«Кодирование текстовой, графической и звуковой информации»

Код-система знаков для представления информации.

Кодирование информации – переход от одной формы представления информации к другой, более удобной для хранения.

Декодирование – процесс обратный кодированию.
Существуют три основных вида кодирования текста:

графический

числовой

символьный

Информация, выраженная с помощью естественных и формальных языков в письменной форме, называется **текстовой информацией**

АЛЕКСАНДР НЕВСКИЙ (1221-1263)

Александр Ярославич Невский - князя Ярослава Всеволодовича.

С 1230 г. получил в княжение Новгород Великий. Вскоре на Русь пришли татары. Озера и болота остановили их и заставили повернуть, не дойдя до Новгорода, но нельзя было поручиться за его свободу в будущем.

Перед молодым князем возникала и другая, более близкая и более серьезная опасность со стороны шведов, ливонцев и Литвы. Борьба с ливонцами и со шведами была, по сути дела, борьбой православного Востока с католическим Западом. В 1237 г. силы ливонцев – тевтонского ордена и меченосцев – объединились против русских. Александр, сыграв в 1239 г. свадьбу с Александрой, дочерью Брячислава Полоцкого, приступил к укреплению западной границы своей области по реке Шелони.

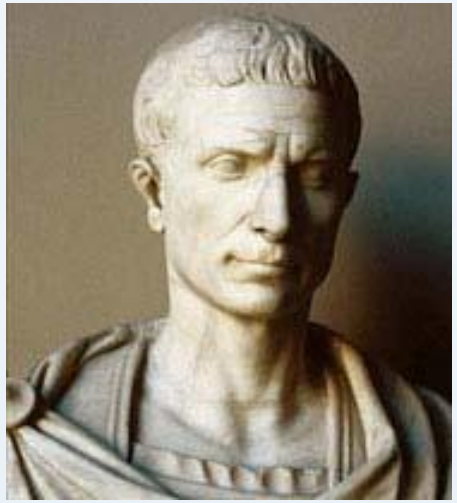
Виды кодирования текста:

Сурдожесты – язык жестов, используемый людьми с нарушениями слуха.

Криптография – это тайнопись, система изменения письма с целью сделать текст непонятным для непросвещенных лиц.

Азбука Морзэ или неравномерный телеграфный код, в котором каждая буква или знак представляет своей комбинацией точек и тире.

Код Цезаря



Юлий Цезарь
(I век до н.э.)

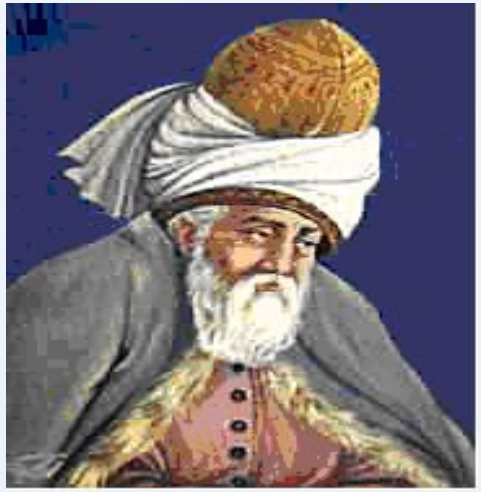
Замени каждую букву шифруемого текста на другую путем смещения в алфавите от исходной буквы на фиксированное количество символов!

Закодируем Б А Й Т –
сместим на 2 символа вправо

Получим: Г В Л Ф

А Б В Г Д Е Ё Ж З И Й К Л М Н О П Р С Т У Ф Х Ц Ч Ш
Щ Ъ Ы Ь Э Ю Я1

Задание:



Руми
1207-1273

Расшифруйте фразу персидского поэта Джалаледдина Руми

«кгнусм ёогкг фесл - тцфхя фзужщз фхгрзх ёогксп»,

закодированную с помощью шифра Цезаря. Известно, что каждая буква исходного текста заменяется **третьей** после нее буквой.

А Б В Г Д Е Ё Ж З И Й К Л М Н О П Р С Т У
Ф Х Ц Ч Ш Щ Ъ Ы Ь Э Ю Я

Кодирование символов

Текстовый файл

- на экране (символы)
- в памяти – двоичные коды

1000001_2	1000010_2	1000011_2	1000100_2
65	66	67	68

! В файле хранятся не изображения символов, а их числовые коды в двоичной системе!

Двоичное кодирование текстовой информации

Для кодирования **1 символа** используется **1 байт** информации.

256 символов → $256=2^8$ → **1 байт**

- **66 букв русского алфавита**
- **52 буквы английского алфавита**
- **0-9 цифры**
- **Знаки препинания**

При обработке текстовой информации в компьютере каждый символ представляется двоичным кодом

1 символ → 8 битов →

→ От 00000000 до 11111111

Присвоение знаку конкретного двоичного кода – это вопрос соглашения, которое фиксируется в кодовой таблице

Кодовая таблица **ASCII**

American **S**tandard **C**ode for **I**nformation **I**nterchange

sp 32	! 33	" 34	# 35	$ 36	% 37	& 38	' 39	(40	) 41	* 42	+ 43	, 44	- 45	. 46	/ 47
0 48	1 49	2 50	3 51	4 52	5 53	6 54	7 55	8 56	9 57	: 58	; 59	< 60	= 61	> 62	? 63
@ 64	A 65	B 66	C 67	D 68	E 69	F 70	G 71	H 72	I 73	J 74	K 75	L 76	M 77	N 78	O 79
P 80	Q 81	R 82	S 83	T 84	U 85	V 86	W 87	X 88	Y 89	Z 90	[91	\ 92	] 93	^ 94	_ 95
` 96	a 97	b 98	c 99	d 100	e 101	f 102	g 103	h 104	i 105	j 106	k 107	l 108	m 109	n 110	o 111
p 112	q 113	r 114	s 115	t 116	u 117	v 118	w 119	x 120	y 121	z 122	{ 123	| 124	} 125	~ 126	

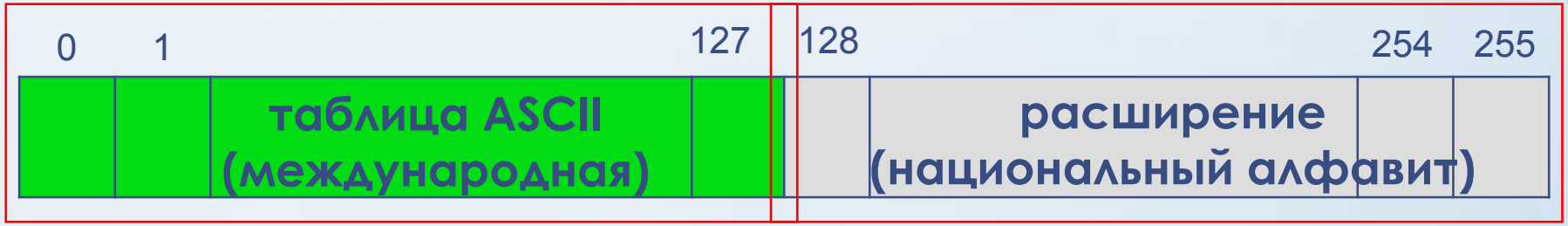
коды от 0 до 32

функциональные клавиши

коды от 33 до 127

буквы английского алфавита, знаки математических операций, знаки препинаний

8-битные кодировки (1 байт на символ)



ASCII = *American Standard Code for Information Interchange*

0-31 управляющие символы:

7 – звонок, 10 – новая строка, 13 – возврат каретки, 27 – Esc.

32 пробел

знаки препинания: **. , : ; ! ?**

специальные знаки: **+ – * / () {} []**

48-57 цифры **0..9**

65-90 заглавные латинские буквы **A-Z**

97-122 строчные латинские буквы **a-z**

Кодовая страница (расширенная таблица ASCII)

для русского языка:

CP-866 для системы *MS DOS*

CP-1251 для системы *Windows* (Интернет)

КОИ8-Р для системы *UNIX* (Интернет)

Таблицы кодировки русскоязычных символов

КОИ-8

─ 128	│ 129	┌ 130	┐ 131	└ 132	┘ 133	├ 134	┤ 135	┬ 136	┴ 137	┼ 138	▀ 139	▄ 140	█ 141	▌ 142	▐ 143
░ 144	▒ 145	▓ 146	⌠ 147	■ 148	∙ 149	√ 150	≈ 151	≤ 152	≥ 153	nbsp 154	⌡ 155	° 156	² 157	· 158	÷ 159
═ 160	║ 161	╒ 162	ё 163	╓ 164	╔ 165	╕ 166	╖ 167	╗ 168	╘ 169	╙ 170	╚ 171	╛ 172	╜ 173	╝ 174	╞ 175
╟ 176	╠ 177	╡ 178	Ё 179	╢ 180	╣ 181	╤ 182	╥ 183	╦ 184	╧ 185	╨ 186	╩ 187	╪ 188	╫ 189	╬ 190	© 191
ю 192	а 193	б 194	ц 195	д 196	е 197	ф 198	г 199	х 200	и 201	й 202	к 203	л 204	м 205	н 206	о 207
п 208	я 209	р 210	с 211	т 212	у 213	ж 214	в 215	ь 216	ы 217	з 218	ш 219	э 220	щ 221	ч 222	ъ 223
Ю 224	А 225	Б 226	Ц 227	Д 228	Е 229	Ф 230	Г 231	Х 232	И 233	Й 234	К 235	Л 236	М 237	Н 238	О 239
П 240	Я 241	Р 242	С 243	Т 244	У 245	Ж 246	В 247	Ь 248	Ы 249	З 250	Ш 251	Э 252	Щ 253	Ч 254	Ъ 255

MAC

А 128	Б 129	В 130	Г 131	Д 132	Е 133	Ж 134	З 135	И 136	Й 137	К 138	Л 139	М 140	Н 141	О 142	П 143
Р 144	С 145	Т 146	У 147	Ф 148	Х 149	Ц 150	Ч 151	Ш 152	Щ 153	Ъ 154	Ы 155	Ь 156	Э 157	Ю 158	Я 159
† 160	° 161	ѝ 162	£ 163	§ 164	• 165	¶ 166	ѝ 167	® 168	© 169	™ 170	Ѓ 171	ѓ 172	Ё̀ 173	ѐ 174	ѐ 175
Ё̀ 176	± 177	≤ 178	≥ 179	ѓ 180	µ 181	ґ 182	Ѐ 183	Ю́ 184	ю́ 185	Я́ 186	я́ 187	Е́ 188	е́ 189	ѝ 190	ѐ 191
ѐ 192	ю̀ 193	¬ 194	√ 195	ƒ 196	≈ 197	Δ 198	« 199	» 200	… 201	nbsp 202	Ѐ 203	ѐ 204	Ѝ 205	ѝ 206	я̀ 207
– 208	— 209	“ 210	” 211	‘ 212	’ 213	÷ 214	„ 215	ю̀ 216	Ы́ 217	У́ 218	у́ 219	№ 220	Ё 221	ё 222	я 223
а 224	б 225	в 226	г 227	д 228	е 229	ж 230	з 231	и 232	й 233	к 234	л 235	м 236	н 237	о 238	п 239
р 240	с 241	т 242	у 243	ф 244	х 245	ц 246	ч 247	ш 248	щ 249	ъ 250	ы 251	ь 252	э 253	ю 254	¤ 255

ISO

Ї 128	Ї 129	Ї 130	Ї 131	Ї 132	Ї 133	Ї 134	Ї 135	Ї 136	Ї 137	Ї 138	Ї 139	Ї 140	Ї 141	Ї 142	Ї 143
Ї 144	Ї 145	Ї 146	Ї 147	Ї 148	Ї 149	Ї 150	Ї 151	Ї 152	Ї 153	Ї 154	Ї 155	Ї 156	Ї 157	Ї 158	Ї 159
nbsp 160	Ё 161	Ђ 162	Ѓ 163	Є 164	S 165	І 166	Ї 167	J 168	Љ 169	Њ 170	Ћ 171	Ќ 172	shy 173	Ў 174	Ц 175
А 176	Б 177	В 178	Г 179	Д 180	Е 181	Ж 182	З 183	И 184	Й 185	К 186	Л 187	М 188	Н 189	О 190	П 191
Р 192	С 193	Т 194	У 195	Ф 196	Х 197	Ц 198	Ч 199	Ш 200	Щ 201	Ъ 202	Ы 203	Ь 204	Э 205	Ю 206	Я 207
а 208	б 209	в 210	г 211	д 212	е 213	ж 214	з 215	и 216	й 217	к 218	л 219	м 220	н 221	о 222	п 223
р 224	с 225	т 226	у 227	ф 228	х 229	ц 230	ч 231	ш 232	щ 233	ъ 234	ы 235	ь 236	э 237	ю 238	я 239
№ 240	ё 241	ђ 242	ѓ 243	є 244	s 245	і 246	ї 247	j 248	љ 249	њ 250	ћ 251	ќ 252	§ 253	ў 254	ц 255

CP1251

Ѓ 128	ѓ 129	‚ 130	ѐ 131	„ 132	… 133	† 134	‡ 135	€ 136	‰ 137	Ѐ 138	‹ 139	ѝ 140	Ѝ 141	Ю́ 142	Ю́ 143
ѐ 144	‘ 145	’ 146	“ 147	” 148	• 149	– 150	— 151	Ё̀ 152	™ 153	ѐ 154	› 155	ю̀ 156	ѝ 157	ю́ 158	ю́ 159
nbsp 160	ю̀ 161	Ы́ 162	Ѐ 163	¤ 164	ы̀ 165	¦ 166	§ 167	Ё 168	© 169	Ю́ 170	« 171	¬ 172	shy 173	® 174	Я́ 175
° 176	± 177	ы́ 178	ѐ 179	′ 180	µ 181	¶ 182	• 183	ё 184	№ 185	ю́ 186	» 187	ѐ 188	ю̀ 189	я̀ 190	я́ 191
А 192	Б 193	В 194	Г 195	Д 196	Е 197	Ж 198	З 199	И 200	Й 201	К 202	Л 203	М 204	Н 205	О 206	П 207
Р 208	С 209	Т 210	У 211	Ф 212	Х 213	Ц 214	Ч 215	Ш 216	Щ 217	Ъ 218	Ы 219	Ь 220	Э 221	Ю 222	Я 223
а 224	б 225	в 226	г 227	д 228	е 229	ж 230	з 231	и 232	й 233	к 234	л 235	м 236	н 237	о 238	п 239
р 240	с 241	т 242	у 243	ф 244	х 245	ц 246	ч 247	ш 248	щ 249	ъ 250	ы 251	ь 252	э 253	ю 254	я 255

Другие кодировочные таблицы для кодирования русских символов

- **КОИ8**
- **CP1251**
- **CP866**
- **Mac**
- **ISO**

1 символ =1 байт

Слово ЭВМ кодируется:	
КОИ-8	252 247 237
CP1251	221 194 204
CP866	157 130 140
Mac	157 130 140
ISO	205 178 188

8-битные кодировки (1 байт на символ)

- 1 байт на символ – файлы небольшого размера!
- просто обрабатывать в программах

- нельзя использовать символы разных кодовых страниц одновременно (русские и французские буквы, и т.п.)
- неясно, в какой кодировке текст (перебор вариантов!)
- для каждой кодировки нужен свой шрифт (изображения символов)

Кодировка Unicode

Идея: объединить все символы в одну таблицу!

Юникод включает практически все современные письменности, в том числе: арабскую, армянскую, бенгальскую, бирманскую, греческую, грузинскую, деванагари, иврит, кириллицу, коптскую, кхмерскую, латинскую, тамильскую, хангыль, хань (Китай, Япония, Корея), чероки, эфиопскую, японскую (катакана, хирагана, кандзи) и другие.

1 символ - 2 байта (16 бит),

которыми можно закодировать

65 536 символов

$$N=2^{16}=65\ 536$$

Таблица 3.2. **Десятичные коды некоторых символов в различных кодировках**

Символ	Windows	MS-DOS	КОИ-8	Mac	ISO	Unicode
А	192	128	225	128	176	1040
В	194	130	247	130	178	1042
М	204	140	237	140	188	1052
Э	221	157	252	157	205	1069
я	255	239	241	223	239	1103

Кодирование графической информации.

Растровое представление графической информации

Графическая информация может быть представлена в аналоговой и дискретной форме

живописное полотно

цифровая фотография

Преобразование изображения из аналоговой (непрерывной) в цифровую (дискретную) форму называется

пространственной дискретизацией

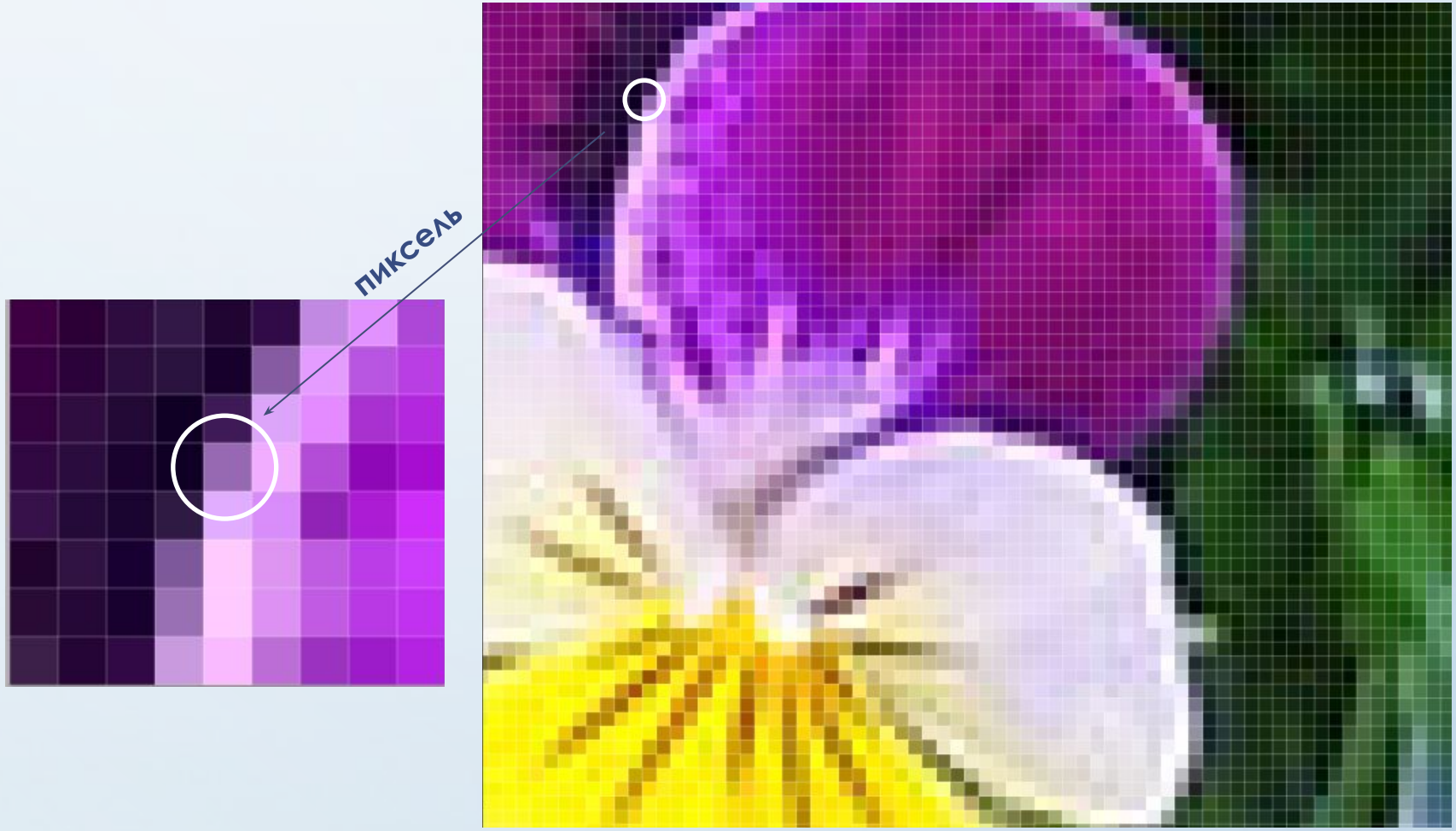
Аналоговая форма

мозаика

сканирование

Дискретная форма

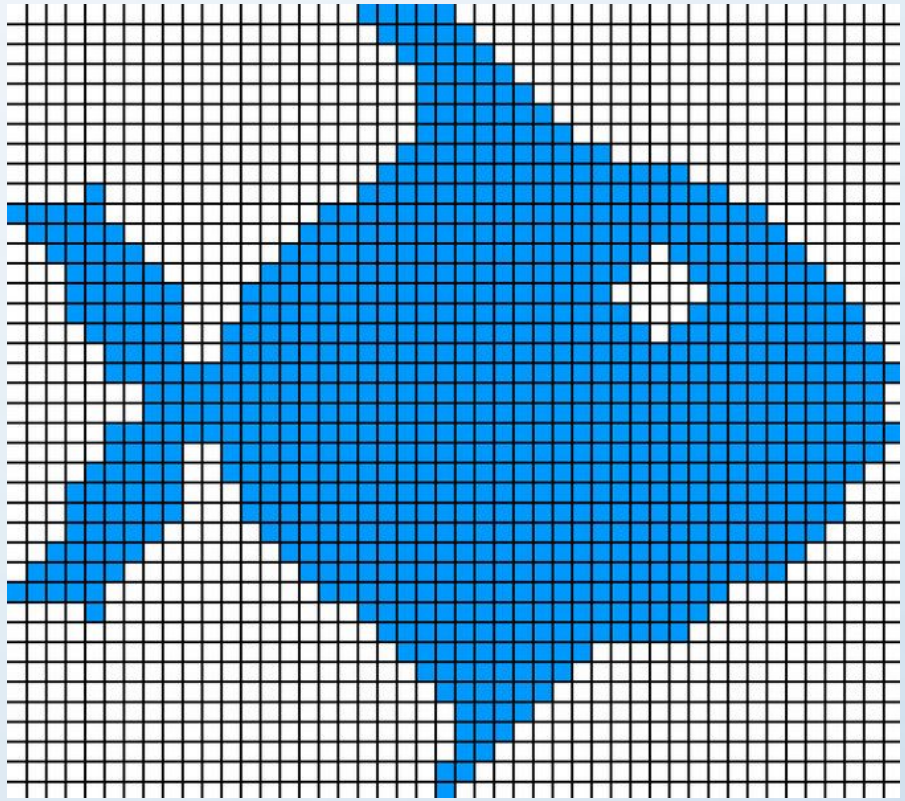
В процессе пространственной дискретизации изображение разбивается на отдельные маленькие фрагменты, точки - **пиксели**



Пиксель – минимальный участок изображения, для которого независимым образом можно задать цвет.

В результате пространственной дискретизации графическая информация представляется в виде растрового изображения.

Разрешающая способность растрового изображения определяется количеством точек по горизонтали и вертикали на единицу длины изображения.



Чем меньше размер точки, тем больше разрешающая способность, а значит, выше качество изображения.

Величина разрешающей способности выражается в dpi (dot per inch – точек на дюйм), т.е. количество точек в полоске изображения длиной один дюйм (1 дюйм=2,54 см.)

В процессе дискретизации используются различные **палитры цветов** (наборы цветов, которые могут принять точки изображения).

Количество цветов **N** в палитре и количество информации **I**, необходимое для кодирования цвета каждой точки, могут быть вычислены по формуле: $N=2^{I}$

Количество информации, которое используется для кодирования цвета точки изображения, называется **глубиной цвета**.

Пример:

Для кодирования черно-белого изображения (без градации серого) используются всего два цвета – черный и белый. По формуле $N=2^I$ можно вычислить, какое количество информации необходимо, чтобы закодировать цвет каждой точки:

$$2=2^{I} \Longrightarrow 2=2^{1} \Longrightarrow I = 1 \text{ бит}$$

Для кодирования одной точки черно-белого изображения достаточно 1 бита.

Глубина цвета и количество цветов в палитре

Глубина цвета, I (битов)	Количество цветов в палитре, N
8	$2^8 = 256$
16	$2^{16} = 65\ 536$
24	$2^{24} = 16\ 777\ 216$

Зная глубину цвета, можно вычислить количество цветов в палитре.

Задачи:

1. Растровый графический файл содержит черно-белое изображение с 16 градациями серого цвета размером 10х10 пикселей. Каков информационный объем этого файла?

<u>Решение:</u> $16 = 2^4$; 10*10*4 = 400 бит

2. 256-цветный рисунок содержит 120 байт информации. Из скольких точек он состоит?

<u>Решение</u>:
120 байт = 120*8 бит; $256 = 2^8$ (8 бит – 1 точка).
120*8/8 = 120

Задание 4

Определите количество цветов в палитре при глубине цвета 16 бит.

- ***Ответ: 65536 цветов***

Задание 5

Цветное (с палитрой из 256 цветов) растровое графическое изображение
имеет размер 10x10 точек.
Какой объем памяти в байтах займет это изображение?

Ответ: 100 байт

- **Задание 6**

- *В процессе преобразования растрового графического изображения количество цветов уменьшилось с 65536 до 16. Во сколько раз уменьшится объем занимаемой памяти?*

Ответ: в 4 раза

Растровые изображения на экране монитора

Качество изображения на экране монитора зависит от величины **пространственного разрешения** и **глубины цвета.**

определяется как произведение количества строк изображения на количество точек в строке (800*600
1024*768
1400*1050 и выше)

характеризует количество цветов, которое могут принимать точки изображения (измеряется в битах)

Формирование растрового изображения на экране монитора

Видеопамять	
Номер точки	Двоичный код цвета точки
1	01010101
2	10101010
…..	
800	11110000
…..	
480 000	11111111

Объем видеопамяти.

Информационный объем требуемой видеопамяти можно рассчитать по формуле:

$$I_n = I \times X \times Y,$$

где In - информационный объем видеопамяти в битах;

X × У - количество точек изображения
(X - количество точек по горизонтали, Y - по вертикали);

I - глубина цвета в битах на точку.

Пример:

Необходимый объем видеопамяти для графического режима с пространственным разрешением 800 х 600 точек и глубиной цвета 24 бита равен:

In = I× X×Y = 24 бита × 800 × 600 =

=11 520 000 бит :8= 1 440 000 байт :1024=

= 1 406,25 Кбайт :1024= 1,37 Мбайт.

Задачи:

1. Рассчитайте объём памяти, необходимый для кодирования рисунка, построенного при графическом разрешении монитора
800x600 с палитрой 32 цвета.

Решение:
$32 = 2^5$
800*600*5 бит = 2400000 бит : 8 : 1024 = 293 Кбайт

2. Какой объем видеопамяти необходим для хранения четырех страниц изображения при условии, что разрешающая способность дисплея 640x480 точек с палитрой 32 цвета?

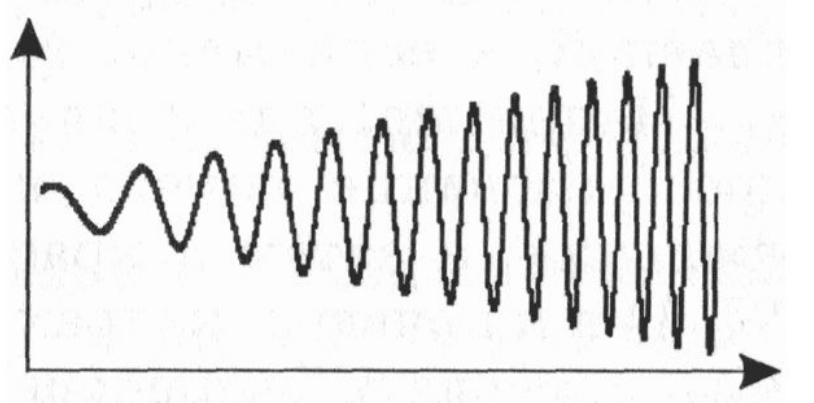
Решение:
640*480*5*4 = 6144000 бит : 8 : 1024 = 750 Кбайт

Задание 7

Достаточно ли видеопамяти объемом 256 Кбайт для работы монитора в режиме 640x480 и палитрой из 16 цветов?

- **Ответ: достаточно**

Кодирование звуковой информации



Способы хранения звука

Звукозапись – процесс сохранения информации о параметрах звуковых волн

Способы хранения

Аналоговый

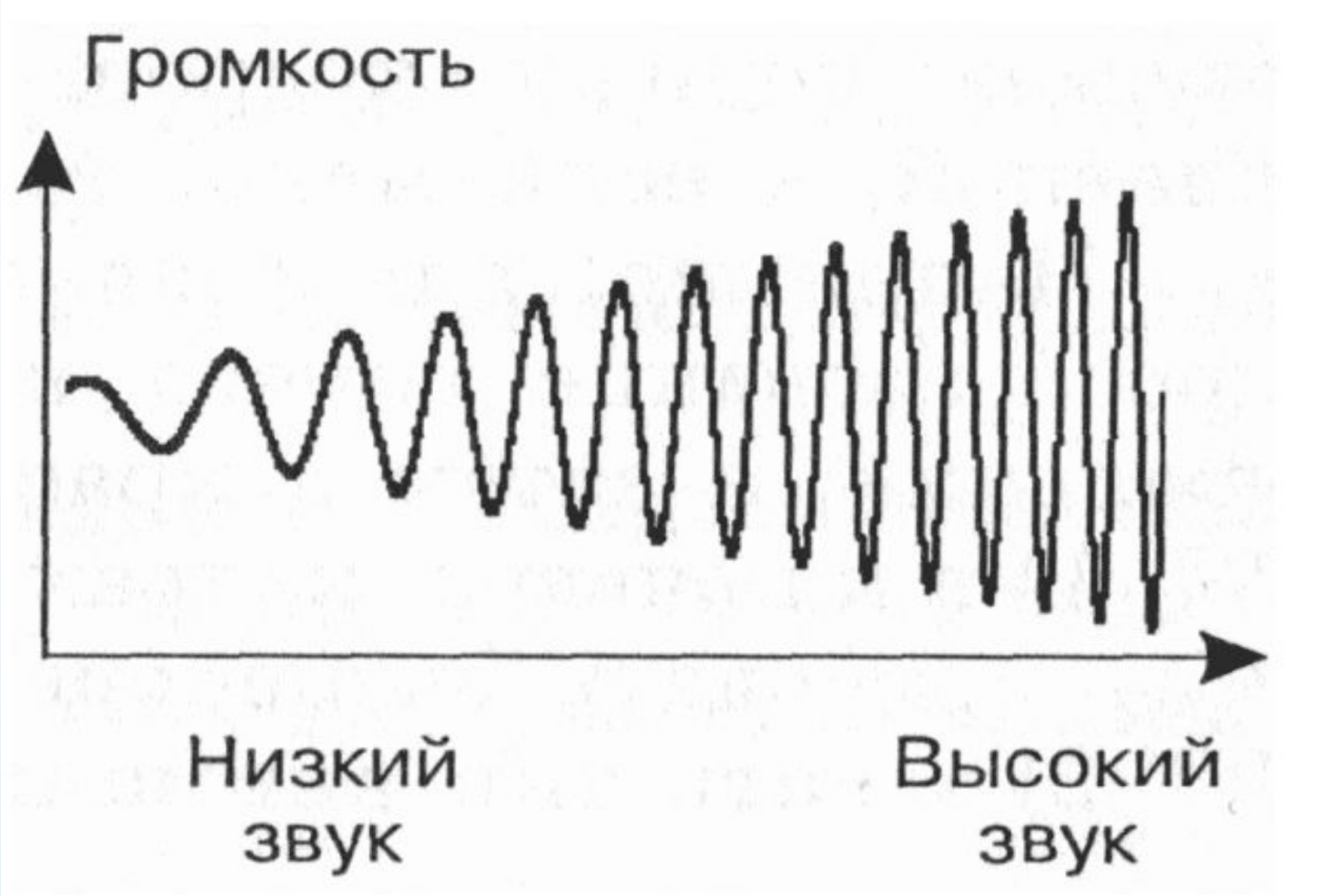
- **грампластинка**
- **магнитная лента**

Цифровой

- **временная дискретизация**
- **квантование**

Звук – это волна с непрерывно меняющейся амплитудой и частотой

Чем больше амплитуда, тем громче звук
Чем больше частота, тем больше тон


Громкость
Низкий звук
Высокий звук

Для измерения громкости звука применяется специальная единица **"децибел"** (дбл)

Некоторые значения уровней шума

Порог слышимости	0 дБ
Шорох листьев, шум слабого ветра	10-20 дБ
Шепот (на задней парте)	20-30 дБ
Разговор средней громкости (в кабинете директора)	50-60 дБ
Автомагистраль с интенсивным движением	80-90 дБ
Авиадвигатели	120-130 дБ
Болевой порог	140 дБ

Кодирование звуковой информации

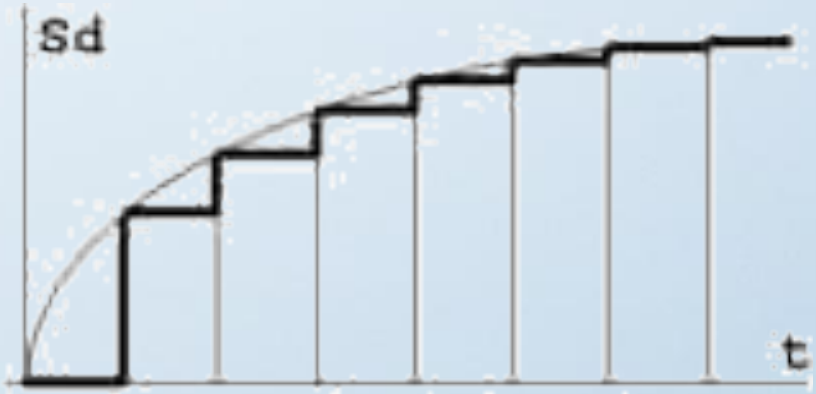
С начала 90-х годов персональные компьютеры получили возможность работать со звуковой информацией. Каждый компьютер, имеющий звуковую плату, микрофон и колонки, может записывать, сохранять и воспроизводить звуковую информацию.

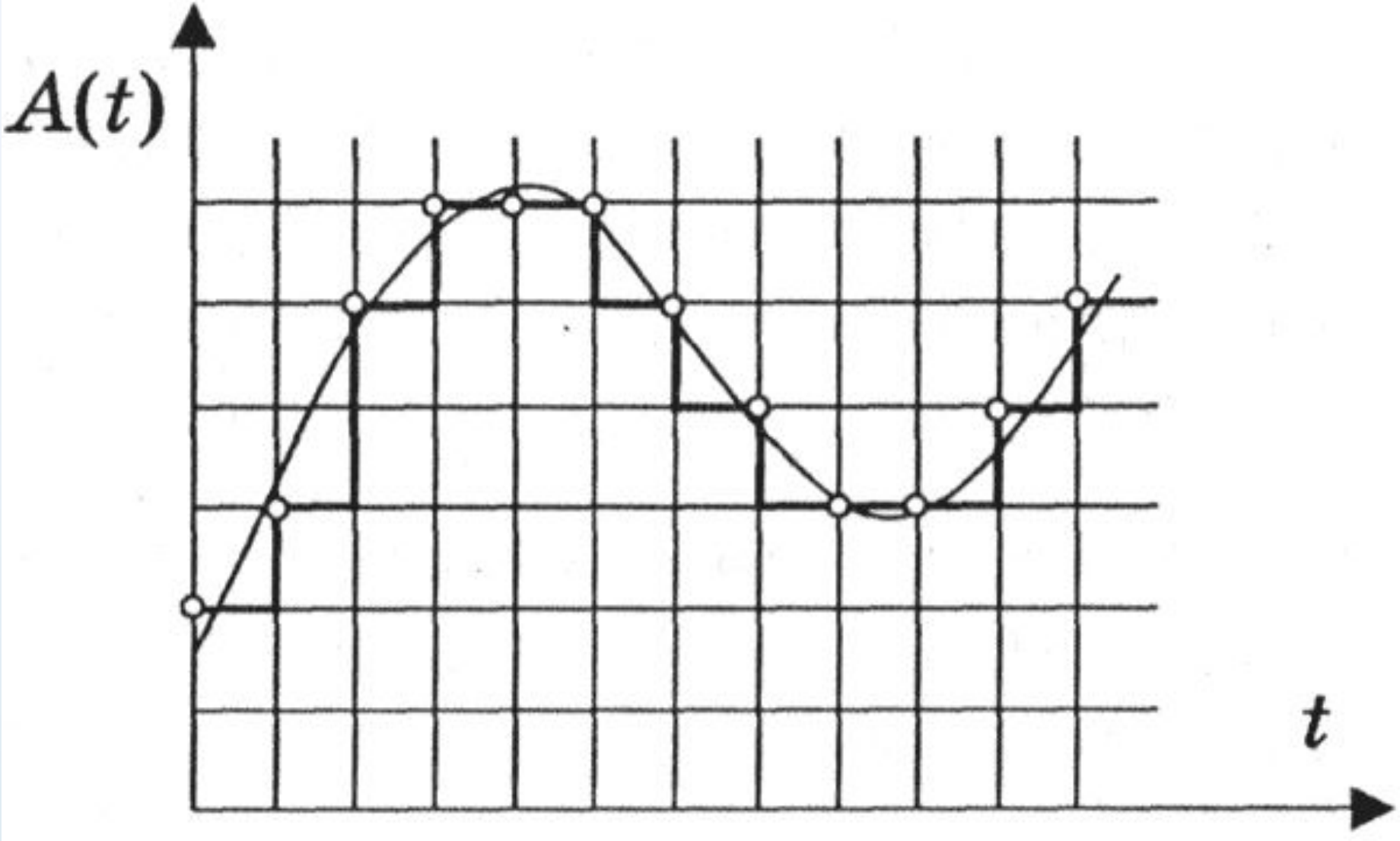
Для того чтобы компьютер мог обрабатывать звук, непрерывный звуковой сигнал должен быть преобразован в цифровую дискретную форму с помощью временной дискретизации.

Дискретизация - это преобразование непрерывных сигналов в набор дискретных значений, каждому из которых присваивается определенный код.

Аналоговый сигнал

Дискретный сигнал




A(t)
t

Характеристика цифрового звука:

1. Частота
2. Глубина

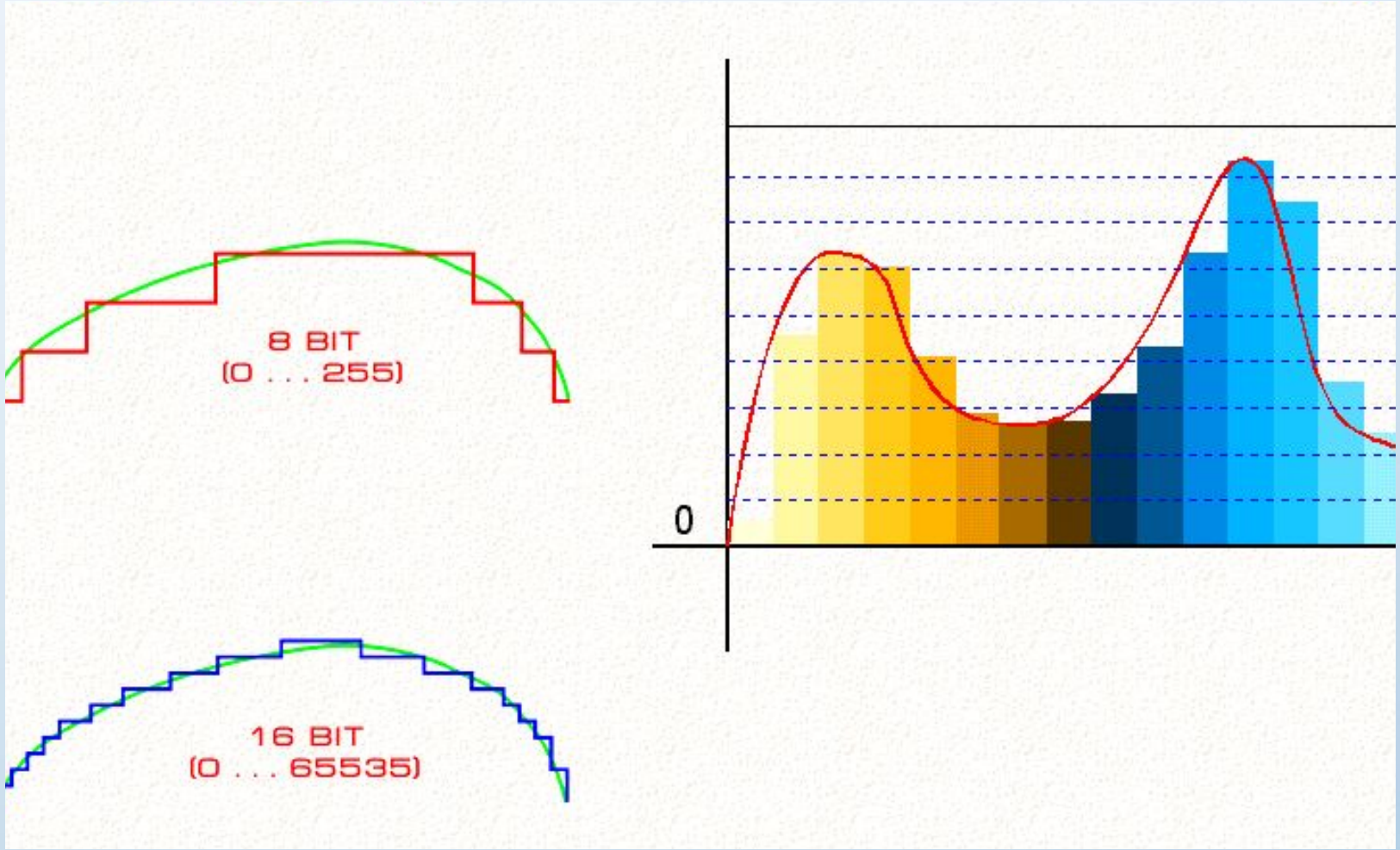
Частота дискретизации звука - это количество измерений громкости звука за одну секунду

Частота дискретизации

- Количество измерений уровней сигнала за 1 секунду.
- Измеряется в Герцах.
- 1 измерение в секунду – 1 Гц
- 1000 измерений в секунду – 1кГц
- Изменяется в диапазоне от 8кГц до 48 кГц

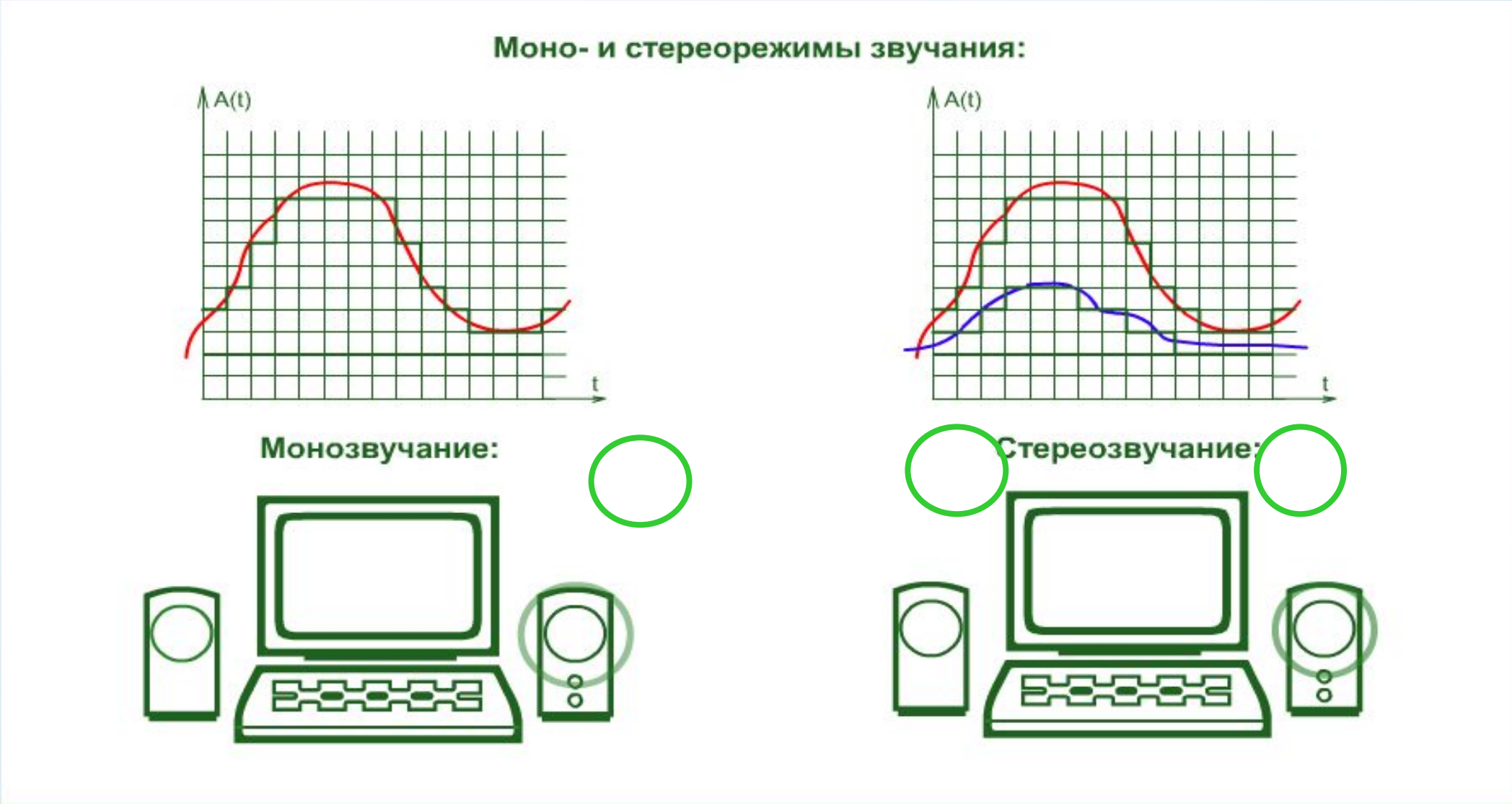
Чем большее количество измерений производится за 1 секунду (чем больше частота дискретизации), тем точнее "лесенка" цифрового звукового сигнала повторяет кривую диалогового сигнала

Глубина (разрядность) кодирования звука - это количество информации, которое необходимо для кодирования дискретных уровней громкости цифрового звука.


8 BIT
(0 . . . 255)
16 BIT
(0 . . . 65535)
0

Если известна глубина кодирования, то количество уровней громкости цифрового звука можно рассчитать по формуле $N = 2^I$.
Пусть глубина кодирования звука составляет 16 битов, тогда количество уровней громкости звука равно:
$N = 2^I = 2^{16} = 65\ 536$.

Режимы


Моно- и стереорежимы звучания:
A(t)
t
A(t)
t
Монозвучание:
Стереозвучание:

Самое низкое качество оцифрованного звука, соответствующее качеству телефонной связи, получается при частоте дискретизации 8000 раз в секунду, глубине дискретизации 8 битов и записи одной звуковой дорожки (режим "моно").

Самое высокое качество оцифрованного звука, соответствующее качеству аудио-CD, достигается при частоте дискретизации 48 000 раз в секунду, глубине дискретизации 16 битов и записи двух звуковых дорожек (режим "стерео").

Объем файла (бит) =
частота (Гц) *
глубина (бит) *
время (сек) *
режим (моно = 1, стерео = 2)

Задача

Определить информационный объем стерео аудио файла длительностью звучания 1 секунда при высоком качестве звука(16 битов, 48 кГц).

Запись условия

T=1 сек

I=16 бит

H= 48 кГц

Стерео - ×2

V=?

Решение

V= T ×I × H × 2

V=1 ×16 × 48 000 × 2=

1536000 бит/8 =192000 байт/1024 = 187,5 Кбайт

Задача

Если глубина кодирования звука составляет 16 битов рассчитайте количество уровней громкости звука

$N = 2^I$

$N = 2^I = 2^{16} = 65\ 536.$

16 бит - глубина звука

44 кГц - частота дискретизации

10 сек - длительность звучания файла

Стерео - режим звучания

Выполняйте вычисления

Введите ответ

бит

Ok

Задание

Подсчитайте объем памяти для хранения звукового файла. Необходимые данные для решения задачи приведены выше.

Задача

Определить информационный объем цифрового аудио файла длительностью звучания которого составляет 10 секунда при частоте дискретизации 22,05 кГц и разрешении 8 битов.

Запись условия

T=10 сек

I=8 бит

H= 22,05 кГц

Моно- ×1

V=?

Решение

V= T ×I × H × 2

V=10 ×8 × 22 050 × 1=

10 × 8 × 22 050 бит/8 = 220500 байт/1024 = 215,332/1024 Кбайт = 0,21 Мбайт