

【Audio】I2S 传输 PCM 音频数据分析总结（一）

1.前言

《I2S 传输 PCM 音频数据分析总结》主要是介绍 PCM 音频数据是如何通过 I2S 进行传输的，主要是分成 2 个部分：

PCM 音频数据介绍

I2S 协议介绍

本篇文章主要是以 PCM 音频数据来介绍。

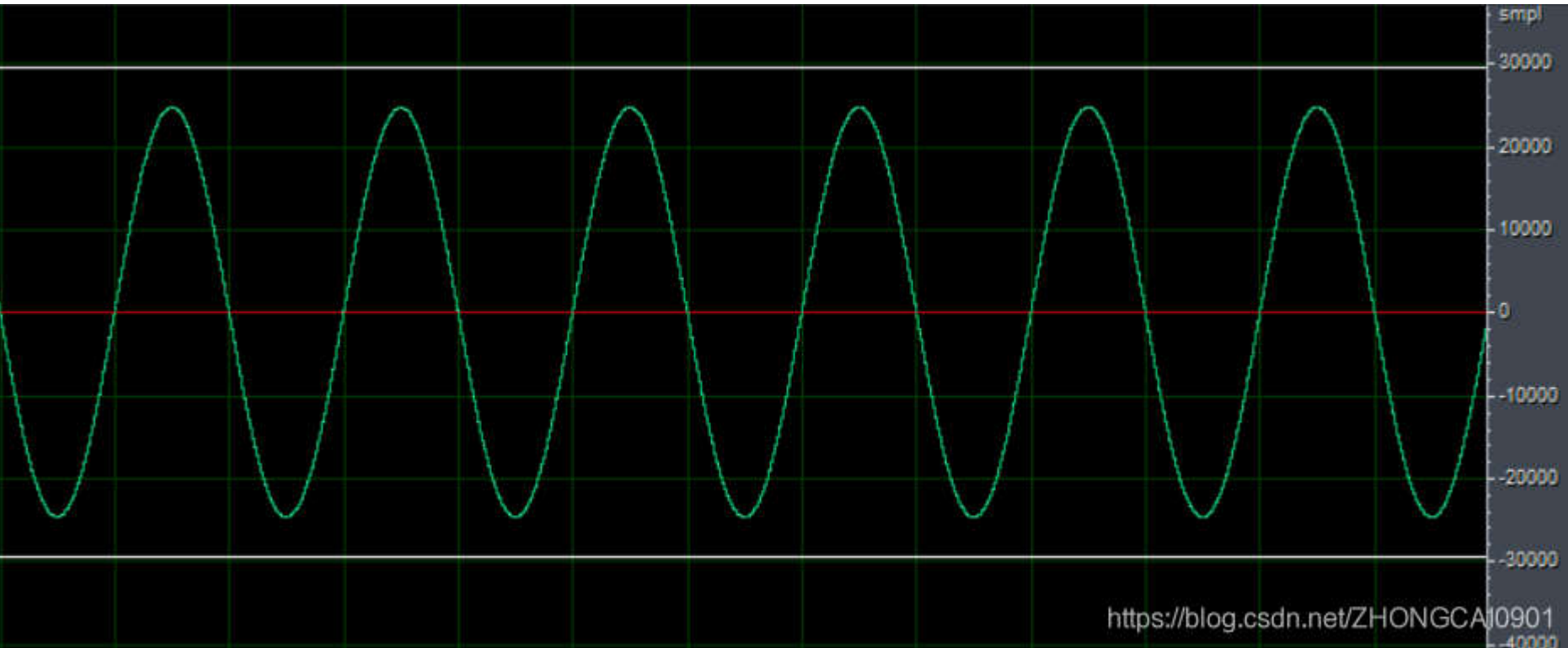
2.PCM 音频简介

PCM 音频（Pulse-Code Modulation）是一种用数字表示采样模拟信号的方法。

要将一段音频模拟信号转换为数字表示，包含如下三个步骤：

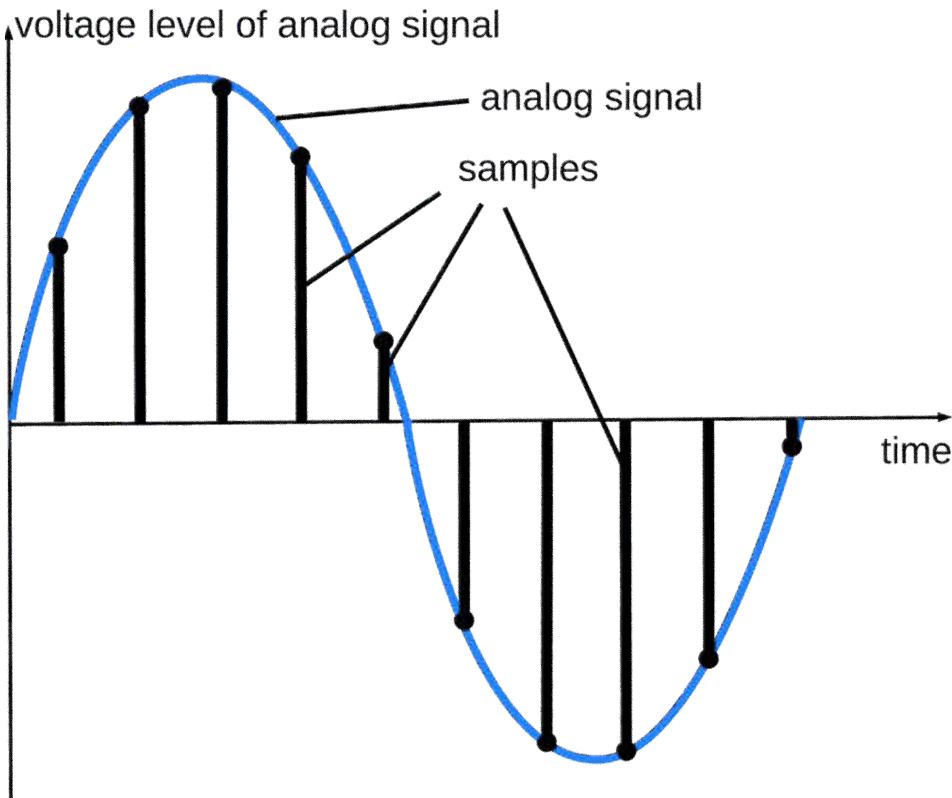
Sampling(采样)、Quantization(量化)、Coding(编码)

下面的截图是 PCM 音频 1KHz 正弦波的波形，播放出来就是“滴”的声音：



3. Sampling（采样）

Sampling(采样)表示处理，实际上就是让采样数据能够完全表示原始信号，且采样数据能够通过重构还原成原始信号的过程。



蓝色曲线：表示音频原始信号。

黑色垂直线段：表示当前时间点对原始信号的一次采样

黑色点：表示采样获取的一系列离散样本。

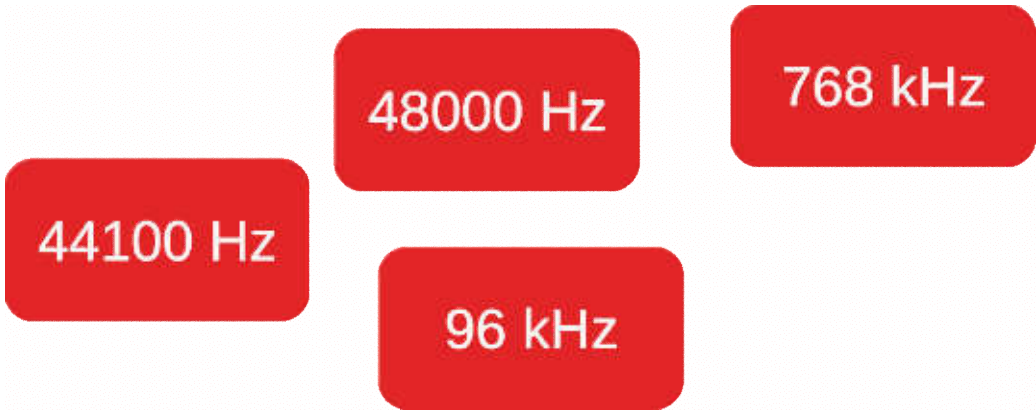
4. 采样率(Sample rate)

采样率（Sample rate）表示每秒采样的数量（单位：Hz）。

对于 PCM 编码音频流，采样率是恒定的。例如：1Hz 表示每秒钟对原始信号采样一次，1KHz 表示每秒钟采样 1000 次。

根据场景的不同，采样率也有所不同，采样率越高，声音的还原程度越高，质量就越好，同时占用空间会变大。

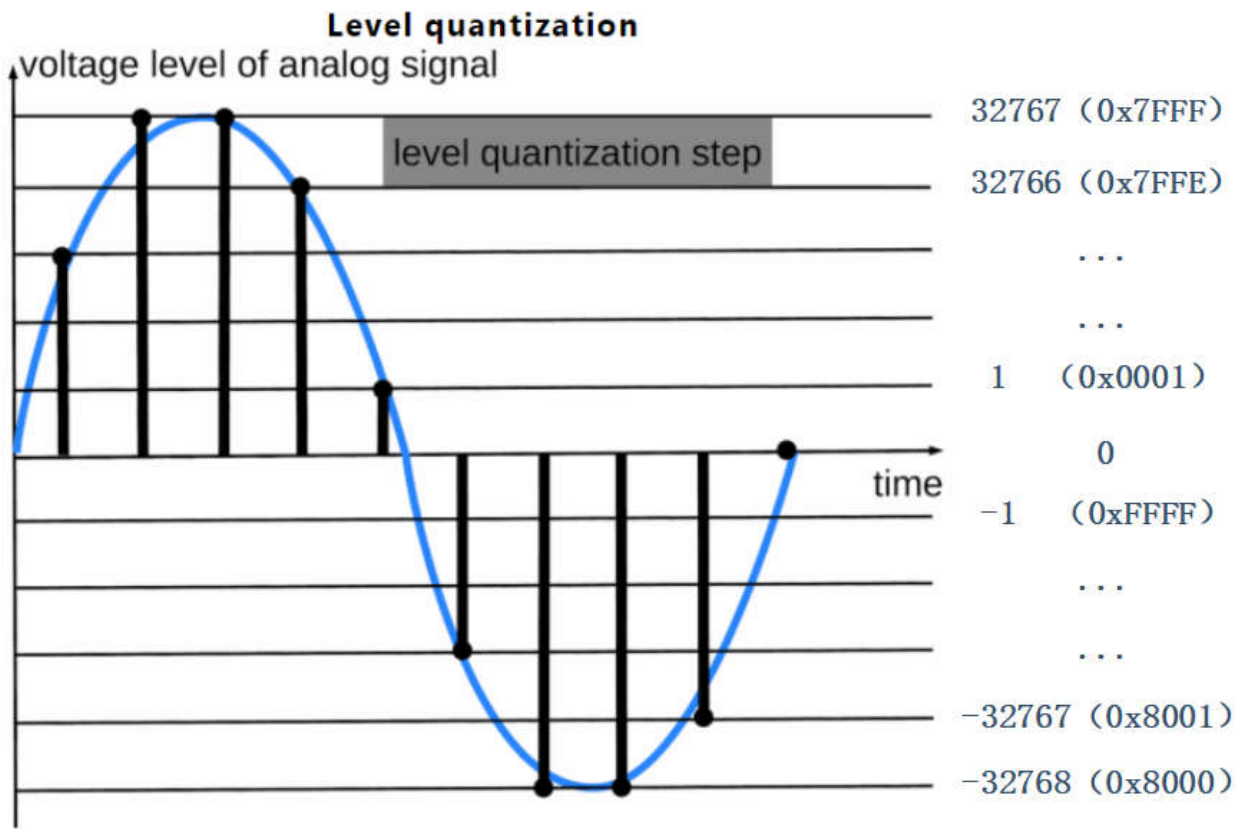
例如：通话时的采样率为 8KHz，常用的媒体采样率有 44KHz、48KHz，对于一些蓝光影片采样率高达 1MHz。



5. Quantization（量化）

原始信号采样后，需要通过量化来描述采样数据的大小。

量化处理过程，就是将时间连续的信号，处理成时间离散的信号，并用实数表示。这些实数将被转换为二进制数用于模拟信号的存储和传输。



上面的图片的例子是 16bit 位深量化过程。bit-depth（位深）：用来描述存储数字信号值的 bit 数，较常用的模拟信号位深如下：

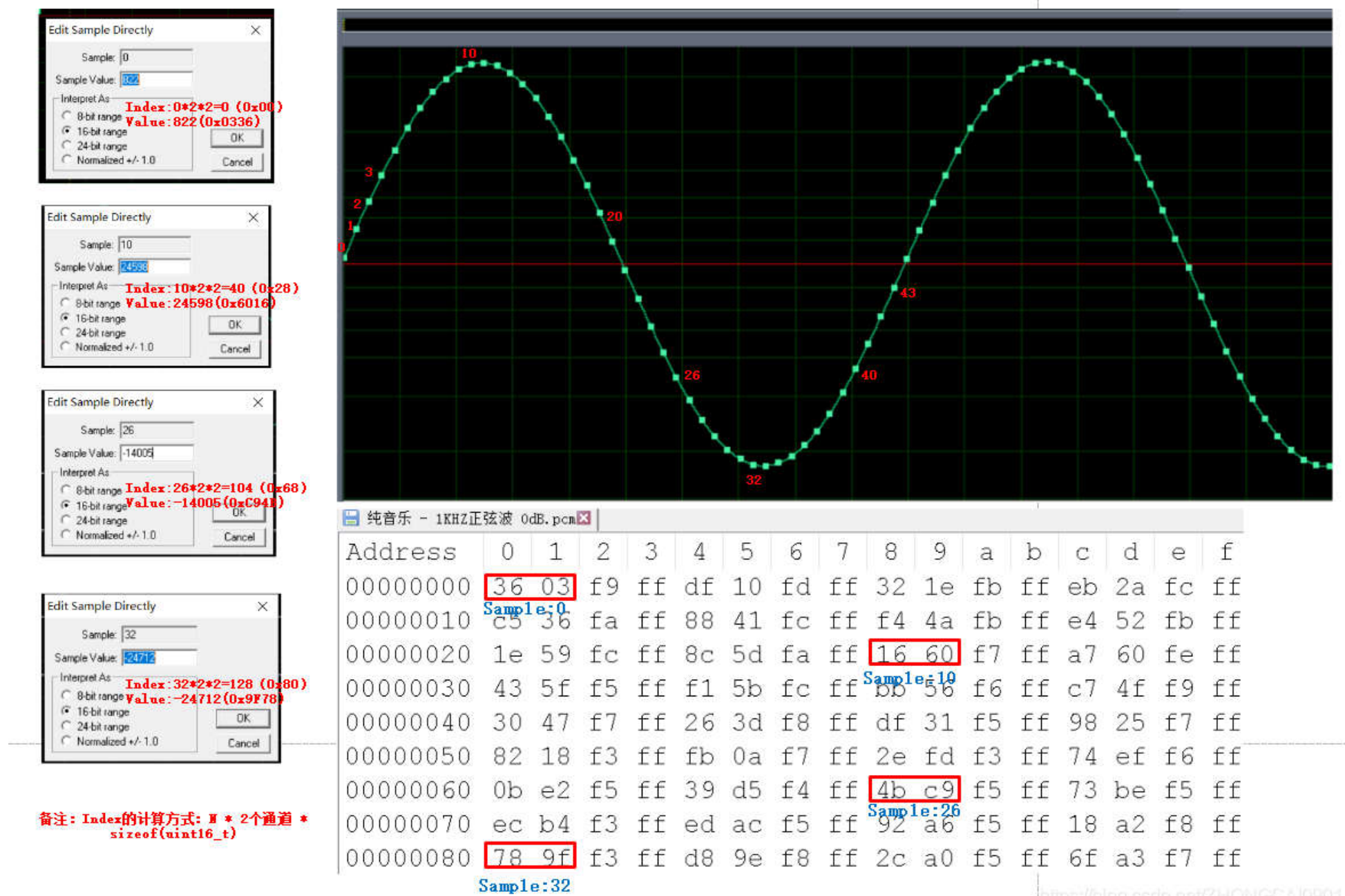
8-bit: $2^8 = 256$ levels，有 256 个等级可以用于衡量真实的模拟信号。

16-bit: $2^{16} = 65,536$ levels，有 65,536 个等级可以用于衡量真实的模拟信号。

24-bit: $2^{24} = 16,666,216$ levels，有 16,666,216 个等级可以用于衡量真实的模拟信号。

可以看到，位深越大，对模拟信号的描述将越真实，对声音的描述更加准确。

下面是实际的采样率 44.1KHz 16bit 双声道正弦波 1KHz 的 PCM 音频数据，量化后存储的数据和波形解析如下：



??? 问题: 为什么一个正弦波周期量化是 44 个数据???

分析:

1 次采样需要的时间: 1 秒/采样率 44.1KHz = 1 / 44100 (秒)

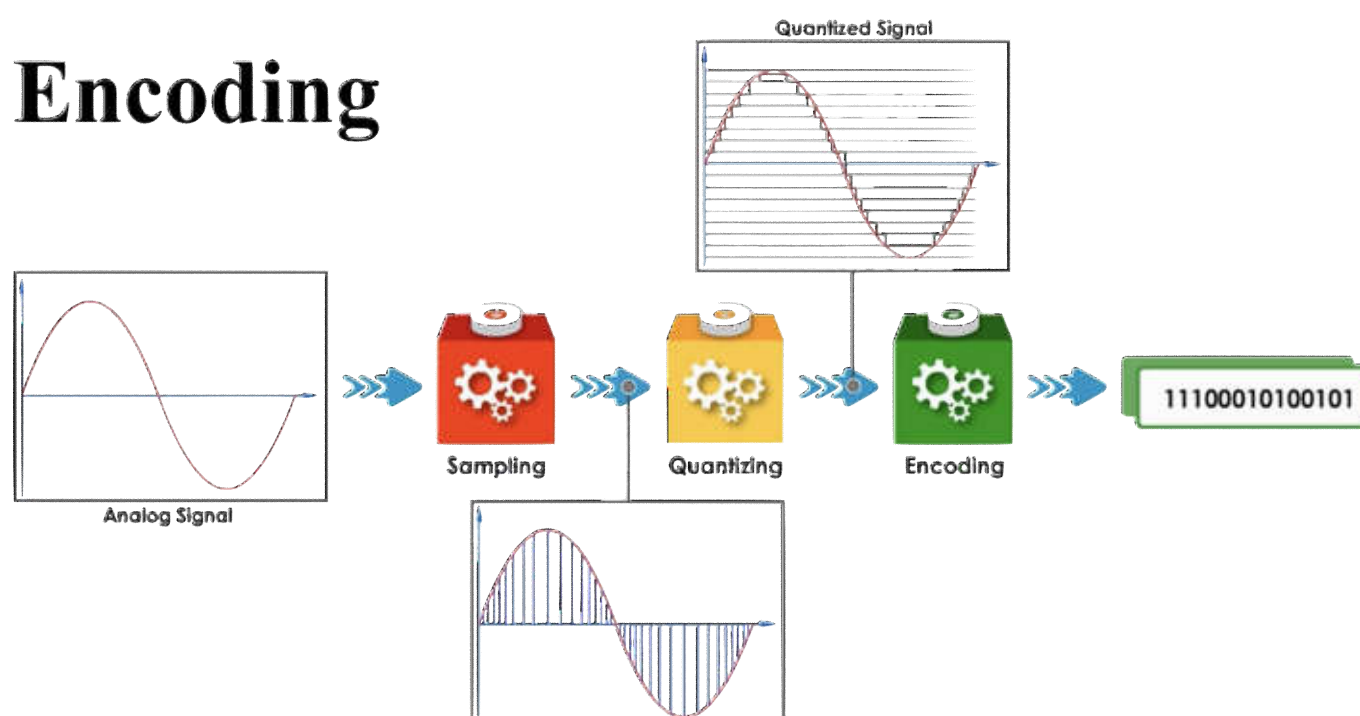
1 个正弦波周期时间: 1 秒/频率 1KHz = 1 / 1000 (秒)

1 个正弦波周期采样次数: 1 个正弦波周期时间 / 1 次采样需要的时间 = (1 / 1000) / (1 / 44100) = 44100 / 1000 = 44.1 次

结果: 如上图显示一个正弦波周期, 采样的次数为 44 次。

备注: 量化会将波形分成多个等份, 值越大对模拟信号描述越真实。采样率是 1 秒钟采样的次数, 值越大对声音的还原程度越高。所以, 声音的质量很大程度上被量化和采样率所影响。

6. Encoding (编码)



在编码这一步, 我们会将时间线上的每个 sample 数据转化为对应的二进制数据。

采样数据经过编码后产生的二进制数据, 就是 PCM 数据。PCM 数据可以直接存储在介质上, 也可以在经过编解码处理后进行存储或传输。

7. PCM 数据常用量化指标

采样率(Sample rate): 每秒钟采样多少次, 以 Hz 为单位。

位深度(Bit-depth): 表示用多少个二进制位来描述采样数据, 一般为 16bit。

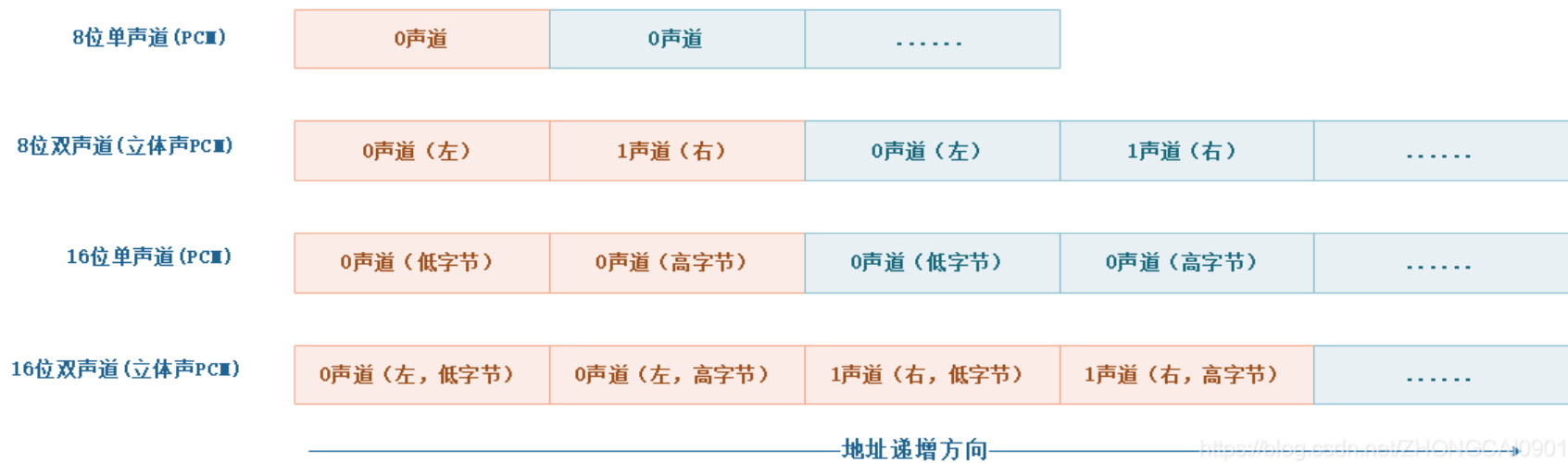
字节序: 表示音频 PCM 数据存储的字节序是大端存储 (big-endian) 还是小端存储 (little-endian), 为了数据处理效率的高效, 通常为小端存储。

声道数 (channel number): 当前 PCM 文件中包含的声道数, 是单声道 (mono)、双声道 (stereo) 等。

采样数据是否有符号 (Sign): 要表达的就是字面上的意思, 需要注意的是, 使用有符号的采样数据不能用无符号的方式播放。

8. PCM 音频数据存储格式

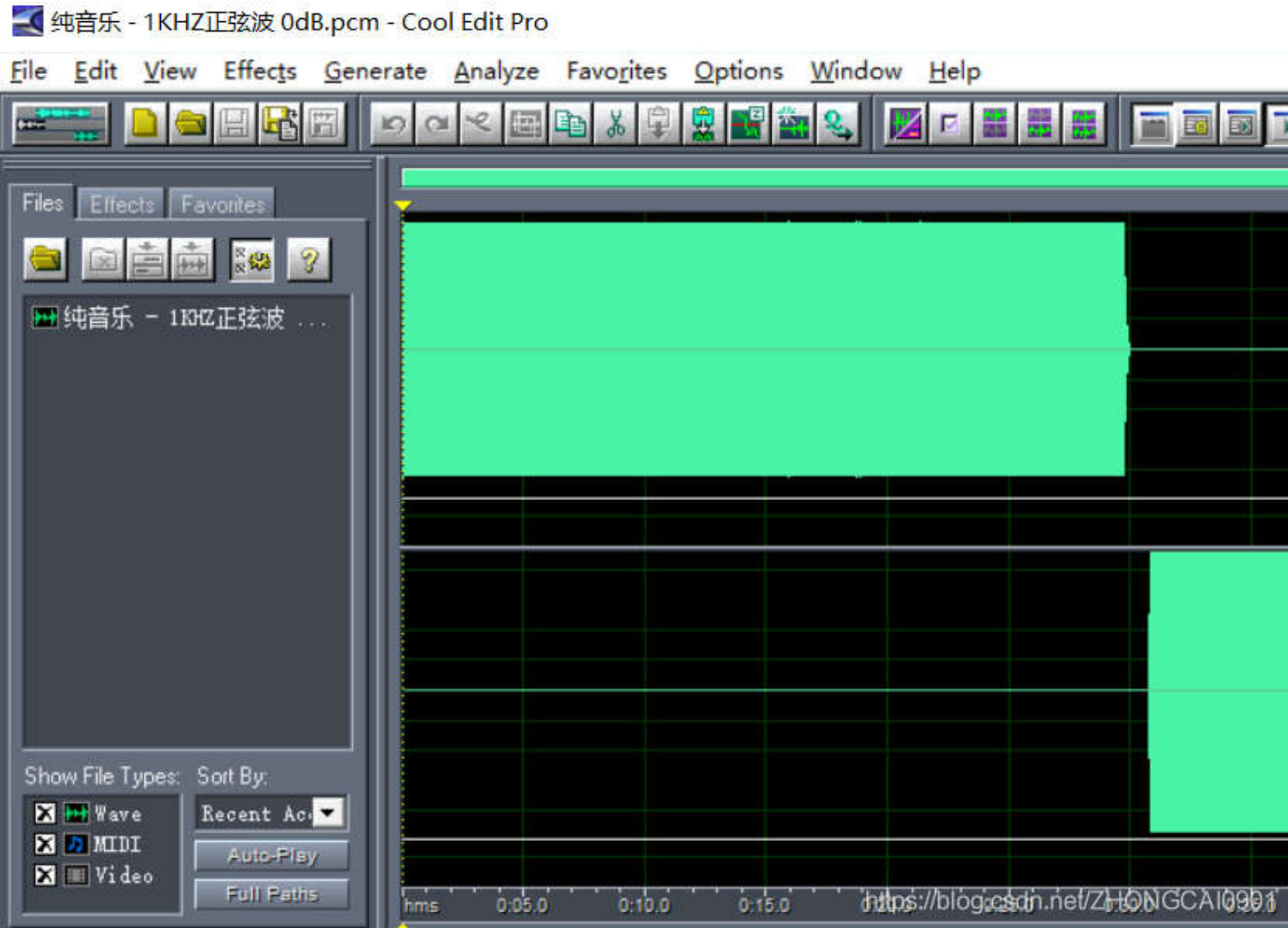
PCM 音频数据排列格式为左右声道每个样本点数据交错排列。



下面是解析一个 16bit 双声道 PCM 的数据, 如下:

Address	0	1	2	3	4	5	6	7	8	9	a	b	c	d	e	f	Dump
00000000	36	03	f9	ff	df	10	fd	ff	32	1e	fb	ff	eb	2a	fc	ff	6.???2.???□□□□□□
00000010	c5	36	fa	ff	88	41	fc	ff	f4	4a	fb	ff	e4	52	fb	ff	??圓?鬚?銻?□□□□□
00000020	1e	59	fc	ff	8c	5d	fa	ff	16	60	f7	ff	a7	60	fe	ff	.Y?宀?.`? ?□□□□□
00000030	43	5f	f5	ff	f1	5b	fc	ff	5b	56	f6	ff	c7	4f	f9	ff	C_?鼻?簾?茂?□□□□□
00000040	30	47	f7	ff	26	3d	f8	ff	0f	81	f5	ff	98	25	f7	ff	0G?&=???.%?□□□□□
00000050	82	18	f3	ff	fb	0a	f7	ff	2e	fd	f3	ff	74	ef	f6	ff	????七秒□□□□□

实际解析的出来的 PCM 波形如下:



从波形上分析 1 声道为无声。

从数据上分析 1 声道, -7(0xFF9)、-3(0xFFD)、-5(0xFFB)等基本接近 0, 所以为无声。

【Audio】I2S 传输 PCM 音频数据分析总结（二）

1. 前言

《I2S 传输 PCM 音频数据分析总结》主要是介绍 PCM 音频数据是如何通过 I2S 进行传输的，主要是分成 2 个部分：

PCM 音频数据介绍

I2S 协议介绍

2. I2S 简介

Inter-IC Sount Bus(I2S)是飞利浦半导体公司(现为恩智浦半导体公司)针对数字音频设备之间的音频数据传输而制定的一种总线标准。在飞利浦公司的 I2S 标准中，既规定了硬件接口规范，也规定了数字音频数据的格式。

3. I2S 特点

I2S 总线只处理音频数据，而其它控制信号，是分开传输的。在比较常用的 CODEC 音频芯片基本都是 I2S 和 I2C 搭配，I2S 传输音频数据，I2C 传输命令控制。主要特点如下：

支持全双工/半双工：全双工是需要 2 个 Data 线，半双工是 1 个 Data 线。

支持主/从模式：主模式需要提供 I2S 通信的 SCK 和 WS 信号。

I2S 支持多声道数据传输：通过 WS 信号控制声道的选择。

下面是《I2S bus specification》协议文档里面的截图，它是典型 I2S 信号时序图，产生 SCK 和 WS 的信号端就是主设备，用 MASTER 表示，具体如下：

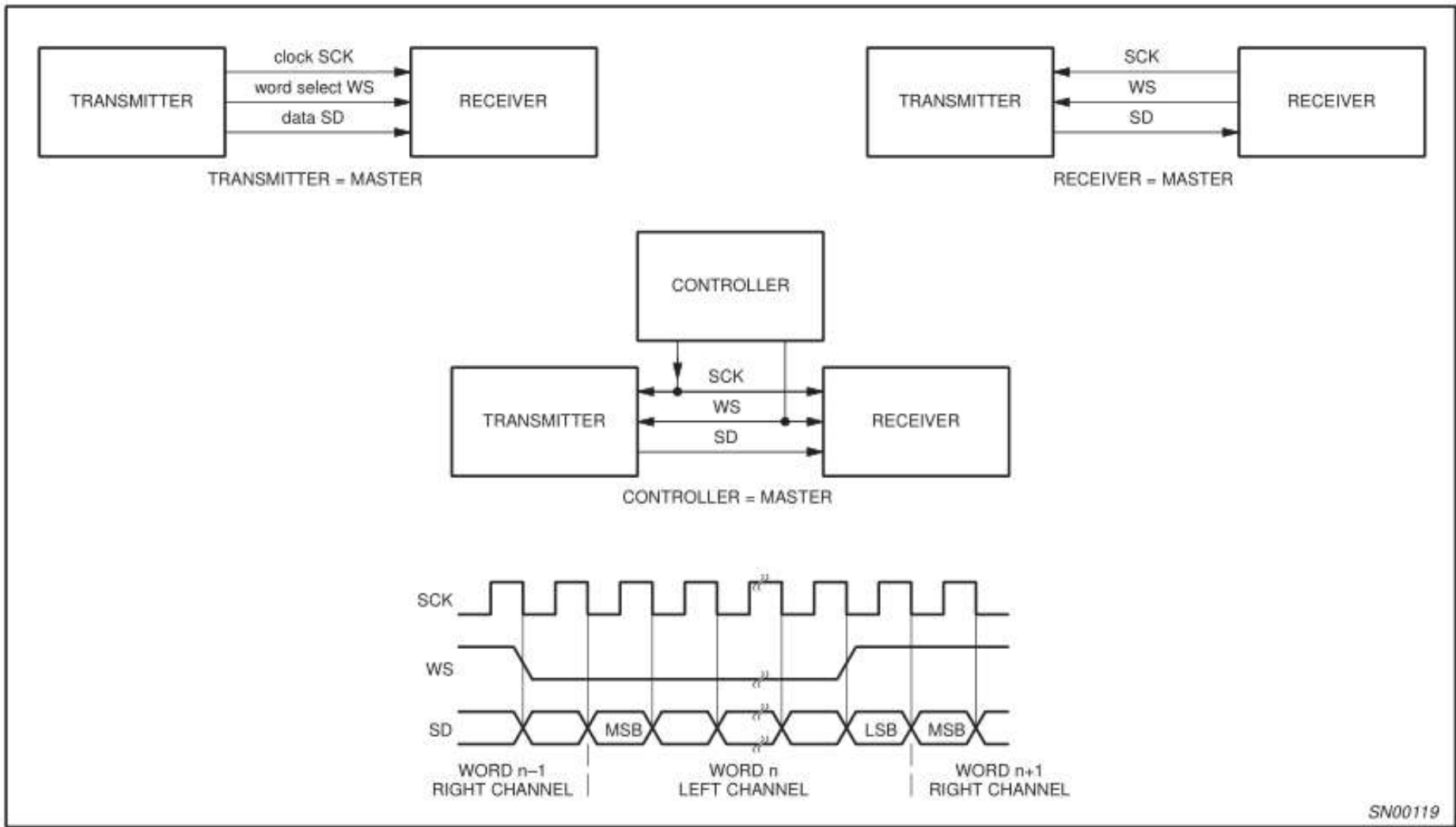


Figure 1. Simple System Configurations and Basic Interface Timing

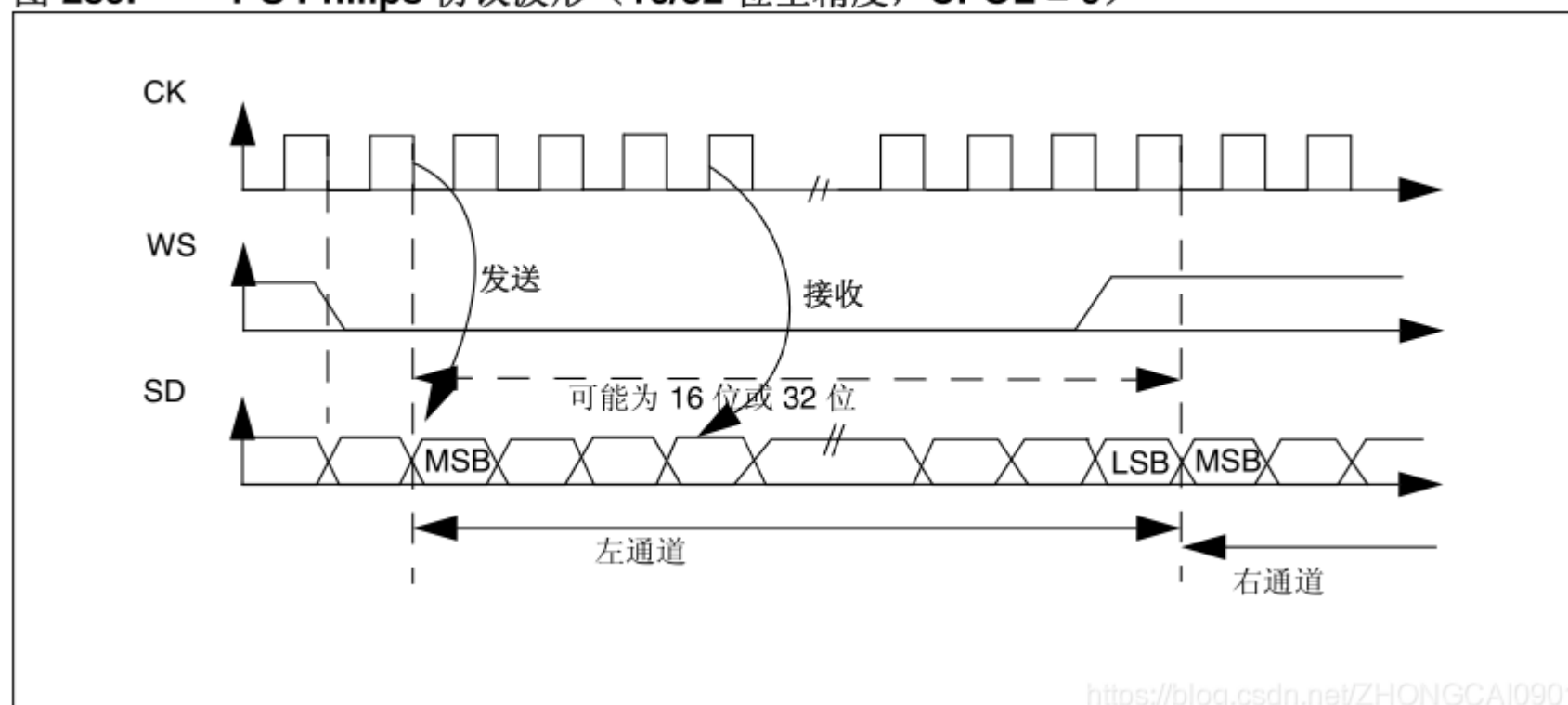
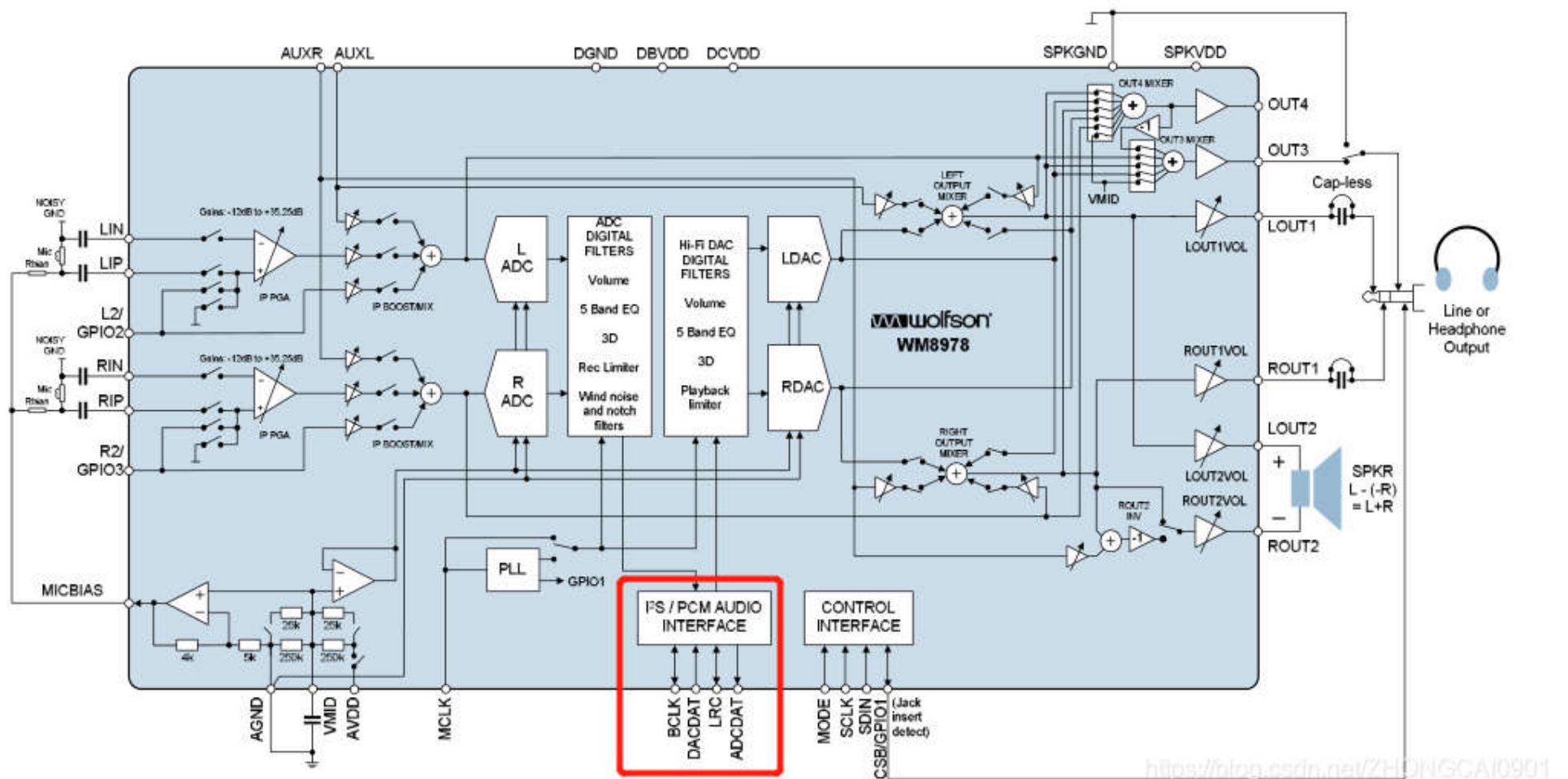
根据上图我们可以知道 I2S 总线一个是三条数据信号线：

BCLK：串行时钟，也叫位时钟，对应数字音频的每一位数据。对应数字音频的每一位数据，SCK 都有 1 个脉冲。SCK 的频率=2×采样频率×采样位数。

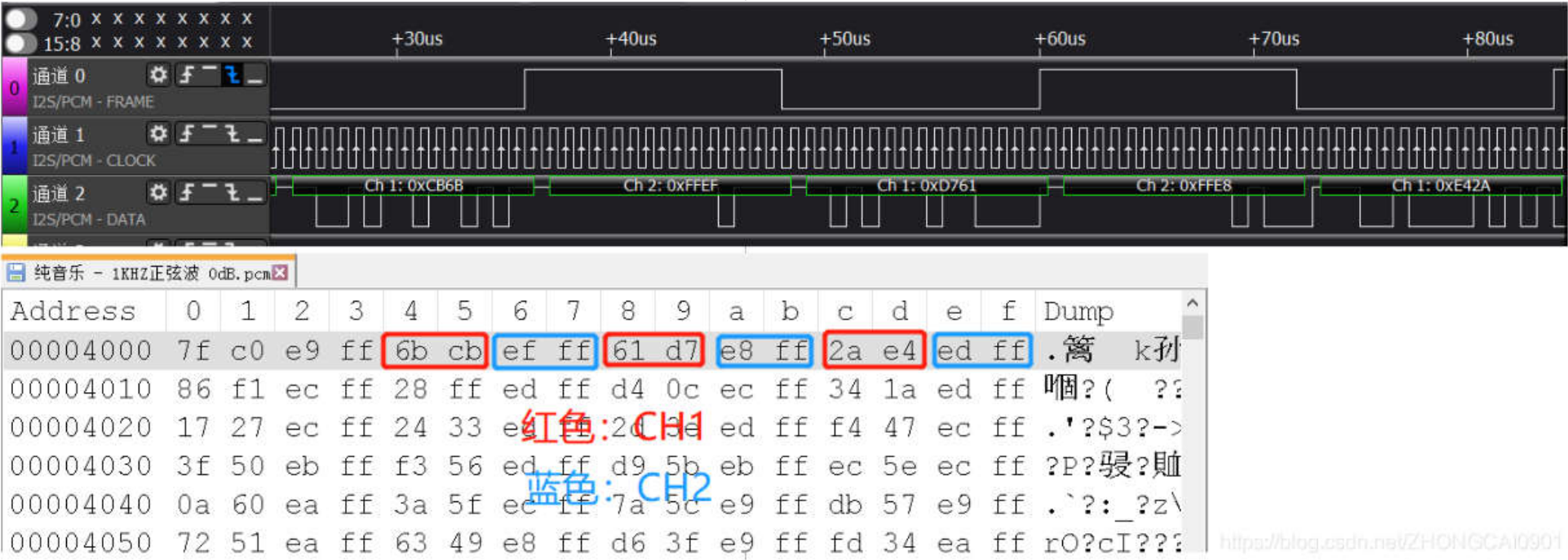
WS：字段（声道）选择，帧时钟 LRCK（左右时钟），用于选择左右声道。为“0”表示正在传输的是左声道的数据，为“1”表示正在传输的是右声道的数据。用于切换左右声道的数据。WS 的频率=采样频率。

SD：串行数据，用二进制补码来表示音频数据（数据传输从高位到低位）。

I2S 总线半双工是 3 条数据线，但是全双工是 4 条数据线。下面的截图是 WM8978 CODEC 音频编解码器芯片，它采用全双工 4 条数据线 I2S 接口，并且通过 I2C 接口进行控制。（备注：后面还会有一篇专门介绍 WM8978 的驱动移植）

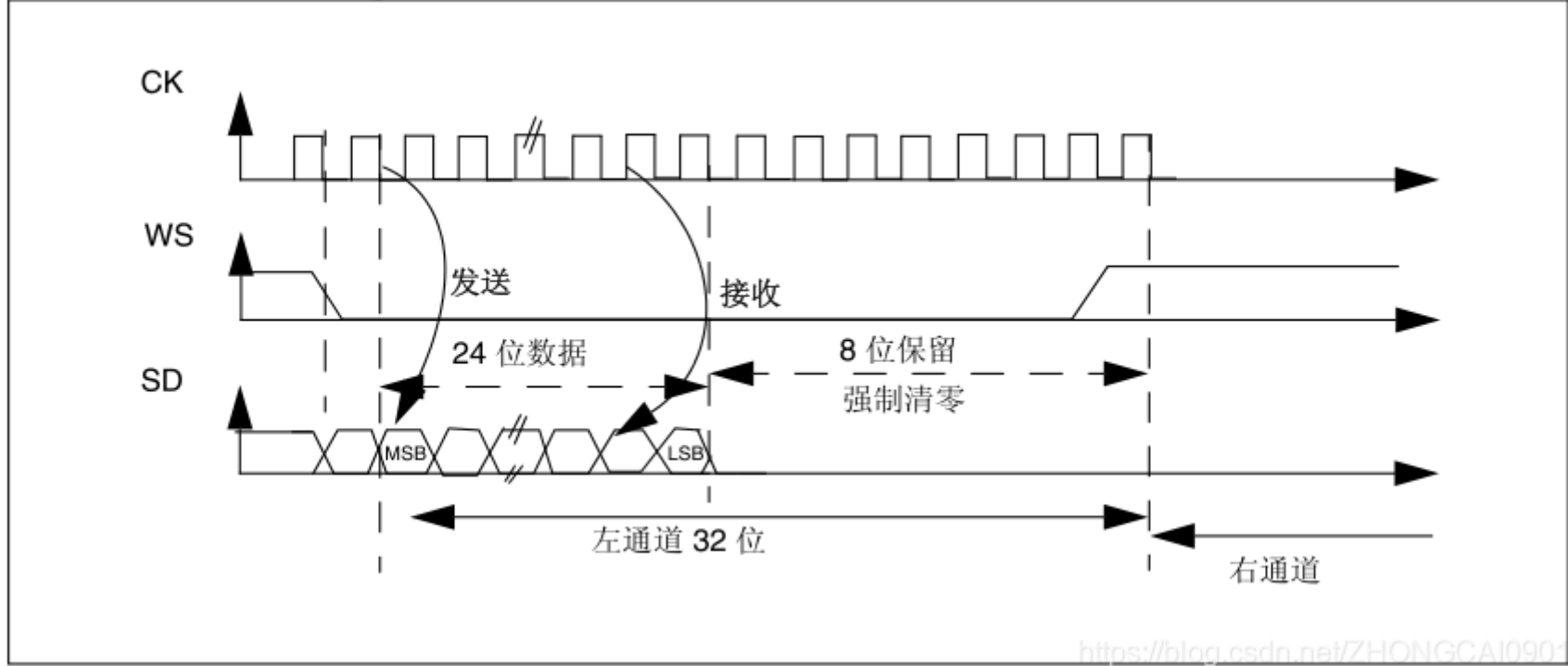


举例：下面是在播放采样率 44.1KHz 16bit 双声道正弦波 1KHz 的 PCM 音频数据时，用逻辑分析仪抓取的 I2S 数据，具体如下：



下面时序图显示了如何发送一个 24bit 的音频数据，24bit 高位先发送，剩余 8bit 强制清零。因为一般 CPU 是字(32Bit)对齐方式，所以这里设计时将剩余的 8bit 用 0 来填充：

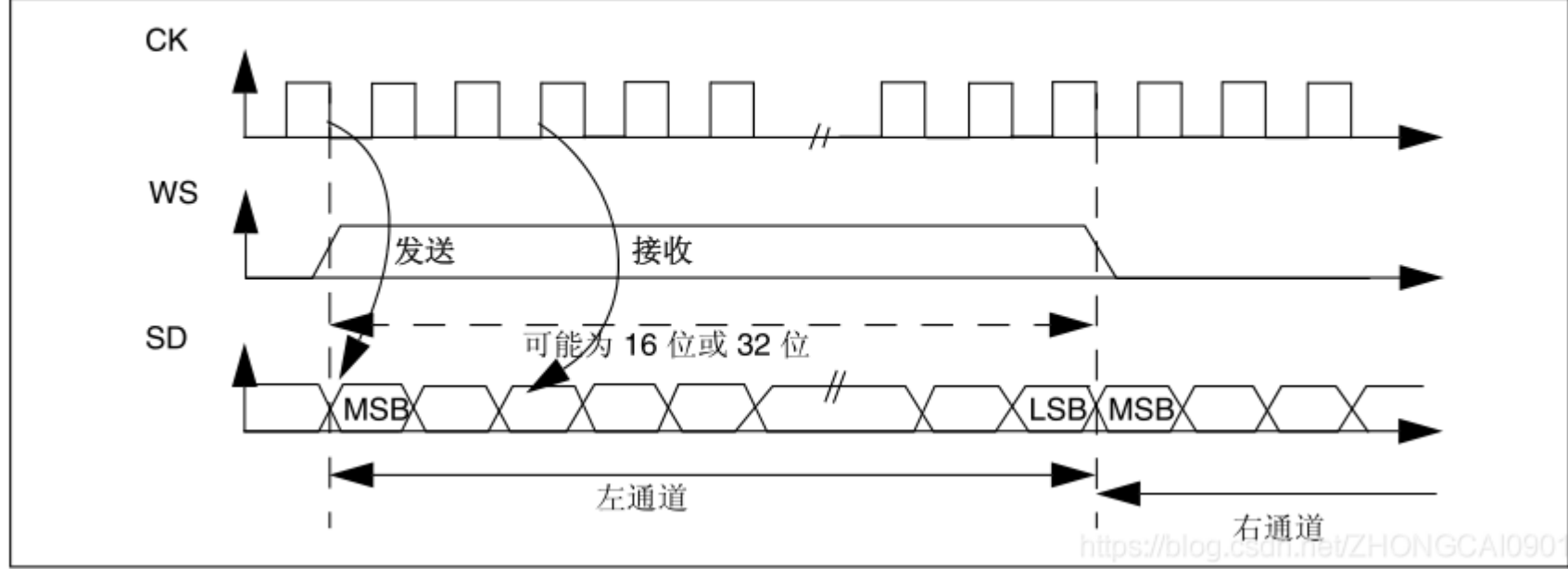
图 290. I²S Philips 标准波形（24 位帧，CPOL = 0）



4.2 MSB 对齐标准

MSB 标准同时生成 WS 信号和第一个数据位（即 MSBit）。这里和 I2S Philips 标准存在的差异，如图所示：

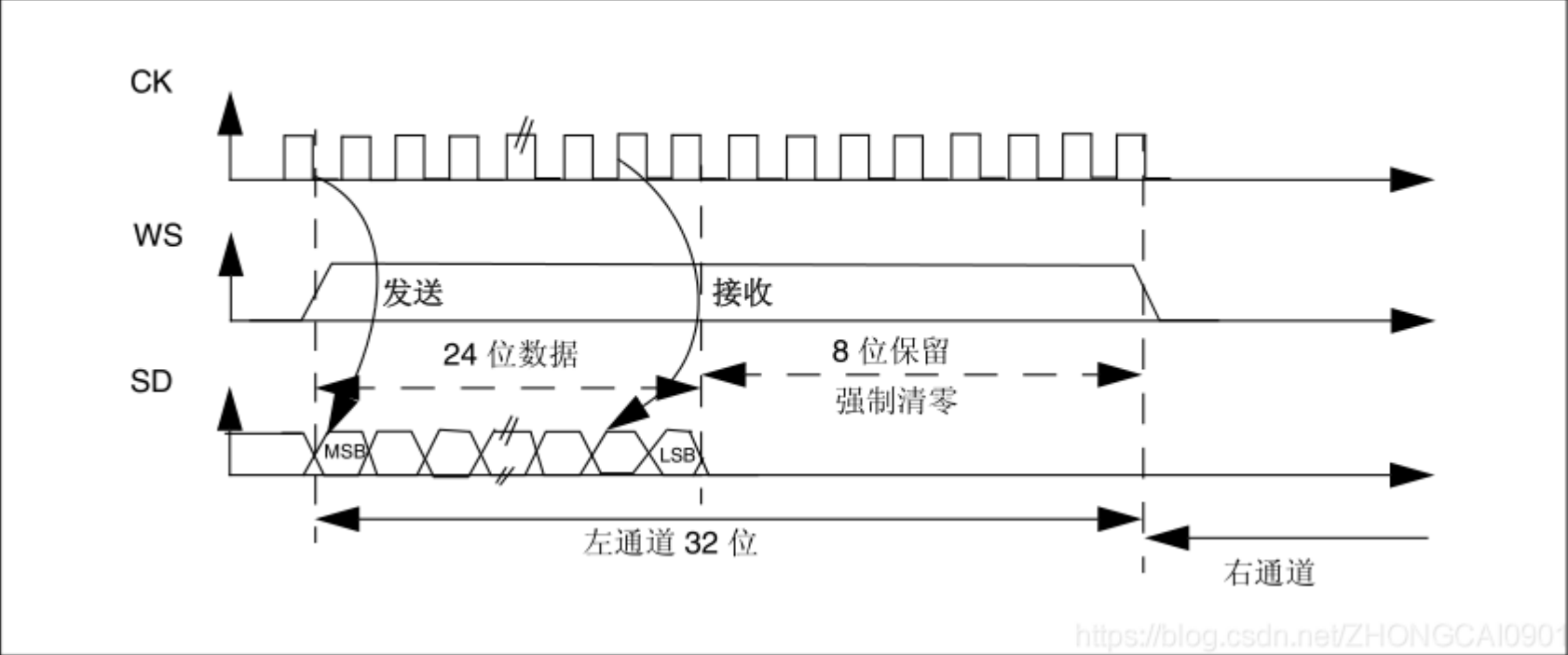
图 295. MSB 对齐的 16 位或 32 位全精度长度，CPOL = 0



发送方在时钟信号的下降沿改变数据；接收方在上升沿读取数据。

下面 MSB 标准时序图显示了如何发送一个 24bit 的音频数据，和前面一样，24bit 高位先发送，剩余 8bit 强制清零。

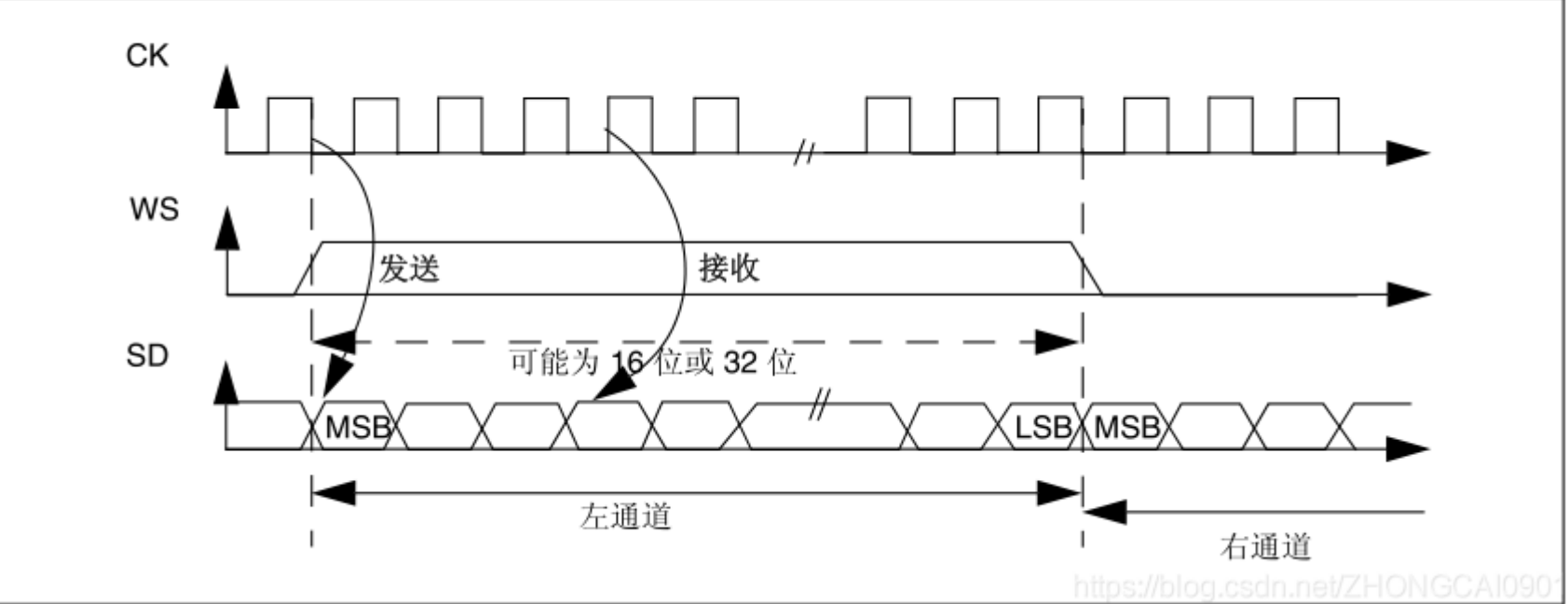
图 296. MSB 对齐的 24 位帧长度，CPOL = 0



4.3 LSB 对齐标准

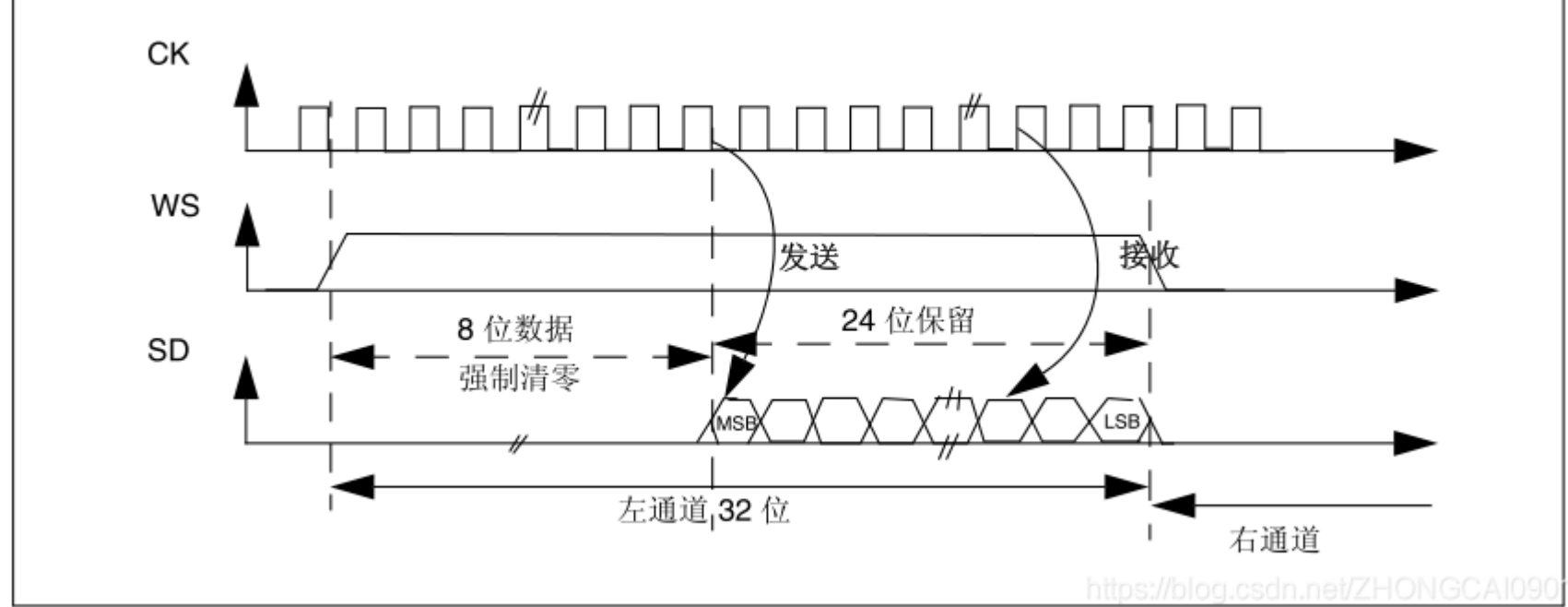
该标准与 MSB 对齐标准类似（对于 16 位和 32 位全精度帧格式，没有任何不同）

图 298. LSB 对齐的 16 位或 32 位全精度，CPOL = 0



下面是 LSB 标准时序图显示了一个 24bit 的音频数据，因为是 LSB 左对齐，所以高 8bit 强制清零先发送，然后 24bit 有效数据从高位再发送。

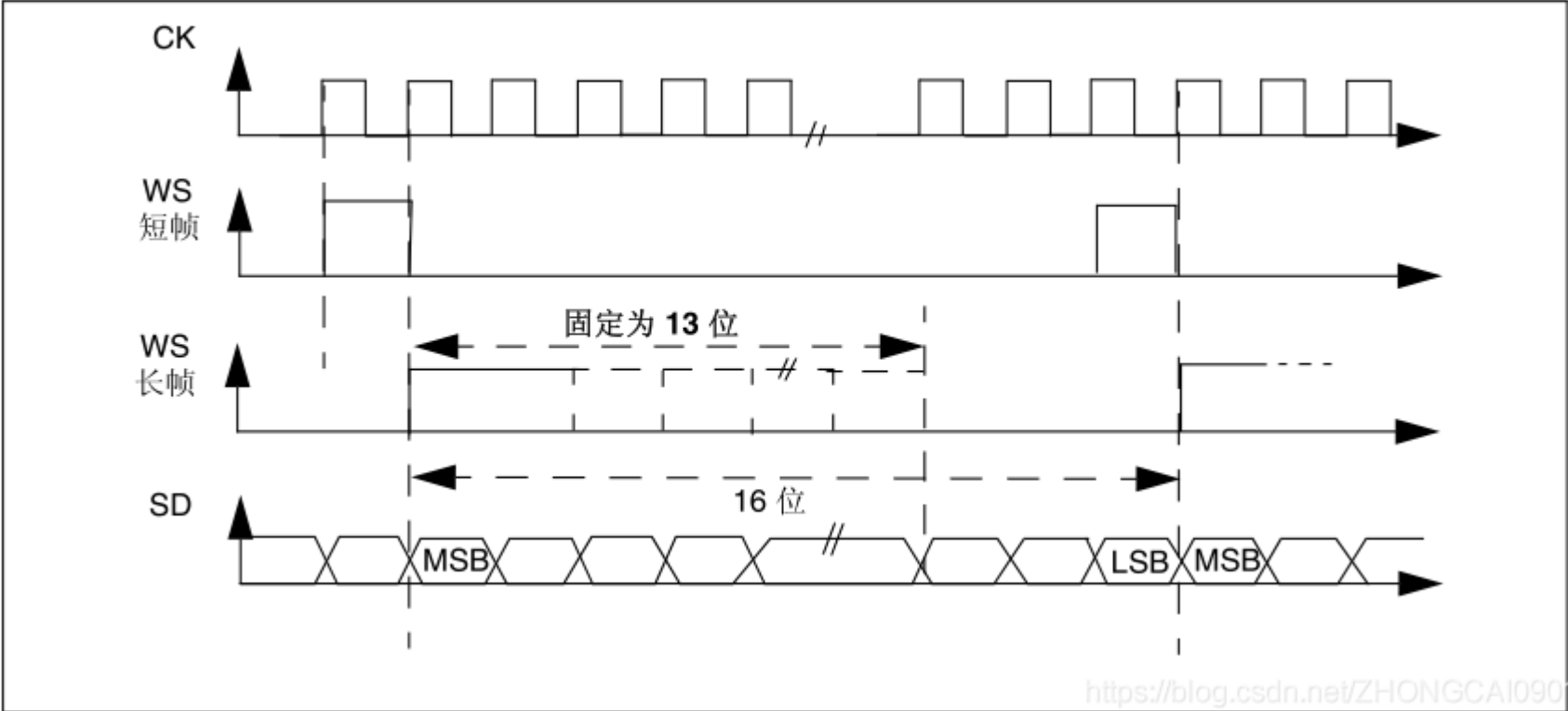
图 299. LSB 对齐的 24 位帧长度，CPOL = 0



4.4 PCM 标准

对于 PCM 标准，无需使用通道信息。有两种 PCM 模式：短帧和长帧。

图 304. PCM 标准波形（16 位）



对于长帧同步，在主模式下会将 WS 信号持续 13 个周期。
对于短帧同步，WS 同步信号的持续时间仅为一个周期。

图 305. PCM 标准波形（16 位扩展到 32 位数据包帧）

