第二十二讲
I^2S 协议
Lecture 22
I^2S Protocol

明玉瑞 Yurui Ming

yrming@gmail.com

声明
Disclaimer

- 本讲义在准备过程中由于时间所限，所用材料来源并未规范标示引用来源。所引材料仅用于教学所用，作者无意侵犯原著者之知识版权，所引材料之知识版权均归原著者所有；若原著者介意之，请联系作者更正及删除。

  The time limit for the preparation of these slides incurs the situation that not all the sources of the used materials (texts or images) are properly referenced or clearly manifested. However, all materials in these slides are solely for teaching and the author is with no intention to infringe the copyright bestowed on the original authors or manufacturers. All credits go to corresponding IP holders. Please address the author for remedy including deletion have you had any concern.

- 本讲特别参考了如下博客内容：

  This lecture specifically refers to the following blog:

  https://dronebotworkshop.com/esp32-i2s/

基础
Basics

- 在物联网的某些应用中，特别是与用户交互的场合，可能会涉及音频的应用。与一般的基于文字的信息交互相比，音频应用可能需要较高的带宽，因此在本讲，我们介绍一种新的外设接口协议，I^2S。

  In some IoT application, audio might be utilized especially in the scenarios which involves users’ interactions. Compared with the cases of text-based communication between human and machines, audio tends to require higher bandwidth, whether for capturing or playback. In this lecture, we introduce another peripheral or bus interface, aka., I^2S protocol.

- I^2S是集成电路间声音协议的英文缩写，由飞利浦半导体于1986年开发。该协议是为了处理数据和声音信息的集成电路之间的兼容性需求而设计的。同期飞利浦还开发了I^2C协议，其都是为了满足类似的需求而构建的。

  I^2S is the abbreviation of the Inter-Integrated Circuit Sound Protocol, which was developed by Phillips Semiconductors in 1986, in order to addressed the need for compatibility between integrated circuits that handled data and sound information. Phillips also developed the I^2C protocol around the period, and both protocols were built to serve similar needs.

基础
Basics

- 为了讲解I²S，下面介绍数字音频的工作原理。众所周知，声音本质上是模拟的，并且在数字声音发展之前，音频设备也是模拟的。尽管相关模拟电子设计往往经过优化，但即使是最好的模拟电子设备也会在信号中产生电噪声和失真。

  In order to understand I²S, in the following we first recap how digital audio works. It is well understood that sound by its own nature is analogue, and, prior to the development of digital sound, audio equipment was also analogue. Although the designs of relevant electronics are impressively optimized, even the best analogue electronics will induce electrical noise and distortion in the signal.

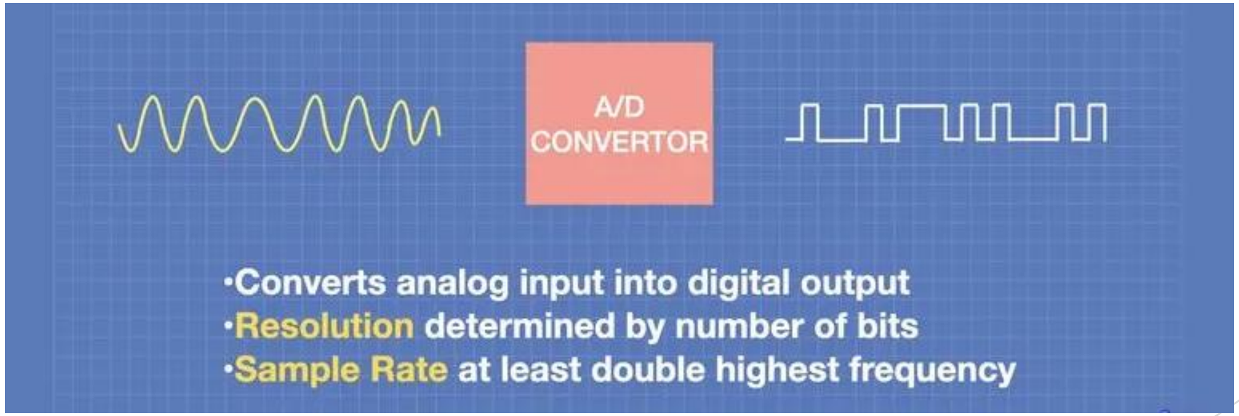
- 尽管麦克风、扬声器等设备似乎必须以模拟的方式进行工作，但作为换能设备，从动能到电能、或电能到动能，并未限制电信号以模拟或数字方式呈现。而数字音频被认为是消除声音失真的一种更好的方法。

  Although literally microphone and speaker seem to be working in the analog way, however, as transducers, which convert momentum energy (vibration) to electric energy or vice versa, put no explicit restrictions that the electric signal must be in digital form or analog form. But digital audio is touted as a way to eliminate those distortions of the sound.

基础
Basics

- 通常对于数字音频，声音被采集后，首先发送到模数转换器以转化为数字表示形式，然后处理、存储或传输而不会出现任何降级。对于输出，数字信号通过等效的数模转换器，重新创建模拟输入，然后放大以驱动扬声器。

  Generally, with digital audio, the sound is captured and sent to an analogue to digital converter (ADC) to create a digital representation of it. Then it can be processed, stored or transmitted without any degradation. For output, the digital signal goes through an equivalent digital to analogue converter (DAC), which recreates the analogue input and is then amplified to drive a speaker.



基础
Basics

- 通常，数字音频信号并不是原始信号的完美表示，信号质量取决于适用于ADC和DAC的两个参数：
  - 分辨率—量化样本使用的位数，更多位等于更好的质量。
  - 采样率—每秒采集多少个样本，通常至少是想要采样的最高频率的两倍。

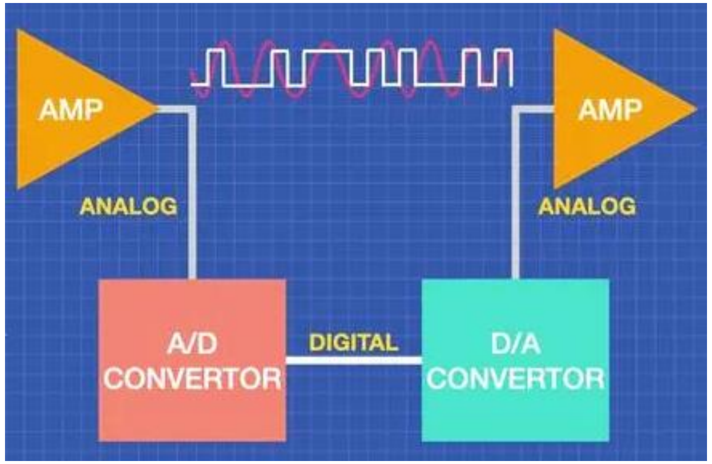
  A digital audio signal is not, of course, a perfect representation of the original signal. The quality of the signal is dependent upon two parameters that apply to both the ADC and DAC:
  - Resolution – The number of bits used in quantizing the sample. More bits equal better quality.
  - Sample Rate – How many samples per second are we taking. This needs to be at least twice as high as the highest frequency we want to sample.
- CD品质的数字音频的分辨率为16位，采样率为44.1kHz，而电话品质的数字音频为8位，采样率为8kHz。通常，分辨率和采样率越高，生成的数字数据文件就越大，传输需要的带宽就越高。

  CD-quality digital audio has a resolution of 16-bits and a sample rate of 44.1 kHz, whereas telephone-quality digital audio is 8-bits and is sampled at 8 kHz. Generally, the higher the resolution and sample rate, the larger the resulting digital data file will be and the higher band width requires to transmit the data.

基础
Basics

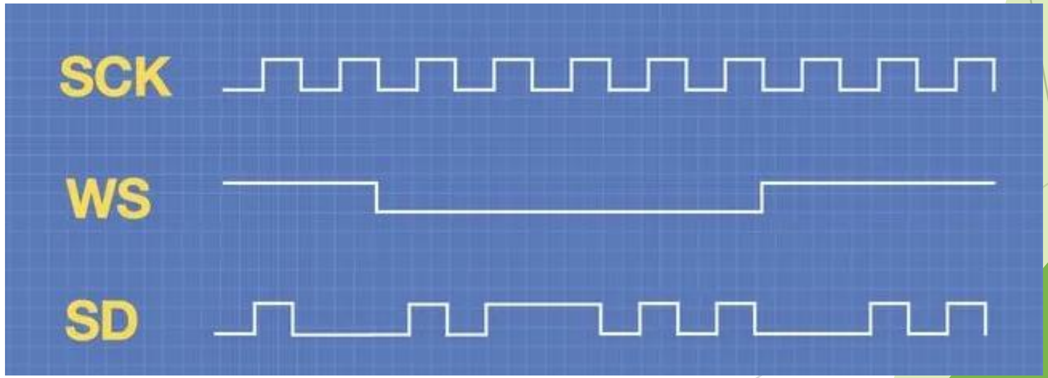
- 当数字音频在集成电路之间传输时，都是以串行格式完成的，常见的一种是脉冲编码调制。I²S可使用任意分辨率和采样率处理数字PCM数据，通过各种信号处理器和均衡器将声音从源设备传递到目的设备。

When digital audio is transmitted between integrated circuits, it is done in a serial format. A common one is the Pulse Code Modulation or PCM. I²S works with digital PCM data, using any resolution and sample rate. It can be used to transmit the sound from source device to destination device, via various signal processors and equalizers in some cases.



I^{2}S协议
I^{2}S Protocol

- I^{2}S协议管理总线上的PCM数据，该总线至少由以下三个连线组成：
  - SCK – 串行时钟线，有时称为“位时钟线”。
  - WS – 字选择，在左音频通道和右音频通道之间进行选择。
  - SD – 串行数据，PCM音频数据。
- The I^{2}S protocol manages PCM data on a bus that consists of at least the following three connection lines:
  - SCK – The Serial Clock Line, sometimes referred to as the “bit clock line”.
  - WS – Word Select, which selects between the Left and Right audio channels.
  - SD – Serial Data, the PCM audio data.



I^2S协议
I^2S Protocol

- 音频信号的质量决定了串行时钟速率，由以下公式确定：

  $$时钟频率 = 采样率 \times 每通道位数 \times 通道数$$

  例如，对于以16 bit采样的双通道MP3音乐（采样频率通常为44.1 kHz），则时钟频率为：

  $$44.1\ \text{kHz} \times 16\ \text{bits} \times 2\ \text{chs} = 1.4112\ \text{MHz}$$

- The quality of the audio signal determines the Serial Clock rate, and is determined with the following formula:

  $$\text{Clock Frequency} = \text{Sample Rate} \times \text{Bits Per Channel} \times \text{Number of Channels}$$

- For example, the music in MP3 format (usually sampled at 44.1 kHz) with 2 channels and quantizing at 16 bits requires the following clock frequency to transmit between modules:

  $$44.1\ \text{kHz} \times 16\ \text{bits} \times 2\ \text{chs} = 1.4112\ \text{MHz}$$

I^2S 设备
I^2S Device

- 连接到I^2S总线的设备可以分为两类：
  - 控制器– 控制SCK和WS信号。
  - 目标设备– 接收SCK和WS信号。

  总线上只能有一个控制器，但是总线上可以有多个目标设备。

- 另外一个概念是音频设备系统，其可以分为三类：
  - 发射器– 发送音频信号。
  - 接收器– 接收音频信号。
  - 控制器– 控制音频信号。

  对音频设备，至少需要一个发射器和接收器，控制器是可选的。

- The devices connected to an I^2S bus can be divided into two categories:
  - Controller – Controls the SCK and WS signals.
  - Target – Receives the SCK and WS signals.

  There can only be one controller on the bus, however, the bus can have multiple targets.

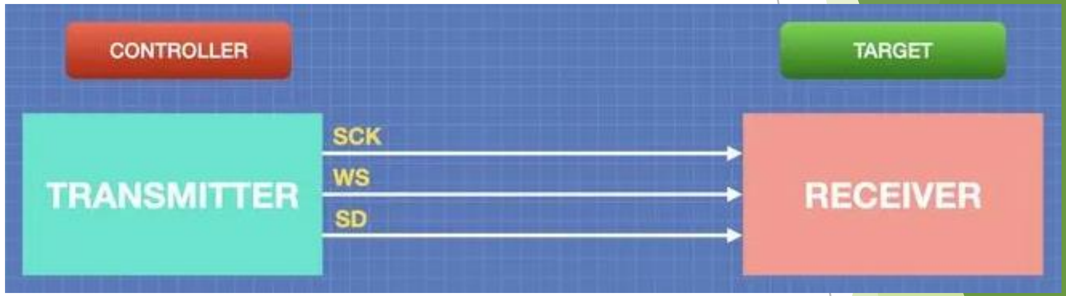
- Another concept is audio devices (or system), they can be divided into three categories:
  - Transmitters – Send audio signals.
  - Receivers – Receive audio signals.
  - Controllers – Control the audio signals

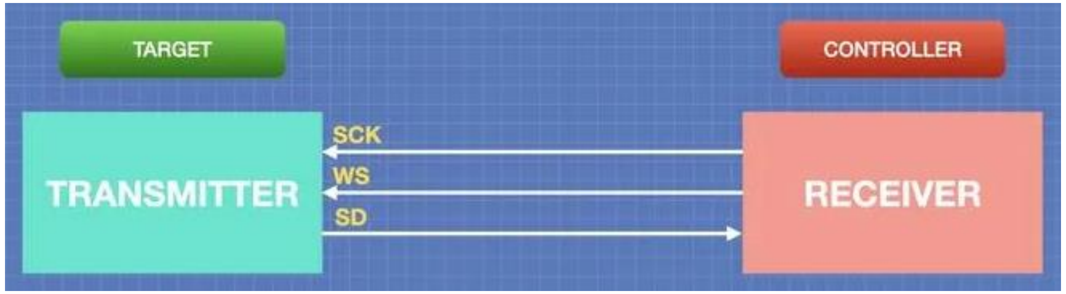
  For audio device, we at least need a Transmitter and Receiver, the Controller is optional.

I²S 设备
I²S Device

- 右图展示了一个简单的拓扑，在此布局中，音频发射器也是I²S控制器，它提供SCK、WS和SD信号；接收器同时也是I²S目标设备。

  A simple topology is illustrated as in the right figure. In this layout, the Audio Transmitter is also the I²S Controller, and it provides the SCK, WS, and SD signals. The Receiver is also the I²S Target.





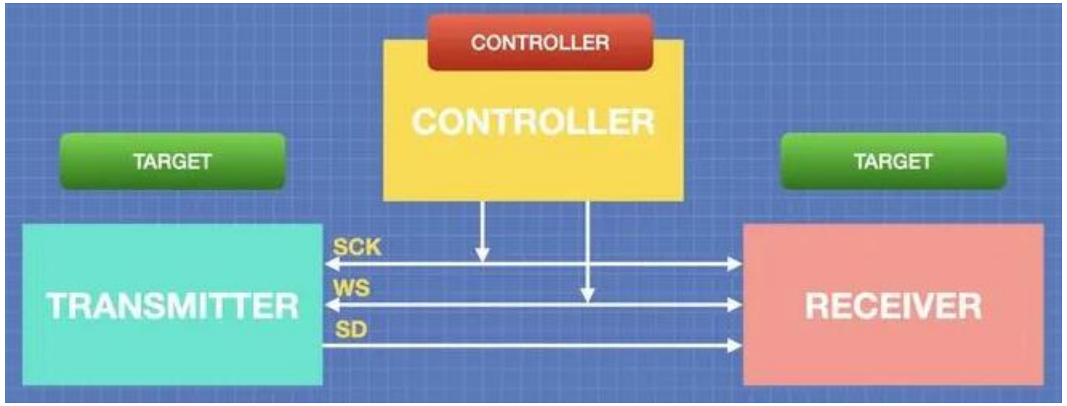
- 虽然上图是一种非常标准的布局，但它不一定是唯一的。左图展示了另一种布局，此时I²S角色颠倒过来，音频接收器充当I²S控制器，提供SCK和WS信号；作为目标设备的音频发射器仍然提供SD（串行数据）。

  While the top-right figure is an outstanding arrangement, it doesn't have to be the only one, the bottom-left layout is just as valid: Here, the I²S roles are reversed, with the audio Receiver also acting as I²S Controller, providing the SCK and WS signals. The audio Transmitter, which is the Target in this arrangement, still provides the SD (Serial Data).

I^2S 设备
I^2S Device

- 包含独立控制器的拓扑如下所示。此时，发射器和接收器都是I^2S的目标设备，即独立的I^2S控制器向两个目标设备提供SCK和WS信息；而发射器提供SD信号。

  A topology which contains a standalone controller is illustrated as below. The Transmitter and Receiver both being I^2S Targets, aka., the separate I^2S controller, and it provides SCK and WS to both Targets. Once again, the Transmitter supplies the SD signal.



- 总之，控制器与目标设备是从总线角度来说；发射器、接收器是从设备功能角度来说的。

  To sum it up, controller and target are categorized from the bus perspective, while transmitter and receiver are from the device functional perspective.

I^2S与ESP32
I^2S and ESP32

- ESP32有两个I^2S外设：I^2S0和I^2S1。每一个都可以配置为控制器或目标设备，每一个都可以作为音频发射器或接收器。每个I^2S控制器都可以在半双工通信模式下运行。因此，两个控制器可以组合起来建立全双工通信。

  The ESP32 has two I^2S peripherals, I^2S0, and I^2S1. Each one can be configured as a Controller or Target, and each one can be an audio Transmitter or Receiver. Each I^2S controller can operate in half-duplex communication mode. Thus, the two controllers can be combined to establish full-duplex communication.

- ESP32支持DMA（直接内存访问）模式，该模式允许传输样本数据，而不需要CPU复制每个数据样本，并且在传输高质量音频时非常有用。

  ESP32 also supports the DMA (Direct Memory Access) mode, this mode allows for streaming sample data without requiring the CPU to copy each data sample, and can be useful when streaming high-quality audio.

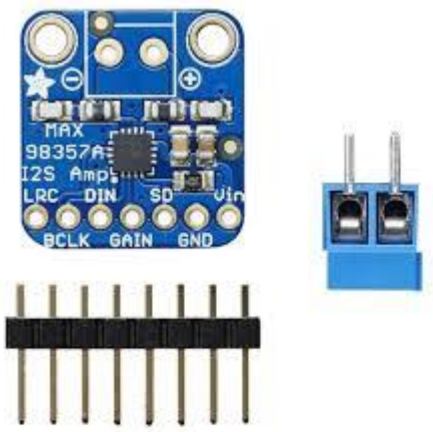
I^2S与ESP32
I^2S and ESP32

- 此外，ESP32允许I^2S0的输出在内部路由到ESP32 DAC的输入，以产生直接模拟输出，而不涉及任何外部I^2S编解码器。
- I^2S外设还支持一种称为“LCD模式”的高级模式，用于通过并行总线进行数据通信。某些LCD和相机模块使用此功能。LCD模式可以在以下模式下运行：
  - LCD主发射方式
  - 相机从机接收方式
  - ADC/DAC模式

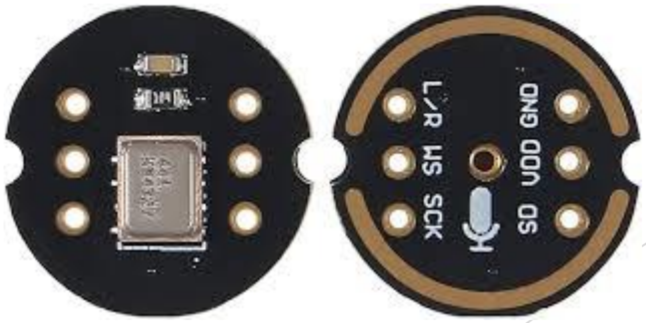
- In addition, ESP32 allows the output of I^2S0 to be internally routed to the input of the ESP32 DAC to produce direct analogue output without involving any external I^2S codecs.
- The I^2S peripherals also support an advanced mode called “LCD mode” for communicating data over a parallel bus. This is used by some LCDs and camera modules. LCD mode can be operated in the following modes:
  - LCD master transmitting mode
  - Camera slave receiving mode
  - ADC/DAC mode

I^{2}S外设
I^{2}S Peripheral

- 支持I^{2}S标准的外围设备范围从小型放大器和麦克风模块到具有 I^{2}S 连接的完整音频系统。
- 我们常用的I^{2}S外设如：
  - INMP441麦克风模块
  - MAX98357A I^{2}S放大器模块
- 下图展示了这些模块的外观。



- Peripherals supporting I^{2}S standard range from small amplifier and microphone modules to complete audio systems with I^{2}S connectivity.
- Commonly-used I^{2}S peripherals include:
  - INMP441 Microphone Module
  - MAX98357A I^{2}S Amplifier Module
- The following figures demonstrates the appearances of these modules:



MP3播放器
MP3 Player

- MAX98357A放大器模块是一款价格低廉但功能强大的音频放大器模块，其具有I²S输入，可向4欧姆负载输出高达3瓦的功率。由于音圈实际上是模块输出滤波电路的一部分，因此，必须使用音圈为4到8欧姆的扬声器。

  The MAX98357A I2S amplifier module is an inexpensive yet powerful audio amplifier module with an I2S input. The device can output up to 3 watts into a 4-ohm load. As the voice coil is actually part of the output filtering circuit in the module, you must use a speaker with a 4 to 8-ohm voice coils.

- 该器件具有如下表所示的引脚排列：

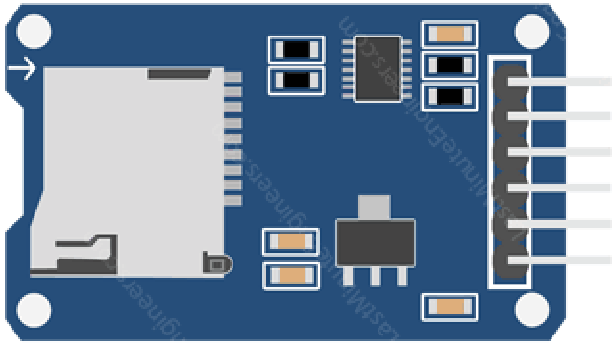
  The device has the following pinout as in the figure below:

Name（名称）	Function（功能）
LRC（左右声道）	I²S WS （选择线）
BCLK（时钟）	I²S SCK （时钟线）
DIN（数据输入）	I²S SD （数据线）
GAIN（增益）	Use Register to set value （通过寄存器设置值）
SD	L - R Output Select （左右声道输出选择）
GND	Ground（地）
Vin	3 - 6 VDC （电源输入）

MP3播放器
MP3 Player

- 由于我们需要存储音乐，因此需要额外的SD卡槽模块。在下图中，我们给出了常用的SD卡槽模块的图示，并在右面列出了相关管脚功能。

  Due to the need to store the music file, we need an extra SD card adaptor module. The following figure shows a commonly-used one and the right table shows the name and functionality of each pin.



Name（名称）	Function（功能）	Mapping（映射）
VCC	pin provides power to the module and recommended to be connected to the 5V pin in case present of voltage regulator（模块电源管脚，当有电压调节模块时，建议接核心板5V输出）	ESP32 VCC
GND	ground pin（地线）	ESP32 GND
MISO	SPI output from the microSD card module（模块输出）	ESP32 GPIO13
MOSI	SPI input to the microSD card module（模块输入）（通过寄存器设置值）	ESP32 GPIO11
SCK	pin accepts clock pulses from the master (an Arduino in our case) to synchronize data transmission（时钟线）	ESP32 GPIO12
CS	control pin that is used to select one (or a set) of slave devices on the SPI bus（片选管脚）	ESP32 GPIO10

MP3播放器
MP3 Player

- MAX98357A支持通过合适的方式配置不同的增益。放大器的增益可配置为低至+3dB 至高至+15dB。默认情况下该板配置为+9dB，读者可以按照右表，使用增益引脚从外部控制，更改增益。

  MAX98357A supports configurable gain which can also be modified in corresponding ways. The gain of the amplifier can be configured from as low as +3dB to as high as +15dB. While the board is configured for +9dB by default, users can change the gain by referring to the right table via controlling the gain pin externally.

Gain (dB)（增益）	Gain Pin Connection（增益管脚连接）
15	Connected to GND through a 100kΩ resistor.（串联100kΩ电阻后接地）
12	Connected to GND（直连到地）
9	Unconnected (Default)（空接）
6	Connected to VDD（直连到电源）
3	Connected to VDD through a 100kΩ resistor（串联100kΩ电阻后接电源）

MP3播放器
MP3 Player

- 针对ESP32 S3，其有内部有两个I²S接口，且这两个接口可被映射到任意闲置的GPIO接口。在本实验中，我们和ESP32的连接方式如右图所示：

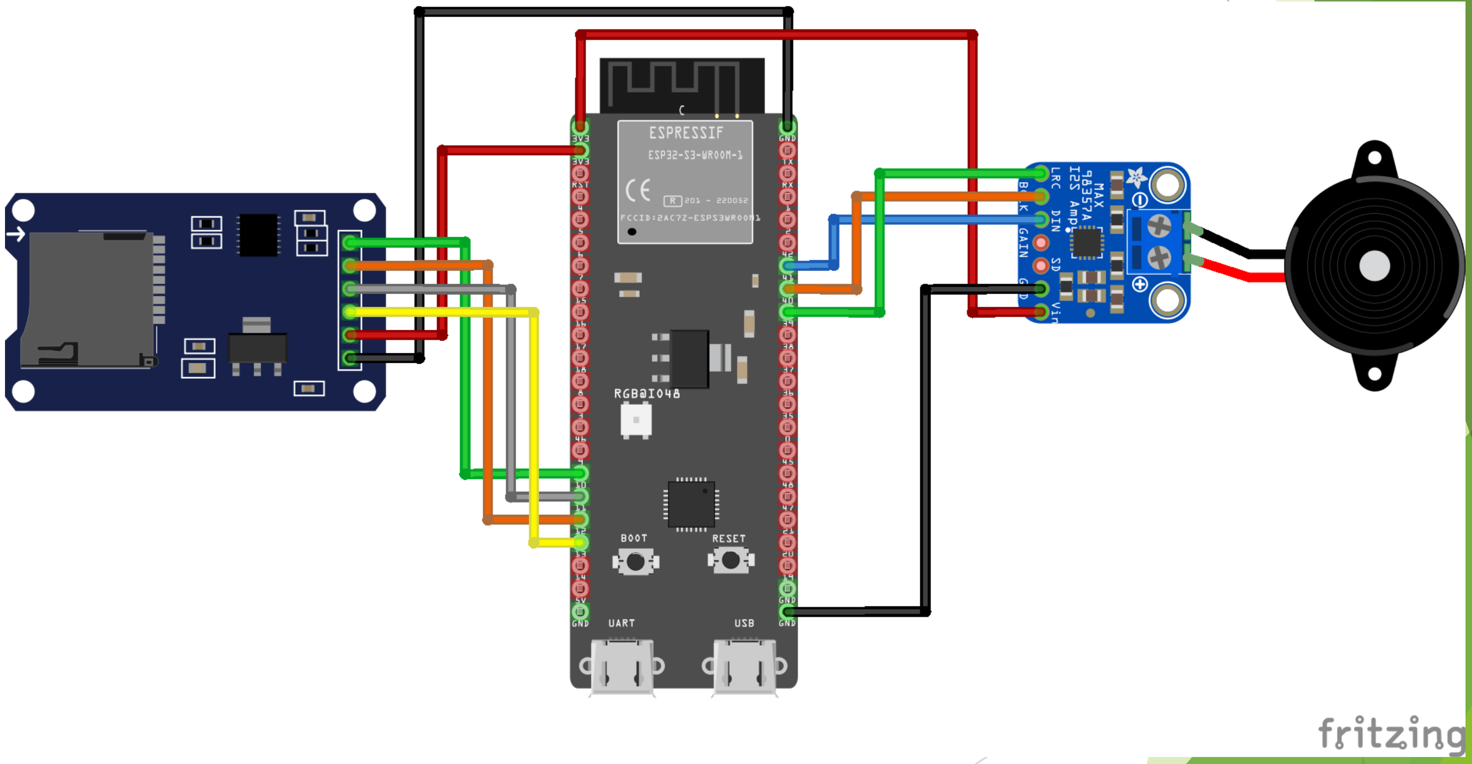
  The ESP32 S3 has two I²S interfaces which can be mapped to any available gpio pins. For this experiment, the connections that need to be made to the ESP32 are shown as the hook table in the right:

ESP32 S3 Pin（ESP32 S3管脚）	I2S Audio Breakout Pin
VUSB/3V3	VDD
GND	GND
GPIO 41	DIN
GPIO 40	BCLK
GPIO 39	LRCLK

MP3播放器
MP3 Player

- 下面是基于Fritzing设计的MP3播放器的连线，读者可以根据自己的实际情况选择合适的GPIO引脚：

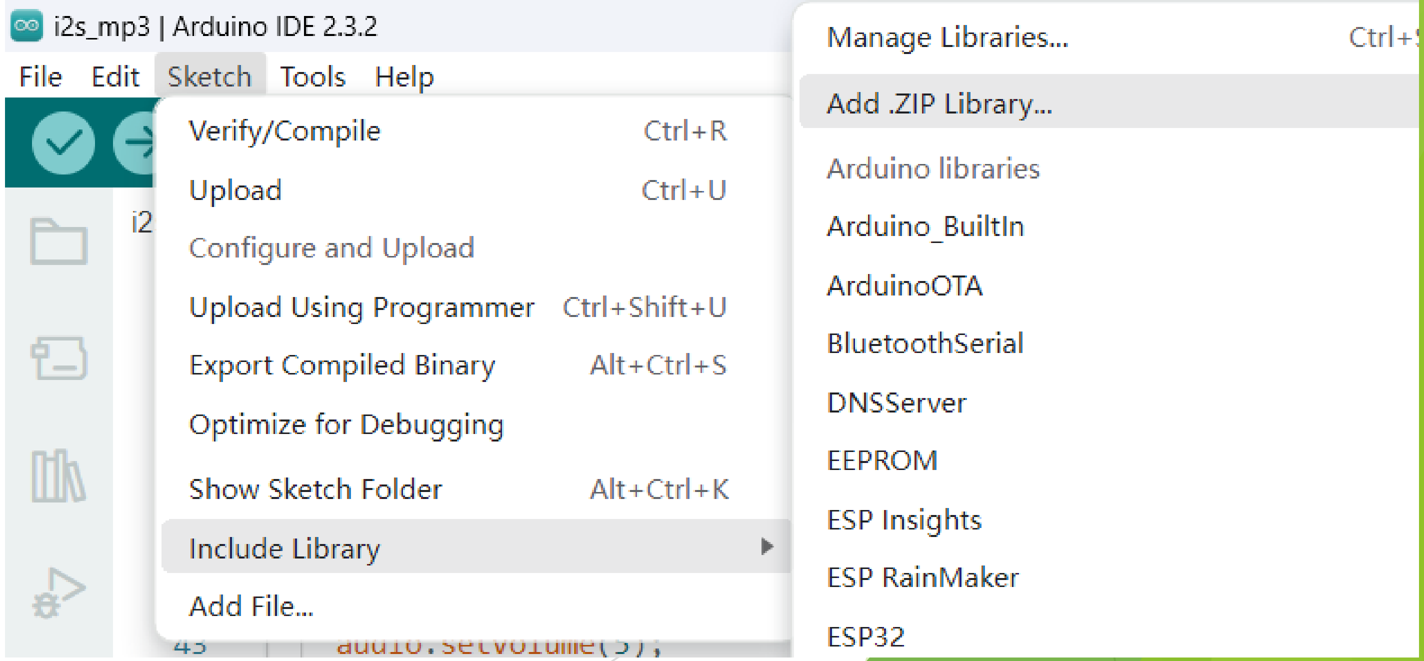
  The following is the hookup diagram for our MP3 player based on Fritzing, and users can adjust the GPIO numbers according to their own situation:



MP3播放器
MP3 Player

- 由于我们完全实现一个解码MP3格式的音乐文件是比较困难的事情，因此推荐基于已有的工作来完成音乐文件的播放，这里我们将使用ESP32-AudioI2S库。但由于该库并不有做成适配Arduino标准安装的形式，需要通过从Github上通过ZIP文件下载下来之后，通过添加Arduino菜单添加.zip文件的方式进行安装。

Due to the difficulty to implement an algorithm in decoding the music file in MP3 format, it is recommended to build our project based on existing libraries. Here we utilize the ESP32-AudioI2S library. However, the implementation of AudioI2S doesn’t conform to the package installation requirement for Arduino, so you need to download it from Github by choosing the zip file, and manually install it via submenu called “Add .ZIP Library …” from the Arduino menu.



MP3播放器
MP3 Player

- 由于涉及到播放，在一定程度上会涉及放大器与扬声器的阻抗匹配问题。阻抗通常用于测量交流电路中的电阻，其以欧姆(Ω)为单位。大多数扬声器的阻抗额定值在4到8欧姆之间，但有些型号的阻抗可低至2欧姆或高达16欧姆。

  Matching the amplifier to the speakers regarding impedance is unavoidable to playback music. Impedance, measured in ohms (Ω), measures electrical resistance in an AC circuit. Most speakers have impedance ratings that range between 4 and 8 ohms, but some models can go as low as 2 ohms or as high as 16 ohms.

- 扬声器阻抗会对音质产生明显影响。阻抗较低的扬声器可以从放大器中吸收更多功率，从而可能产生更高的音量或更大的动态范围。但是，阻抗过低可能会使放大器承受压力，从而可能导致失真甚至损坏。阻抗较高的扬声器可能会发出较小的音量。但它对放大器的压力也较小，从而可以产生更干净、无失真的声音。

  Speaker impedance can have a noticeable impact on sound quality. A speaker with a lower impedance can draw more power from an amplifier, which might translate to higher volume levels or a greater dynamic range. However, too low an impedance may strain an amplifier, potentially leading to distortion or even damage. Conversely, a speaker with a higher impedance might deliver less volume for a given amplifier setting. But it also puts less strain on the amplifier, which could result in cleaner, distortion-free sound.

MP3播放器
MP3 Player

- 扬声器有多种类型，最常见的动圈式扬声器由一个连接到线圈的圆锥体组成。该线圈悬浮在磁场中，流过线圈的音频信号引起的电流变化会导致线圈移动，从而导致锥体移动。这导致将电音频信号转换成声音。而静场扬声器类型采用与动圈式扬声器类型完全不同的原理，其是一种利用施加在悬浮在静电场中的膜上的力来产生声音的扬声器。

  There are several different loudspeaker types. The moving coil type of loudspeaker is the type that is most commonly seen. It consists of a cone attached to a coil that is held within a magnetic field. The coil is suspended within a magnetic field and this means that the variations in current flow resulting from the electrical audio signal cause the coil, and hence the cone to move. This results in the loudspeaker converting the electrical audio signal into sound. The electrostatic loudspeaker type is uses a totally different principle to that of the moving coil and horn loudspeaker types. Instead the electrostatic loudspeaker is one in which sound is generated by utilising the force exerted on a membrane suspended in an electrostatic field.

- 如前所述，由于在这里扬声器的线圈会组成输出的滤波电路，因此这里要用到动圈式扬声器。

  As aforementioned, due to the fact that the coil which constitutes the speaker will be part of the filtering circuit for output, here we must use the moving coil ones.

MP3播放器
MP3 Player

- 当成功导入库之后，我们可以利用库，进行MP3文件的播放。通常的处理流程如右所示：

  After we import the library, we can play our MP3 file according to 自 the following procedures:

MP3播放器
MP3 Player

- 右图是我们构建的MP3播放器的实物图：

  The figure on the right shows the of layout of the MP3 player we wired: