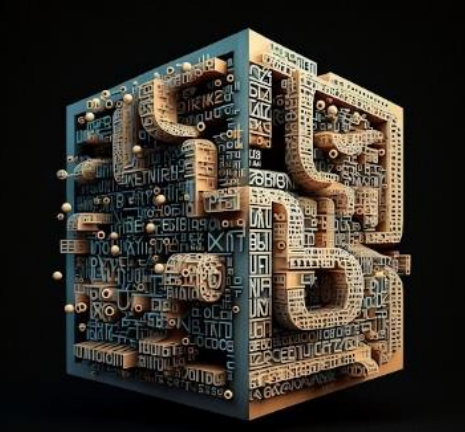


Міністерство освіти та науки України
Харківський національний економічний
університет імені Семена Кузнеця

НАВЧАЛЬНО – НАУКОВИЙ ІНСТИТУТ
ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ
кафедра: Мультимедійних систем і технологій

дисципліна
"ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ"
викладач: доцент, к.т.н.
Токарєв Володимир Володимирович



Тема. Теоретичні основи мультимедіа. Фізичні та психофізіологічні основи звуку.



Association for
Computing Machinery

Advancing Computing as a Science & Profession

Основна література

1. Конспект лекцій з дисципліни «ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ».
2. Анісімов А.В. Інформаційні системи та бази даних: Навчальний посібник для студентів факультету комп'ютерних наук та кібернетики. / Анісімов А.В., Кулябко П.П. – Київ. – 2021. – 110 с.
3. Антоненко В. М. Сучасні інформаційні системи і технології: управління знаннями : навч. посібник / В. М. Антоненко, С. Д. Мамченко, Ю. В. Рогушина. – Ірпінь : нац. університет ДПС України, 2019. – 212 с.
4. Павлиш В. А. Основи інформаційних технологій і систем: Навчальний посібник. / Павлиш В. А., Гліненко Л. К. - Львів: Видавництво Львівської політехніки, 2021. – 500 с.

Цілі дисципліни

Після вивчення дисципліни студенти мають:

Знати:

- основні поняття інформації, даних та інформаційних процесів;
- сучасні архітектури інформаційних систем і технологій;
- принципи функціонування апаратного та програмного забезпечення;
- методи аналізу, зберігання та обробки інформації.

Вміти:

- застосовувати інформаційні технології для розв'язання практичних задач професійної діяльності.



ВСТУП

Інформаційні технології є фундаментальною складовою сучасного цифрового суспільства. Вони визначають розвиток науки, економіки, промисловості та освіти.

Метою даної лекції є формування системного розуміння принципів функціонування інформаційних систем, методів обробки даних та ролі цифрових технологій у професійній діяльності фахівця. У ході лекції розглядаються базові поняття інформації, процеси зберігання та передавання даних, а також сучасні тенденції розвитку ІТ.

Сучасний фахівець у будь-якій галузі діяльності працює в умовах цифрового середовища, де обробка, зберігання, передавання та аналіз даних здійснюються із застосуванням обчислювальних систем різного рівня складності — від персональних пристроїв до глобальних хмарних інфраструктур.

Розвиток інформаційних технологій ґрунтується на досягненнях:

- теорії інформації;
- алгоритміки та математичного моделювання;
- мікроелектроніки та обчислювальної техніки;
- комп'ютерних мереж;
- програмної інженерії.

Особливого значення набуває інтеграція ІТ з такими напрямками, як:

- штучний інтелект;
- великі дані (Big Data);
- кібербезпека;
- хмарні обчислення;
- інтернет речей (IoT).

ВСТУП

Мультимедіа як технологічна концепція сформувалася на перетині кількох напрямів:

- комп'ютерної графіки;
- цифрової обробки сигналів (DSP);
- телекомунікацій та інтерактивних інформаційних систем.

Початковий етап розвитку мультимедіа був пов'язаний із поєднанням тексту, статичних зображень та звуку в межах персонального комп'ютера.

У 1990-х роках розвиток CD-ROM та перших графічних інтерфейсів (GUI) дозволив інтегрувати:

- текстову інформацію;
- растрову та векторну графіку;
- аудіофайли;
- відеофрагменти низької роздільної здатності.

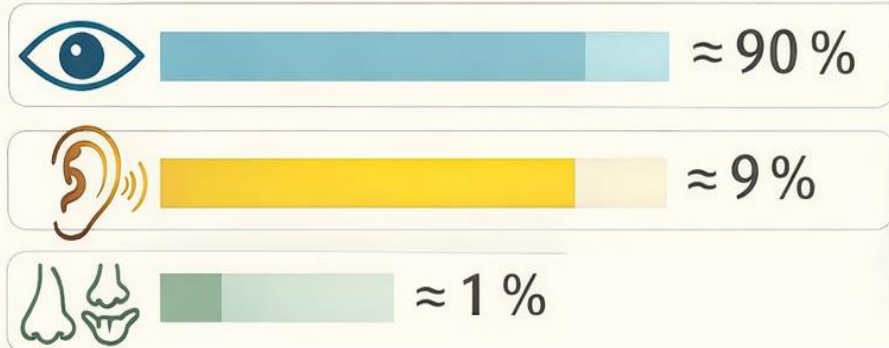
Подальший розвиток Інтернету трансформував мультимедіа в мережеву модель розповсюдження контенту.

Розглянемо фізичні та психофізіологічні основи звуку

Всебічне сприйняття навколишнього світу забезпечується зоровими та аудіальними сигналами.

Наведений розподіл демонструє важливу роль слуху в отриманні інформації:

- при бл. **90 %** інформації надходить через зір;
- при бл. **9 %** – через слух;
- близько **1 %** – через інші органи чуття.



Слух є невід'ємною частиною **комунікації та сприйняття аудіоінформації**. Звукові хвилі, що поширюються в усіх напрямках, дозволяють людині отримувати сигнали небезпеки навіть у стані сну. Звукове середовище відіграє ключову роль у нашій культурі й побуті.

ВИНИКНЕННЯ ТА ПОШИРЕННЯ ЗВУКУ

Поняття **ЗВУК** можна розглянути з двох принципово різних позицій:

- **Звук як фізичне явище** — це коливання часток пружного середовища, що хвилеподібно розповсюджується, зокрема — повітря, води, метал, тощо.
- **Звук як фізіологічне явище** — це специфічне відчуття, що викликається дією звукових хвиль, що поширюються у повітряному середовищі, на орган слуху людини.



Фізичне явище



Фізіологічне явище

В **ISO 18405:2017** - (терміни підводної акустики) **sound** визначається як фізичне явище в середовищі, пов'язане з акустичними збуреннями (у стандарті далі через зв'язані терміни розкриваються звуковий тиск, швидкість частинок, усунення частинок тощо).

Важливо: **ISO**-доступні публічні фрагменти часто дають визначення через зв'язані величини (sound pressure, particle velocity і т.інш.), і це нормально для ІТ стандартів.

Фізичне визначення

1.Звук – це пружні коливання (зміни тиску та/або швидкості частинок) у середовищі, що поширюються у вигляді механічної хвилі.

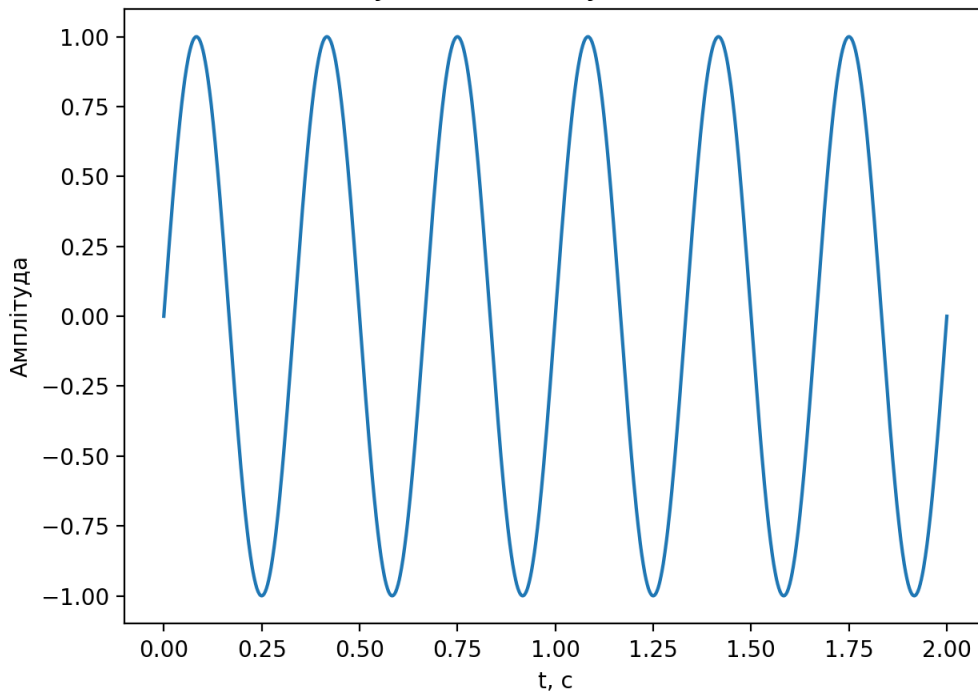
2.Шум - це звук, небажаний (який заважає) для людини/системи або випадкова/широкосмугова складова сигналу, що погіршує сприйняття або точність вимірювань.

Приклади:

А. Білий шум в аудіотракті (рівномірна спектральна густина за частотою в ідеалізації).

В. Транспорний шум як тло для оцінки акустичного комфорту.

Синусоїдальний звуковий сигнал



Це зображення показує синусоїдальний звуковий сигнал – найпростіший приклад тонального (гармонічного) звуку.

Сигнал описується формулою:

$$x(t) = A \sin(2\pi f t + \varphi)$$

де:

$x(t)$ — значення сигналу у часі (амплітуда);

- A — амплітуда (максимальне відхилення);
- f — частота (Гц), скільки коливань за секунду;
- t — час (с);
- φ — початкова фаза (зсув за часом).

На графіці видно регулярне повторення "вгору-вниз", що притаманно чистого тону (наприклад, камертон).

Підпис осі X: час, **t** (с).

Що означає: по горизонталі йде час у секундах. На малюнку показаний інтервал приблизно **0...2** с.

Підпис осі Y: амплітуда сигналу, **x(t)** (умов. од.).

Що означає: по вертикалі - величина сигналу (напруга, тиск, зсув тощо). Якщо одиниці не задані (як часто буває у навчальних схемах), пишуть умовні одиниці.

Що означає графік

1) Амплітуда A. На графіку максимум близько +1, мінімум близько -1. Значить амплітуда:

$$A \approx 1$$

2) Період T та частота f.

Період **T** — відстань між сусідніми піками (або западинами).

Частота **f**:

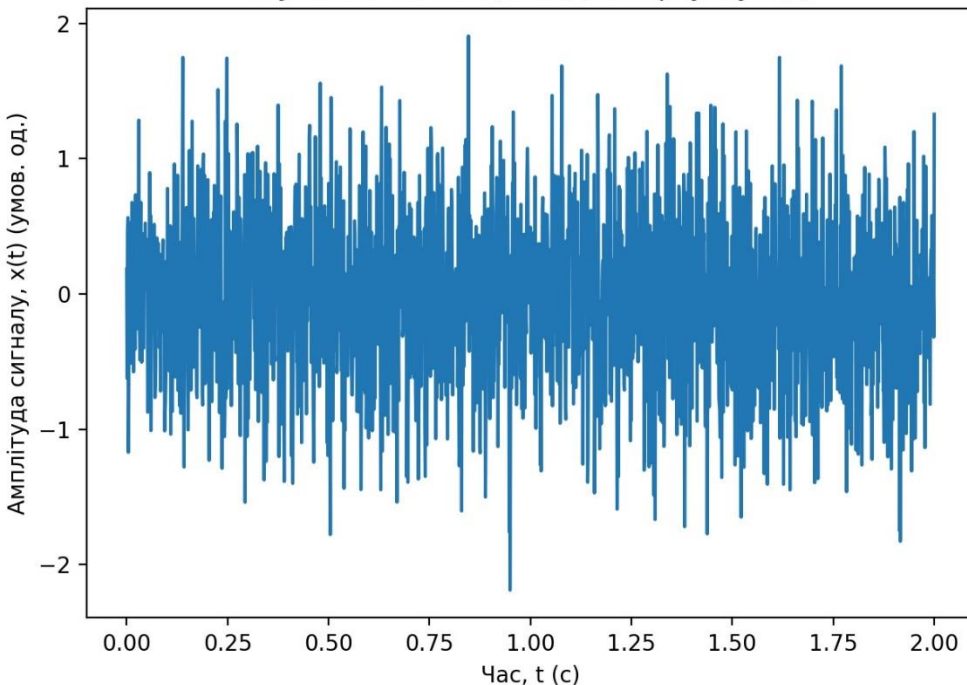
$$f = \frac{1}{T}$$

якщо на інтервалі **2** секунди видно приблизно **6** повних коливань, то:

$$f = \frac{6}{2} = 3\text{Гц}$$

Шумовий сигнал (для наочного порівняння)

Шумовий сигнал (випадкові флуктуації)



Це зображення показує **шумовий сигнал** – приклад не тонального, а випадкового (стохастичного) коливального процесу.

На графіку представлено тимчасову реалізацію випадкового сигналу $x(t)$. Значення амплітуди змінюються неперіодично, без стійкого малюнку, що повторюється, як у синуса.

Такі сигнали моделюють:

- фонові шуми (аудіотракт, мікрофонний підсилювач);
- вітровий/транспортний шум;
- шуми датчиків (тепловий, електронний);
- перешкоди під час запису/передачі звуку.

Ключова ознака: немає виділеної частоти, немає регулярного періоду T .

Підпис осі X: час, t (с).

Що означає: за часом видно, як змінюється сигнал. На графіку показаний інтервал приблизно $0...2$ с.

Підпис осі Y: амплітуда сигналу, $x(t)$ (умов. од.).

Що означає: по вертикалі - значення сигналу на даний момент часу. Одиниці "умовні", тому що це модель/приклад (у реальних вимірах тут може бути: **Па** - для тиску, **В** - для напруги, **м/с** - для швидкості частинок і т. п.).

3. Ударна хвиля (shock wave) - ISO - 2041-2012:

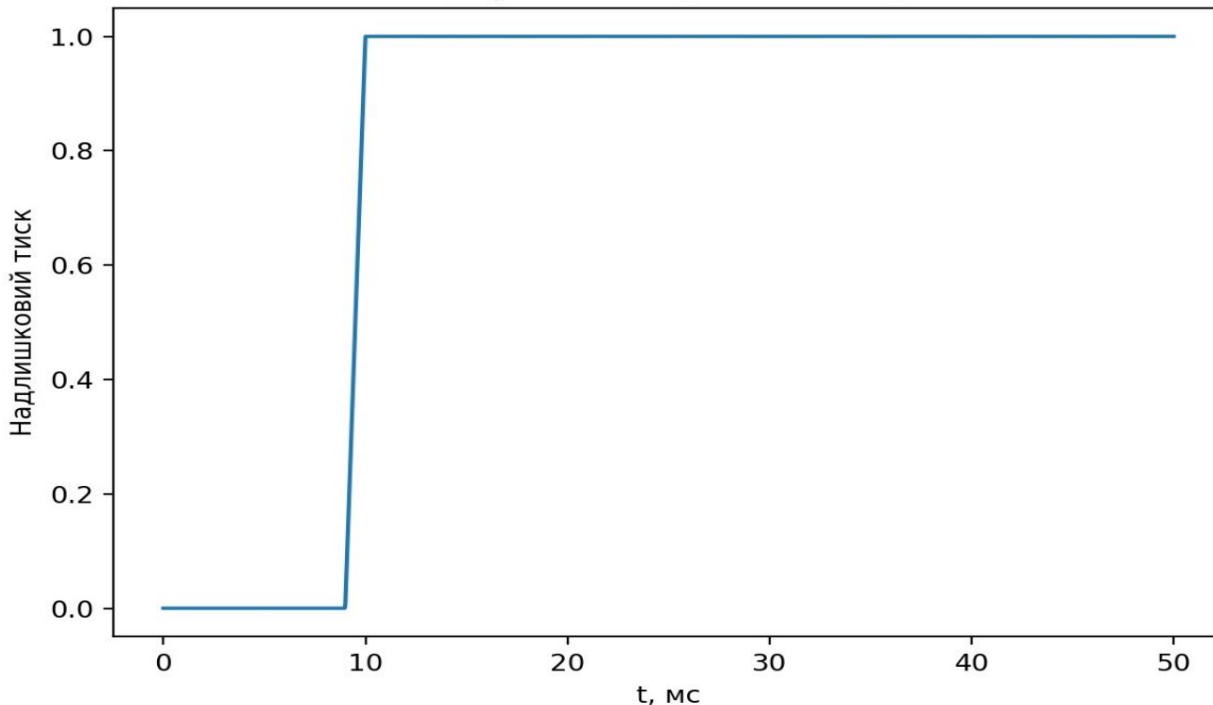
- тимчасовий сигнал, що описує ударний імпульс (переміщення/тиск/інш. змінної);

- пов'язаний з поширенням удару в середовищі або конструкції; і зазначено, що у рідині/газі зазвичай є фронт, де тиск різко зростає.

Фізичний зміст

Ударна хвиля – це нелінійна хвиля з різким фронтом (майже розривом) параметрів (тиску, щільності, температури, швидкості потоку). Часто виникає при надзвукових рухах чи вибухах.

Ударна хвиля (схематично)



На малюнку представлена схематична тимчасова форма ударної хвилі - різка стрибкоподібна зміна тиску в середовищі.

Графік показує залежність надлишкового тиску від часу:

$$p'(t) = p(t) - p_0$$

де:

$p'(t)$ - надлишковий тиск;

$p(t)$ - поточний тиск;

p_0 - атмосферний тиск.

До моменту $t \approx 10$ мс тиск дорівнює нулю (немає обурення).

У момент фронту відбувається стрибок тиску - характерна ознака ударної хвилі.

Вісь X (горизонтальна) - година, t (мс).

де:

- t - час;

- **мс** - мілісекунди.

Це тимчасова еволюція тиску.

Вісь Y (вертикальна) - надлишковий тиск $p'(t)$, Па.

де:

$p'(t)$ - надлишковий тиск;

Па - паскалі.

(На схемі використовуються умовні одиниці, але насправді вимірюється в Па.).

На відміну від звичайної звукової хвилі:

-звичайна хвиля → плавна синусоїда;

-ударна хвиля → майже вертикальний фронт.

Чому фронт «вертикальний»? Це наслідок нелінійності. Швидкість поширення залежить від тиску. Рівняння швидкості:

$$c = \sqrt{\gamma \frac{p}{\rho}}$$

де:

c - швидкість звуку;

γ - показник адіабати;

p - тиск;

ρ - щільність.

При великому тиску передня частина хвилі рухається швидше за задню → виникає «стиск» фронту → формується стрибок.

Де виникає ударна хвиля

- вибух;
- надзвуковий літак (звуковий удар);
- детонація;
- удар метеорита;
- сильний електричний розряд.

Фронти хвиль від точкового джерела (сферична розбіжність, 2D-схема).

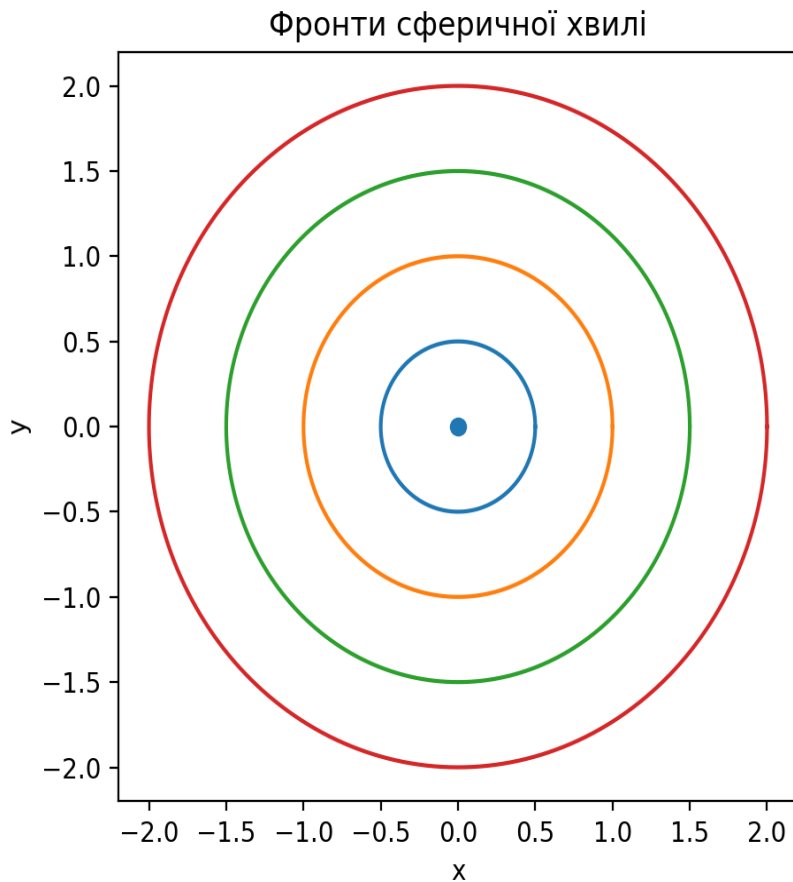


Схема фронтів сферичної звукової хвилі, показана в **2D**-перетині площиною **x-y**.

Крапка в центрі - джерело звуку (точковий випромінювач).

Концентричні кола - це фронт хвилі (хвильові поверхні) у різні моменти часу.

Фронт хвилі - геометричне місце точок, до яких обурення дійшло в той самий момент часу.

Насправді для точкового джерела в однорідному середовищі хвиля поширюється сферично:

$$\mathbf{r} \cdot (\mathbf{t}) = \mathbf{c} \cdot \mathbf{t}$$

де:

- r - відстань від джерела (радіус фронту), м;
- c - швидкість звуку серед, м/с;
- t - час, с.

На площині ми бачимо перетин сфер тому вони виглядають як кола.

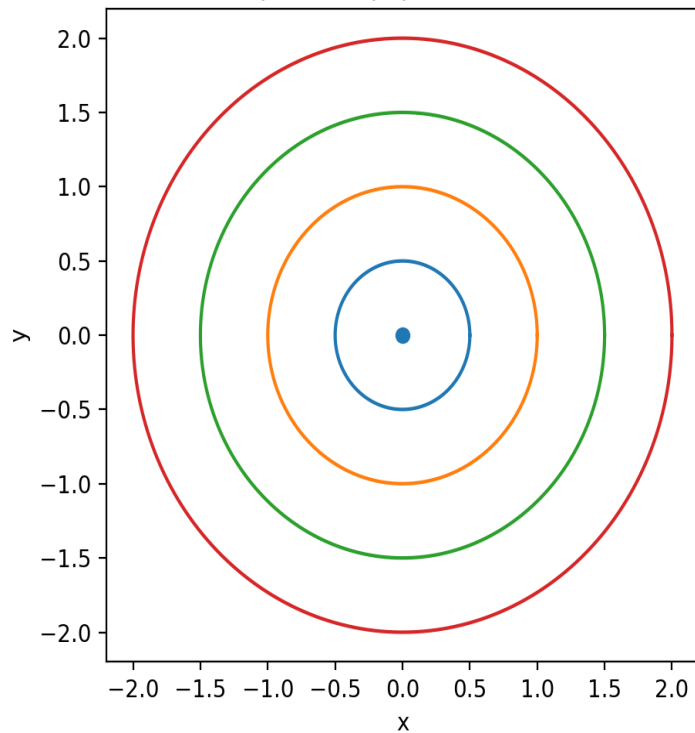
Що означає кожне коло

Якщо кола мають радіуси: r_1, r_2, r_3, r_4 , то вони відповідають часам:

$$t_i = \frac{r_i}{c}$$

Саме зовнішнє коло \rightarrow пізніший момент часу, фронт вже "пішов далі".

Фронти сферичної хвилі



Як “читати” рисунок

У центрі – джерело.

Кола - положення фронту в різні моменти.

Кожне коло означає однаковий час приходу хвилі на дану відстань.

Чим далі — тим слабше (енергія розсіюється сферою).

Вісь X (горизонтальна):

x - (умов. од.) або x, м (якщо задаєте реальні метри).

Вісь Y (вертикальна):

y - (умов. од.) або y, м (якщо задаєте реальні метри).



Association for
Computing Machinery

Advancing Computing as a Science & Profession

ШВИДКІСТЬ ЗВУКУ У РІЗНИХ СЕРЕДОВИЩАХ

Середовище	Густина, ρ (кг/м ³)	Швидкість звуку, c (м/с)
Тверді матеріали		
Сталь	≈ 7850	$\approx 5000-6000$
Деревина (уздовж волокон)	$\approx 600-800$	$\approx 3300-4000$
Гума	≈ 1100	$\approx 50-200$
Рідини		
Вода (20°C)	≈ 1000	≈ 1480
Спирт (етанол)	≈ 789	≈ 1140
Гази		
Азот	≈ 1.25	≈ 334
Повітря (20°C)	≈ 1.20	≈ 343
Водень	≈ 0.09	≈ 1270

Висновок

Швидкість звуку: Тверді тіла > Рідини > Гази

Причина:

модуль пружності зростає значно швидше, ніж густина.



ЧАСТОТА ЗВУКУ

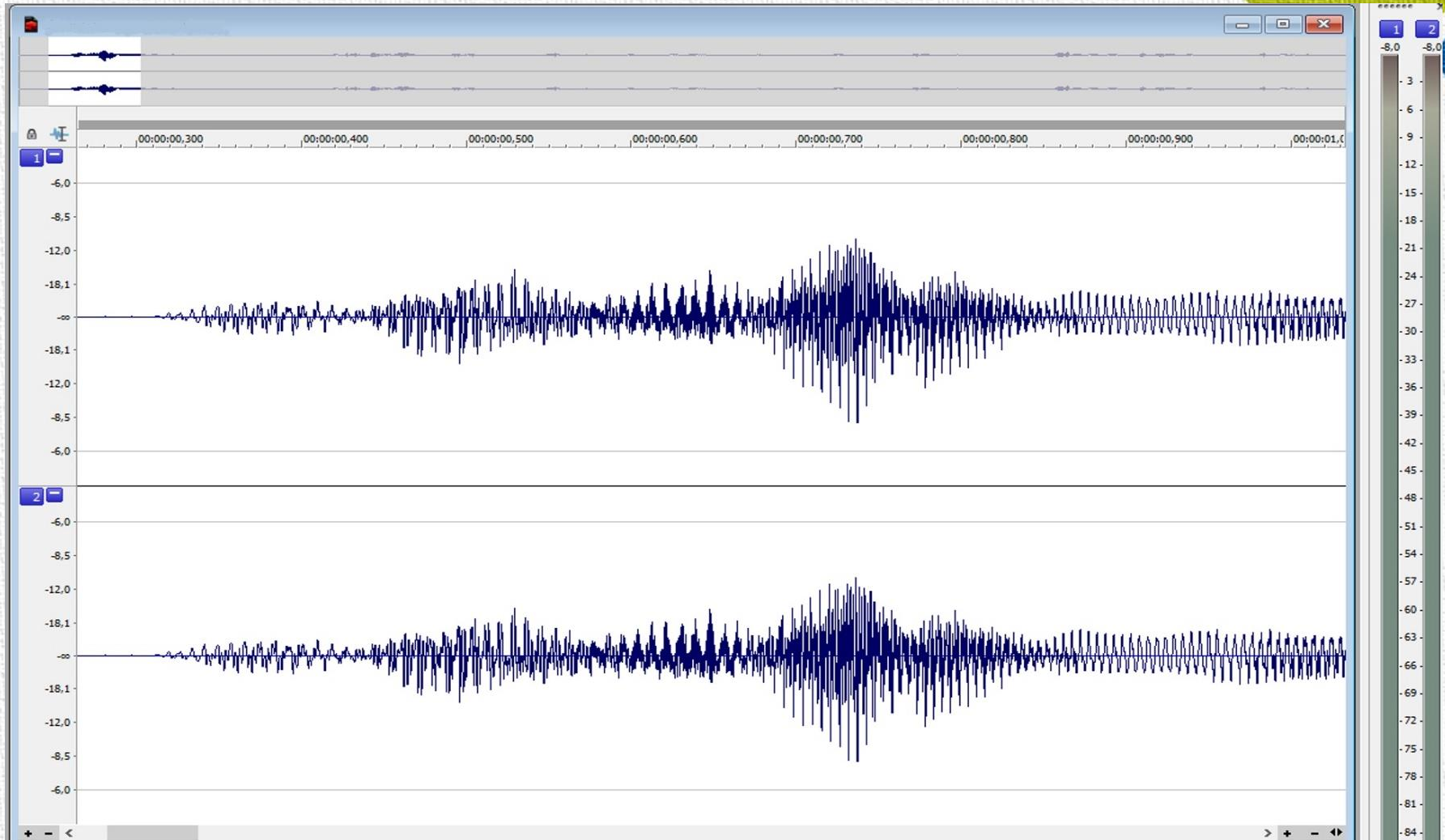
Людина в ідеальних умовах здатна сприймати звук у діапазоні частот приблизно від 16–20 Гц до 20 кГц. Частота визначає теоретичну можливість сприйняття звуку, тоді як практичне його сприйняття залежить також від амплітуди (інтенсивності) сигналу. Усе, що має частоту менше 16–20 Гц, називається **інфразвуком**, а все, що має частоту понад 20 кГц, - **ультразвуком**. Основний чутний діапазон людини становить приблизно 20 Гц – 20 кГц. У музичній акустиці умовною межею між низькими та високими тонами часто вважається частота близько 1000 Гц. Людина сприймає не абсолютне значення частоти, а її відносні співвідношення. Ми легко визначаємо, який звук вищий або нижчий за частотою. Мелодію можна виконати в різних октавах, і вона залишиться впізнаваною.



ЧАСТОТА ЗВУКУ

Основну інформацію про звукові коливання людський мозок отримує в діапазоні частот приблизно до 4–5 кГц. Це цілком логічно, оскільки основні життєво важливі для людини звуки - мовлення, голоси людей і тварин, природні шуми (вода, вітер тощо) - мають найбільшу інформативність саме в цьому частотному інтервалі. Частоти вище 5 кГц відіграють допоміжну роль. Вони не є визначальними для розбірливості мовлення, проте істотно впливають на суб'єктивне сприйняття якості звуку - створюють відчуття «прозорості», «деталізації» та «яскравості» звучання. Низькі частоти (до приблизно 250–300 Гц) формують основу звукового сигналу, забезпечують повноту та масивність звучання. Середні частоти (приблизно 300 Гц – 4 кГц) є найбільш інформативними для розбірливості мовлення. Високі частоти (вище 4–5 кГц) відповідають за деталізацію та суб'єктивну якість звучання.

Приклад звукового сигналу зображено вікно аудіоредактора



Що саме зображено.

Дві хвильові форми - це два канали стерео. Верхня доріжка - лівий канал (**Left**). Нижня доріжка - правий канал (**Right**). Форми майже ідентичні.

Оси та шкали

Горизонтальна вісь - Час. У верхній частині видно часову шкалу з позначеннями типу:

00:00:00,300, 00:00:00,400 тощо.

Це означає, що відображено фрагмент тривалістю приблизно **1** секунда.

Вертикальна вісь — Амплітуда / Рівень сигналу. Зліва видно шкалу в децибелах (dB): -6 dB - 12 dB - 18 dB. У цифровому аудіо це зазвичай dBFS (decibels relative to full scale).

0 dBFS — максимальний рівень (межа без перевантаження). Чим ближче сигнал до **0 dBFS**, тим він гучніший. Якщо хвиля «зрізається» зверху або знизу - це кліпінг (перевантаження).

Частотні діапазони різних джерел звуку

1. Навушники.

Смуга частот, яку здатні відтворювати якісні навушники: 20 Гц – 20 кГц (у професійних моделях іноді 5 Гц – 40 кГц, однак це перевищує межі чутності людини).

2. Студійні монітори або стереосистеми.

Смуга частот відтворення: 40 Гц – 20 кГц (у високоякісних акустичних системах – 30 Гц – 22 кГц).

Рояль (фортепіано).

Частотний діапазон звучання:

27,5 Гц – 4186 Гц (\approx 4,2 кГц)

Найнижча нота - (27,5 Гц).

Найвища - (4186 Гц).

Жіночі голоси

Тип голосу	Частотний діапазон
Контральто	175 – 700 Гц
Мецо-сопрано	220 – 880 Гц
Сопрано	260 – 1050 Гц
Колоратурне сопрано	260 – 1400 Гц



Association for
Computing Machinery

Advancing Computing as a Science & Profession

Частотні діапазони різних джерел звуку

Чоловічі голоси

Тип голосу	Частотний діапазон
Бас	80 – 330 Гц
Баритон	100 – 400 Гц
Тенор	130 – 520 Гц

Важливе уточнення.

- ◆ Наведені значення стосуються основної частоти.
- ◆ Реальний спектр голосу значно ширший через наявність обертонів (гармонік).
- ◆ Форманти мовлення можуть досягати 3–5 кГц і більше.

Психофізіологічні параметри звуку

До основних психофізіологічних параметрів звуку належать:

тон (висота тону), тембр та гучність.

1.Тон (основний тон).

Тон - це частотна складова звукового сигналу, яка домінує за амплітудою та визначає основну частоту коливань. Основна частота **f₀** називається основним тоном, а кратні їй частоти - обертонами (гармоніками).

2.Висота звуку.

Висота звуку - це суб'єктивна характеристика, що дозволяє розташовувати звуки на шкалі від низьких до високих. Висота безпосередньо пов'язана з частотою коливань: збільшення частоти → звук сприймається як вищий; зменшення частоти → звук сприймається як нижчий. Людина визначає висоту шляхом порівняння двох тонів.

Психофізіологічні параметри звуку

До основних психофізіологічних параметрів звуку належать:
тон (висота тону), тембр та гучність.

3. Тембр звуку.

Тембр - це якісна характеристика звуку, яка дозволяє відрізнити звучання різних джерел за однакової частоти та гучності. Тембр визначається: спектральним складом сигналу; кількістю та інтенсивністю обертонів; співвідношенням амплітуд гармонік; часовою формою сигналу (атакою, спадом, резонансом). Саме завдяки тембру ми розрізняємо, наприклад, звук скрипки і фортепіано при однаковій ноті.

4. Гучність звуку

Гучність — це суб'єктивна оцінка інтенсивності звуку, тобто міра сили слухового відчуття.

Між фізичною інтенсивністю звуку та відчуттям гучності існує тісний, але **нелінійний** зв'язок. На сприйняття гучності впливають:

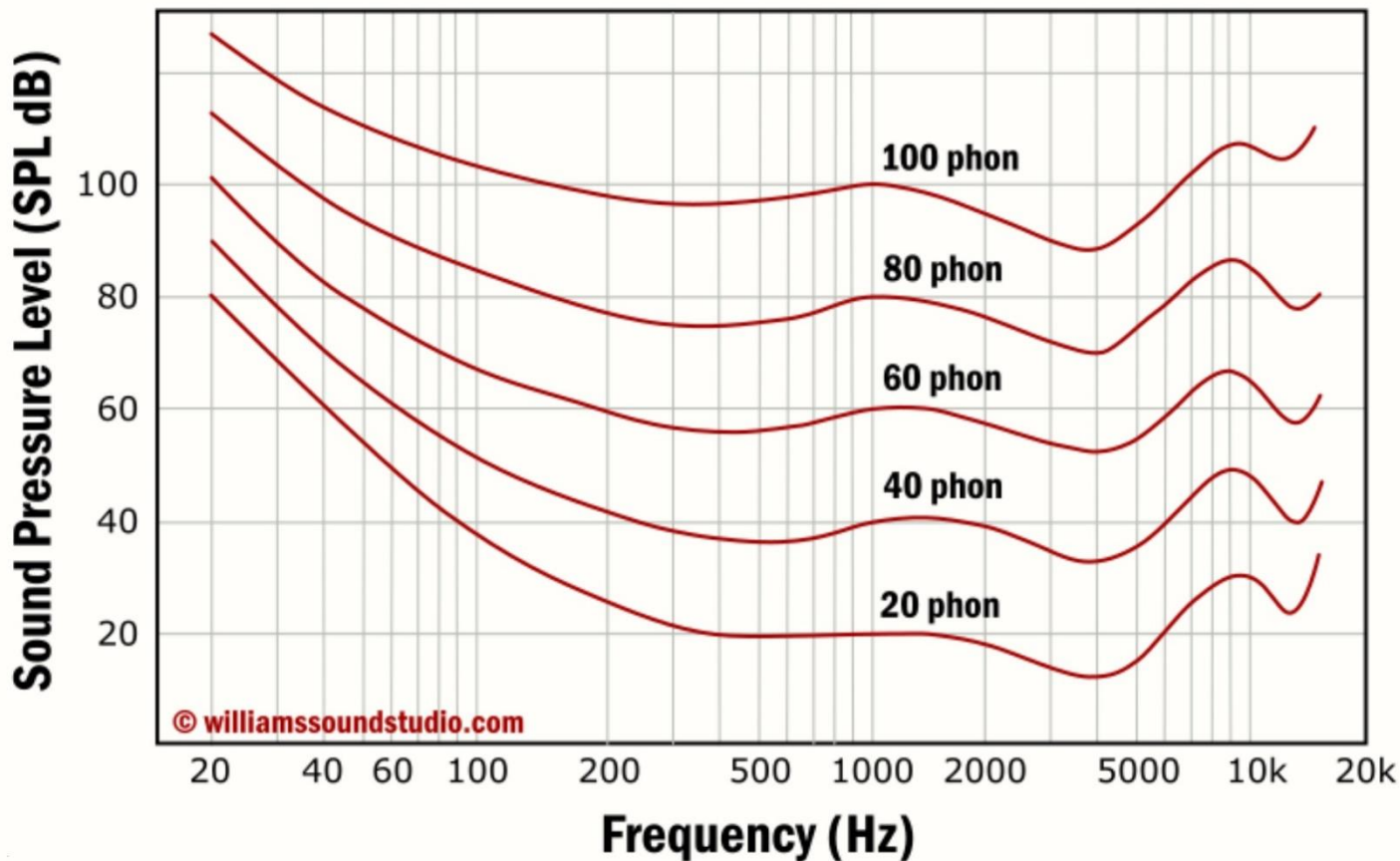
- звуковий тиск;
- частота сигналу;
- тривалість звучання;
- спектральний склад;
- індивідуальні особливості слуху.

Гучність вимірюють у фонах (**phon**) та сонах (**sone**), тоді як фізичний рівень звукового тиску — у децибелах (**dB SPL**).

Рівні гучності різних джерел звуку

Джерело звуку	Орієнтовний рівень гучності, дБ (SPL)
Акустична тиша (безлунна камера)	0–10 дБ
Тихий шепіт, цокання наручного годинника	20–30 дБ
Шелест листя, тиха житлова кімната	30–40 дБ
Сільська місцевість далеко від дороги, бібліотека	40 дБ
Тиха житлова територія, парк, спокійна розмова	40–50 дБ
Розмова середньої гучності, тихий офіс	50–60 дБ
Нормальна розмова на відстані 1 м	60–65 дБ
Жвава вулиця, телефонний дзвінок	70–80 дБ
Гучний будильник, вантажний автомобіль, мотоцикл	80–90 дБ
Гучний крик, відбійний молоток	100–110 дБ
Поїзд метро, фен, ковальський цех, гучне виробництво	90–110 дБ
Рок-музика, крики дитини, гелікоптер (на близькій відстані)	100–120 дБ

Fletcher Munson Equal Loudness Contours



На графіку показані криві рівної гучності (Fletcher-Munson/Equal Loudness Contours). Це класичний психоакустичний графік, який демонструє: який рівень звукового тиску потрібен на різних частотах, щоб людина сприймала звук однаково гучним.

Кожна червона крива позначена значенням phon (20, 40, 60, 80, 100 phon).

Одна крива = однакова суб'єктивна гучність по всій частотній осі.

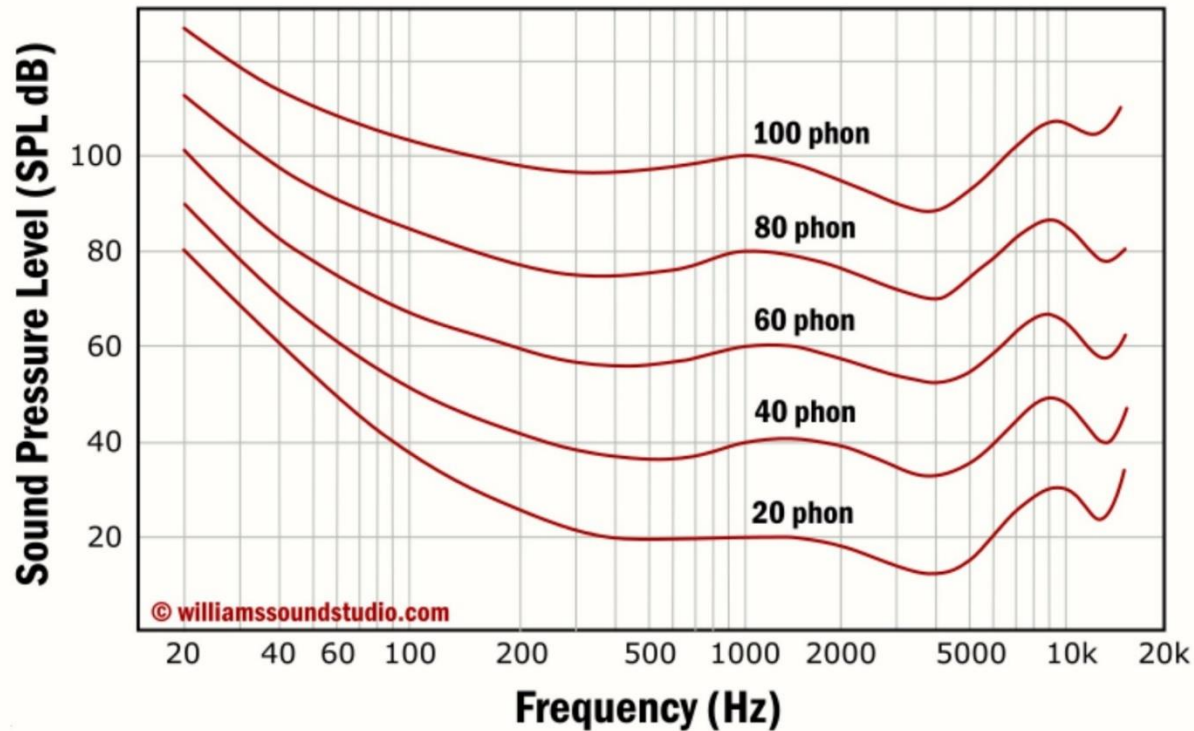
Приклад: крива 40 phon означає, що звук на 100 Гц і звук на 1 кГц сприйматимуться однаково гучними, якщо їх рівні SPL відповідають точкам цієї кривої.

Вісь X (горизонтальна). Frequency (Hz) → Частота, f (Гц). Шкала логарифмічна (20 Гц...20 кГц). Це важливо: відстані по осі X відповідають кратним змінам частоти, а не лінійним.

Вісь Y (вертикальна). Sound Pressure Level (SPL, dB) → Рівень звукового тиску.

Що таке фон (фон). Phon – одиниця рівня гучності (loudness level).

Fletcher Munson Equal Loudness Contours



Крок 1. Виберіть криву (наприклад, 60 phon) Це рівень однакової гучності.

Крок 2. Виберіть частоту (наприклад, 100 Гц) На осі X знайдіть 100 Гц.

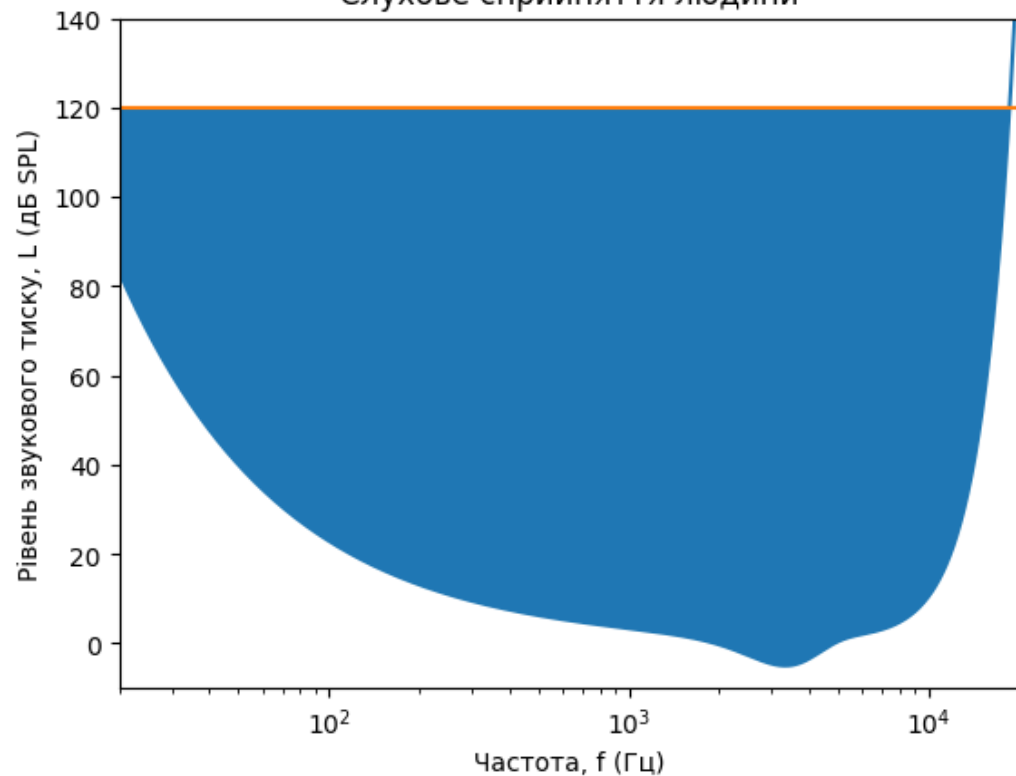
Крок 3. Підніміться до перетину з кривою 60 phon.

І потім рахуйте по осі Y рівень SPL. Інтерпретація.

Якщо на 100 Гц потрібно приблизно $\sim 75\text{--}80$ dB SPL, щоб було 60 phon, то на 1 кГц достатньо 60 dB SPL (бо 60 phon = 60 dB на 1 кГц).

Висновок: на низьких частотах вухо менш чутливе - потрібно набагато більший SPL.

Слухове сприйняття людини



На графіку зображено область слухового сприйняття людини в координатах:

Вісь X (горизонтальна) — частота f , Гц (логарифмічна шкала від 20 Гц до 20 кГц);
Вісь Y (вертикальна) — рівень звукового тиску L , дБ SPL (Sound Pressure Level).

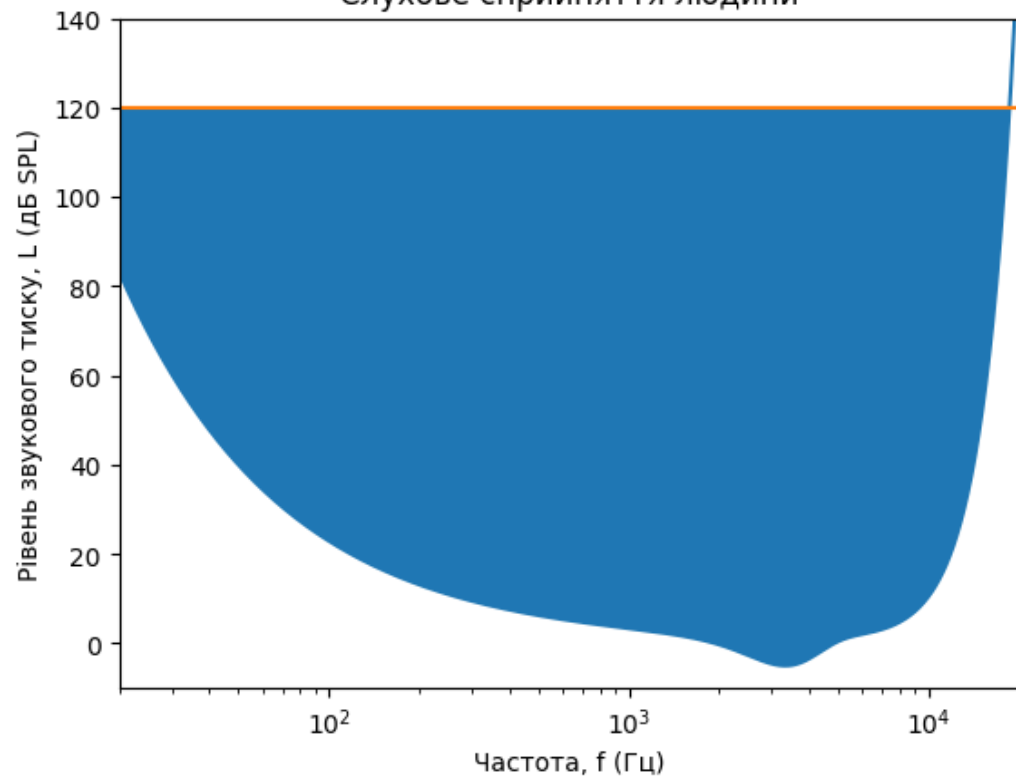
1. Нижня крива - Поріг чутності. Це мінімальний рівень звукового тиску, який людина здатна почути на різних частотах.

2. Верхня горизонтальна лінія - Поріг больового відчуття.

Вище цього рівня звук викликає:

- біль;
- можливе пошкодження слуху;
- акустичну травму.

Слухове сприйняття людини



Заштрихована область - область слухового сприйняття.

Це фізично допустима зона між:

- порогом чутності;
- порогом болю.

Саме в цій області знаходяться всі звуки, які людина може чути без больового відчуття.

Практичні наслідки.

Аудіосистеми повинні враховувати нерівномірну чутливість вуха.

При однаковому рівні дБ низькі частоти здаються тихішими. Саме тому існують:

- еквайзери;
- шумова-корекція;
- спеціалізовані фільтри.

Аналіз звукових сигналів

Аналіз звукових сигналів здійснюється з метою виявлення закономірностей у структурі звукової хвилі, що формується різними джерелами. Отримані дані є необхідними під час: проектування та оцінювання якості аудіоапаратури; розроблення систем кодування та стиснення сигналів; вибору методів цифрової обробки; створення систем автоматичного розпізнавання мовлення; розв'язання інших прикладних і наукових задач. Для аналізу звукових сигналів застосовують такі основні підходи:

- Амплітудно-часовий аналіз — дослідження сигналу в часовій області (time domain).
- Спектральний аналіз — дослідження сигналу в частотній області (frequency domain).

Аналіз за часом

Звуковий сигнал у часовій області описується зміною амплітуди звукового тиску в часі.

Розвиток звуку складається з чотирьох фаз:

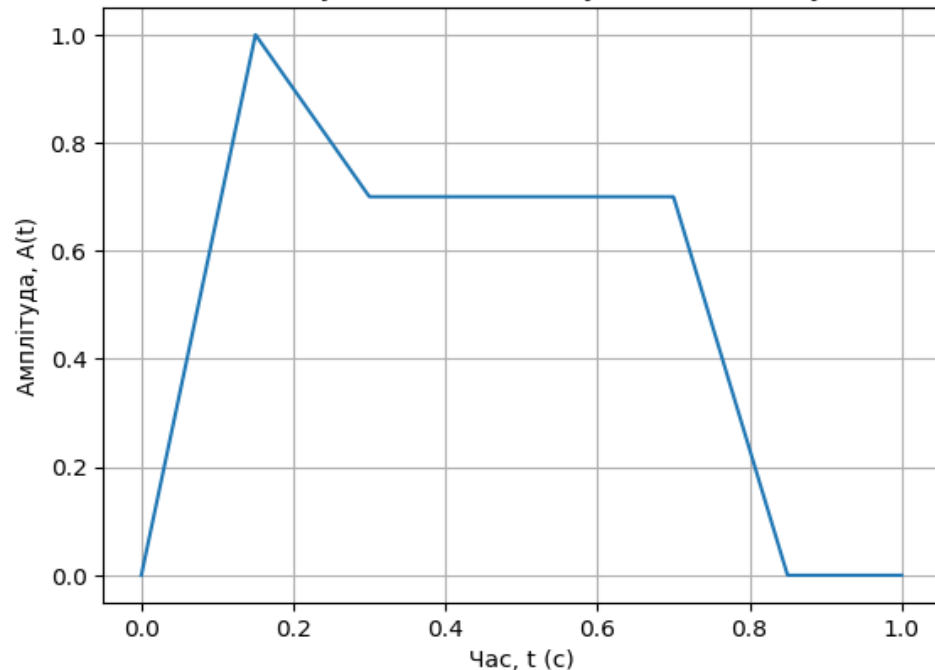
1. Атака (Attack) — наростання амплітуди від нуля до максимуму.

2. Спад (Decay) — зменшення амплітуди до рівня стабілізації.

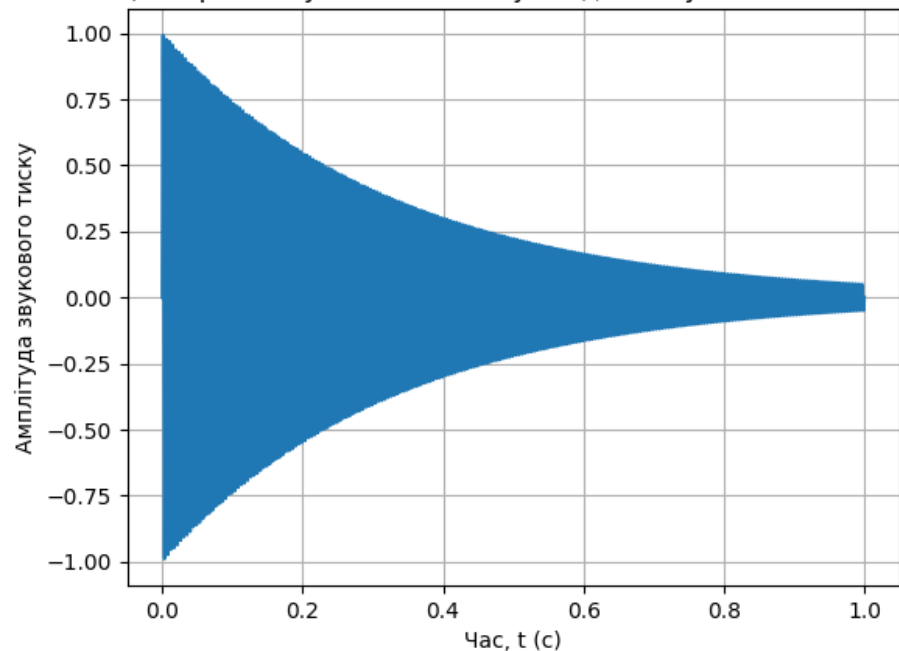
3. Стабілізація (Sustain) — підтримання сталого рівня.

4. Згасання (Release) — спад амплітуди до нуля.

Типова амплітудна огинаюча звукового сигналу (ADSR)



Осцилограма звукового сигналу (модель звучання віолончелі)



Association for
Computing Machinery

Advancing Computing as a Science & Profession

Спектральний аналіз

Звукові сигнали складнішої структури можна характеризувати спектром, тобто набором синусоїдальних складових (гармонік), сума яких утворює результуючу звукову хвилю, що відтворює вихідний сигнал.

Отримання даних про звуковий сигнал у вигляді спектра для подальшого аналізу та обробки набуло дуже широкого поширення, можливо навіть більшого, ніж амплітудно-часове подання сигналу.

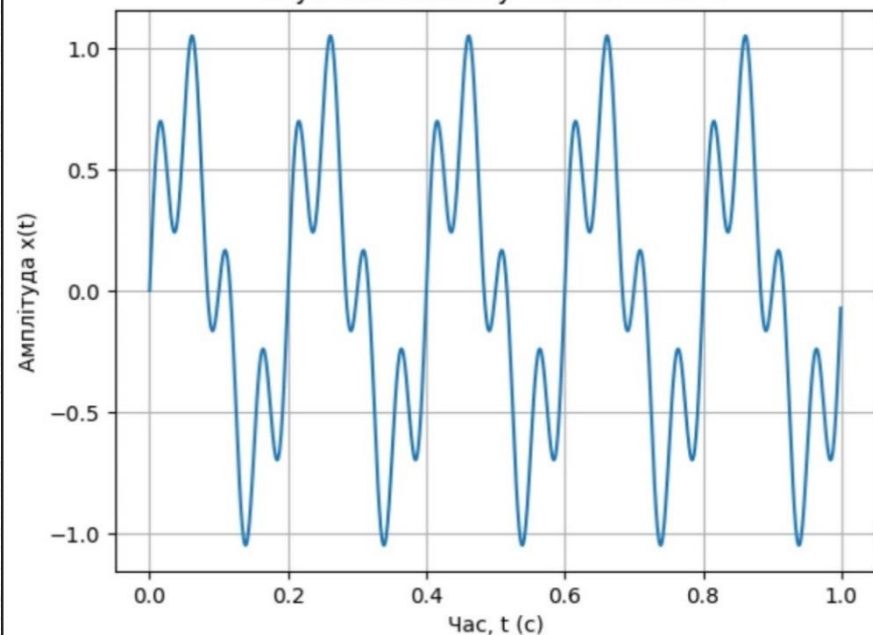
Для цього застосовують:

1. Амплітудно-часовий аналіз — дослідження сигналу у часовій області $x(t)$;

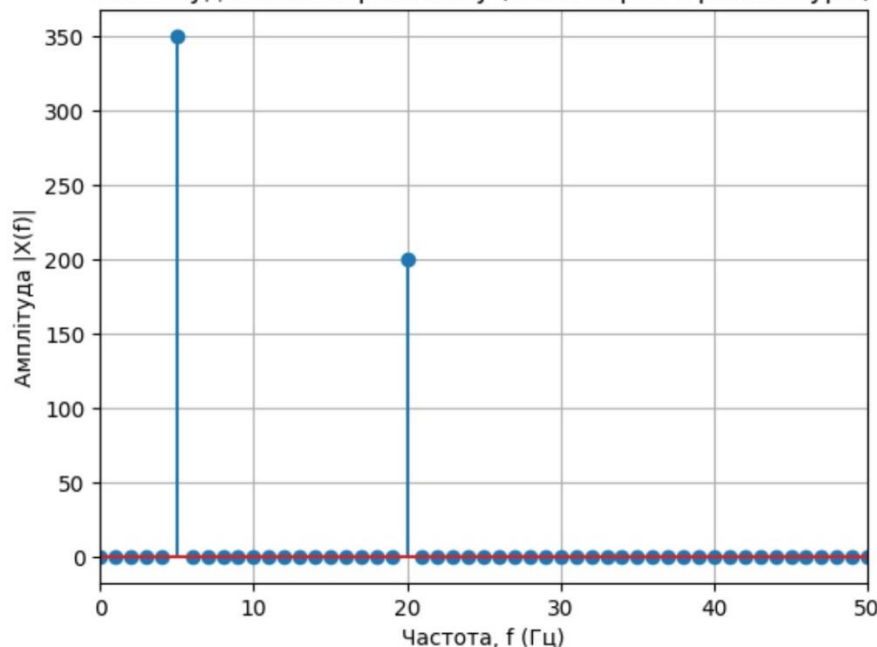
2. Спектральний аналіз — дослідження сигналу у частотній області $X(f)$ за допомогою

перетворення Фур'є.

Звуковий сигнал у часовій області



Амплітудний спектр сигналу (після перетворення Фур'є)



1.Звуковий сигнал у часовій області

Верхній графік показує сигнал суми двох гармонік:

- 5 Гц** - основна складова;
- 20 Гц** - додаткова гармоніка.

По осі:

- X** — час **t**, с;
- Y** — амплітуда **x(t)**.

Це типове представлення сигналу у часовій області — осцилограма.

2. Амплітудний спектр (частотна область)

Нижній графік — це результат перетворення Фур'є:

По осі:

- X** — частота **f**, Гц;
- Y** — амплітуда спектра **X(f)**.

Ми бачимо два виражені піки: при 5 Гц при 20 Гц. Це означає, що сигнал складається саме з цих частот.

3.Фізичний сенс

Перетворення Фур'є дозволяє перейти від опису:

Амплітуда ↔ **Час**

до опису:

Амплітуда ↔ **Частота**

Тобто замість того, щоб дивитися "як змінюється сигнал у часі", ми бачимо які частоти присутні, з якою амплітудою, яка енергетична структура

Маскування звуків

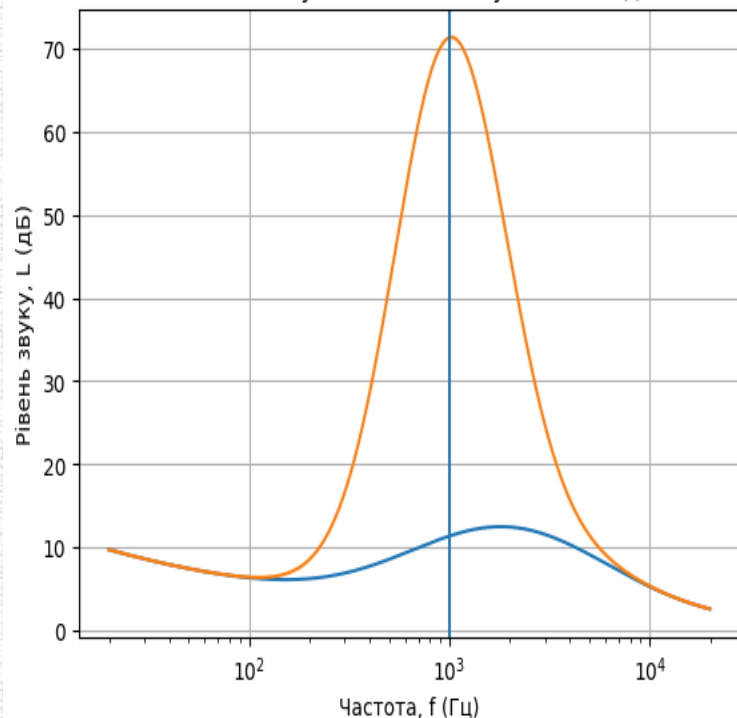
При одночасному впливі на слух двох звуків один із них може не сприйматися на тлі іншого. Це явище називається **частотним маскуванням**.

Існує також **тимчасове маскування**, коли гучний звук маскує тихіші звуки, що передують або слідують за ним у часі.

Механізм маскування пов'язаний із явищем **адаптації слуху**. Під час тривалого впливу звуку великої інтенсивності відчутна гучність поступово зменшується внаслідок зниження чутливості слухового аналізатора. Після припинення дії інтенсивного звуку чутливість слуху поступово відновлюється.

Ефект частотного маскування

Частотне маскування (психоакустична модель)



Якщо слабкий тон має частоту поблизу **1 кГц** він не буде почутий, навіть якщо перевищує поріг чутності в тиші.

Частотне маскування використовується в:

- аудиокомпресії;
- шумопригніченні;
- акустичному дизайні;
- психоакустичних моделях.

Алгоритми видаляють компоненти, які все одно замасковані.

Особливості розповсюдження звукових хвиль

На цьому слайді перелічено основні фізичні явища, які виникають під час поширення звуку в середовищі. Нижче — детальне пояснення кожного пункту.

-**суперпозиція хвиль кількох джерел.** Йдеться про принцип суперпозиції: якщо в одній точці простору одночасно діють кілька звукових хвиль, результуюче коливання дорівнює сумі їх миттєвих зміщень або тисків.

-**інтерференція. Інтерференція** — це наслідок суперпозиції когерентних хвиль. Можливі два випадки:

Конструктивна інтерференція — підсилення (фази співпадають).

Деструктивна інтерференція — ослаблення або повне гасіння (фази протилежні).

-**відображення.** Звукова хвиля, досягаючи межі розділу середовищ, частково або повністю відбивається.

-**заломлення.** Заломлення виникає при переході хвилі з одного середовища в інше зі зміною швидкості поширення.

-**поглинання.** Частина енергії звукової хвилі переходить у тепло через в'язкість і внутрішнє тертя.

-**розсіювання.** Розсіювання відбувається, коли хвиля зустрічається з нерівностями або об'єктами розміром, співрозмірним з довжиною хвилі. Енергія розподіляється в різних напрямках.

Особливості розповсюдження звукових хвиль

Дифракція - здатність звукових хвиль огинати перешкоди.

Резонанс - ефект різкого зростання амплітуди вимушених коливань якоїсь пружної системи при близькому наближенні чи повному збігу частоти вимушених коливань із частотою цієї системи.

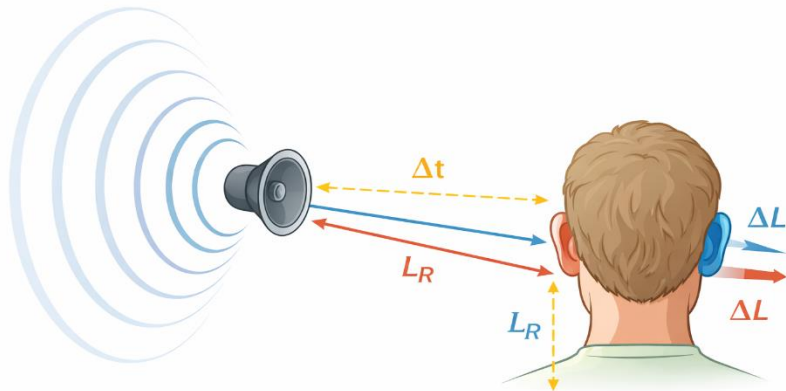
Сприйняття просторового звуку, 3D-звук

Людина здатна визначати, звідки чути звук. Цю здатність називають бінауральним ефектом. У створенні цього ефекту беруть участь два явища: запізнення та згасання.

На низьких частотах (**300-1000** Гц) велику роль грає запізнення сигналу, що сприймається лівим та правим вухом. Людині вдається навіть приблизно визначати кут напрямку, де знаходиться джерело звуку.

На більш високих частотах (**> 1000** Гц) через порівняно швидке загасання звукових хвиль напрямок на джерело визначається мозком людини шляхом аналізу гучності звуку (амплітуди звукових хвиль, що доходять до лівого та правого вух будуть різними). І в першому, і в другому випадку важливо те, що звук проходить різну відстань на шляху до лівого та правого

Визначення напрямку звуку



На слайді зображено схему визначення напрямку звуку (бінауральна локалізація).

Фізичне джерело розташоване зліва від слухача, тому відстані до лівого і правого вуха різні.

Це призводить до:

- різниці часу приходу сигналу;
- різниці рівня звукового тиску.

Саме ці механізми дозволяють мозку визначати напрямок джерела. Зображення демонструє принцип бінауральної локалізації звуку, що базується на:

- геометричній різниці відстаней;
- різниці часу приходу;
- різниці інтенсивностей;
- акустичному затіненні головою.

Саме ці механізми лежать в основі:

- просторового аудіо;
- 3D-звуку;
- систем Dolby Atmos;
- HRTF-моделювання у мультимедійних технологіях.

Декілька фактів про звуки

Рівень звуку вимірюють у децибелах (дБ). Больовий поріг слуху людини настає приблизно на рівні **120–130** дБ. Рівні понад **150–160** дБ можуть спричинити тяжкі ушкодження слухового апарату, а ще вищі значення (понад **180–200** дБ) становлять загрозу для життя. Звук і шум — не тотожні поняття.

Звук — це механічні коливання пружного середовища, що сприймаються органами слуху людини або тварин.

Шум — це безладне поєднання звукових коливань різних частот і амплітуд, яке зазвичай не має чіткої гармонічної структури. Власний голос у записі сприймається інакше, ніж під час мовлення. Це пояснюється тим, що ми чуємо себе двома шляхами:

-через повітряну провідність (зовнішній слуховий канал, барабанна перетинка, середнє вухо);

-через кісткову провідність (передавання коливань через тканини черепа безпосередньо до внутрішнього вуха).

Кісткова провідність підсилює низькі частоти, тому власний голос здається нижчим і глибшим. Деякі люди можуть чути внутрішні фізіологічні звуки — наприклад, звук руху очних яблук або власне дихання. Це пов'язано з підвищеною чутливістю внутрішнього вуха або з особливостями провідності звуку в організмі.

Декілька фактів про звуки

Під час дії інтенсивного шуму відбувається адаптація слуху. Гучні звуки активують захисні механізми слухової системи, що тимчасово знижують її чутливість. Завдяки цьому людина сприймає власний крик менш гучним, ніж оточуючі, оскільки мозок частково пригнічує сприйняття прогнозованих звукових сигналів. Шум моря, який ми нібито чуємо через морську раковину, насправді є не «звуком моря», а результатом посилення навколишніх шумів і внутрішніх звуків організму (зокрема кровотоку). Подібний ефект можна відтворити, приклавши до вуха звичайну чашку або інший порожнистий предмет. Люди з порушенням слуху в окремих випадках здатні сприймати звуки завдяки кістковій провідності. Відомий приклад — Людвіг ван Бетховен, який після втрати слуху продовжував створювати музику. Композитор прикладав до рояля кінець тростини, а інший кінець затискав зубами. Таким чином коливання передавалися через кістки черепа до внутрішнього вуха, яке частково зберігало функціональність. Звук «а» вважається одним із найпоширеніших голосних звуків у мовах світу. Загальна кількість мов на планеті становить приблизно 6–7 тисяч. Найбільша кількість носіїв припадає на китайську, іспанську, англійську, гінді, арабську та інші мови.

Декілька фактів про звуки

Нормою слуху вважається здатність людини чути тиху розмовну мову на відстані приблизно 5–6 метрів. Погіршення слуху часто відбувається поступово, тому людина може не одразу помітити зниження чутливості. На ранніх етапах проблему інколи можна скоригувати або уповільнити її розвиток. Однак за відсутності лікування процес може стати незворотним.