

DOI: <https://doi.org/10.36910/6775-2524-0560-2020-41-24>

УДК 004.522

Дмитренко Тарас Васильович, асистент

<https://orcid.org/0000-0002-2339-2903>

Національний авіаційний університет

## МЕТОДИКА ОБРОБКИ АУДІО-СИГНАЛІВ ЗА ДОПОМОГОЮ АЛГОРИТМІВ НА БАЗІ МОВИ ПРОГРАМУВАННЯ PYTHON

Дмитренко Т. В. Методика обробки аудіо-сигналів за допомогою алгоритмів на базі мови програмування Python. Проведено аналіз можливості застосування інтерпретовано об'єктно-орієнтованої мови програмування при обробці масивів аудіоданих як цифрового способу представлення звукових сигналів. Продемонстровано принцип використання з зазначеною метою мови програмування високого рівня зі строгою динамічною типізацією Python. Визначено особливості застосування у даній галузі таких модулів (python-бібліотек) як: NumPy, SciPy та Matplotlib. Наведено методи обробки та модифікації масивів аудіоданих з метою їх подальшого застосування у мультимедійних комп'ютерних мережах. Побудовано математичну модель обробки аудіо-даних, ефективність якої перевірено на базі відповідних програмних алгоритмів. Показано можливість вирішення актуальних задач та дослідження теоретичних аспектів проблеми області обробки аудіо-даних шляхом використання інтерпретовано об'єктно-орієнтованої мови програмування і спеціалізованих бібліотек з відкритим кодом.

**Ключові слова:** цифрова обробка аудіо-даних, високорівнева мова програмування Python, python-бібліотеки, NumPy, SciPy, Wave, Matplotlib.

Дмитренко Т. В. Методика обработки аудио сигналов с помощью алгоритмов на базе языка программирования Python. Проведен анализ возможности применения интерпретируемого объектно-ориентированного языка программирования при обработке массивов аудиоданных как цифрового способа представления звуковых сигналов. Продемонстрирован принцип использования с указанной целью языка программирования высокого уровня со строгой динамической типизацией Python. Определены особенности применения в данной области таких модулей (python библиотек) как: NumPy, SciPy и Matplotlib. Приведены методы обработки и модификации массивов аудиоданных с целью их дальнейшего применения в мультимедийных компьютерных сетях. Построена математическая модель обработки аудиоданных, эффективность которой проверена на базе соответствующих программных алгоритмов. Показана возможность решения актуальных задач и исследования теоретических аспектов проблем в области обработки аудиоданных путем использования интерпретируемого объектно-ориентированного языка программирования и специализированных библиотек с открытым кодом.

**Ключевые слова:** цифровая обработка аудиоданных, высокоуровневый язык программирования Python, python-библиотеки, NumPy, SciPy, Matplotlib.

Dmitrenko T. V. A technique for processing audio signals using algorithms based on the Python programming language. The analysis of the possibility of using an interpreted object-oriented programming language when processing arrays of audio data as a digital way of representing audio signals is carried out. The principle of using the high-level programming language Python with strong dynamic typing for the indicated purpose is demonstrated. The features of application in this area of epy modules (python libraries) NumPy, SciPy and Matplotlib have been determined. Methods for processing and modifying audio data arrays are presented for further use in multimedia computer networks. A mathematical model of audio data processing has been built, the efficiency of which has been tested on the basis of appropriate software algorithms. The possibility of solving urgent problems and studying the theoretical aspects of problems in the area of audio data processing by using an interpreted object-oriented programming language and specialized open source libraries is shown.

**Key words:** digital processing of audio data, high-level programming language Python, python libraries, NumPy, SciPy, Matplotlib.

### 1. Вступ

На сьогоднішній день в області інформаційних технологій (ІТ) найбільш актуальною проблемою є побудова ефективних алгоритмів обробки великих об'ємів даних. Для виділення зазначеної галузі як методології по роботі з окремим класом задач було запропоновано ввести таке поняття як «великі дані» (Big Data), яке відповідає опису структурованих та неструктурованих масивів даних настільки великих розмірів, що стандартні методи аналізу і обробки на автоматичному та напівавтоматичному рівні до них не можуть бути застосовані. Вирішення проблеми розробки алгоритмів ефективної роботи з великими даними також відповідає актуальному завданню впровадження концепції машино-машинної взаємодії (machinetomachine, M2M), що, у свою чергу, базується на системах обробки великих обсягів мультимедійних даних у режимі реального часу і використовується у сучасних системах доступу, безпеки, моніторингу та ін.

Одним з найбільш важливих завдань, що можуть бути представлені у галузі роботи з великими даними, є обробка масивів аудіо-даних. Наявність відповідного програмного забезпечення як вузькоспеціалізованого інструментарію, що може бути використано за умови наявності підготовленого оператора, очевидно, не вирішує поставлене завдання. Зазначені модулі програмного забезпечення характеризуються налаштованістю під конкретні задачі, а також закритим кодом і, відповідно, вони не можуть бути модифіковані та включені до M2M-системи аналізу. Слід окремо зазначити, що вимоги по розробці сучасних систем машинного аналізу великих даних вказують на необхідність побудови гнучких алгоритмів і масштабованих апаратно-програмних комплексів, що модифікуються

відповідно проведеного аналізу та внесенню змін у формулювання базового завдання. Для цього необхідно також провести вдосконалення роботи з даними на рівні «людина-машина» (person to machine, P2M), зокрема, на рівні візуального представлення процесу обробки аудіо-сигналів, що надає можливість побудувати математичну модель та ефективно модифікувати її у процесі виконання практичних задач.

Таким чином, **постановка проблеми** даного дослідження включатиме у себе наступні складові:

- ☐ побудова узагальненої математичної моделі процесу обробки аудіо-сигналів, що може бути використана при роботі з широким класом задач;
- ☐ розробка методологічних рекомендацій по візуальному представленню процесу обробки аудіо-сигналів;
- ☐ вирішення задачі обробки великих об'ємів аудіо-даних.

**Аналіз останніх досліджень і публікацій** присвячених проблемі обробки великих обсягів аудіо-даних включав у себе як дослідження по розробці стандартних програмних алгоритмів [1-3], так і дослідження по побудові нейромережових методик обробки [4, 5]. Проведений аналіз вказав на те, що стандартні програмні засоби обробки характеризуються меншим навантаженням на обчислювальний ресурс апаратно-програмного комплексу, тому для ряду практичних задач зазначений підхід може бути розглянуто як пріоритетний. Було зазначено, що ефективні підходи по обробці масивів аудіо-даних було реалізовано на базі спеціалізованих бібліотек та додатків процедурно-імперативних мов програмування загального призначення [6, 7], а також таких програмних інструментів обробки матриць числових даних як MATLAB, Scilab і «Wolfram Mathematica» [8-10]. Розглянуті підходи надають можливість побудувати математичний апарат та провести ефективну обробку аудіо-сигналів з візуалізацією окремих складових процесу, у тих межах, що були представлені розробником. **Виділення невирішеної частини проблеми**, таким чином, полягає у відсутності відкритого програмного коду, який можна модифікувати алгоритми, розширити функціонал відповідно до поставленої задачі та впровадити отриманий код у загальну систему машинного аналізу великих масивів аудіо-даних.

Відповідно до проведеного аналізу **мета дослідження** полягає у побудові математичної моделі процесу обробки аудіо-сигналів та розробці методології побудови алгоритмів машинної обробки великих об'ємів аудіо-даних, що включає у себе етапи візуалізації складових процесу.

## 2. Методологія дослідження

Розглянемо та формалізуємо процес цифрового запису та обробки аудіо-даних. Запропонована модель (рис. 1) включає у себе наступні складові: параметри вхідного потоку та зовнішнього середовища, у якому відбувається процес аудіо запису, параметри програмно-апаратного комплексу системи аудіо-запису та набір процедур, що виконується оператором, система обробки масиву аудіо-даних, результати обробки та візуалізація процесу обробки.

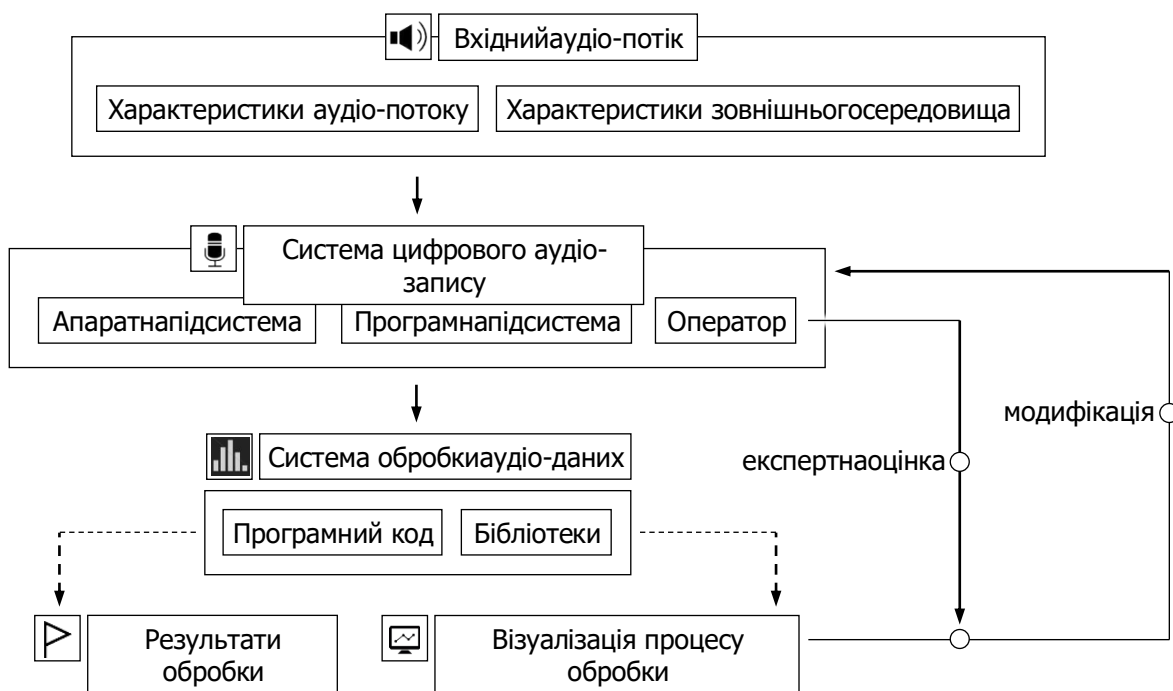


Рисунок 1. Базова схема системи цифрового аудіо-запису і обробки і аудіо-даних.

Характерно, що представлений узагальнений алгоритм обробки і аудіо-даних є нелінійним: як можна побачити на основі отриманих результатів та візуалізації процесу обробки проводиться експертна оцінка і оптимізуються система цифрового аудіо-запису та алгоритми роботи за вхідними даними. Таким чином, система реєстрації і обробки аудіо-даних може бути модифікована та вдосконалена відповідно до уточнення поставленого завдання, зміни зовнішніх умов або розширенню апаратно-програмного комплексу загальної системи.

У рамках даного дослідження головним завданням є пошук оптимального програмного інструментарію на основі якого можна реалізувати схему ефективної обробки великих обсягів великих масивів аудіо-даних у режимі реального часу. При виборі спеціалізованого програмного забезпечення необхідно враховувати наступні ключові характеристики:

- ☐ наявність відкритого програмного коду з можливістю його модифікації та включення у систему аналізу і обробки;
- ☐ рівень підтримки аналітичних операцій, обмеження розрядної сітки при розрахунках та кількість символічно-аналітичних операцій;
- ☐ інтерфейс роботи з програмним кодом, можливість оперативного внесення правок;
- ☐ стабільність роботи програмного забезпечення при високому рівні навантаження на обчислювальний ресурс системи та низькій оптимізації програмного коду;
- ☐ можливість інтеграції засобів обробки з сучасними технологіями та розподіленими масштабованими системами, такими як веб-фреймворки, хмарні сервіси і мобільні додатки.

Такими чином у рамках дослідження перевагу було надано інтерпретовано об'єктно-орієнтованій мові програмування високого рівня зі строгою динамічною типізацією Pythona відповідним спеціалізованим бібліотеками з відкритим кодом NumPy, Matplotlib і SciPy, що дозволяють поєднати потужність і гнучкість системи обробки масивів даних. Pythone мультиплатформеним засобом програмування з широким набором засобів розробки, що оптимальним чином може бути інтегрована до коду написаними на іншій мові програмування, веб-фреймворку та програмного додатку [13]. Характерними особливостями є представлення коду програми у вигляді скриптів, підтримка всіх парадигм програмування (зокрема, об'єктно-орієнтовного програмування), наявність інтерпретатора, що конвертує мову високого рівня у машинний бінарний код, а також підсистема автоматичного керування ресурсами пам'яті.

NumPy і SciPy є бібліотеками Python, які використовуються для проведення математичного аналізу. Бібліотека NumPy містить інформацію про роботу з масивами та операції сортування і індексація. Бібліотека SciPy, у свою чергу, є числовим кодом, що розширює можливість роботи у області лінійної алгебри, перетворення Фур'є та ін. [12].

Matplotlib є пакетом Python, що використовується для візуалізації даних. Це кроссплатформенна бібліотека для створення 2D графіків на основі масивів даних. Matplotlib базується на бібліотеці NumPy, як числовому розширенні для Python [11]. У рамках застосування зазначеного пакету розробнику пропонується об'єктно-орієнтований інтерфейс прикладного програмування (application programming interface, API) який додає графічні функції на базі інструментарію «Python GUI», PyQt і wxPython of Tkinter. Програмний пакет Matplotlib використовує процедурний інтерфейс PyLab, що є близьким до MATLAB з відкритим вихідним кодом [11].

Слід окремо зазначити, що зазначені бібліотеки Python працюють з цифровим аудіо-сигналом. Таким чином якість обробки аудіоданих значною мірою залежить від процедур дискретизації та квантування по рівням вхідного аудіо-сигналу, що має бути включено до аналізу при проведенні модифікації системи реєстрації і обробки даних.

### 3. Результати дослідження

Відповідно до запропонованої методології дослідження на базі мови програмування Python було побудовано базовий алгоритм роботи з масивами аудіоданих, що складався з наступних елементів:

- ☐ підключення бібліотек NumPy і SciPy для роботи з математичними функціями;
- ☐ підключення програмного пакету Matplotlib для графічного відображення процесу реєстрації і обробки аудіо-даних;
- ☐ вибір довжини вхідного аудіо-сигналу
- ☐ моделювання масиву даних вхідного аудіо-сигналу на базі функції `np.linspace`, що формує одномірний масив визначеної кількості елементів які рівномірно розподіляються вздовж інтервалу довжини сигналу;
- ☐ математичне представлення вхідного сигналу та його графічне відображення.

Лістинг програмного коду було суміщено з результатами моделювання (рис. 2).

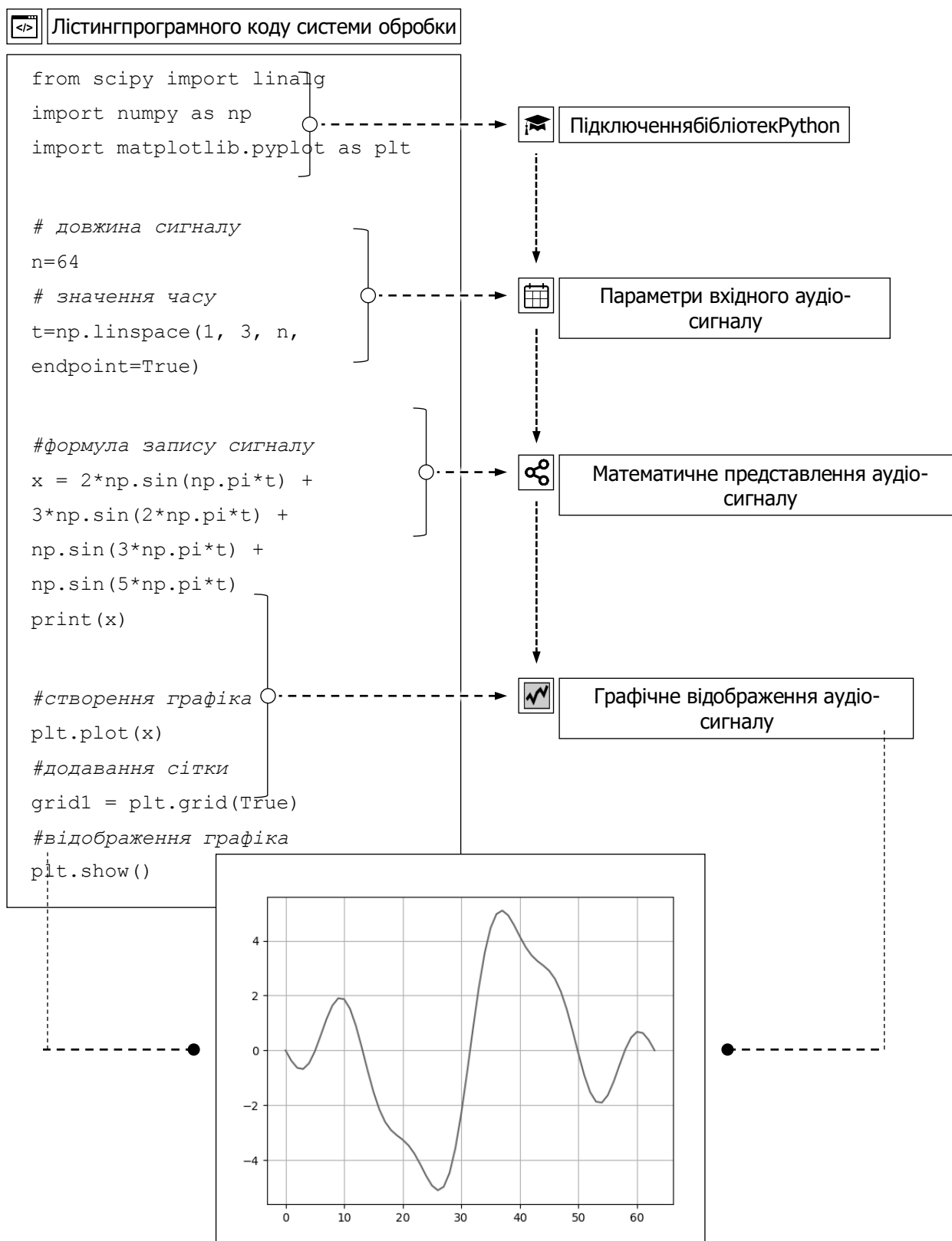


Рисунок 2. Моделювання аудіо-сигналу на баз функції `np.linspace`.

Також, аналогічним чином було побудовано синусоїдальний сигнал з заданими значеннями частоти, фази і амплітуди за стандартною формулою  $y = A \sin(\omega t + \varphi) = A \sin(2\pi f t + \varphi)$ , а також з заданим значенням частоти дискретизації (рис. 3). У наведеному прикладі значення амплітуди складало  $A = 1$ , а фази  $-\varphi = 0$ .

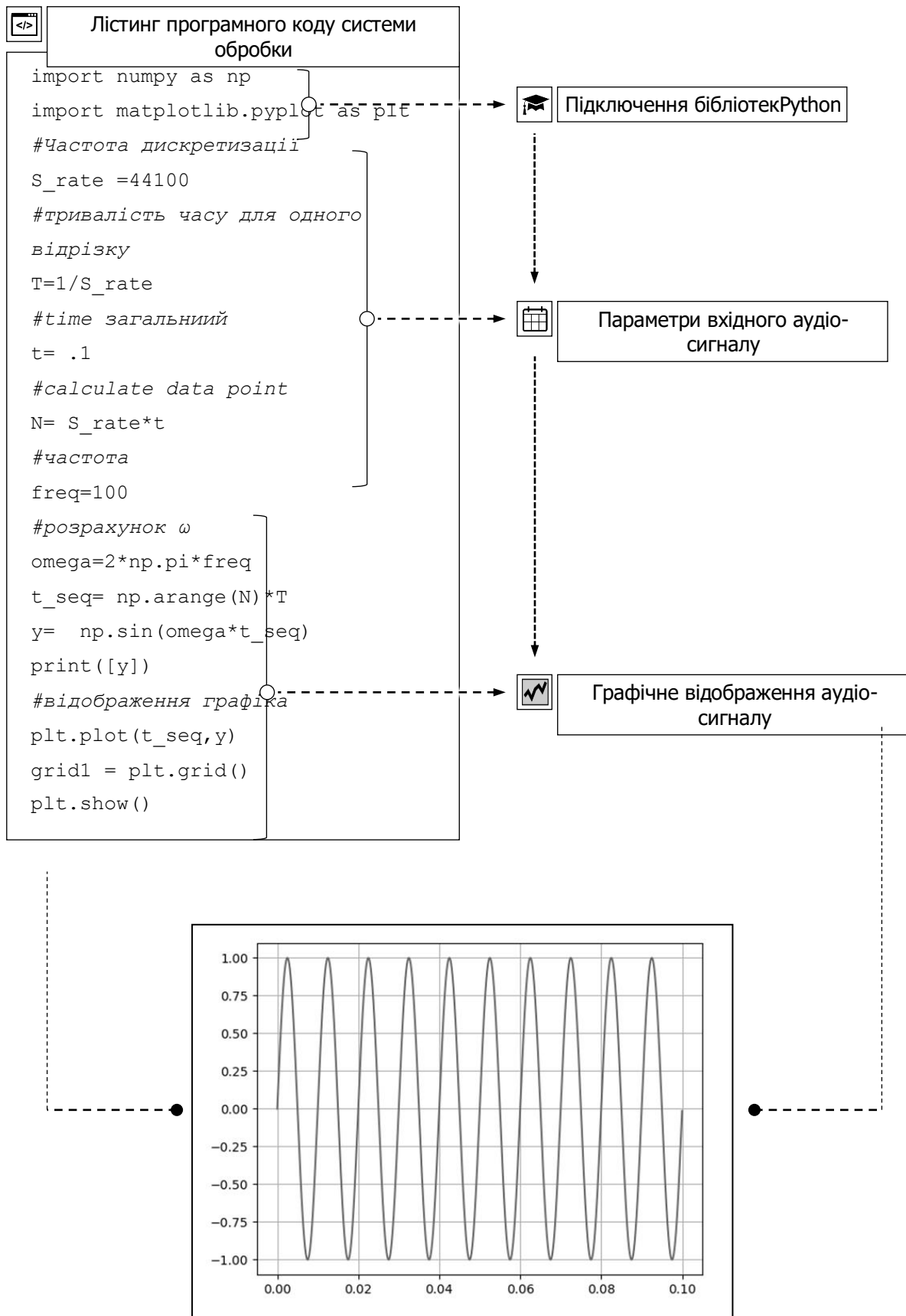


Рисунок 3. Моделювання синусоїдального сигналу з заданими значеннями частоти, фази і амплітуди.

Вказаний підхід дозволяє не тільки змоделювати аудіо-сигнал будь-якого рівня складності для оцінки ефективності системи обробки, але й працювати з наявним масивом аудіо-даних (рис. 4).

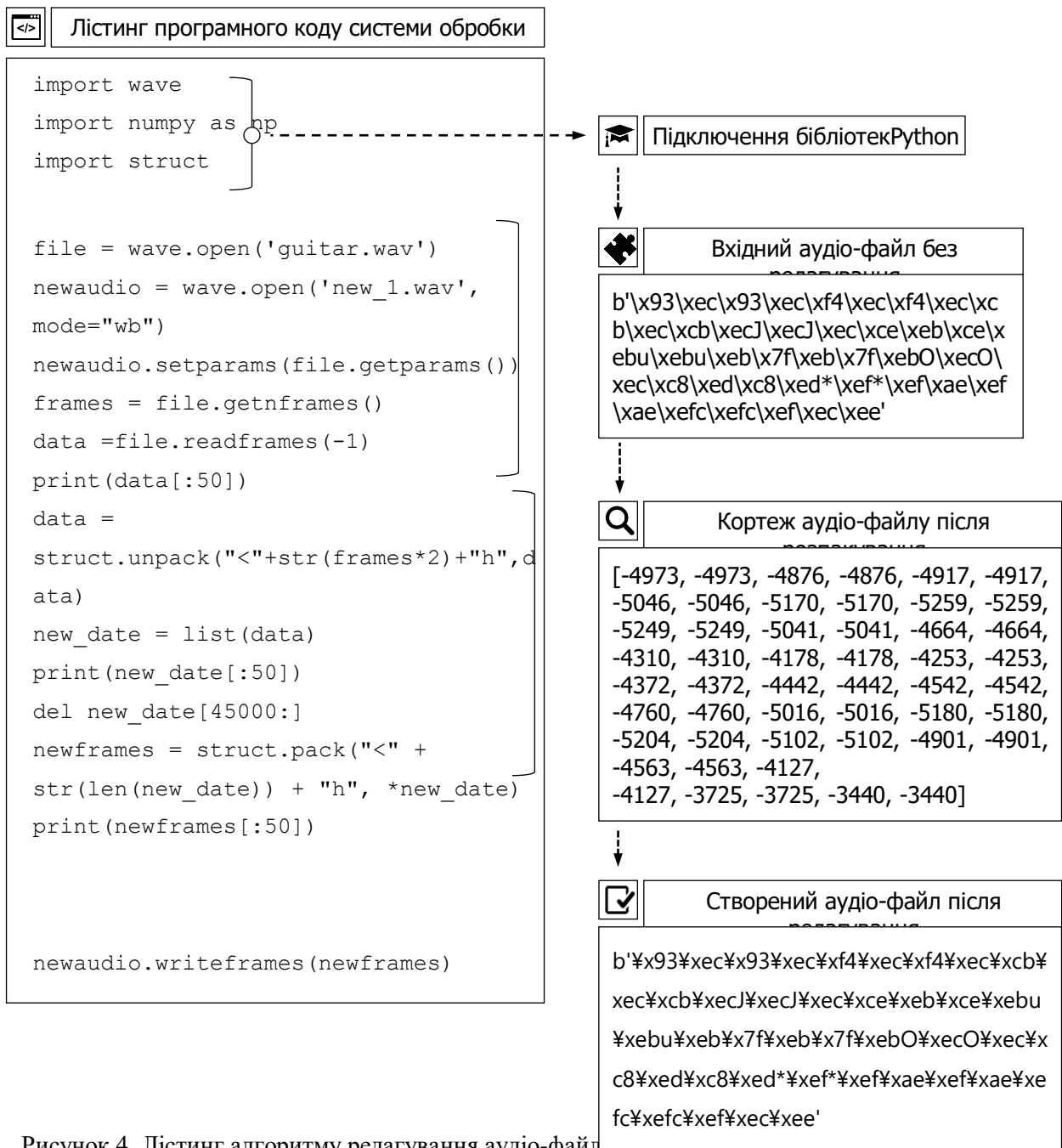


Рисунок 4. Лістинг алгоритму редагування аудіо-файлу з перетворення бінарних даних

У рамках даного дослідження було завантажено фрагмент аудіо-файлу з відкритої бази даних. На першому етапі було необхідно отримати блок даних аудіо-файлу у вигляді бінарних чисел. Для цього використовувалася команда `file=wave.open("")`, що відкриває аудіо-файл і команда `newaudio=wave.open("", mode="wb")` і створює новий аудіо-файл у відповідній директорії. У свою чергу, команда `newaudio.setparams(file.getparams())` переносить налаштування і характеристики із базового аудіо-файлу у новий. Командою `data=file.readframes(-1)` було отримано значення частоти в шістнадцятковій системі числення та виведено на екран перші 50 значень функцією `print(data[:50])`. За допомогою команди `data=struct.unpack("<"+str(frames*2)+"h",data)` було виконано розпакування значень частоти, які далі переводились у список, над якими згідно наявного інструментарію, можна виконувати широкий набір процедур. Для перегляду отриманого результату зі списку було видалено всі елементи після зазначеного функцією `del new_date[45000:]`. Після виконання відповідних маніпуляцій зі списками останнім етапом було заповнено дані у послідовний кортеж `newframes=struct.pack("<" + str(len(new_date)) + "h", *new_date)` та записано відповідний аудіо-файл `new_1.wav` за допомогою функції `newaudio.writeframes(newframes)`.

#### 4. Висновки та перспективи подальших досліджень

Проведено дослідження показало ефективність застосування інтерпретовано об'єктно орієнтованої мови програмування Python та бібліотек з відкритим кодом NumPy, SciPy і Matplotlib при генеруванні і обробці масивів аудіоданих, а також візуалізації процесу обробки і вдосконаленню системи аналізу вхідних даних. Представлена модель та алгоритми на її основі можуть бути використані при роботі з широким класом задач, зокрема **перспективи подальших досліджень** у даній галузі полягають у впровадженні компактного апаратно-програмного комплексу обробки аудіосигналів на базі мікрокомп'ютерів. Також, слід зазначити, що представлена схема системи цифрового аудіо-запису і обробки і аудіо-даних вказує на можливість масштабування та оптимізації систем роботи з аудіо-файлами у режимі реального часу. Запропонований підхід надає можливість формалізувати параметри підсистем та поставленого завдання на рівні аргументів цільових функцій ефективності і часу обробки аудіо-даних, з подальшим зведенням задачі оптимізації до математичної задачі пошуку екстремумів цільових функцій.

#### References.

1. BjörnSchuller. Intelligent audio analysis/ BjörnSchuller. — Berlin: Springer, 2013. — 345 p. — ISBN: 978-3-642-36805-09.
2. W. Jackson, "The Algorithms of Digital Audio: Audio Processing", Digital Audio Editing Fundamentals, vol. 1, pp. 55-64, 2015, doi:10.1007/978-1-4842-1648-4\_8.
3. W. Jackson, "The Data Footprint of Digital Audio: Compression", Digital Audio Editing Fundamentals, vol. 2, pp. 107-117, 2015, doi:10.1007/978-1-4842-1648-4\_12.
4. X. Wu, F. Tian and J. Liu, "Efficient spectrum multiplexing using wavelet packet modulation and channel estimation based on ANNs", International Conference on Audio, Language and Image Processing, NY, 2008, doi:10.1109/icalip.2008.4590189.
5. D. Riordan, J. Walsh, and P. Doody, "Modelling of the Human Perception of Sound using ANNs" 25th IET Irish Signals & Systems Conference 2014 and 2014 China-Ireland International Conference on Information and Communities Technologies (ISSC 2014/CICT 2014), 2014, doi:10.1049/cp.2014.0668.
6. Y.D.Pra, & F. Fontana, "Programming Real-Time Sound in Python", Applied Sciences, vol.10, no 12, 4214-4221, 2020, doi:10.3390/app10124214.
7. G. Schuller, "The Python Perceptual Audio Coder", Filter Banks and Audio Coding, vol. 3, pp. 155-160, 2020, doi:10.1007/978-3-030-51249-1\_6.
8. C. Howard and B. Cazzolato, "Sound Absorption in a Lined Duct", Acoustic Analyses Using Matlab and Ansys, vol 11, 360-407, 2014, doi:10.1201/b17825-12.
9. S.L. Campbell, J. Chancelier & R. Nikoukhah "Modeling and Simulation in Scilab", Modeling and Simulation in Scilab/Scicos with ScicosLab vol. 4, no 4, pp. 73-106, 2014, doi:10.1007/978-1-4419-5527-2\_3.
10. I. Rouno, "Fourier Sound Synthesis", Wolfram Demonstrations Project, vol. 3, 2007. doi:10.3840/001114.
11. John Hunter Darren Dale Eric Firing Michael Droettboom and the matplotlib development team, September 16, 2020. Режим доступу: <https://matplotlib.org/3.3.2/Matplotlib.pdf>, <https://matplotlib.org/3.3.2/contents.html>.
12. SciPy community, Scipy 1.5.3 Reference Guide, October 17, 2020. Режим доступу: <https://docs.scipy.org/doc/scipy-1.5.3/scipy-ref-1.5.3.pdf>
13. R. Smallshire & A. Bingham. Python /R. Smallshire. — Culemborg: Van Duuren Media, 2019. — 352 p., ISBN-10 : 1788293185.