第五讲 椭圆曲线密码学

前言

对大多数初学者而言，密码学都是研究通过各种运算实现加密、解密的，而椭圆曲线是几何中要研究的，两个八竿子打不着的东西之间怎么会联系在一起呢？今天我们就来介绍一下什么是椭圆曲线密码学。

欧氏几何

这一切要从几何开始说起。我们都知道几何学的鼻祖是欧几里得，这个生活在2000多年前的哲人在《几何原本》中创立了欧氏几何。欧几里得最伟大的贡献不仅是欧氏几何的内容本身，而且是他首创了公理化的数学研究方法。在此之前的数学研究都是基于直觉的，也就是在直觉上认为某些命题正确的，继而推导出其它不那么直观的结论。而在公理化的数学研究方法中，需要从一理论体系中提炼出若干无法由其它命题证明的命题，称之为公理，然后在公理集上推演出整个理论体系中的其它结论，如此便使得数学研究变得更加准确和严谨。欧式几何本身也正是如此。欧式几何的公理共包括以下五条。

1. 任意两个点可以通过一条直线连接；
2. 任意线段可以无限延长成一条直线；
3. 给定任意线段可以以其一端为圆心，该线段为半径作一个圆；
4. 所有直角都全等；
5. 若两条直线都与第三条直线相交，并且在同一边的内角之和小于两个直角和，则这两条直线在这一边必定相交，否则不相交。

第五条公理又称为平行公理。由平行公理可以推导出以下结论。

> *通过一个不再直线上的点，有且仅有一条不与该直线相交的直线。*

事实上，这个结论即给出了平行线的定义，即**平行线不相交。**

欧氏几何的所有结论都可以由这五条公理推导得出，也就是说整个欧式几何都是建立在这五条公理的基础之上的。

黎曼几何

现在我们做一个大胆的创想：把公理改一改，岂不是可以得到和欧式几何不同的另外一套几何？

这种几何就是非欧几何。

例如罗巴切夫斯基和鲍耶创建的罗氏几何，和黎曼创建的黎曼几何。

第一个这么做的人是叫黎曼，黎曼19世纪世纪的数学家，前段时间还刚刚刷了屏，因为他提出的著名的黎曼猜想被证明了。

言归正传，黎曼把第五条公理改了，改成了

> *任意两条直线都相交于无穷远点。*

也就是说，**平行线不存在了。**

你肯定会想，“嗯？那怎么可能呢？平行线明明是有的，你看我随手就能画出两条平行线，他们无论怎么延长都不可能相交。”

可是你怎么知道他们无限延长之后不会相交呢？

那是因为我们小时候老师告诉我们，他们无限延长之后也不会相交，因为他们是平行线，所以这种固有认识就固定在我们的脑子里了。

可实际上，当他们无限延长之后，究竟是什么样子，我们谁也不知道，谁也不能站在无穷远处看一看他们到底有没有相交。

所以他们到底有没有相交，只能取决于我们的想象。

再此之前，你都是想象它们不相交。**现在，你何不大胆想象一下，在我们的新几何中，任意无限延长的直线都相交于一个无穷远点。**

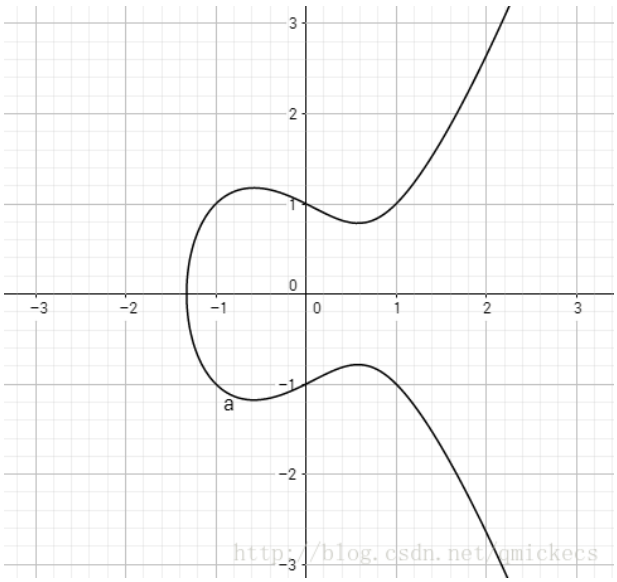
这就像是我们都生活在一个大气球的表面上，气球的尾巴被拉的很长很长，一直到无穷远的地方，我们在这个表面上画出的任何一条无限延展的直线，都一定会经过气球的尾巴，也就是哪个无穷远点。

这个新几何就是黎曼几何。

黎曼几何椭圆曲线

实数域椭圆曲线

现在我们考虑一个方程：$y^2 = x^3 - x + 1$，满足这个方程的(x, y)描述的坐标点显然会形成一条曲线。



这条曲线会向着右上方和右下方无限延伸。这也没什么新奇的，只不过，如果这是在**黎曼几何**中，那么这条无限延展的曲线一定会经过所谓的**无穷远点**，这个无穷远点，我们暂且用符号0来表示。

这条曲线就称为椭圆曲线，当然这条椭圆曲线上有无穷多个点。

> *这里先问个题外话，你知道为什么$2 + 3 = 5$吗?*
>
> *你可能会说，这还用问吗？加法就是这么定义的!*
>
> *没错，加法就是这么定义的。根据加法的定义，给它两个整数，它会输出另一个整数。*
>
> *既然可以定义整数上的运算，那我们是不是也可以定义这区曲线上的点之间的加法运算呢？给它两个点A和B，可以得到另外一个点D。*

那么点和点之间的加法运算是怎么定义的呢?

当一条直线与椭圆曲线相交于A、B、C三个点时，有

$$A + B + C = 0 \tag{1}$$

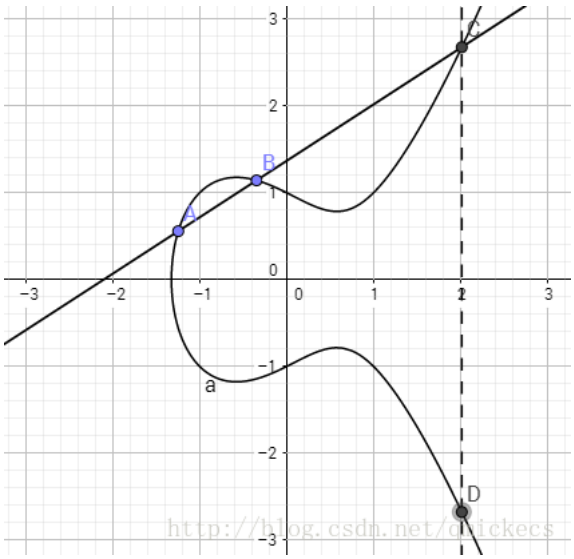
因此有

$$A + B = -C = D \tag{2}$$

其中点D是与点C沿x轴对称的那个点。

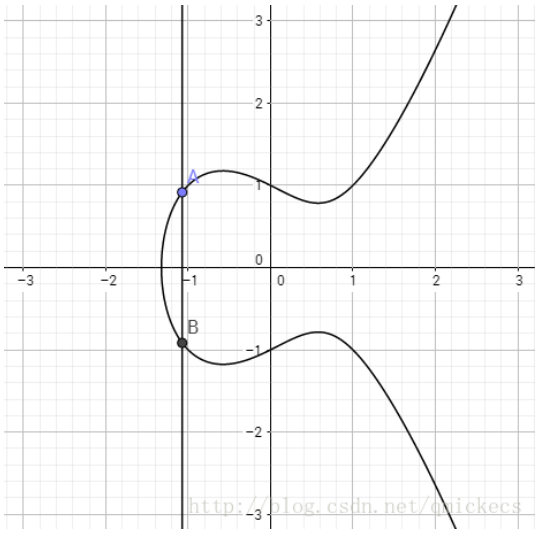
也就是说我们将点A与点B的和定义为点D。

因此当要计算点A和点B的和的时候，首先将点A和点B相连形成一条直线，那么这条直线必然与该椭圆曲线相交于某点点C，再取点C沿x轴对称的点便得到点D。

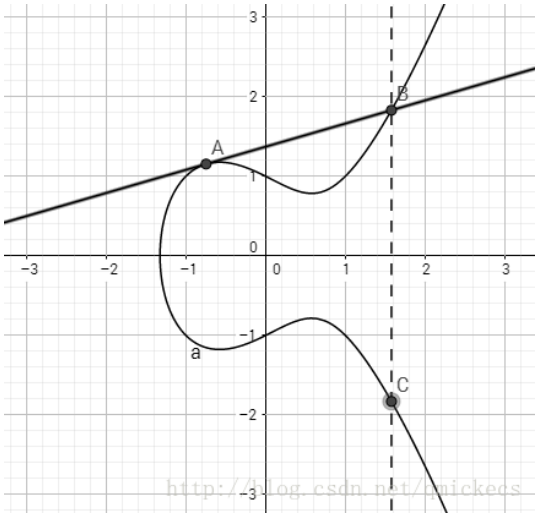


加法运算中有这样几种情况值得专门分析一下。

- 点A和点B相加的结果与顺序无关，即$A+B=B+A$。根据点加法的定义，这个结论是显而易见的。
- 若两个点沿x轴对称，它们怎么相加？例如在上图中，点C和D沿x轴对称，它们的连线与椭圆曲线看起来没有第三个交点啊。别忘了，这是黎曼几何。也就是说直线CD和椭圆曲线还有一个公共点——无穷远点。无穷远点在沿x轴翻转得到的当然还是无穷远点，所以$C+D$等于无穷远点。也就说两个沿x轴对称的点相加等于无穷远点。
- 任意一个点（例如点A）与无穷远点0怎么相加？从上面的分析，我们知道直线AB经过无穷远点0（看下面的图），所以经过点A和无穷远点0的直线当然也经过点B，所以点A与无穷远点的连线与曲线相交于B，而B沿x轴翻转后又会得到点A。所以A与无穷远点相加还是等于A。即$A+0=A$。



- 任意一个点和自己怎么加？一个点A当然无法定位一条直线，但是你可以想象当点A'无限接近于点A时，A和A'定位的直线是与A相切的。所以，要求$A+A$，我们可以画出它的切线，当然，切线也与椭圆曲线相交于某一点，然后我们再将该点B沿x轴翻转得到一个点C，这个点就是$A+A$。



有限域椭圆曲线

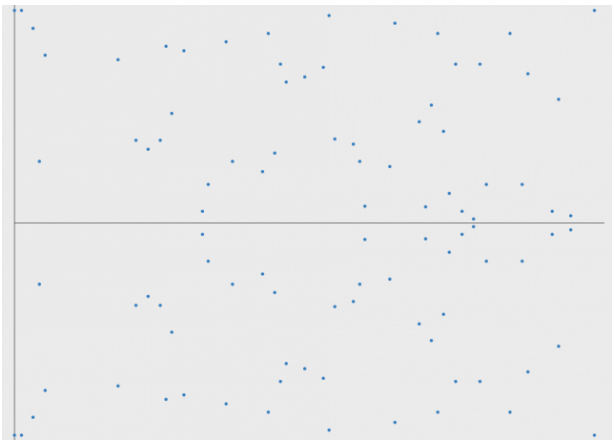
上述实数域上的椭圆曲线和加法的定义还可以扩展到整数有限域上。

例如，我们定义椭圆曲线

$$y^2 \equiv x^3 - x + 1 \pmod{97} \tag{3}$$

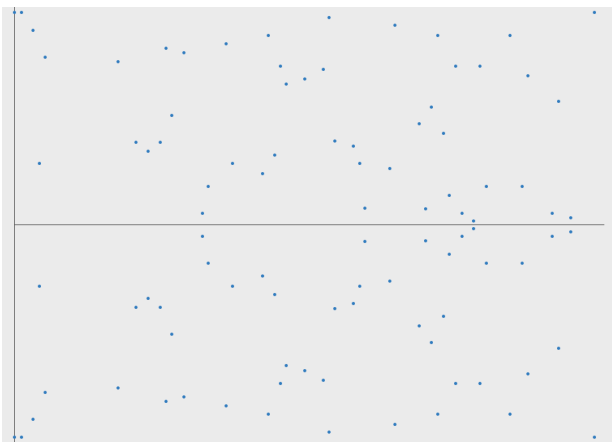
其中x和y都是小于97，大于等于0的整数。

所以该曲线是只有有限多个点，如下图所示。



尽管只有有限多个点，但同样可在这些点上以相同的方式定义点和点之间的加法运算。

例如，点A和点B相加得到点C。



由于椭圆曲线E及其上的加法$+$具有以下性质，所以椭圆曲线及其加法$< E, + >$构成了一种抽象代数结构——阿贝尔群。所谓阿贝尔群，就是同时具备以下性质的集合和运算。

- **运算封闭**：$\forall A, B \in E, A + B \in E$；
- **存在幺元**：$\exists 0 \in E, \forall x \in E, x + 0 = x$；
- **任意元素都有逆元**：$\forall x \in E, \exists y \in E, x + y = 0$；
- **运算满足交换律**：$\forall A, B \in E, A + B = B + A$。
- **运算满足结合律**：$\forall A, B, C \in E, (A + B) + C = A + (B + C)$。

标量乘法

我们还可以在椭圆曲线上定义**标量乘法**。

既然可以计算$A + A$，当然也可以计算$A + A + A, A + A + A + A, \cdots$。如此一来我们就可以定义

$$
\begin{aligned}
A \times 2 &= A + A && (4)\\
A \times 3 &= A + A + A && (5)\\
A \times 4 &= A + A + A + A && (6)\\
&\cdots && (7)
\end{aligned}
$$

如果将点和点之间的加法类比成整数之间的乘法，那么点的标量乘法不就相当于整数上的幂模运算吗？

反过来，点之间的除法运算不就相当于证书上的离散对数运算吗？

运算类比	椭圆曲线	整数
标量乘法 vs. 幂	$A \times 4 = B$	$A^4 \bmod M = B$
除法 vs. 离散对数	$4 = \frac{B}{A}$	$4 \equiv \operatorname{ind}_A B \pmod M$

椭圆曲线密码学

椭圆曲线上的标量乘法和整数上的幂模运算不光在形式上相似，而且在性质上也相近，它们都是一种单向陷门函数。

因此，所有那些基于离散对数的公钥密码通通都存在椭圆曲线上的版本。例如我们之前讲过的DH密钥交换和DSA数字签名方案，对应的分别有椭圆曲线上的DH密钥交换——ECDH和椭圆曲线上的DSA——ECDSA。

基于大整数素因子分解和离散对数的公钥密码用的好好的，为什么要整这么复杂的椭圆曲线密码呢？

这是因为，在同等安全性下，ECC使用的密钥要短得多，因此计算量也小得多。反过来说，就是ECC可以以相对较小的计算量获得较高的安全性。下表是达到相同的安全性，RSA、DSA和ECDSA所需的密钥位数。

基于大整数素因子分解的RSA	基于离散对数的DSA	基于椭圆曲线的公钥密码
512	512	112
1024	1024	160
2048	2048	224
3072	3072	256
7680	7680	384
15360	15360	512

例如，比特币中所使用的数字签名方案就是ECDSA，比特币种所使用的椭圆曲线参数如下。

- 椭圆曲线方程：$y^2 \equiv x^3 + ax + b \bmod p$；
- $p = 2^{256} - 2^{32} - 2^9 - 2^8 - 2^7 - 2^6 - 2^4 - 1$
- $a = 0$
- $b = 7$
- 起始点：
  $G = (0\text{x}79\text{BE}667\text{EF}9\text{DCBBAC}55\text{A}06295\text{CE}870\text{B}07029\text{BFCDB}2\text{DCE}28\text{D}959\text{F}2815\text{B}16\text{F}81798, 0\text{x}483\text{ada}7726\text{a}3\text{c}4655\text{da}4\text{fbfc}0\text{e}11$
- 起始点G的阶：
  $n = 0\text{xFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFEBAAEDCE}6\text{AF}48\text{A}03\text{BBFD}25\text{E}8\text{CD}0364141$

小结

通过这几讲，我们学习了RSA、DH密钥交换、DSA数字签名方案，了解了椭圆曲线密码学是怎么回事，现在应该已经清楚了公钥密码是如何在区块链中解决账户唯一性和身份认证问题的。

从下一讲开始，我们将目光转移到记账上。

当一笔交易被大部分用户认为合法有效时，该由谁记录下来？

该怎样记才能让别人相信这个账本没有被恶意篡改过？

我们为您一一解答。