

وسائط متعددة المحاضرة التاسعة

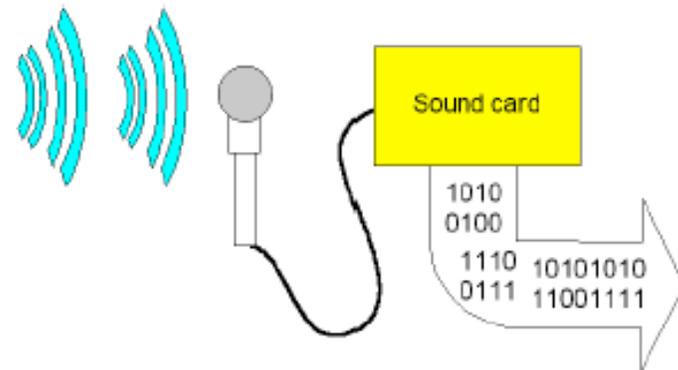
أ.د. فادي غصنه

د. أحمد أحمد

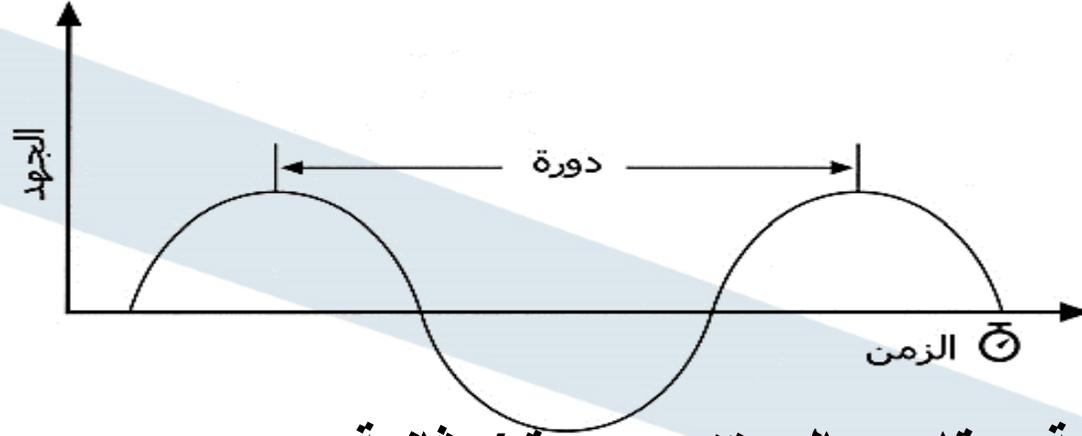


الصوت

- الصوت: عبارة عن اضطراب (تغير) فيزيائي في الوسط (**medium**) حيث ينتقل هذا الاضطراب (التغير) على شكل موجة متغيرة مع الزمن. يمكن نقل الصوت (نشر الصوت) عبر أوساط عديدة (هواء، مواد صلبة) وبالتالي يمكن قياس عن طريق قياس مستوى الضغط (التغير) عند موضع ما (**location**).
- الإشارات الصوتية: إشارات تمثيلية مستمرة متغيرة مع الزمن.



الصوت



• يعتبر الصوت كإشارة حيث يعتمد على :

- 1- التردد **frequency** : عدد النبضات في الثانية ويقاس بالهرتز = دورة / ثانية
- 2- المطال **Amplitude** : مقدار الإزاحة لموجة الصوت انطلاقاً من موجة الضغط الجوي في منطقة انتشار الصوت .

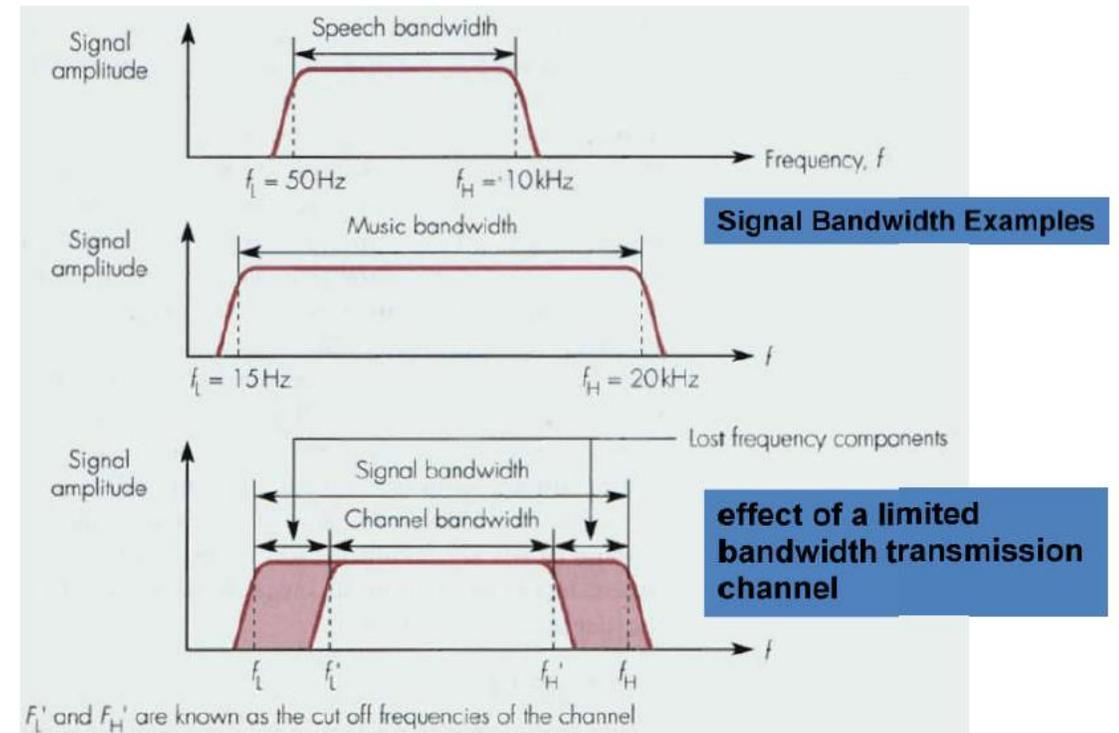
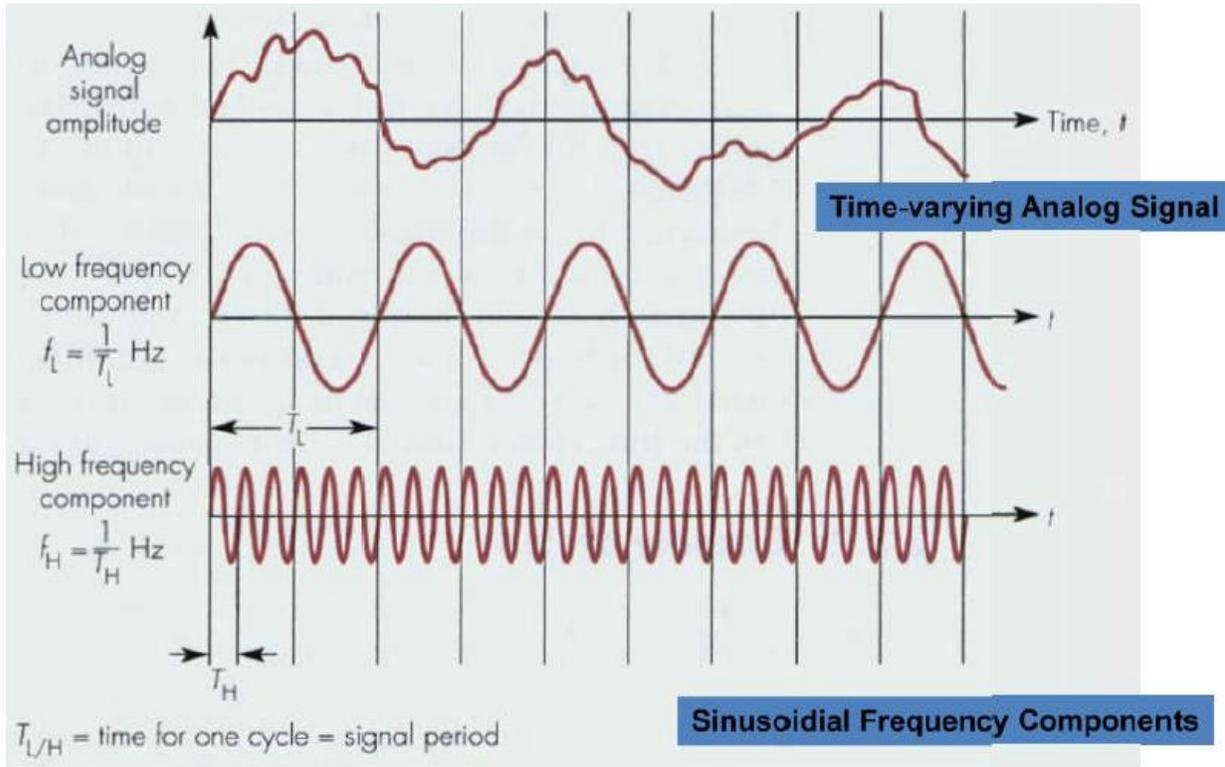
• تتحسس الأذن البشرية لمجال واسع من الترددات الصوتية يمتد بين (20 HZ → 22KHZ) وتكون ذات حساسية قصوى للترددات الصوتية بين (2KHZ → 4KHZ) ، بينما الصوت الصادر عن الإنسان فيمتد في المجال (500HZ → 2KHZ) ، وهنا نميز مفهومين :

- 1- المجال تحت الصوتي **Infra Sonic** : (0 → 20 HZ).
- 2- المجال فوق الصوتي **Ultra Sonic** : (20 KHZ → 1 GHZ).



الصوت

Signal Properties



الصوت

- تتطلب جودة الصوت في القرص المضغوط (CD) أخذ عينات بمعدل ١٦ بت عند 44.1 KHz، و من الممكن حتى معدلات صوتية أعلى (مثال 24 bits بتردد 96KHz).
- في حالة الكلام: 10.025 KHz، أما النظام التلفوني: 8 KHz.
- جودة الصوت المنخفضة: 10.05 KHz (WWW Audio, AM Radio)
- الحد الأعلى للسمع من قبل الكائن البشري هي 20-22 KHz.
- بعد أخذ العينات تأتي عملية التكميم، 8 bit تعطي من 0-255، و 16 bits تعطي من 0-65535 مستوى.
- يحتاج الصوت مساحة تخزين كبيرة في حال عدم الضغط.
- مثال: دقيقة واحدة من الصوت بجودة قرص مضغوط أحادي (غير مضغوط) = ٥ دقائق. صوت بجودة قرص مضغوط ستيريو (غير مضغوط) = ١٠ دقائق.



بعض تطبيقات الصوت

- توليف الصوت الرقمي: إنتاج بعض الأصوات.
- مؤثرات الصوت الرقمية: تغيير الأصوات باستخدام بعض المؤثرات القياسية.
- **MIDI: (Musical Instrument Digital Interface)** (واجهة رقمية للألات الموسيقية).
- توليف الصوت والتحكم في المؤثرات وضغطها.
- ملفات الصوت الشائعة ما يلي: **.wav** ، **.ai** ، **.au**.
- ضغط الصوت:
- **DPCM**: طريقة ضغط صوت بسيطة وواسعة الانتشار و هي تعديل شفرة دلتا النبضي التكيفي (**Adaptive Delta Pulse Code Modulation**). بناءً على العينات السابقة، تتنبأ هذه الطريقة بالعينات التالية وترمز الفرق بين القيمة الفعلية والقيمة المتوقعة.



طبيعة الصوت

- يمكن تمييز الموجات الصوتية بالسمات التالية: الدور (Period) – التردد (Frequency) – عرض الحزمة (Bandwidth) - مستوى الصوت (Loudness) - النطاق الديناميكي (Dynamic range).

• **Loudness = volume**

- للصوت شدة مختلفة، تحمل قدراً من الطاقة تقاس بالواط (watt).
- شدة الصوت: مقدار الطاقة المنقولة عبر مساحة متر مربع واحد (1 m^2), عمودية على اتجاه انتشار الصوت.
- عتبة الإحساس: إذا كانت شدة الصوت 1 watt/m^2 , فقد نبدأ بالشعور بالصوت، و هذا يكافئ (120 dB).
- عتبة السمع: إذا كانت الشدة $10^{-12} \text{ watt/m}^2$, فقد نتمكن من سماعه

فقط (0 dB) <https://manara.edu.sy>



طبيعة الصوت

- يمكن تمييز الموجات الصوتية بالسمات التالية: الدور (Period) – التردد (Frequency) – عرض الحزمة (Bandwidth) - مستوى الصوت (Loudness) - النطاق الديناميكي (Dynamic range).

• Loudness = volume

- النطاق الديناميكي: يعني التغير في مستويات الصوت.

▪ يعبر عن شدة الصوت بـ: **Sound Pressure Level (SPL) = 10 Log (P1/P2)**

حيث أن **P1**: مستوى ضغط الصوت، **P2**: المستوى $10 \log \frac{10^{-12}}{10^{-12}} = 0 \text{ dB}$

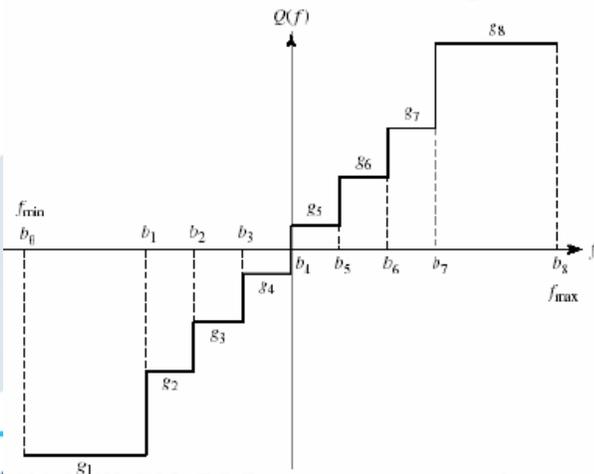
▪ النقطة المرجعية لـ . ديسيبل هي عتبة السمع البشري

▪ مستوى الانتقال إلى 1 watt/m^2 **Dynamic Range: $10 \log \frac{1}{10^{-12}} = 120 \text{ dB}$**



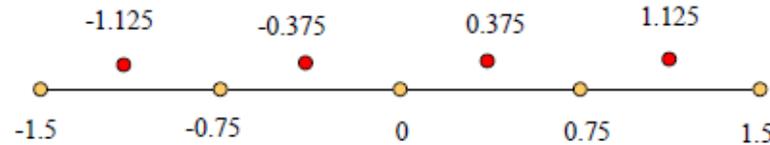
التكميم (Quantization)

- يقسم المجال الكلي للمعطيات الى L مجال جزئي متساوي بطول Q (يعرف بمجال التكميم أو حجم خطوة التكميم)
 - يتم تنسيب قيمة العينة المأخوذة i الى قيمة منتصف المجال الجزئي.
- يوجد نوعان من المكيمات:
 - التكميم الموحد: تكون فيه مجالات التكميم متباعدة بشكل متساوٍ.
 - التكميم غير الموحد: التباعد بين المستويات غير متساوي. يكون عادة لوغاريتمي. تكون الإشارات الصوتية الحقيقية (الكلام والموسيقى) أكثر تركيزاً بالقرب من الاصفر. مجال التكميم أصغر بالقرب من الصفر.



التكميم (Quantization)

- مثال عن التكميم الموحد:
- ليكن لديك لتسلسل التالي $\{1.2, -0.2, -0.5, 0.4, 0.89, 1.3 \dots\}$ ، قم بتكميمه باستخدام مُكَمّ موحد في نطاق $(-1.5, 1.5)$ مع 4 مستويات، واكتب التسلسل المُكَمّ.
- الحل:



$$L = 4, Q = 3/4 = 0.75$$

Yellow dots indicate the partition levels (boundaries between separate quantization intervals)
Red dots indicate the reconstruction levels (middle of each interval)

1.2 fall between 0.75 and 1.5, and hence is quantized to 1.125

السلسلة المكّمة: $\{1.125, -0.375, -0.375, 0.375, 1.125, 1.125\}$



نسبة الإشارة إلى ضجيج التكميم (SQNR)

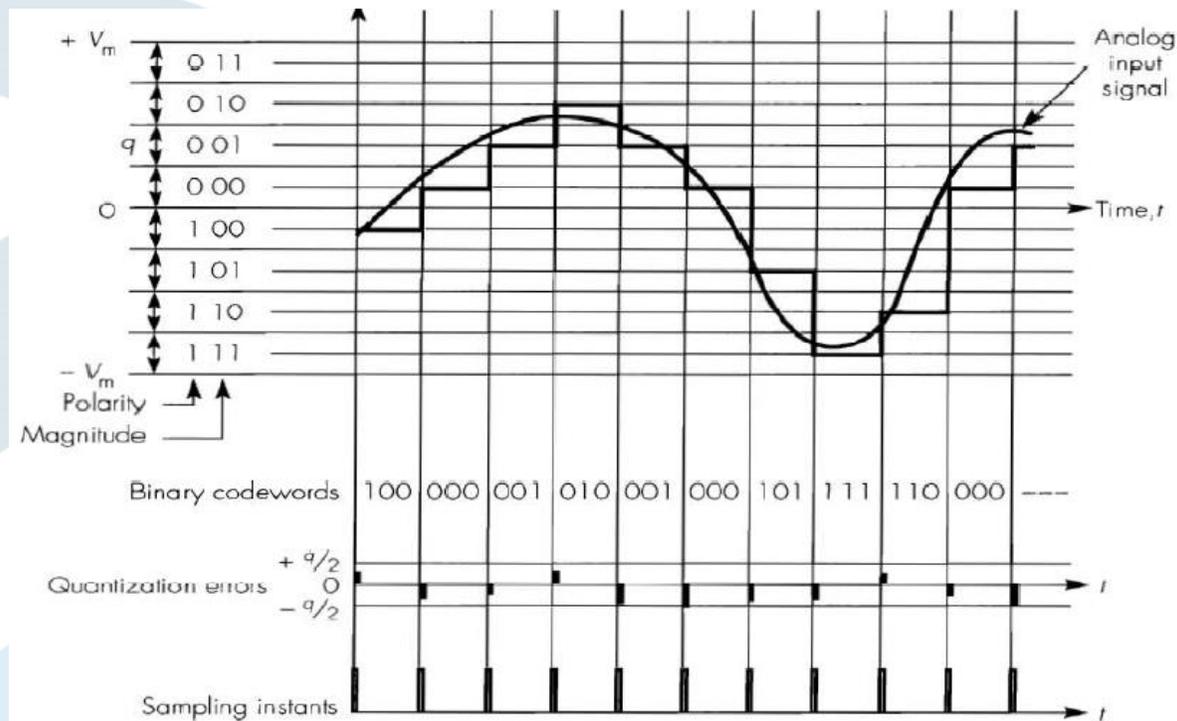
- ضجيج التكميم هو الفرق بين قيمة الإشارة التناظرية، عند أخذ عينات محدد، وقيمة مجال التكميم الزمني الأقرب.
- على فرض لدينا **N bits** لكل عينة مأخوذة، سيكون في هذه الحالة عدد مستويات التكميم الممكنة: $1 - 2^{N-1}$ to $2^{N-1} - 1$.
- إذا كانت قيم الإشارة التمثيلية الفعلية ضمن النطاق من: $-V_{max}$ to $+V_{max}$ ، يمثل كل مستوى كمية جهداً قدره: $\frac{V_{max}}{2^{N-1}}$ or $\frac{2V_{max}}{2^N}$ ، في هذه الحالة تحسب **SNR** :

$$SNR(dB) = 20 \log_{10} \left(\frac{V_{signal}}{V_{quant-noise}} \right) = 20 \log_{10} \left(\frac{2^{N-1}}{\frac{1}{2}} \right) = 20 * N * \log_{10}(2) = 6,02 * N \text{ (dB)}$$

وهذا يعني أن كل بت يضيف حوالي ٦ ديسيبل من الدقة، وبالتالي فإن ١٦ بت تمكن



نسبة الإشارة إلى ضجيج التكميم (SQNR)



مثال عن (SQNR)

- تتمتع الإشارة التمثيلية بنطاق ديناميكي يبلغ 40 dB . المطلوب: تحديد حجم ضجيج التكميم بالنسبة إلى الحد الأدنى لسعة الإشارة، إذا كان المُكَمَّم يستخدم: **6 bits** أولاً ثم **10 bits** ثانياً، تفترض أن الحد الأقصى للضجيج هو نصف خطوة التكميم. علق على ما إذا كان مستوى ضجيج التكميم مقبولاً أم لا.

$$D = 20 \log_{10} \left(\frac{V_{max}}{V_{min}} \right) \Rightarrow 40 = 20 \log_{10} \left(\frac{V_{max}}{V_{min}} \right) \Rightarrow V_{min} = \frac{V_{max}}{100}, \text{Quantization Error} = \mp \frac{Q}{2} = \mp \frac{V_{max}}{2^B}$$

$$B = 6 \text{ Bits} \Rightarrow \frac{Q}{2} = \mp \frac{V_{max}}{2^6} = \mp \frac{V_{max}}{64}, \text{ضجيج التكميم هو أكبر من } V_{min} \text{ و هو غير مقبول.}$$

$$B = 10 \text{ Bits} \Rightarrow \frac{Q}{2} = \mp \frac{V_{max}}{2^{10}} = \mp \frac{V_{max}}{1024}, \text{ضجيج التكميم هو أصغر من } V_{min}, \text{ و له تأثير أقل بكثير.}$$



الجودة مقابل حجم الملف

- يعتمد حجم التسجيل الرقمي على معدل أخذ العينات والدقة وعدد القنوات:

$$S = R * \left(\frac{b}{8}\right) * C * D$$

حجم الملف S bytes ، معدل أخذ العينات R samples/second ، الدقة b bit ، قنوات C : 1 التسجيل الأحادي،
٢ التسجيل ستيريو ، مدة التسجيل D seconds .
معدل أخذ عينات أعلى ودقة أعلى يعطي جودة أعلى ولكن حجم ملف أكبر.

- مثال: إذا سجلنا 10 seconds من الموسيقى الاستريو بمعدل $44.1kHz$ و 16 bits ، فسيكون الحجم:

$$S = 44100 * (16/8) * 2 * 10 = 1\,764\,000 \text{ bytes} = 1.68 \text{ Mbytes.}$$

ملف صوت عالي الجودة، يمكن تقليل الحجم عن طريق الضغط



مثال ١

- عرض نطاق إشارة الكلام هو **10 kHz** \Rightarrow **50 Hz** وإشارة الموسيقى: **15Hz**
 - **20KHz** \Rightarrow المطلوب رقمنة هذه الإشارات باستخدام نايكويست.
 - ✓ ما هو معدل البت الناتج عن الكلام إذا تم استخدام 12 bits/sample ؟
 - يجب أخذ عينات من إشارة الكلام بتردد ٢٠ كيلوهرتز (معياري نايكويست).
 - $The \text{ bit rate} = 12 * 20 = 240 \text{ Kbps.}$
 - ✓ قم بنفس الشيء بالنسبة لإشارة الموسيقى باستخدام 16 bits/sample
 - يجب أخذ عينات من إشارة الموسيقى بتردد ٤٠ كيلوهرتز.
 - $the \text{ bit rate} = 16 * 40 = 640 \text{ Kbps.}$
 - ✓ ما هي مساحة التخزين التي تحتاجها للاستماع إلى موسيقى استريو لمدة ١٠ دقائق؟
 - تتطلب **10** دقائق من الموسيقى الأحادية $384 \text{ Mbits} = 10 * 60 * 640K$. أما بالنسبة للموسيقى الاستريو ، $96 \text{ Mbytes} = 48 * 2 =$



مثال ٢

■ الحد الأقصى للتردد العالي للسمع البشري $\cong 20 \text{ kHz}$ ، ولكن أظهرت الدراسات أن الكلام المفهوم يتطلب تردد فقط حتى 4 kHz .

✓ برّر لماذا معدل أخذ العينات للقرص المضغوط الصوتي هو 44.1 kHz . ما هو معدل نيكويست للنقل الصوتي الموثوق؟

• الأذن هي المُستقبل النهائي للصوت الرقمي المُقدّم. ولأنها لا تستطيع سماع ترددات تزيد عن 20 K ، يجب أخذ عينة من الإشارة الأصلية عند 40 KHz على الأقل لالتقاط جميع ترددات الـ 20 KHz . مع بعض المساحة الإضافية للنطاقات الأعلى، تم أخذ عينات صوتية من القرص المضغوط بمعدل 44.1 KHz . معدل أخذ العينات الموثوق به للتواصل الكلامي هو $2 * 4 \text{ kHz} = 8 \text{ kHz}$.

✓ لماذا تعتقد أن صوت الأشخاص على الهاتف يختلف عن صوتهم في الواقع؟

• يتم أخذ عينات من أصوات الأشخاص وإعادة بنائها من خلال خطوط الهاتف، مما يؤدي إلى فقدان الترددات العالية، وهذا يؤدي إلى صوت مختلف عن الصوت الحقيقي.



مثال ٢

✓ لنفترض أن الكلام المفهوم يتطلب 7 bits/sample . إذا صُمم نظام الهاتف لتلبية متطلبات الكلام بدقة (وهذا صحيح)، فما هو أقصى معدل بت مسموح به عبر خطوط الهاتف؟

• الحد الأقصى لمعدل البت عبر خطوط الهاتف $56 \text{ Kbps} = 2 * 4 * 7$.

✓ تستخدم الأقراص المضغوطة 16 b/sample . ما هو معدل بت الموسيقى الصادرة من القرص المضغوط؟

• معدل بت صوت القرص المضغوط $705.6 \text{ Kbps} = 44.1 * 16$ أحادي،
 1.411 Mbps ستيريو.



صوت بجودة الأقراص المضغوطة (CD-Quality)

- أقراص **CD-ROM** هي وسائط تخزين رقمية للموسيقى ستريو (**stereophonic**) وغيرها من تدفقات المعطيات المتعلقة بأنظمة الوسائط المتعددة.
- **CD Digital Audio (CD-DA)**: هو المعيار المرتبط بالأجهزة التي تستخدم هذا النوع من الأقراص.
- تتمتع الموسيقى بنطاق ترددي مسموع يتراوح من $20 \text{ kHz} \Rightarrow 15 \text{ Hz}$ ، وبالتالي فإن معدل أخذ العينات الأدنى هو 40 ksps .
- يتطلب تسجيل الموسيقى (**stereo**) قناتين منفصلتين (معدل البت الإجمالي المطلوب هو ضعف المعدل المطلوب للموسيقى الأحادية **(mono)**).

✓ $\text{Bit rate per channel} = \text{Sampling rate} * \text{bits per sample}$

• $44100 * 16 = 705.6 \text{ kbps}$

✓ $\text{Total bit rate} = 2 * 705.6 = 1.411 \text{ Mbps}$



صوت بجودة الأقراص المضغوطة (CD-Quality)

■ مثال:

■ بافتراض استخدام معيار **CD-DA**، قم باستنتاج:

✓ سعة تخزين قرص **CD-ROM** لتخزين ٦٠ دقيقة من الوسائط المتعددة.

• تؤدي عملية التحويل الرقمي **CD-DA** إلى معدل بت يبلغ $1.411Mbps$ ،
 $\Rightarrow 60 \text{ min} = 1.411 * 60 * 60 \text{ Mbits} = 5079.6 \text{ Mbits or } 634.95 \text{ Mbytes}$

✓ الوقت اللازم لنقل جزء بطول ٣٠ ثانية من الوسائط المتعددة باستخدام قناة إرسال بمعدل 64 kbps & 1.5 Mbps .

• ٣٠ ثانية من الوسائط المتعددة، ($= 1.411 \times 30 = 42.33 \text{ Mbits}$)،
ومن هنا الوقت لنقل هذه البيانات:

• At 64 kbps : $\frac{42.33 \cdot 10^6}{64 \cdot 10^3} = 661.4 \text{ s (about 11 min)}$

• At 1.5 Mbps : $\frac{42.33 \cdot 10^6}{1.5 \cdot 10^6} = 28.22 \text{ s}$



ترميز الصوت

- تتلخص عملية ترميز الصوت بعملية التكميم و من ثم التحويل (**transformation**).
- تستخدم عملية الترميز (**μ -law**) من أجل الصوت، و لكن يوجد تقنية أخرى و هي **PCM**
- (**Pulse Code Modulation**)
تتلخص عملية **PCM** بما يلي:
- الفلترية: للحد من الترددات العالية للإشارة التي تؤثر على معدل أخذ العينات (يجب عدم تشويه الإشارة).
- عملية أخذ العينات (التقطيع).
- التكميم بـ **L** مستوى.



ترميز الصوت

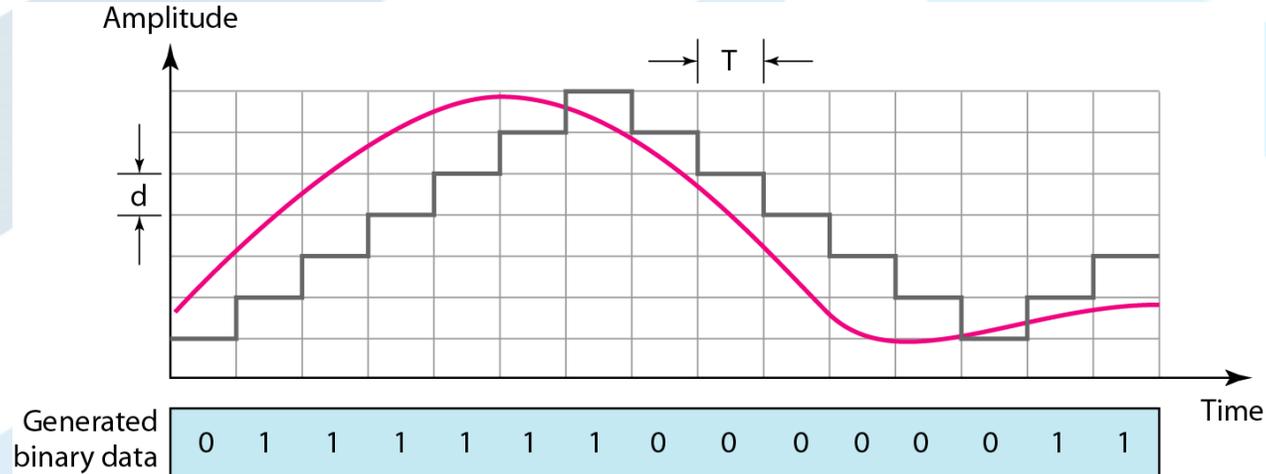
■ مثال:

- ✓ نريد تحويل اشارة الصوت البشري الى اشارة رقمية:
- يحتوي الصوت البشري عادةً على ترددات تتراوح بين 0 to 4 kHz لذا، يُحسب معدل أخذ العينات ومعدل البت كما يلي:
 - $Sampling\ rate = 4000 * 2 = 8000\ samples/sec$
 - $Bit\ rate = 8000 * 8 = 64000\ bps.$
- افترض أن تطبيق ما يجمع المعطيات المرمزة كل $20\ ms$ لتنفيذ قطع كبيرة من المعطيات، ما هو حجم كل جزء (قطعة)؟
 - $the\ chunk\ size = 64000 * 20 * 10^{-3} = 1280\ bits = 160\ bytes$



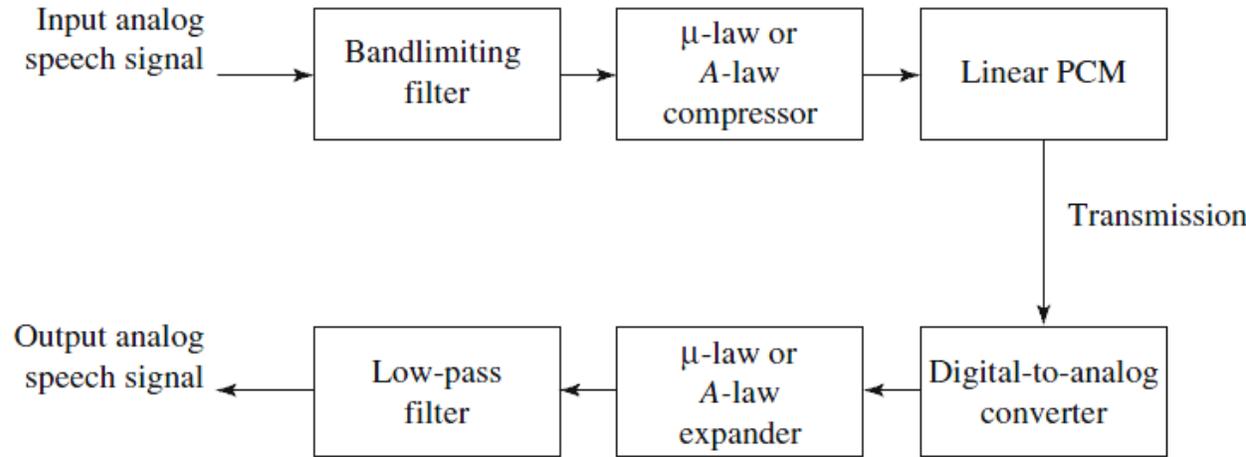
التعديل دلتا Delta

- يقيس **delta** الفرق بين النبضات، إذا كانت قيمة النبضة عند اللحظة t_{n+1} أكبر من قيمة النبضة عند اللحظة t_n يتم اعتماد القيمة (1)، أي الإشارة تميل الى القيمة الموجبة. أما إذا كانت قيمة النبضة أقل، أي سالبة، فيستخدم (0).
- يعتبر استخدام **Delta** جيد عندما تكون التغيرات صغيرة في قيم الإشارة بين العينات، بينما إذا كانت التغيرات في القيم كبيرة، فسوف يؤدي هذا إلى أخطاء كبيرة.



التعديل Delta PCM (DPCM)

- بدلاً من استخدام بت واحد للإشارة إلى الفروقات الإيجابية والسلبية، يُمكننا استخدام المزيد من البتات. بشكل آخر، تكميم الفرق من خلال عدة مستويات، و كلما زاد عدد البتات، زادت المستويات و بالتالي تزداد الدقة.
- مثال عن PCM in Speech Compression:



PCM signal encoding and decoding





جامعة
المنارة
MANARA UNIVERSITY

الصوت

المستوى النسبي	SPL	الصوت
10,000,000x	140	صوت مسدس (7 م)
	130	صوت صفارة الحريق (30 م)
1,000,000x	120	إقلاع طائرة نفاثة (60م) حفلة موسيقية (3م)
	110	
100,000x	100	موسيقى كلاسيكية عالية
	90	اردحام مروري (2م)
10,000x	80	مقصورة الركاب في طائرة نفاثة
	70	محادثة طبيعية (1م)
1,000x	60	
	50	منزل ريفي طبيعي (ليلاً)
100x	40	قاعة هادئة
	30	همس هادئ (2م)
10x	20	حفيف الأشجار
	10	
المستوى المرجعي	0	عتبة السمع
	Decibels	

علاقة شدة الصوت بالأصوات الشائعة



نهاية المحاضرة أسئلة؟؟

