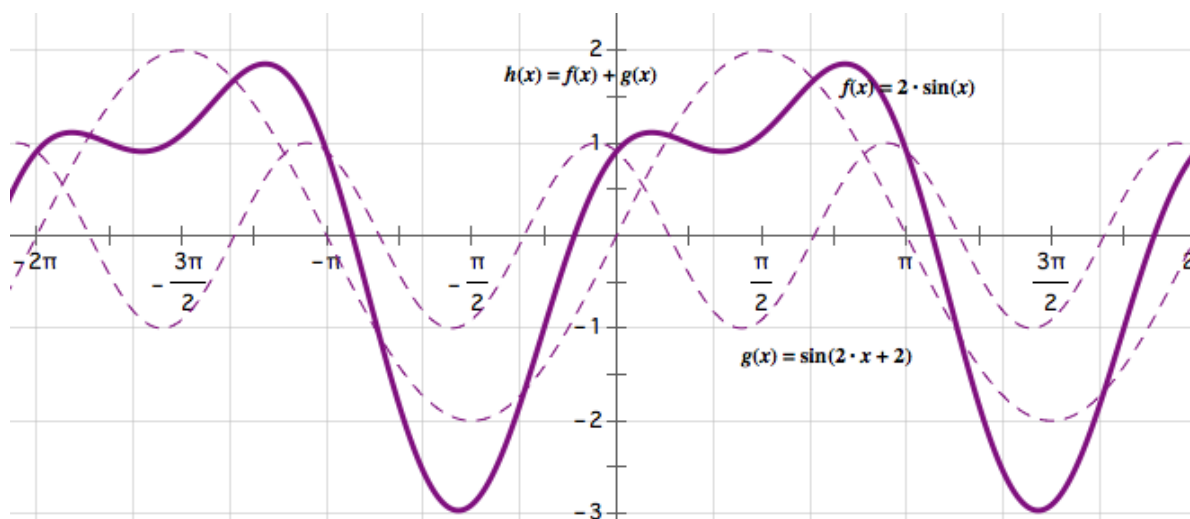


# 音频合成原理

讲音频之前，我们先了解下波，波是连续性的，两个波又可以互相叠加 形成新的波，

## 1.1 波的合成



波可以合成，也可以分解

## 1.2 波的分解



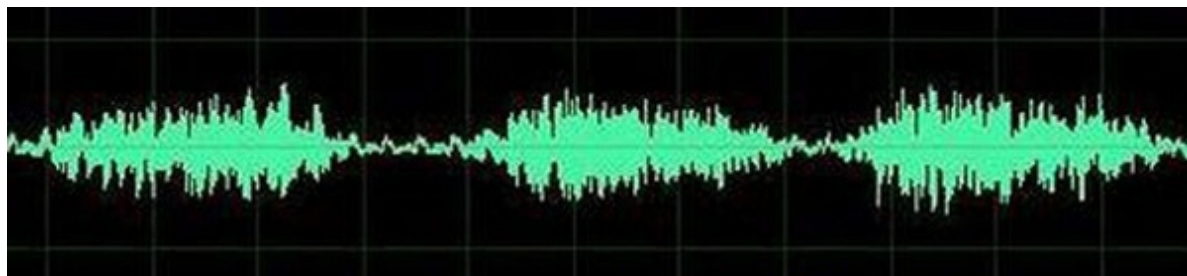
傅里叶分析不仅仅是一个数学工具，更是一种可以彻底颠覆一个人以前世界观的思维模式。但不幸的是，傅里叶分析的公式看起来太复杂了，所以很多大一新生上来就懵圈并从此对它深恶痛绝。老实说，这么有意思的东西居然成了大学里的杀手课程，不得不归咎于编教材的人实在是太严肃了。（您把教材写得好玩一点会死吗？会死吗？）所以我一直想写一个有意思的文章来解释傅里叶分析，有可能的话高中生都能看懂的那种。所以，不管读到这里的您从事何种工作，我保证您都能看懂，并且一定将体会到通过傅里叶分析看到世界另一个样子时的快感。至于对于已经有一定基础的朋友，也希望不要看到会的地方就急忙往后翻，仔细读一定会有新的发现。

## 2.1 什么是频域

从我们出生，我们看到的世界都以时间贯穿，股票的走势、人的身高、汽车的轨迹都会随着时间发生改变。这种以时间作为参照来观察动态世界的方法我们称其为时域分析。而我们也想当然的认为，世间万物都在随着时间不停的改变，并且永远不会静止下来。但如果我告诉你，用另一种方法来观察世界的话，你会发现**世界是永恒不变的**，你会不会觉得我疯了？我没有疯，这个静止的世界就叫做频域。

先举一个公式上并非很恰当，但意义上再贴切不过的例子：

在你的理解中，一段音乐是什么呢？



作者：Heinrich（知乎ID）  
微博：@花生油工人  
转载请注明出处



这是我们对音乐最普遍的理解，一个随着时间变化的震动。但我相信对于乐器小能手们来说，音乐更直观的理解是这样的：



作者：Heinrich（知乎ID）  
微博：@花生油工人  
转载请注明出处



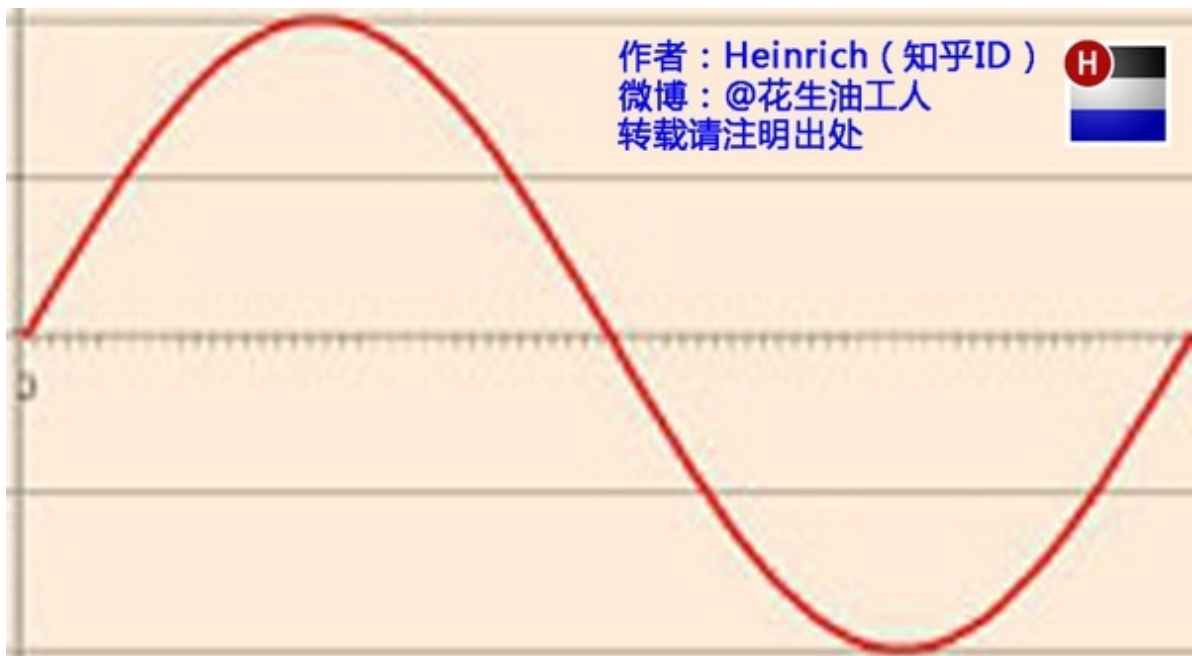
好的！下课，同学们再见。

是的，其实这一段写到这里已经可以结束了。上图是音乐在时域的样子，而下图则是音乐在频域的样子。所以频域这一概念对大家都不陌生，只是从来没意识到而已。

现在我们可以回过头来重新看看一开始那句痴人说梦般的话：世界是永恒的。

将以上两图简化：

时域：



频域：



在时域，我们观察到钢琴的琴弦一会上一会下的摆动，就如同一支股票的走势；而在频域，只有那一个永恒的音符。

所以

**你眼中看似落叶纷飞变化无常的世界，实际只是躺在上帝怀中一份早已谱好的乐章。**

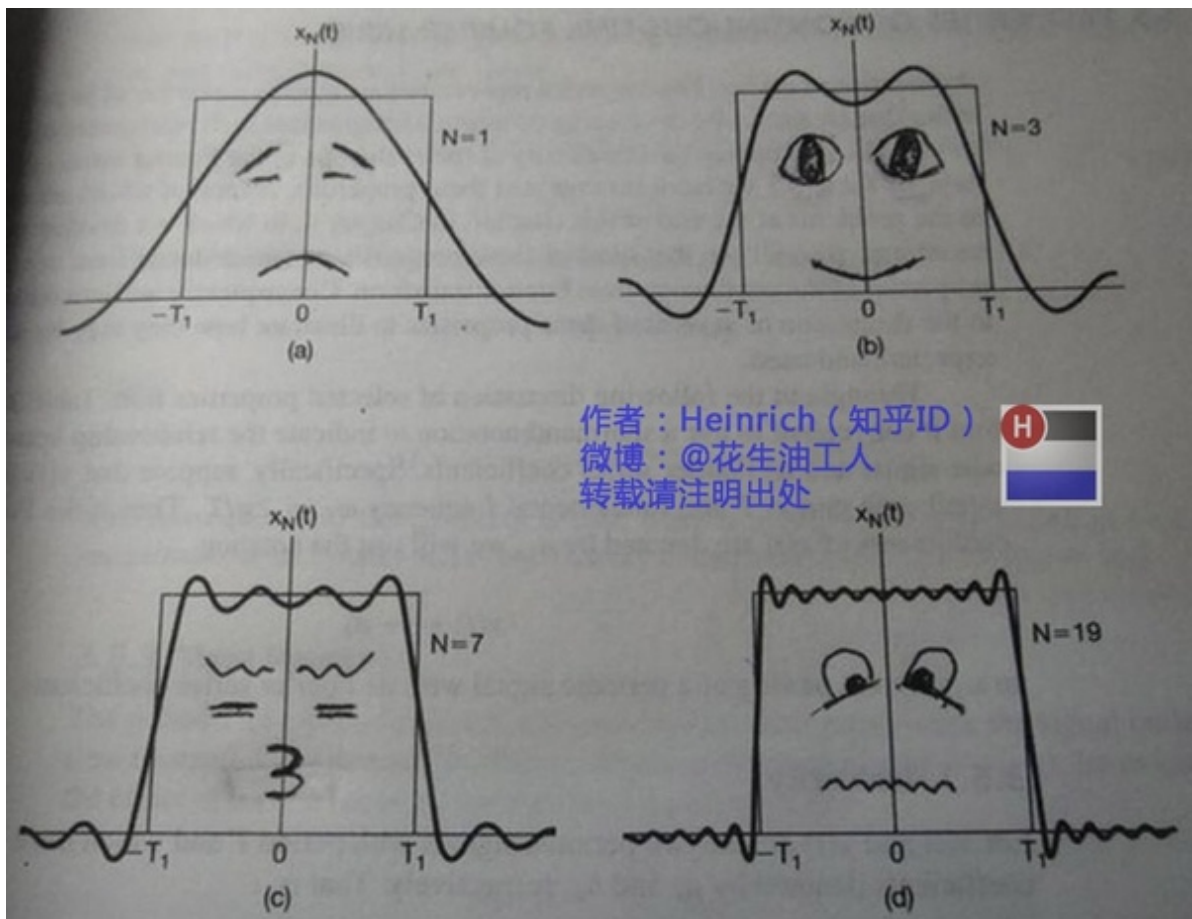
抱歉，这不是一句鸡汤文，而是黑板上确凿的公式：傅里叶同学告诉我们，任何周期函数，都可以看作是不同的振幅，不同相位正弦波的叠加。在第一个例子里我们可以理解为，利用对不同琴键不同力度，不同时间点的敲击，可以组合出任何一首乐曲。

而贯穿时域与频域的方法之一，就是传中说的傅里叶分析。傅里叶分析可分为傅里叶级数（Fourier Serie）和傅里叶变换(Fourier Transformation)，我们从简单的开始谈起。

### 3.1、傅里叶级数(Fourier Series)的频谱

还是举个栗子并且有图有真相才好理解。

如果我说我能用前面说的正弦曲线波叠加出一个带 90 度角的矩形波来，你会相信吗？你不会，就像当年的我一样。但是看看下图：



作者：Heinrich ( 知乎ID )  
微博：@花生油工人  
转载请注明出处



第一幅图是一个郁闷的正弦波  $\cos(x)$

第二幅图是 2 个卖萌的正弦波的叠加  $\cos(x) + a \cdot \cos(3x)$

第三幅图是 4 个发春的正弦波的叠加

第四幅图是 10 个便秘的正弦波的叠加

随着正弦波数量逐渐的增长，他们最终会叠加成一个标准的矩形，大家从中体会到了什么道理？

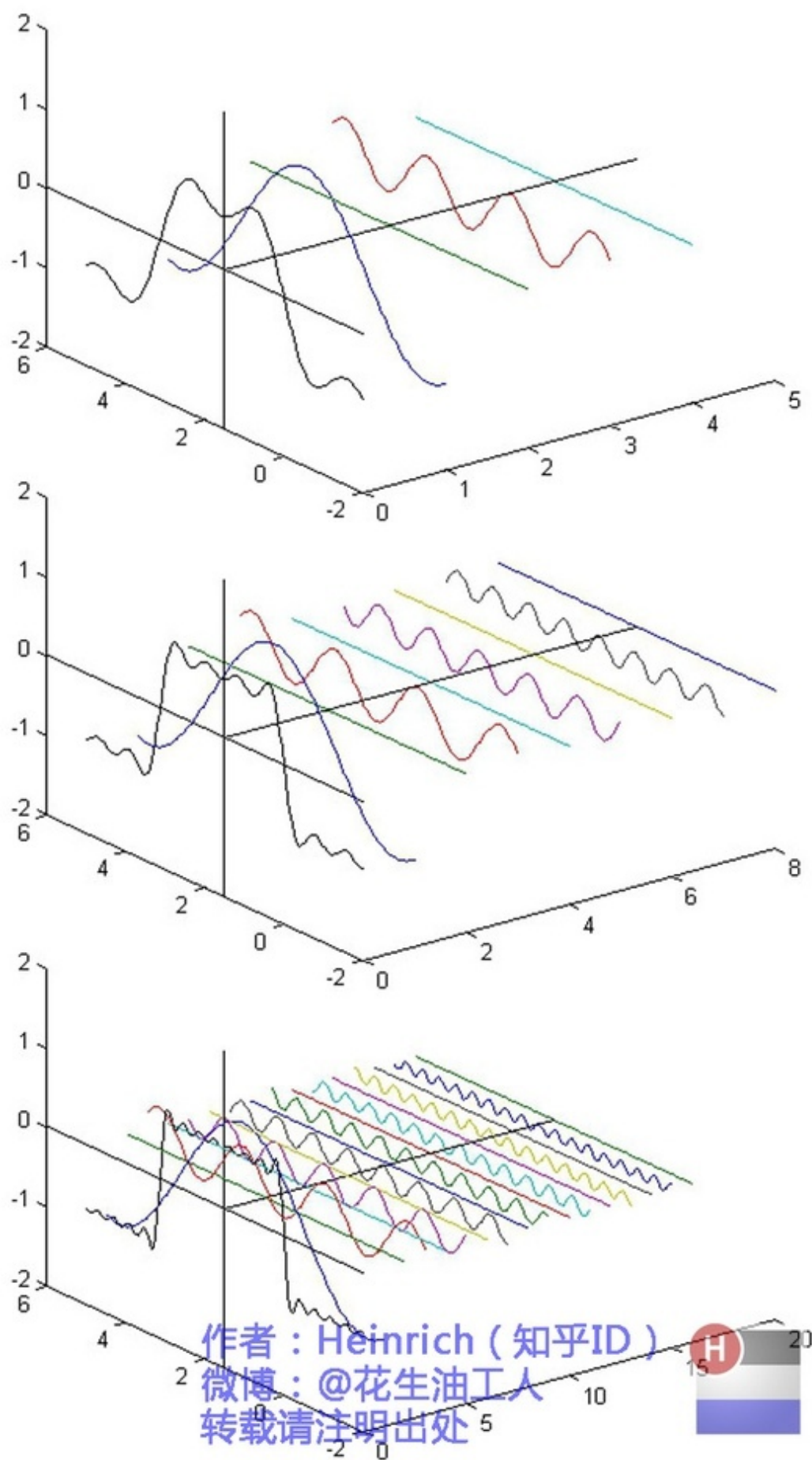
(只要努力，弯的都能掰直！)

随着叠加的递增，所有正弦波中上升的部分逐渐让原本缓慢增加的曲线不断变陡，而所有正弦波中下降的部分又抵消了上升到最高处时继续上升的部分使其变为水平线。一个矩形就这么叠加而成了。但是要多少个正弦波叠加起来才能形成一个标准 90 度角的矩形波呢？不幸的告诉大家，答案是无穷多个。

(上帝：我能让你们猜着我？)

不仅仅是矩形，你能想到的任何波形都是可以如此方法用正弦波叠加起来的。这是没有接触过傅里叶分析的人在直觉上的第一个难点，但是一旦接受了这样的设定，游戏就开始有意思起来了。

还是上图的正弦波累加成矩形波，我们换一个角度来看：



在这几幅图中，最前面黑色的线就是所有正弦波叠加而成的总和，也就是越来越接近矩形波的那个图形。而后面依不同颜色排列而成的正弦波就是组合为矩形波的各个分量。这些正弦波按照频率从低到高从前向后排列开来，而每一个波的振幅都是不同的。一定有细心的读者发现了，每两个正弦波之间都还有一条直线，那并不是分割线，而是振幅为 0 的正弦波！也就是说，为了组成特殊的曲线，有些正弦波



成分是不需要的。

这里，不同频率的正弦波我们成为频率分量。

**好了，关键的地方来了！！**

如果我们把第一个频率最低的频率分量看作“1”，我们就有了构建频域的最基本单元。

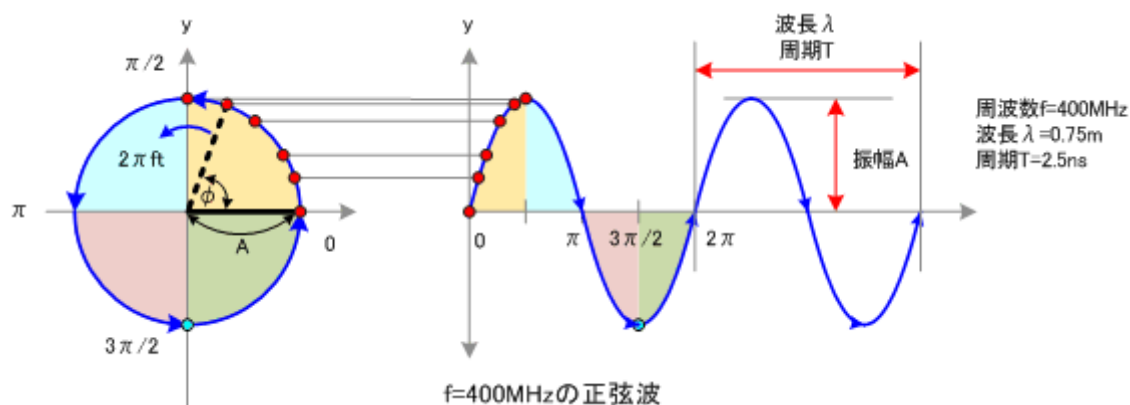
对于我们最常见的有理数轴，数字“1”就是有理数轴的基本单元。

(好吧，数学称法为——基。在那个年代，这个字还没有其他奇怪的解释，后面还有正交基这样的词汇我会说吗?)

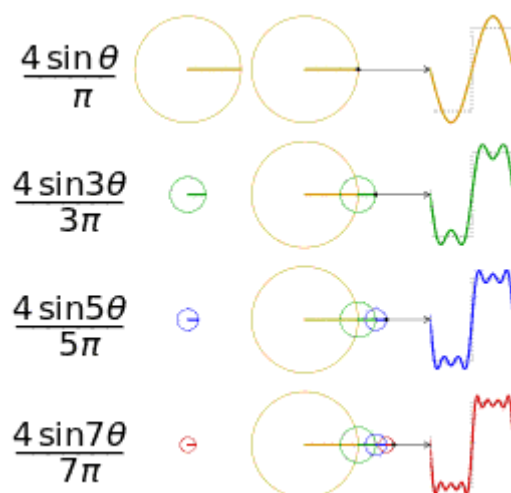
时域的基本单元就是“1 秒”，如果我们将一个角频率为 $\omega_0$ 的正弦波  $\cos(\omega_0 t)$  看作基础，那么频域的基本单元就是 $\omega_0$ 。

有了“1”，还要有“0”才能构成世界，那么频域的“0”是什么呢？ $\cos(0t)$  就是一个周期无限长的正弦波，也就是一条直线！所以在频域，0 频率也被称为直流分量，在傅里叶级数的叠加中，它仅仅影响全部波形相对于数轴整体向上或是向下而不改变波的形状。

接下来，让我们回到初中，回忆一下已经死去的八戒，啊不，已经死去的老师是怎么定义正弦波的吧。

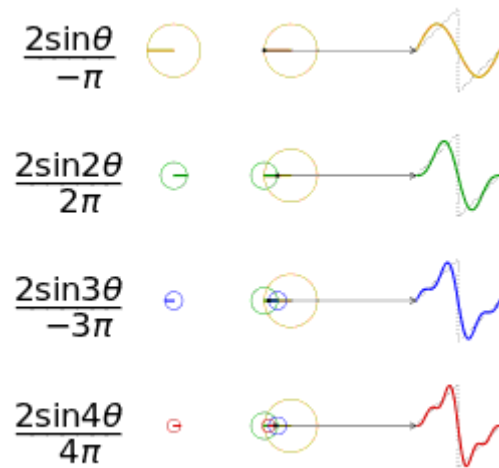


正弦波就是一个圆周运动在一条直线上的投影。所以频域的基本单元也可以理解为一个始终在旋转的圆



想看动图的同学请戳这里：

[File:Fourier series square wave circles animation.gif](#)

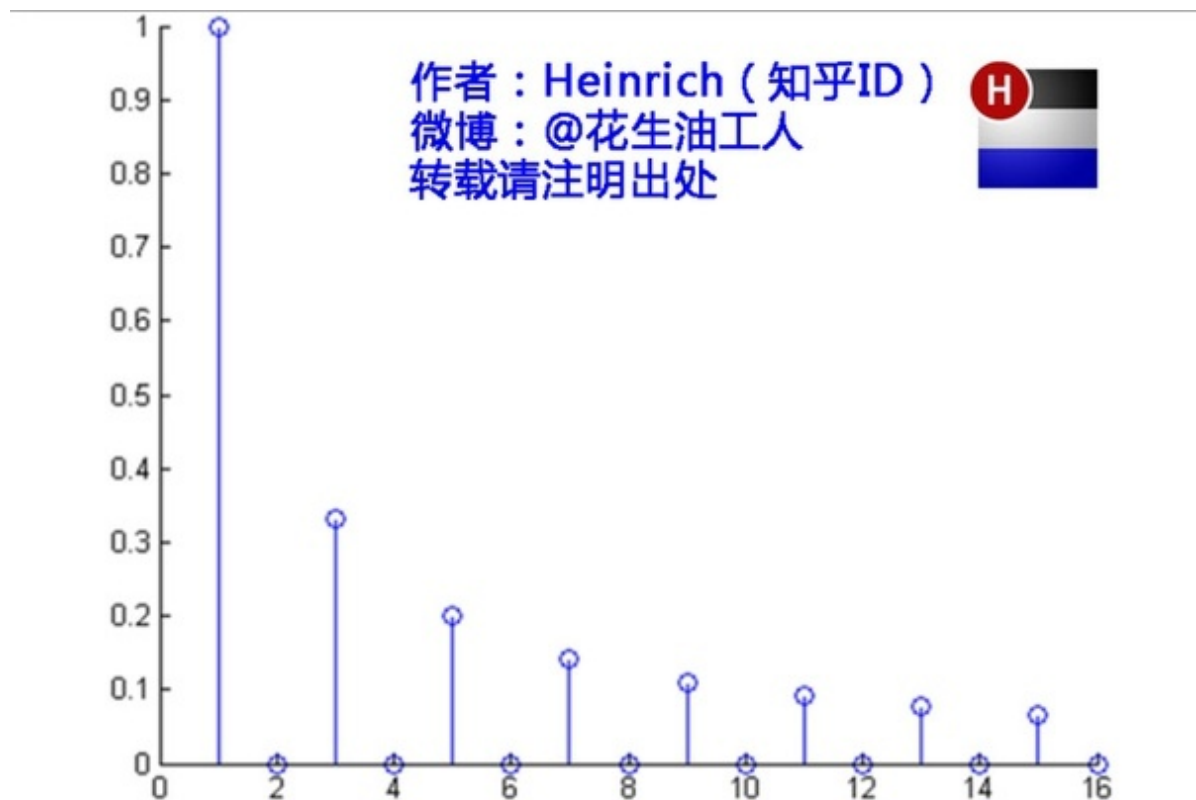


以及这里：

[File:Fourier series sawtooth wave circles animation.gif](#)

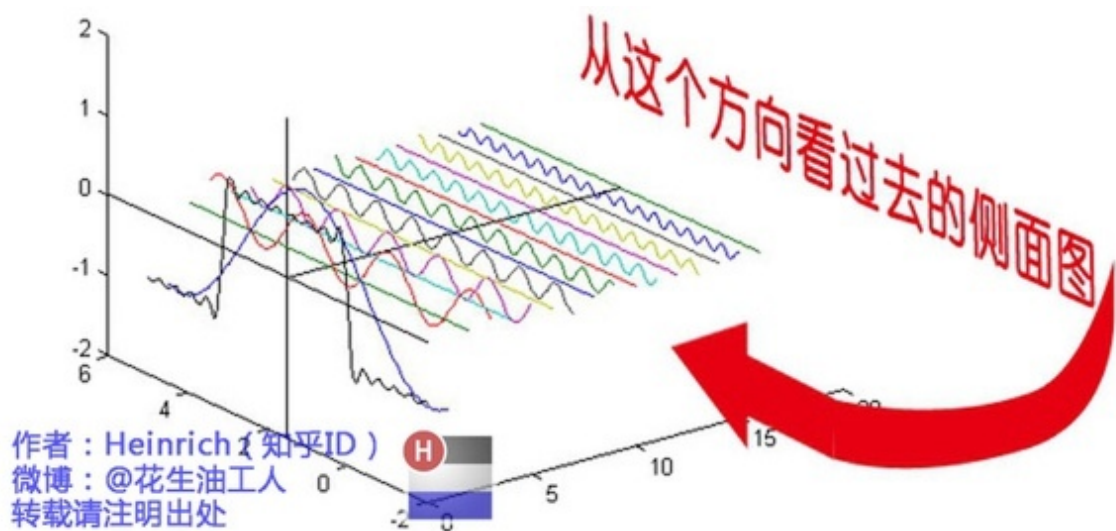
点出去的朋友不要被 wiki 拐跑了，wiki 写的哪有这里的文章这么没节操是不是。

介绍完了频域的基本组成单元，我们就可以看一看一个矩形波，在频域里的另一个模样了：

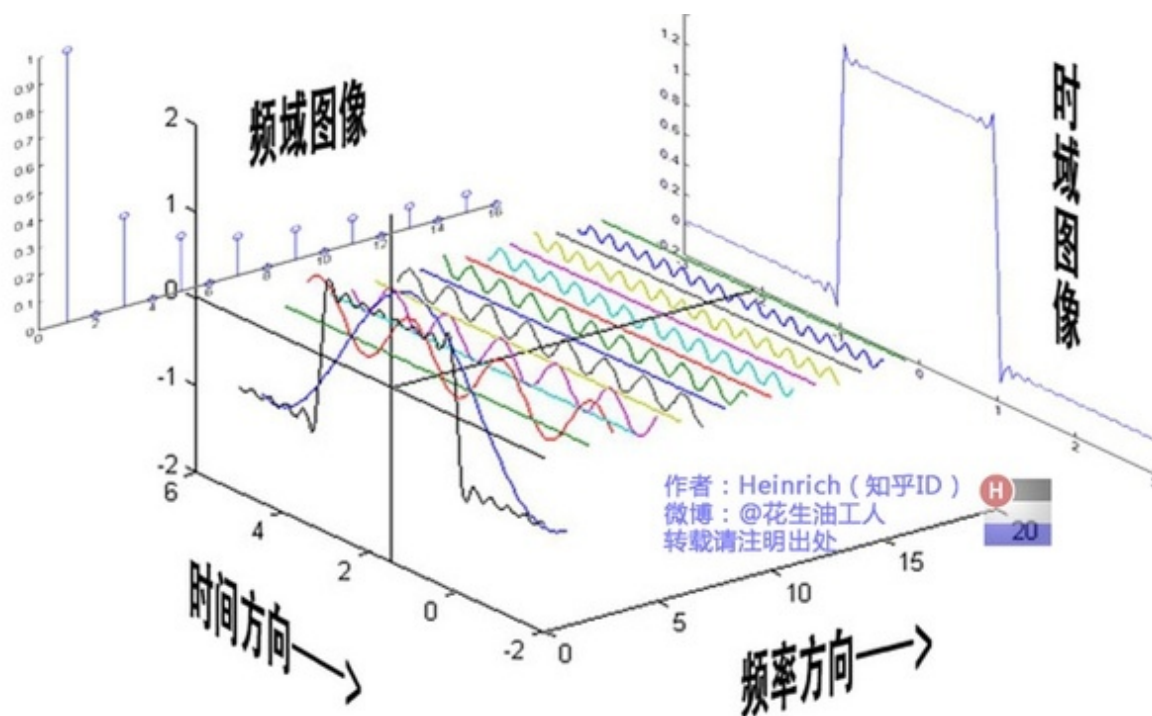


这是什么奇怪的东西？

这就是矩形波在频域的样子，是不是完全认不出来了？教科书一般就给到这里然后留给了读者无穷的遐想，以及无穷的吐槽，其实教科书只要补一张图就足够了：频域图像，也就是俗称的频谱，就是——



再清楚一点：



可以发现，在频谱中，偶数项的振幅都是0，也就对应了图中的彩色直线。振幅为0的正弦波。



动图请戳：

[File:Fourier series and transform.gif](#)



老实说，在我学傅里叶变换时，维基的这个图还没有出现，那时我就想到了这种表达方法，而且，后面还会加入维基没有表示出来的另一个谱——相位谱。

但是在讲相位谱之前，我们先回顾一下刚刚的这个例子究竟意味着什么。记得前面说过的那句“世界是静止的”吗？估计好多人对这句话都已经吐槽半天了。想象一下，世界上每一个看似混乱的表象，实际都是一条时间轴上不规则的曲线，但实际这些曲线都是由这些无穷无尽的正弦波组成。我们看似不规律的事情反而是规律的正弦波在时域上的投影，而正弦波又是一个旋转的圆在直线上的投影。那么你的脑海中会产生一个什么画面呢？

我们眼中的世界就像皮影戏的大幕布，幕布的后面有无数的齿轮，大齿轮带动小齿轮，小齿轮再带动更小的。在最外面的小齿轮上有一个小人——那就是我们自己。我们只看到这个小人毫无规律的在幕布前表演，却无法预测他下一步会去哪。而幕布后面的齿轮却永远一直那样不停的旋转，永不停歇。这样说来有些宿命论的感觉。说实话，这种对人生的描绘是我一个朋友在我们都是高中生的时候感叹的，当时想想似懂非懂，直到有一天我学到了傅里叶级数.....

### 三、傅里叶级数（Fourier Series）的相位谱

上一章的关键词是：从侧面看。这一章的关键词是：从下面看。

在这一章最开始，我想先回答很多人的一个问题：傅里叶分析究竟是干什么用的？这段相对比较枯燥，已经知道了的同学可以直接跳到下一个分割线。

先说一个最直接的用途。无论听广播还是看电视，我们一定对一个词不陌生——频道。频道频道，就是频率的通道，不同的频道就是将不同的频率作为一个通道来进行信息传输。下面大家尝试一件事：

先在纸上画一个 $\sin(x)$ ，不一定标准，意思差不多就行。不是很难吧。

好，接下去画一个 $\sin(3x) + \sin(5x)$ 的图形。

别说标准不标准了，曲线什么时候上升什么时候下降你都不一定画的对吧？

好，画不出来不要紧，我把 $\sin(3x) + \sin(5x)$ 的曲线给你，但是前提是你不知道这个曲线的方程式，现在需要你从图里拿出去，看看剩下的是什么。这基本是不可能做到的。

但是在频域呢？则简单的很，无非就是几条竖线而已。

所以很多在时域看似不可能做到的数学操作，在频域相反很容易。这就是需要傅里叶变换的地方。尤其是从某条曲线中去除一些特定的频率成分，这在工程上称为滤波，是信号处理最重要的概念之一，只有在频域才能轻松的做到。

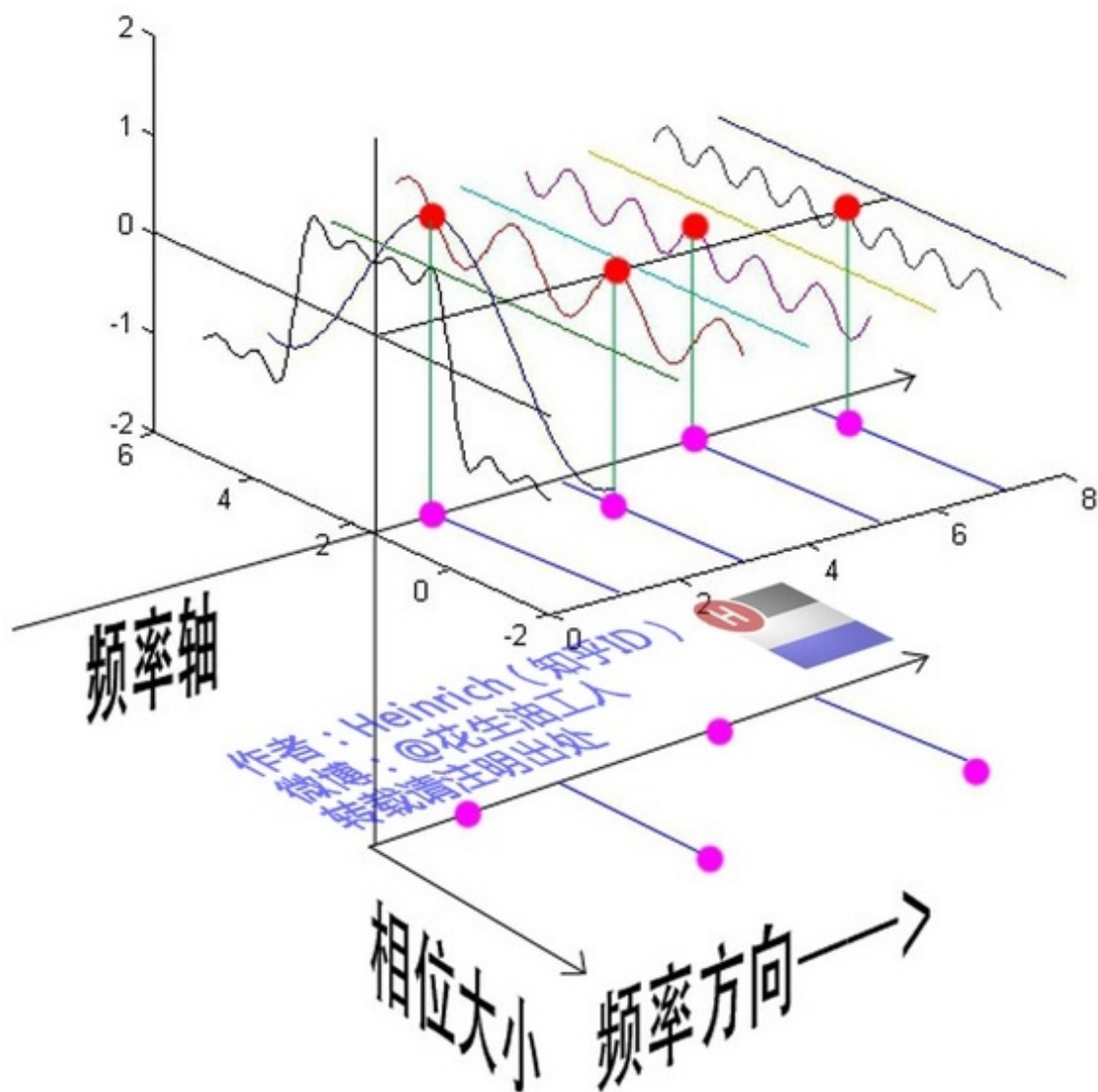
再说一个更重要，但是稍微复杂一点的用途——求解微分方程。（这段有点难度，看不懂的可以直接跳过这段）微分方程的重要性不用我过多介绍了。各行各业都用的到。但是求解微分方程却是一件相当麻烦的事情。因为除了要计算加减乘除，还要计算微分积分。而傅里叶变换则可以让微分和积分在频域中变为乘法和除法，大学数学瞬间变小学算术有没有。

傅里叶分析当然还有其他更重要的用途，我们随着讲随着提。

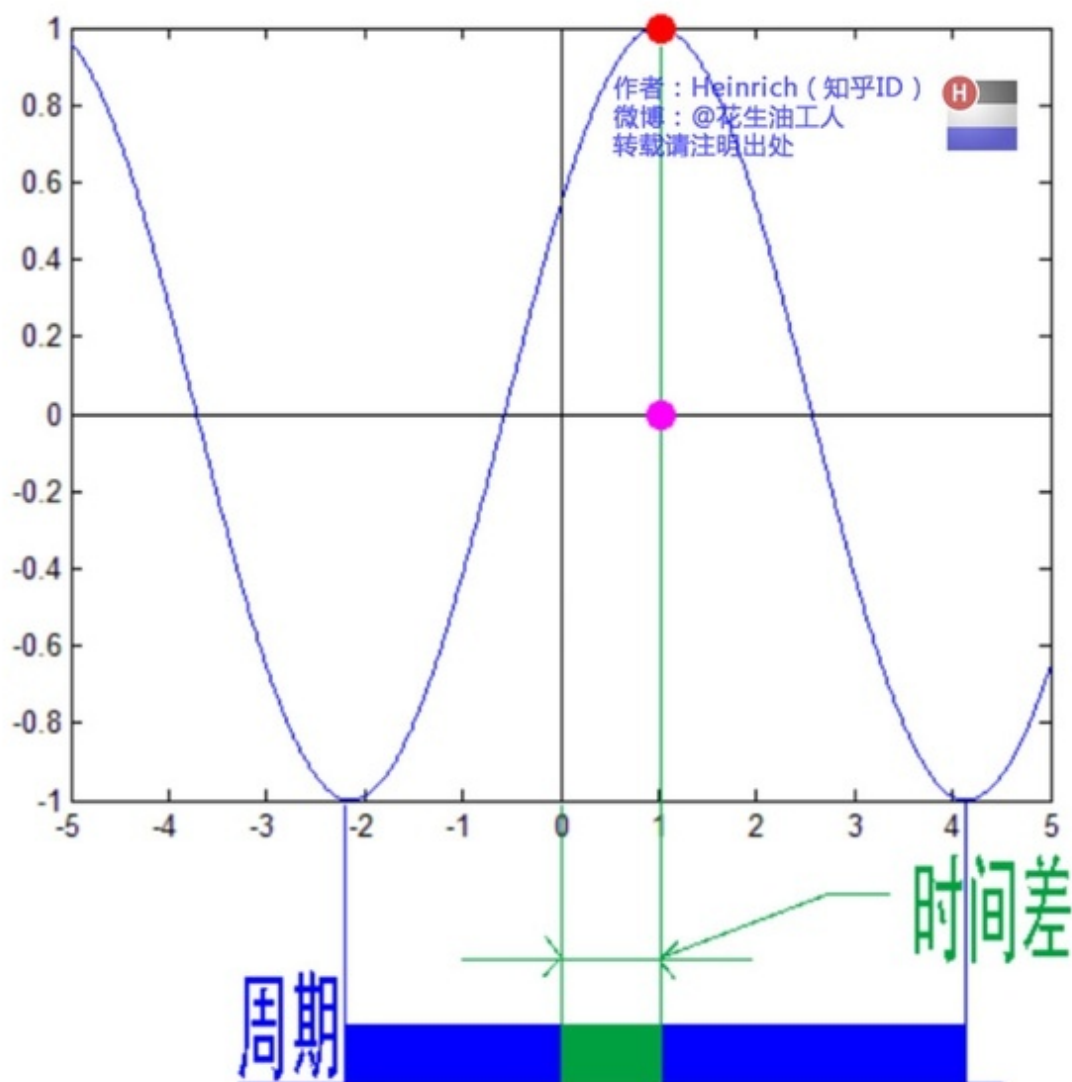
---

下面我们继续说相位谱：

通过时域到频域的变换，我们得到了一个从侧面看的频谱，但是这个频谱并没有包含时域中全部的信息。因为频谱只代表每一个对应的正弦波的振幅是多少，而没有提到相位。基础的正弦波 $A \cdot \sin(\omega t + \theta)$ 中，振幅，频率，相位缺一不可，不同相位决定了波的位置，所以对于频域分析，仅仅有频谱（振幅谱）是不够的，我们还需要一个相位谱。那么这个相位谱在哪呢？我们看下图，这次为了避免图片太混乱，我们用7个波叠加的图。



鉴于正弦波是周期的，我们需要设定一个用来标记正弦波位置的东西。在图中就是那些小红点。小红点是距离频率轴最近的波峰，而这个波峰所处的位置离频率轴有多远呢？为了看的更清楚，我们将红色的点投影到下平面，投影点我们用粉色点来表示。当然，这些粉色的点只标注了波峰距离频率轴的距离，并不是相位。

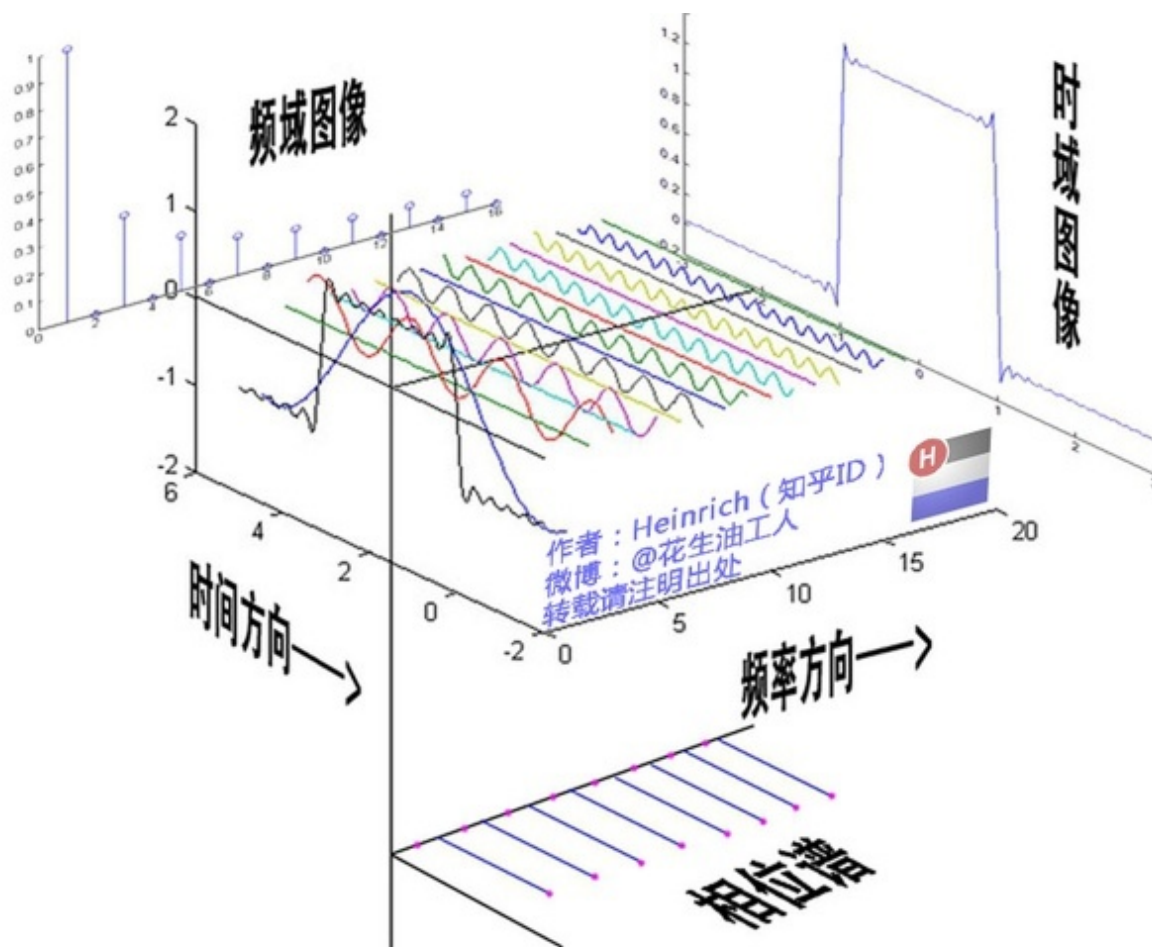


这里需要纠正一个概念：时间差并不是相位差。如果将全部周期看作 $2\pi$ 或者 $360^\circ$ 的话，相位差则是时间差在一个周期中所占的比例。我们将时间差除周期再乘 $2\pi$ ，就得到了相位差。

在完整的立体图中，我们将投影得到的时间差依次除以所在频率的周期，就得到了最下面的相位谱。所以，频谱是从侧面看，相位谱是从下面看。下次偷看女生裙底被发现的话，可以告诉她：“对不起，我只是想看看你的相位谱。”

注意到，相位谱中的相位除了0，就是 $\pi$ 。因为 $\cos(t + \pi) = -\cos(t)$ ，所以实际上相位为 $\pi$ 的波只是上下翻转了而已。对于周期方波的傅里叶级数，这样的相位谱已经是很简单的了。另外值得注意的是，由于 $\cos(t + 2\pi) = \cos(t)$ ，所以相位差是周期的， $\pi$ 和 $3\pi$ ， $5\pi$ ， $7\pi$ 都是相同的相位。人为定义相位谱的值域为 $(-\pi, \pi]$ ，所以图中的相位差均为 $\pi$ 。

最后来一张大集合：



## 四、傅里叶变换 (Fourier Transformation)

相信通过前面三章，大家对频域以及傅里叶级数都有了一个全新的认识。但是文章在一开始关于钢琴琴谱的例子我曾说过，这个栗子是一个公式错误，但是概念典型的例子。所谓的公式错误在哪里呢？

傅里叶级数的本质是将一个周期的信号分解成无限多分开的（离散的）正弦波，但是宇宙似乎并不是周期的。曾经在学数字信号处理的时候写过一首打油诗：

**往昔连续非周期，**

**回忆周期不连续，**

**任你ZT、DFT，**

**还原不回去。**

（请无视我渣一样的文学水平……）

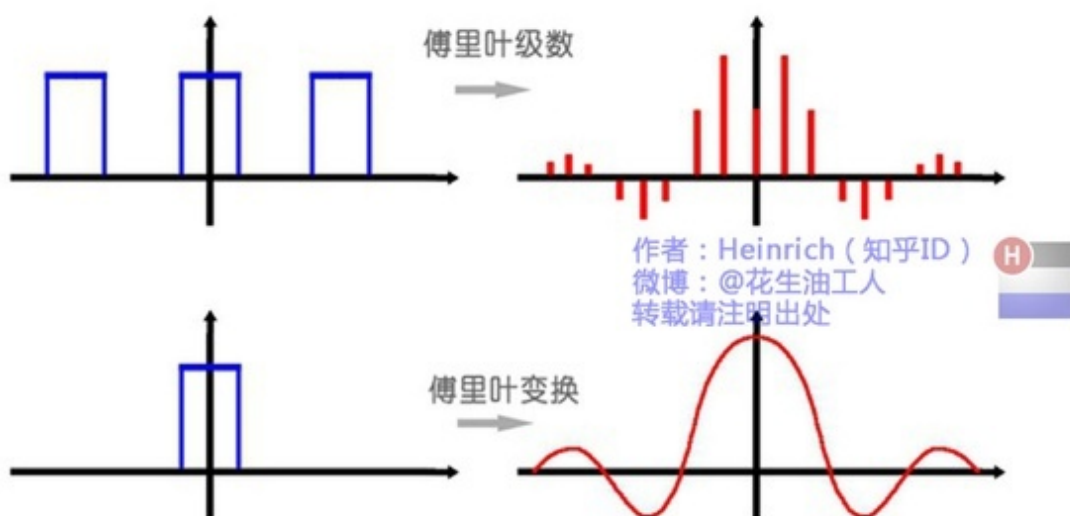
在这个世界上，有的事情一期一会，永不再来，并且时间始终不曾停息地将那些刻骨铭心的往昔连续的标记在时间点上。但是这些事情往往又成为了我们格外宝贵的回忆，在我们大脑里隔一段时间就会周期性的蹦出来一下，可惜这些回忆都是零散的片段，往往只有最幸福的回忆，而平淡的回忆则逐渐被我们忘却。因为，往昔是一个连续的非周期信号，而回忆是一个周期离散信号。

是否有一种数学工具将连续非周期信号变换为周期离散信号呢？抱歉，真没有。

比如傅里叶级数，在时域是一个周期且连续的函数，而在频域是一个非周期离散的函数。这句话比较绕嘴，实在看着费事可以干脆回忆第一章的图片。

而在我接下去要讲的傅里叶变换，则是将一个时域非周期的连续信号，转换为一个在频域非周期的连续信号。

算了，还是上一张图方便大家理解吧：



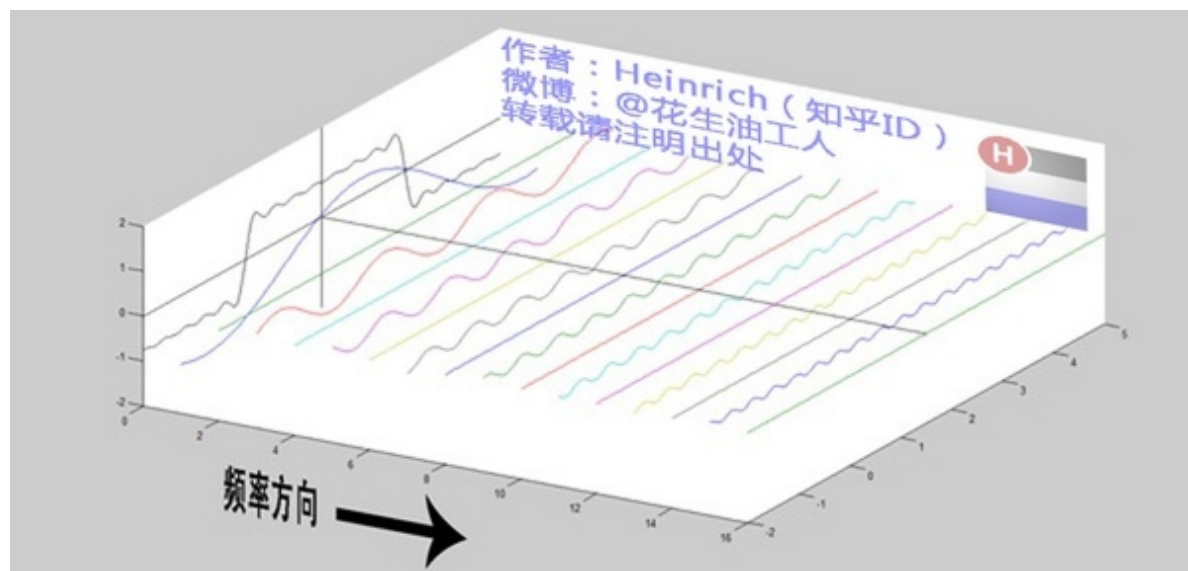
或者我们也可以换一个角度理解：傅里叶变换实际上是对一个周期无限大的函数进行傅里叶变换。

所以说，钢琴谱其实并非一个连续的频谱，而是很多在时间上离散的频率，但是这样的贴切的比喻真的是很难找出第二个来了。

因此在傅里叶变换在频域上就从离散谱变成了连续谱。那么连续谱是什么样子呢？

### 你见过大海么？

为了方便大家对比，我们这次从另一个角度来看频谱，还是傅里叶级数中用到最多的那幅图，我们从频率较高的方向看。

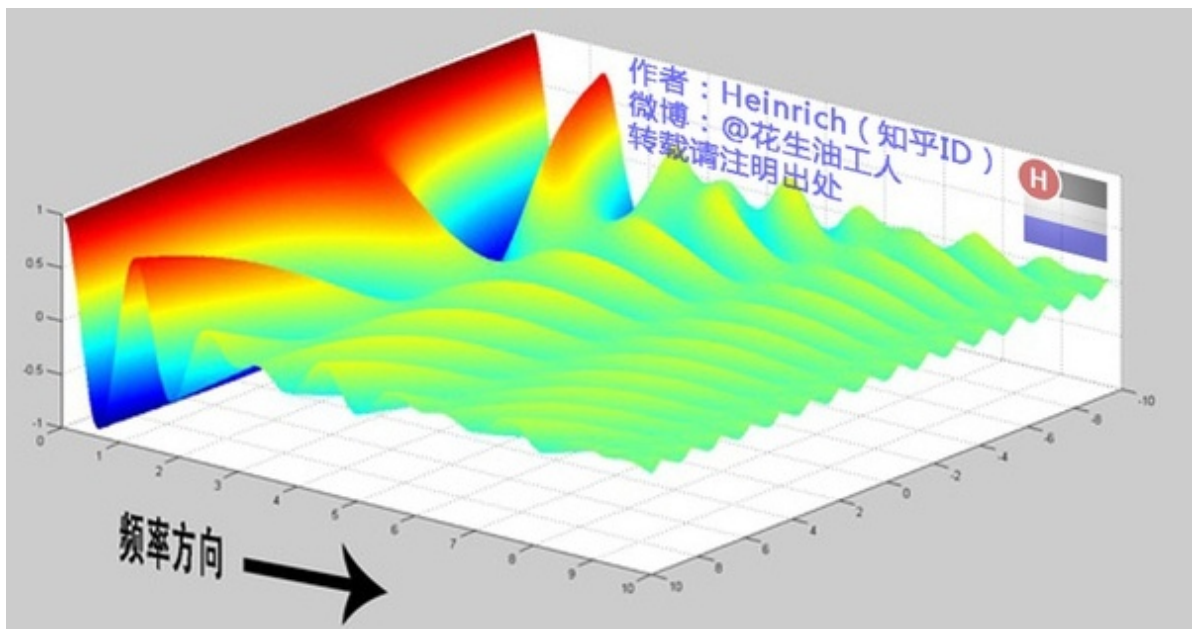


以上是离散谱，那么连续谱是什么样子呢？

尽情的发挥你的想象，想象这些离散的正弦波离得越来越近，逐渐变得连续.....

直到变得像波涛起伏的大海：





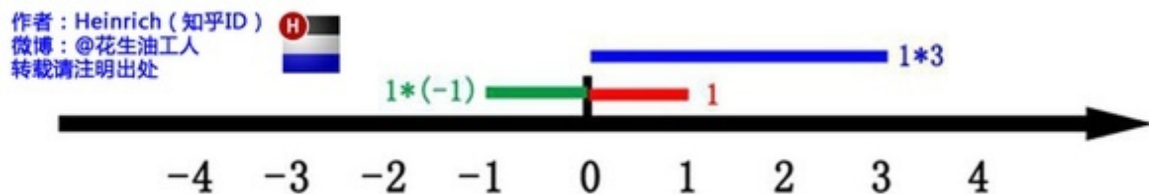
很抱歉，为了能让这些波浪更清晰的看到，我没有选用正确的计算参数，而是选择了一些让图片更美观的参数，不然这图看起来就像屎一样了。

不过通过这样两幅图去比较，大家应该可以理解如何从离散谱变成了连续谱的了吧？原来离散谱的叠加，变成了连续谱的累积。所以在计算上也从求和符号变成了积分符号。

不过，这个故事还没有讲完，接下去，我保证让你看到一幅比上图更美丽壮观的图片，但是这里需要介绍到一个数学工具才能故事继续，这个工具就是——

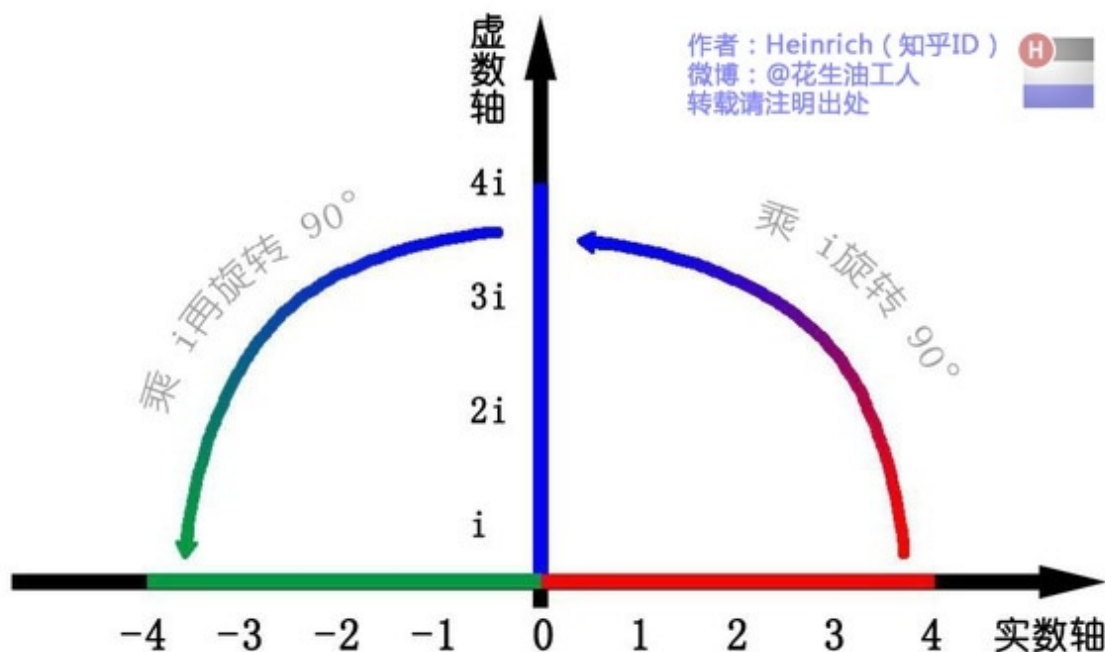
## 五、宇宙耍帅第一公式：欧拉公式

虚数*i*这个概念大家在高中就接触过，但那时我们只知道它是-1 的平方根，可是它真正的意义是什么呢？



这里有一条数轴，在数轴上有一个红色的线段，它的长度是1。当它乘以 3 的时候，它的长度发生了变化，变成了蓝色的线段，而当它乘以-1 的时候，就变成了绿色的线段，或者说线段在数轴上围绕原点旋转了 180 度。

我们知道乘-1 其实就是乘了两次 *i* 使线段旋转了 180 度，那么乘一次 *i* 呢——答案很简单——旋转了 90 度。



同时，我们获得了一个垂直的虚数轴。实数轴与虚数轴共同构成了一个复数的平面，也称复平面。这样我们就了解到，乘虚数 $i$ 的一个功能——旋转。

现在，就有请宇宙第一耍帅公式欧拉公式隆重登场——

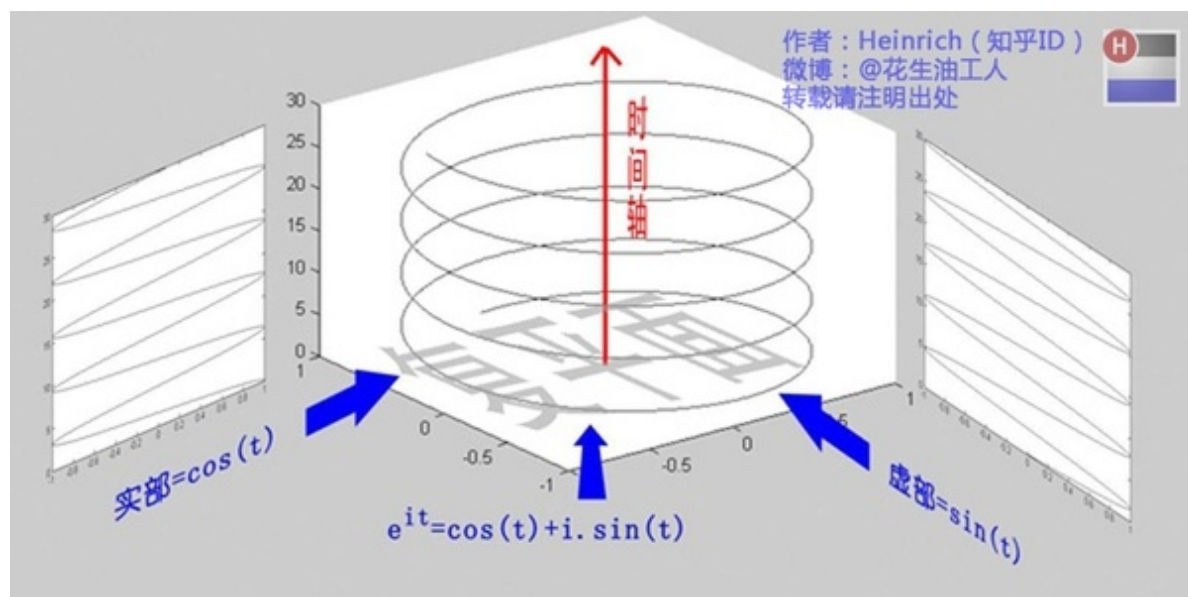
$$e^{ix} = \cos x + i \sin x$$

这个公式在数学领域的意义要远大于傅里叶分析，但是乘它为宇宙第一耍帅公式是因为它的特殊形式——当 $x$ 等于 $\pi$ 的时候。

$$e^{i\pi} + 1 = 0$$

经常有理工科的学生为了跟妹子表现自己的学术功底，用这个公式来给妹子解释数学之美：“石榴姐你看，这个公式里既有自然底数 $e$ ，自然数 $1$ 和 $0$ ，虚数 $i$ 还有圆周率 $\pi$ ，它是这么简洁，这么美丽啊！”但是姑娘们心里往往只有一句话：“臭屌丝……”

这个公式关键的作用，是将正弦波统一成了简单的指数形式。我们来看看图像上的涵义：



欧拉公式所描绘的，是一个随着时间变化，在复平面上做圆周运动的点，随着时间的改变，在时间轴上就成了一条螺旋线。如果只看它的实数部分，也就是螺旋线在左侧的投影，就是一个最基础的余弦函数。而右侧的投影则是一个正弦函数。

关于复数更深的理解，大家可以参考：

## [复数的物理意义是什么？](#)

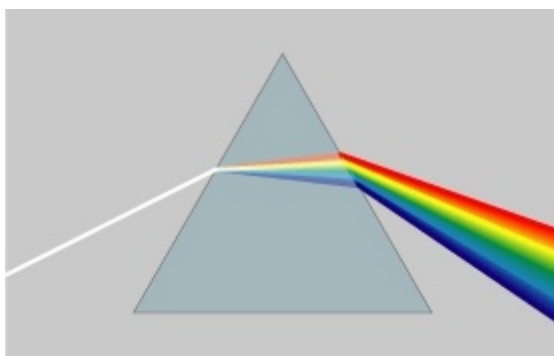
这里不需要讲的太复杂，足够让大家理解后面的内容就可以了。

## 六、指数形式的傅里叶变换

有了欧拉公式的帮助，我们便知道：正弦波的叠加，也可以理解为螺旋线的叠加在实数空间的投影。而螺旋线的叠加如果用一个形象的栗子来理解是什么呢？

### 光波

高中时我们就学过，自然光是由不同颜色的光叠加而成的，而最著名的实验就是牛顿师傅的三棱镜实验：



所以其实我们在很早就接触到了光的频谱，只是并没有了解频谱更重要的意义。

但不同的是，傅里叶变换出来的频谱不仅仅是可见光这样频率范围有限的叠加，而是频率从 0 到无穷所有频率的组合。

这里，我们可以用两种方法来理解正弦波：

第一种前面已经讲过了，就是螺旋线在实轴的投影。

另一种需要借助欧拉公式的另一种形式去理解：

$$\begin{aligned}e^{it} &= \cos(t) + i.\sin(t) \\ e^{-it} &= \cos(t) - i.\sin(t)\end{aligned}$$

将以上两式相加再除2，得到：

$$\cos(t) = \frac{e^{it} + e^{-it}}{2}$$

这个式子可以怎么理解呢？

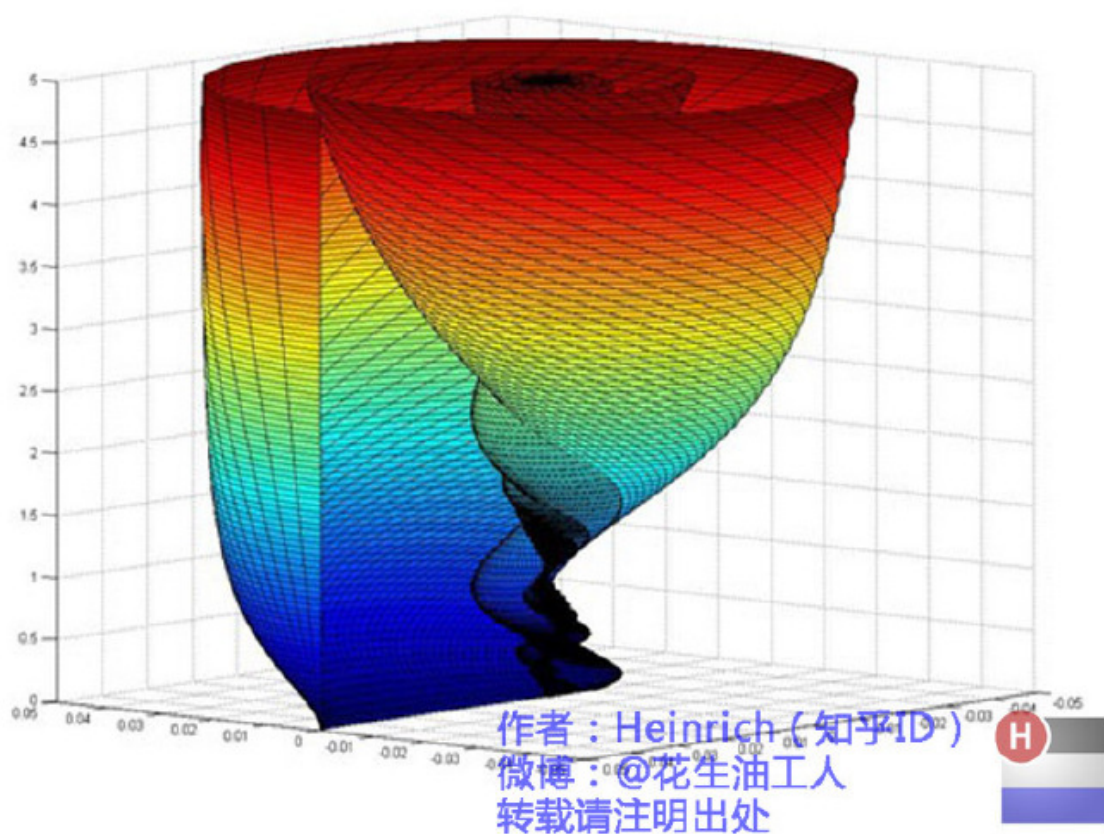
我们刚才讲过， $e^{it}$ 可以理解为一个逆时针旋转的螺旋线，那么 $e^{-it}$ 则可以理解为一个顺时针旋转的螺旋线。而  $\cos(t)$ 则是这两条旋转方向不同的螺旋线叠加的一半，因为这两条螺旋线的虚数部分相互抵消掉了！

举个例子的话，就是极化方向不同的两束光波，磁场抵消，电场加倍。

这里，逆时针旋转的我们称为正频率，而顺时针旋转的我们称为负频率（注意不是复频率）。

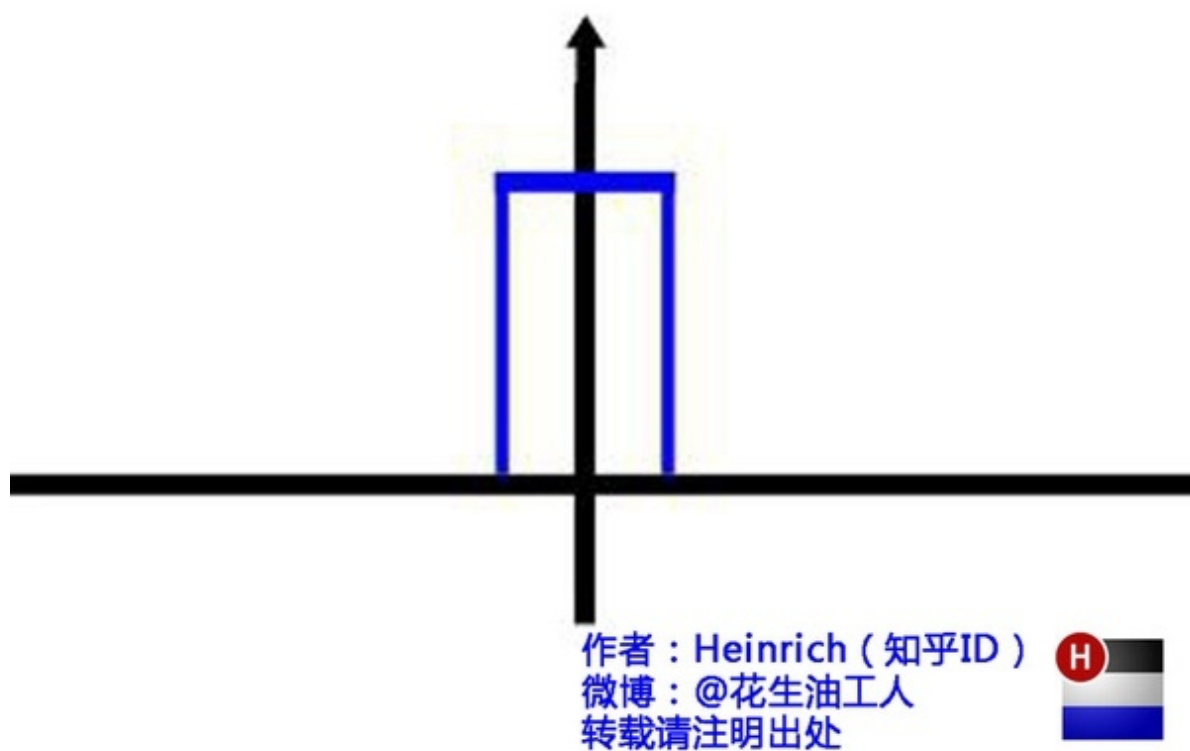
好了，刚才我们已经看到了大海——连续的傅里叶变换频谱，现在想一想，连续的螺旋线会是什么样子：

想象一下再往下翻：



是不是很漂亮？

你猜猜，这个图形在时域是什么样子？



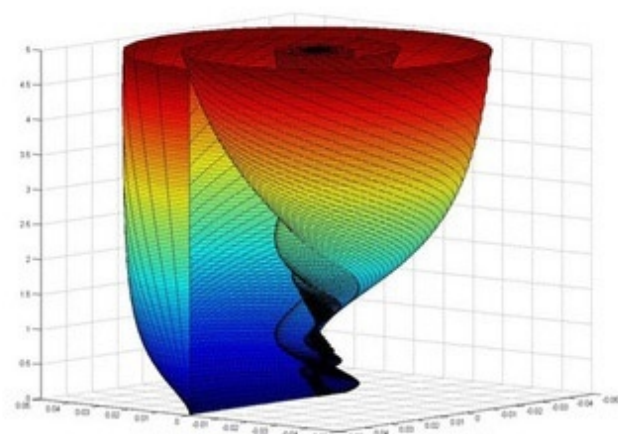
哈哈，是不是觉得被狠狠扇了一个耳光。数学就是这么一个把简单的问题搞得很复杂的东西。

顺便说一句，那个像大海螺一样的图，为了方便观看，我仅仅展示了其中正频率的部分，负频率的部分没有显示出来。

如果你认真去看，海螺图上的每一条螺旋线都是可以清楚的看到的，每一条螺旋线都有着不同的振幅（旋转半径），频率（旋转周期）以及相位。而将所有螺旋线连成平面，就是这幅海螺图了。

好了，讲到这里，相信大家对傅里叶变换以及傅里叶级数都有了一个形象的理解了，我们最后用一张图来总结一下：

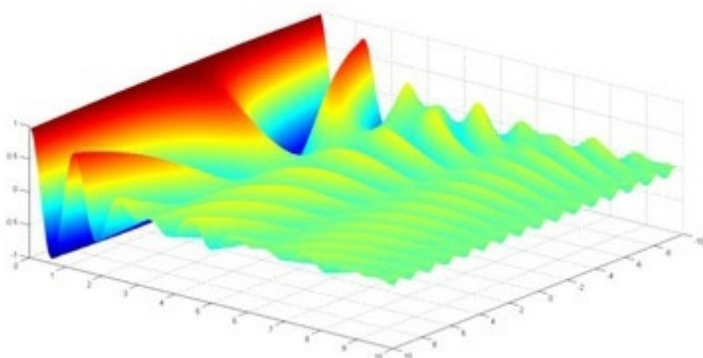




复频域



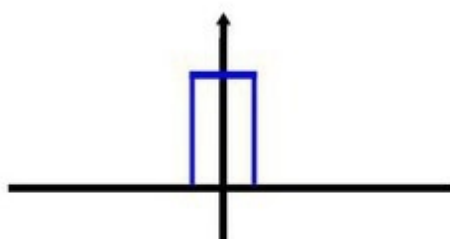
投影到实数空间



频域



各频率成分累积



时域

作者：Heinrich（知乎ID）  
 微博：@花生油工人  
 转载请注明出处



好了，傅里叶的故事终于讲完了，下面来讲讲我的故事：

这篇文章第一次被卸下来的地方你们绝对猜不到在哪，是在一张高数考试的卷子上。当时为了刷分，我重修了高数（上），但是后来时间紧压根没复习，所以我就抱着裸考的心态去了考场。但是到了考场我突然意识到，无论如何我都不会比上次考的更好了，所以干脆写一些自己对于数学的想法吧。于是用了一个小时左右的时间在试卷上洋洋洒洒写了本文的第一草稿。

你们猜我的了多少分？

6分

没错，就是这个数字。而这6分的成绩是因为最后我实在无聊，把选择题全部填上了C，应该是中了两道，得到了这宝贵的6分。说真的，我很希望那张卷子还在，但是应该不太可能了。

那么你们猜猜我第一次信号与系统考了多少分呢？

45分

没错，刚刚够参加补考的。但是我心一横没去考，决定重修。因为那个学期在忙其他事情，学习真的就抛在脑后了。但是我知道这是一门很重要的课，无论如何我要吃透它。说真的，信号与系统这门课几乎是大部分工科课程的基础，尤其是通信专业。

在重修的过程中，我仔细分析了每一个公式，试图给这个公式以一个直观的理解。虽然我知道对于研究数学的人来说，这样的学习方法完全没有前途可言，因为随着概念愈加抽象，维度越来越高，这种图像或者模型理解法将完全丧失作用。但是对于一个工科生来说，足够了。

后来来了德国，这边学校要求我重修信号与系统时，我彻底无语了。但是没办法，德国人有时对中国人就是有种藐视，觉得你的教育不靠谱。所以没办法，再来一遍吧。

这次，我考了满分，而及格率只有一半。

老实说，数学工具对于工科生和对于理科生来说，意义是完全不同的。工科生只要理解了，会用，会查，就足够了。但是很多高校却将这些重要的数学课程教给数学系的老师去教。这样就出现一个问题，数学老师讲得天花乱坠，又是推理又是证明，但是学生心里就只有一句话：学这货到底干嘛用的？

缺少了目标的教育是彻底的失败。

在开始学习一门数学工具的时候，学生完全不知道这个工具的作用，现实涵义。而教材上有只有晦涩难懂，定语就二十几个字的概念以及看了就眼晕的公式。能学出兴趣来就怪了！

好在我很幸运，遇到了大连海事大学的吴楠老师。他的课全程来看是两条线索，一条从上而下，一条从下而上。先将本门课程的意义，然后指出这门课程中会遇到哪样的问题，让学生知道自己学习的某种知识在现实中扮演的角色。然后再从基础讲起，梳理知识树，直到延伸到另一条线索中提出的问题，完美的衔接在一起！

这样的教学模式，我想才是大学里应该出现的。

最后，写给所有给我点赞并留言的同学。真的谢谢大家的支持，也很抱歉不能一一回复。因为知乎专栏的留言要逐次加载，为了看到最后一条要点很多次加载。当然我都坚持看完了，只是没办法一一回复。

本文只是介绍了一种对傅里叶分析新颖的理解方法，对于求学，还是要踏踏实实弄清楚公式和概念，学习，真的没有捷径。但至少通过本文，我希望能让这条漫长的路变得有意思一些。