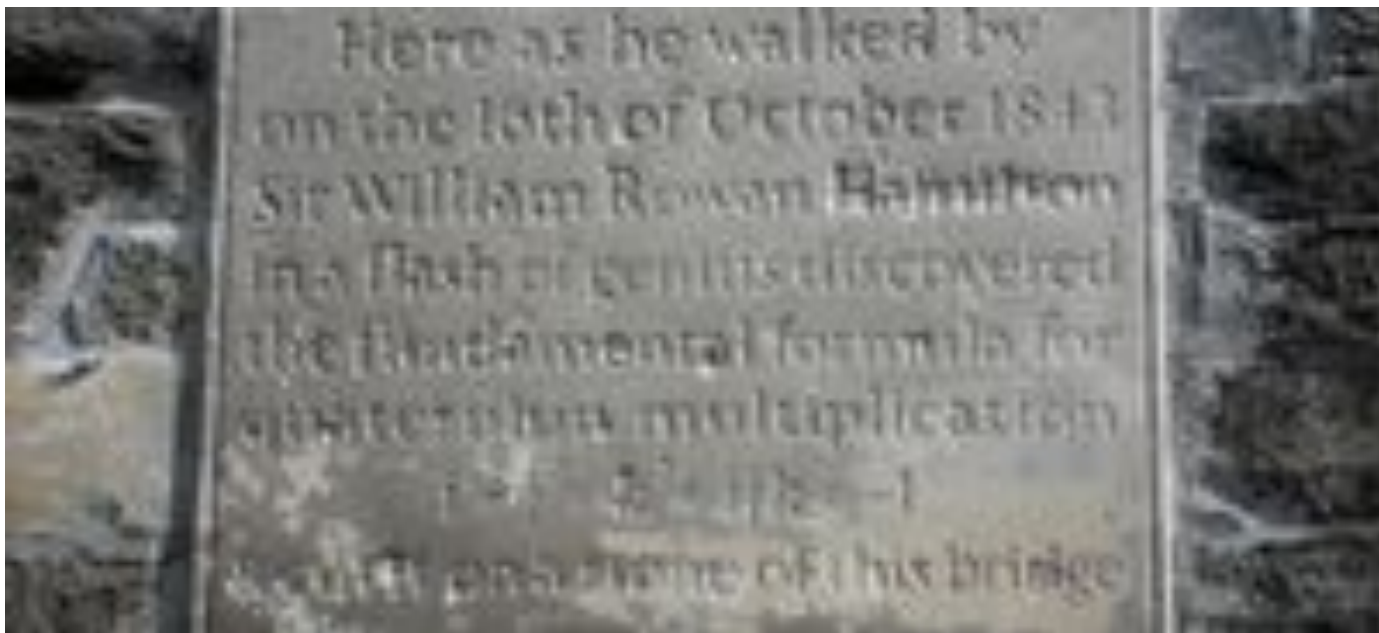


zhuanlan.zhihu.com /p/27471300

四元数——基本概念



杨智为

四元数系列:

[四元数——基本概念](#)

[四元数——旋转](#)

[四元数应用——转矩阵、Slerp插值与万向节](#)

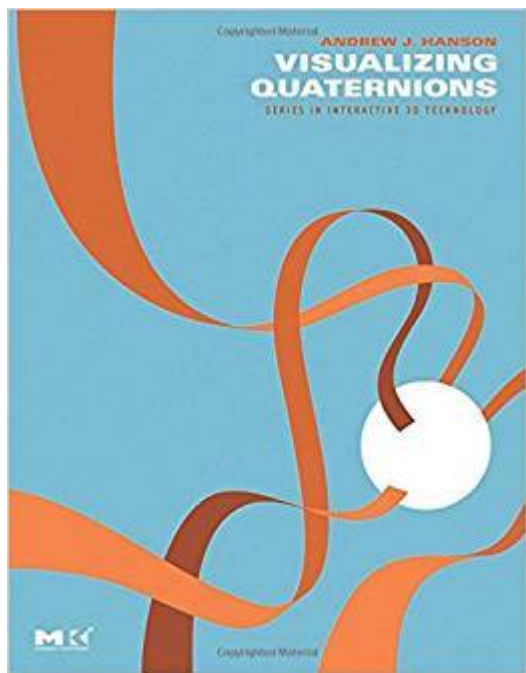
[四元数应用——顺序无关的旋转混合](#)

导语：相比矩阵，用四元数处理3D旋转的优势是毋庸置疑的，但由于概念复杂，难于理解，一直是我头疼的问题。最近忙里偷闲，翻了些材料，感觉醍醐灌顶、豁然开朗，这里写点心得，也就抛个砖，引个玉，希望能帮到和我一样困惑的人。

0.先甩干货

不得不说，目前国内了解这方面知识的渠道少之又少。知乎上也有类似问题——[如何形象地理解四元数](#)，干货私货一大堆。这里先推荐两个比较有用的资料，算是相辅相成。搞懂了这些，我觉得针对图形学来说就差不多够用了。

首先推荐一本书《**Visualizing Quaternions**》。这书算是本数学书，不过里面也有不少实现代码，估计受众就是程序员。本书顾名思义，主要介绍如何将四元数可视化，所以全书主要重心是研究四元数各种几何意义，虽然提到旋转，但并没有牵扯一个核心问题，如何用单位四元数去旋转一个三维向量，好在介绍这个的文章一抓一大把。这本书也不需要都看，我是挑着看的，个人觉得通读**30%-40%**的内容就够用了，其余部分比较复杂，内容也比较深，可以有需求针对翻阅。顺带一提国内此书价格一般在**1000+**，而美亚二手书一般**60-70**刀，喜欢收实体书的可以参考一下。

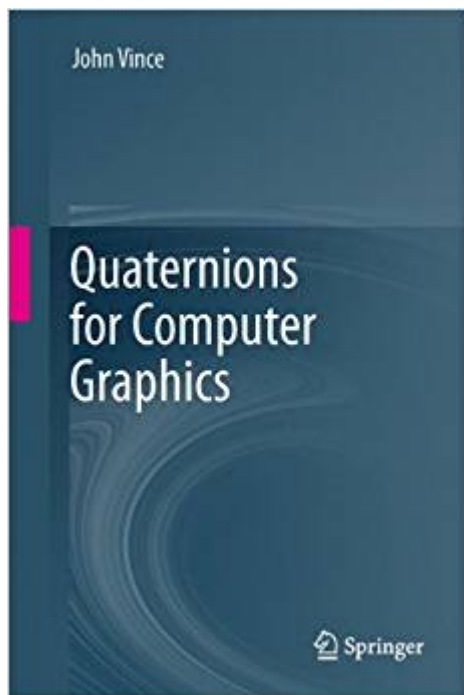


下面推荐一篇文章《**Understanding Quaternions**》，这篇文章相比

大多数技术博客来说，写的已经非常全面详尽。但是对于复数和四元数的关联，四元数的指数对数运算，还有一些四元数的理解显然由于篇幅原因没有说的太清晰，个人觉得这部分参考上面推荐的书就行了，该文章最核心的部分在于对旋转的描述，通过实例，极为清晰地介绍了四元数怎么去旋转一个三维向量。下面给出文章的链接。

英文原版：[Understanding Quaternions](#) 中文翻译：[理解四元数](#)

最后推荐另一本书《**Quaternions For Computer Graphics**》。上面的文章主要引用自这本书，这本书对于四元数如何应用于旋转，从原理到实践都说的非常详细，正好弥补第一本书籍的不足。而且《**Understanding Quaternions**》有部分也说的不是很明白，但都能在这本书上面找到对应的解释。



这里说一下，其实上面两本读物就已经够了。本文拟了这篇文章，更多是给自己写个归纳笔记，可能大多以概念体会为主，而不是传统意义上的手把手教程。所以，尽量还是以推荐的干货读物为主，许多基础的公式推算本文是不涉及的。

1.为什么要用四元数

可能四元数的由来大家都看过很多遍。很久以前，一位老者坐在大桥边上，看着过往船只，突然灵光一闪，在桥边石碑上洋洋洒洒刻上几行大字，四元数诞生了！故事大家都爱听，那么为什么我们需要四元数？一种说法是解决向量乘法，我们知道向量之间乘法有内积和外积，但这两个运算均不完美，即不满足群的条件（当然四元数诞生的时候也还没有内积外积的说法）。那向量之间是否存在这样一个非常完美的乘法，于是三维空间无法解决的问题就映射到四维空间。这便是四元数诞生的契机。

那么问题又来了，既然四元数只是为了解决矩阵乘法，那为什么我们现在要用四元数进行旋转，甚至替代了欧拉角、轴角等形式？首先，四元数并不是生来为了解决三维旋转，而是它的性质非常有利于表达旋转信息（后面会详述），所以了解四元数的性质要先于了解四元数在旋转中的应用。至于四元数替代欧拉角等形式，就需要牵扯到一些别的知识点，我先罗列一下四元数相比其他形式的优点：

- 解决万向节死锁（Gimbal Lock）问题
- 仅需存储4个浮点数，相比矩阵更加轻量
- 四元数无论是求逆、串联等操作，相比矩阵更加高效

所以综合考虑，现在主流游戏或动画引擎都会以缩放向量+旋转四元数+平移向量的形式进行存储角色的运动数据。

2.四元数是怎么想出来的

平庸的教程会直接提出四元数的定义、运算规则等等，然后读者不知所云。相反，较为系统的教程一般会从复数（Complex Number）进行引导，逐步提出四元数的定义，这样会让读者更容易理解，在脑中也更好形成画面。那复数与四元数之间的关系、以及如何从复数这样一个概念扩展到四元数是我们需要理清的一个思路。

先说几个概念。

空间中的子空间：一般而言，空间（维度 >2 ）都存在更低维的子空间，比如二维空间中一维子空间，也就是直线；三维空间中的一维子空间和二维子空间，也就是直线和面。当超过三维的概念我们就很难去想象是什么样子，但四维空间一定会存在三维子空间或二维子空间。我们这里会加一个前缀超（hyper）来形容这些空间，比如高维空间中的平面我们称为超平面。

空间和子空间的映射：我们将二维空间表示为 (x,y) ，当 $y=0$ 时，其实可以看成是一维的，只不过它表示成 $(x,0)$ 这种形式。推到四维， (w,x,y,z) ，当 $w=0$ 时， $(0,x,y,z)$ 就是一个三维子空间，这也是为什么我们可以用单位四元数对三维向量进行操作，其实我们是将三维向量映射到四维的三维子空间（ $w=0$ ，这种形式也成纯四元数），然后对其进行旋转，最终得到的向量结果依然是这个三维子空间中的，因而可以映射回三维空间。

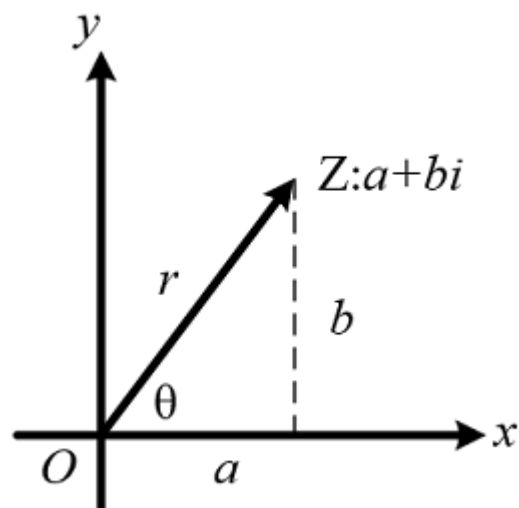
广义球：这里的球是广义上的。我们在二维平面上，广义球其实指代circle，三维空间上就是我们认知上的球，称为two-sphere，而四维空间中广义球其实是一个超球（hyper-sphere），又称为three-sphere。单位向量其实就是广义球上面的点，而单位四元数也就是three-sphere上面的点。

约束与特征向量：空间中的一点由 x, y, z 等参数来表示，一般来说参数的数量与维数相等，二维空间的点用 $\{x, y\}$ 参数，四维空间的点用 $\{x, y, z, w\}$ 参数。但是对于空间的点加以约束，则该会减少参数的数量，比如三维空间的点在某一单位球面上，原本三个参数 $\{x, y, z\}$ 才能表达的点现在只需要两个参数 $\{u, v\}$ 就可以表达。如果 $\{u, v\}$ 是单位向量，也可以称 $\{u, v\}$ 是 $\{x, y, z\}$ 的特征向量。

上述概念给大家一个思路，四元数这样一个东西并不是一蹴而就的，从空间来说，它与我们熟知的低维空间本质上是没有什么区别的，或者说是有很大共性的。四元数的很多特性都是从低维拓展而来的，更具体的说是从复数这一概念拓展的。

那什么是复数？

复数概念很简单，它其实就是为了满足数学家的强迫症， -1 的平方根要有意义。看似很没有意义的复数，在实际中却用处非凡。我们首先看一下它的几何表达形式，令 x 为实部， y 为虚部，则 $x+iy$ 可以表示为下图。



复数的提出将原本一维的数值范围直接加了一个维度，数据量的增幅可以说是巨大的。那我们从极坐标的角度去思考复数， $x+i \cdot y$ 可以写成 $r \cdot \cos\theta + i \cdot r \cdot \sin\theta$ ，我们只考虑单位复数， $|r|=1$ ，也就得到 $\cos\theta + i \cdot \sin\theta$ 。说到这，就要提一个代数学中常用的公式——欧拉公式，基本上复数的所有运算推到都会用到这个公式。

$$\begin{aligned} z &= x + iy \\ &= r \cos \theta + ir \sin \theta \\ &= r e^{i\theta} \end{aligned}$$

接下来我先不说复数的乘法，单位复数乘法可以达到一个二维旋转的效果，有兴趣的可以在这里画一画，推一推。我先介绍如何从复数推导出四元数，正如复数是有一个实部和一个虚部组成的，那我们将一个虚部换成三个虚部，即两两相交 $\{i, j, k\}$ 。

$$\begin{aligned} e^{i \cdot \hat{n} \theta / 2} &= \cos(\theta / 2) + i \cdot \hat{n} \sin(\theta / 2) \\ &= q_0 + i \cdot q \\ &= q_0 + i q_1 + j q_2 + k q_3, \end{aligned}$$

其中 \hat{n} 为三维的单位向量， $i^2=j^2=k^2=i \cdot j \cdot k=-1$ 。这便是四元数的常规表达形式，不过单位四元数是有一大堆的约束的，并不是所有四维向量都是四元数。

这里再开一个脑洞，可能大家听说四元数、八元数，但是存在三元数、五元数？其实从客观上说，三元数、四元数、 n 元数都是存在的（也就是复数存在无穷个等级），但并不是所有的数系都满足模运算。而且随着维度的提高，特性会逐步牺牲，相比复数运算，四元数牺牲了交换律；相比四元数运算，八元数牺牲了结合律。当然也还存在十六元数，不过十六元数的特性将更少，在数学中也就更加没有意义，所以目前就四元数的应用最

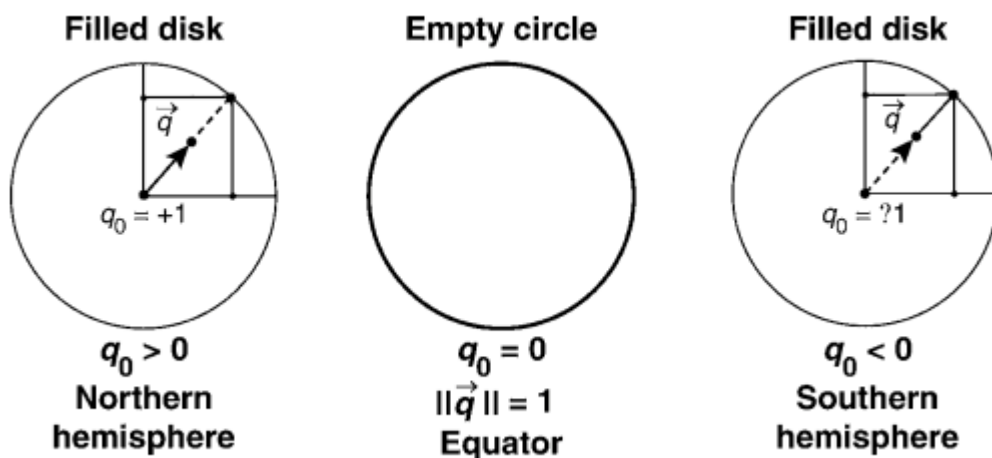
为广泛。

3. 如何去理解四元数

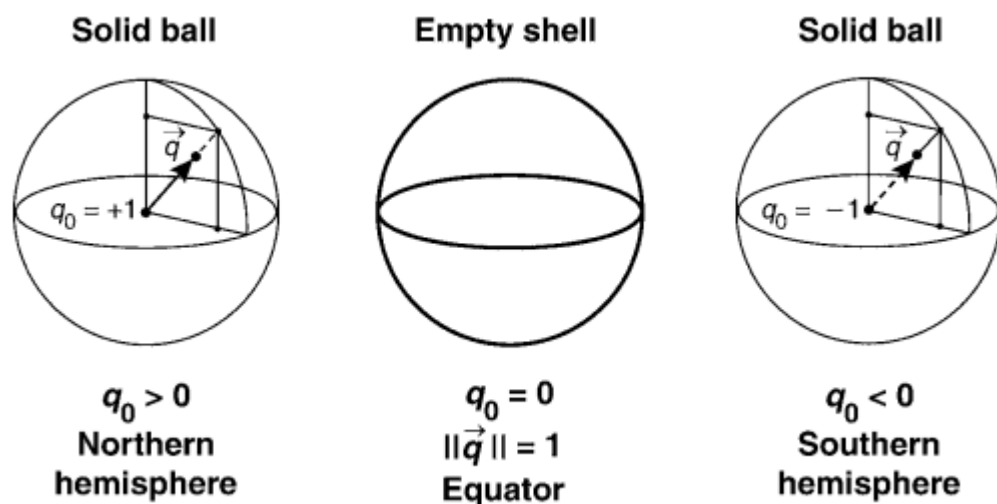
现在我们总结一下已经得出的四元数特性：

- 四元数（以后不特指四元数=单位四元数）是四维空间中一个超球上面的点，满足 $w^2+x^2+y^2+z^2=1$ ；而纯四元数是四维空间在 $w=0$ 时的一个子空间的点，形式为 $\{0, \mathbf{q}\}$ ，特别注意的是纯四元数与四元数是不同的概念。
- 四元数是复数虚部扩展的结果，复数的虚部为1个，而四元数虚部有3个，且两两互相正交，其中实部是 $\cos\theta/2$ ，而虚部为一个单位轴乘以 $\sin\theta/2$ 。
- 四元数自由度并没有四个维度，由于存在 $w^2+x^2+y^2+z^2=1$ 这个约束，它的自由度其实只有3，且每个四元数可以对应一个特征向量，即 \mathbf{n} 。但请记住四元数并不是与特征向量一一对应的，后文会有说。

由于四元数存在于四维空间，所以如何利用低维信息去理解高维信息就显得尤为重要。我们这里先用三维举个例子，三维的球用代数表示为 $x^2+y^2+z^2=1$ ，虽然球上面的点是由 x, y, z 三个参数来确定，但实际上我们只需要两个。假设取 x 和 z 表示，其中 y 可以通过 x 和 z 进行求解。那么，我们将 y 轴信息给隐去，只看投影平面，如下图所示。这张图的意思是，整个球在 XOZ 平面上投影是一个圆，当球面一点投影在圆上时， $y=0$ ；投影的位置位于圆内时，则分别两种情况， $y>0$ 处于北半球， $y<0$ 处于南半球。所以我们仅通过投影后的圆即可还原出整个球体。



让我们推广到四维， $w^2+x^2+y^2+z^2=1$ 中取 x, y 和 z 来表示超球。如下图所示，四维空间投影到三维超平面（ $w=0$ ）可能是一个two-sphere。当投影点在整个two-sphere的边缘时， w 一定为0，值得一提的是在这个空间内的四元数是一个纯四元数。当投影点落在two-sphere的内部时，也分为两种情况， $w>0$ 和 $w<0$ 。但是我们可以发现这两种情况下对应的特征向量是一样的，所以我们将旋转矩阵向四元数转换时，是有两个对应值的，四元数的范围是2倍覆盖于3D旋转（2:1 mapping）。



上面两段写的相对来说比较玄乎，但我觉得可以对理解四维空间和四元数相当有帮助。本文表述能力有限，如果实在看不懂的，可以翻阅《Visualizing Quaternions》第八章，个人觉得这是书中最为精彩的部分之一。

4. 四元数的“乘法”运算

前文提到了四元数之间的“乘法”规则是满足群的性质，那究竟是怎样的乘法，又是怎么满足的，接下来我们一起来看看。

由于四元数有 i, j, k 三个虚部，所以得满足 $i^2=j^2=k^2=i \cdot j \cdot k=-1$ 这个条件。这里令 \star 为“乘法”操作符，则 $p \star q$ 的公式如下，具体的推导步骤我就不写了。

$$\begin{aligned}
 p \star q &= (p_0, p_1, p_2, p_3) \star (q_0, q_1, q_2, q_3) \\
 &= \begin{bmatrix} p_0 q_0 - p_1 q_1 - p_2 q_2 - p_3 q_3 \\ p_1 q_0 + p_0 q_1 + p_2 q_3 - p_3 q_2 \\ p_2 q_0 + p_0 q_2 + p_3 q_1 - p_1 q_3 \\ p_3 q_0 + p_0 q_3 + p_1 q_2 - p_2 q_1 \end{bmatrix} \\
 &= (p_0 q_0 - \mathbf{p} \cdot \mathbf{q}, p_0 \mathbf{q} + q_0 \mathbf{p} + \mathbf{p} \times \mathbf{q})
 \end{aligned}$$

我们首先针对单位四元数，按照群的性质一条条过：

- 封闭性：易证明， \mathbf{p} 和 \mathbf{p} 的共轭相乘即可， $|\mathbf{p} \star \mathbf{q}|=1$ 。
- 结合律：这条也很好证明，只要证明 $(\mathbf{p} \star \mathbf{q}) \star \mathbf{r} = \mathbf{p} \star (\mathbf{q} \star \mathbf{r})$ 。

以上两个证明这里就不详述了，毕竟知乎也不好编辑公式。我们可以换种思路，从几何的角度思考，由于单位四元数旋转后还是单位四元数，所以满足封闭性。同时，类比矩阵乘法，单位四元数也应该满足结合律。

- 单位元素： $\mathbf{e}=(1,0,0,0)$ ，这也是四元数的一个初始值（相当于单位矩阵）。带入上面的公式可知， $\mathbf{p} \star (1,0,0,0)$ 还是等于 \mathbf{p} 。

- 逆元素：p存在一单位四元数，★操作的结果为e，具体参考下面公式，可以看出逆元素就是其共轭除以模的平方。

$$q^* q q^{-1} = q^*$$

$$|q|^2 q^{-1} = q^*$$

如果这里是复数，可能还会满足交换律，就是一个阿贝尔群。但四元数肯定不满足交换律，无论从代数的角度，还是从几何角度。到这里，基本也就介绍完了四元数的乘法运算，如果用其指数形式（exp形式），证明这些性质就会更简单。

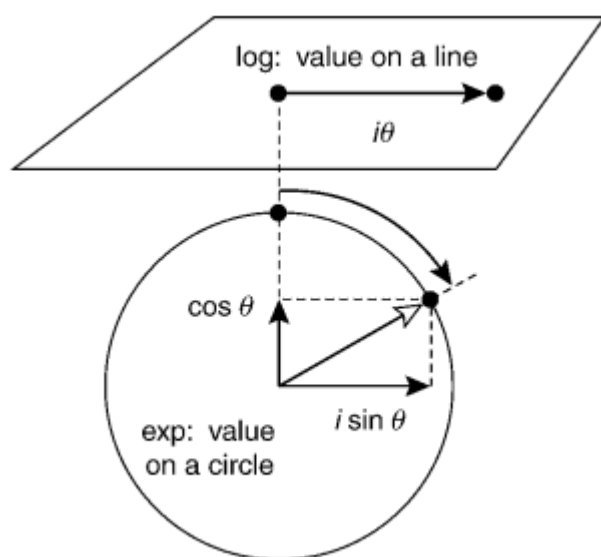
5. 四元数的指数、对数与内积

根据欧拉公式可知，四元数的指数形式只是换了一种表达方式，都是表达同一个东西。而其对数形式也就是四元数指数形式的指数，如下图公式所示。这对公式在计算四元数插值的时候很有用，这个后文再说。

$$e^{(0, \hat{n} \frac{\theta}{2})} = \left(\cos \frac{\theta}{2}, \hat{n} \sin \frac{\theta}{2} \right)$$

$$\log \left(\cos \frac{\theta}{2}, \hat{n} \sin \frac{\theta}{2} \right) = \left(0, \hat{n} \frac{\theta}{2} \right)$$

那么对数和指数的关系是什么呢？我们参考下图，其实对数的变化量其实就是角度的大小，所以用下图可以很直观的看出四元数指数形式和对数形式之间的关系。



最后介绍一下四元数的内积。我们先看三维空间或二维空间，如果两个单位向量算内积，内积的结果为两个向量的夹角。四元数也是如此，两个单位四元数的内积就是其夹角，这个结论推广到n维也是成立，原因很简单，因为三点共面，即使n维空间，两个向量所在的空间也仅是二维子空间。至于外积，由于对于复数和四元

数来说，外积的意义均不是很明显，这里就不做过多讨论。

关于旋转部分就下次再说吧！