. Genel-geçer standart ise 200 sentlik tempere tam sestir. Çoğu koşulda bu ayar yeterlidir, ama oktavlı glissandolar yapılacaksa 1200 sentlik ses sahası seçilmelidir.
2. MIDI notaları aynı kanalda (aynı TRACK içinde) peşpeşe çalınırken, her notanın farklı miktarda bir perde bükme değeri varsa, bir notadan diğerine geçişte istenmeyen “cırıklama” sesleri duyulabilir. Bunun sebebi, nota süresinin sonuna gelindiği halde MIDI çalgısının o nota daha sönümlenmeden diğer notaya geçmesi ve bir sonraki notaya ait perde bükme değerinin ilk notaya isabet etmesidir. Bunu önlemenin yolu, peşpeşe olan bu tür notaları ayrı kanallara dağıtmaktır.
3. Bununla birlikte, tek bir çalgı değil de orkestrasyon sözkonusu ise, o vakit 16 kanal sınırları zorlayabilir. Bu durumda notaları stratejik olarak daha az sayıda kanala optimal dağıtmanın yolları denenmelidir.

4. Her ne kadar ihtimal dışı gibi görünse de, polifoni sınırını aşmamaya dikkat etmek gerekir. MIDI protokolü, ürün markasına göre, kanal başına 32-ses veya 64-ses polifoni imkanı sunar. Bu, aynı zamanda 32 veya 64 notadan daha fazla notanın (çalınan bir nota susturulmadan) seslendirilemeyeceği anlamına gelir.
5. Standart MIDI dosyası formatı iki türdür: Type 0 = *single track*; Type 1 = *multitrack*. Adlarından da anlaşılacağı gibi, ilk tür bütün MIDI patikalarının ayrı ayrı ele alındığı bir formattır ve genellikle okunabilen dosya biçimi budur. İkinci tür ise nadiren karşımıza çıkar ve her ardışıklayıcının okuyamadığı, 16 kanalın her patikasının birarada paketlenildiği bir dosya biçimi demektir. Bu formatları birbirine dönüştürmek için özel program gerekir.

MIDI konusu üzerine daha da çok konuşulabilir: PATCH yahut PROGRAM denilen 128 “enstruman gediği” olduğundan; bazı yazılım ile donanımların bu yetersizliği gidermek için gedik başına 128 BANK şeklinde alt-gedikler koyduklarından ve bu sayede teorik olarak 16384 farklı enstruman depolanabileceğinden; 10. Kanallın genellikle *drum kit* denilen vurmaları ayrıldığından; *velocity layers* denilen uygulama sayesinde 0 ila 127 bandındaki tuş dokunma şiddetlerinin bir enstrumanın (*pianissimo*’dan *fortissimo*’ya kadar) farklı dinamik tınlarını tetiklemeye hasredilebilecek tarzda gruplandırılmasının mümkün olduğundan; *continuous controllers* (“sürekli kontrol mesajları”) ile volüm, panning, modülasyon gibi efektlerin ayarlanmasının kolaylığından, vs... bahsedilebilir. Ancak, bundan sonrası, bilgisayarda MIDI ve AUDIO programlara aşinalık ve Ses Teknolojilerine ilginin sürmesi ile koşut seyredecektir.

Resimler, grafikler ve şekiller internetten edinilmiştir.

Ufuk Önen, 2008, SES KAYIT VE MÜZİK TEKNOLOJİLERİ (3. Basım), Çitlenbik Yayınları, İstanbul.