ECM

2024-12-16

### pollard p-1

如果一个整数的所有素因子都不大于 ，我们称这个数为 光滑数。

假设 是 光滑 的，那么我们记

则

可以证明 ，那么

假设 ，此时我们有一定概率得到 （或者 ）有

但此时由于 是未知的，我们很难给出其一个确定的上界，这里我们计算序列

其中

这里我们按上式逐个求出 ，每步我们仅需要做快速幂操作即可（做 次模乘操作），故总的时间复杂度为

# 设 p-1 为 k-Smooth 数  
def pollard\_Fac\_n(n):  
 a = 2  
 k = 2  
 while True:  
 a = int(pow(a, k, n))  
 q = gcd(a-1, n)  
 if q > 1 and q != n:  
 print(f'[+] find p = {n//q}')  
 print(f'[+] find q = {q}')   
 break  
 k += 1

可以看到上述算法的局限性是要求 相对较小，当 较大时，往往是不满足的。故我们有了一个基于 p-1 方法的改进方法 ECM，其只需要满足 中存在一个数较为光滑即可（这一点将在算法原理中进行证明）。

### 一些数学的知识

首先我们给出一些群论的基础知识和结论以便于理解后面的椭圆曲线，对于结论我们这里不给出详细证明。

#### 群论知识补充

##### 群的性质

椭圆曲线上的点形成一个阿贝尔群(Abelian group)，这意味着点加法和点倍增操作遵循群的性质。具体来说，群的性质包括：

* 封闭性：对于群中的任意两个元素 和 ，它们的和 也是群中的一个元素。
* 结合律：对于群中的任意三个元素 和 ，有
* 单位元：存在一个元素 (无穷远点)，使得对于群中的任意元素 ，有 。
* 逆元：对于群中的任意元素，存在一个元素 ，使得 。

##### 拉格朗日定理

拉格朗日定理是群论中的一个重要定理，它指出在一个有限群中，任何子群的阶（即子群中元素的个数）必须整除整个群的阶。具体来说，如果 是一个有限群，其阶为 ，那么对于 中的任何元素 ，存在一个正整数 使得 ，且 是 的因子。

#### 椭圆曲线

##### 椭圆曲线的基本形式

椭圆曲线 (Elliptic Curve) 是一门古老而内容丰富的数学分支，涉及到较多深奥的算术理论。最早的椭圆曲线方程可追溯到古希腊时期丢番图 (Diophantus，约公元 200-300 年) 出版的著作《算术》一书，书中讨论的基本问题就是整数多项式方程的有理数解。

设 是域，域 上的 Weierstrass 方程 可由如下二元三次方程给出：

其中 。

设 ，即 为实数域，上述 Weierstrass 方程可通过恒等变换化简为

我们记形如上式的椭圆曲线为

**恒等变换过程**

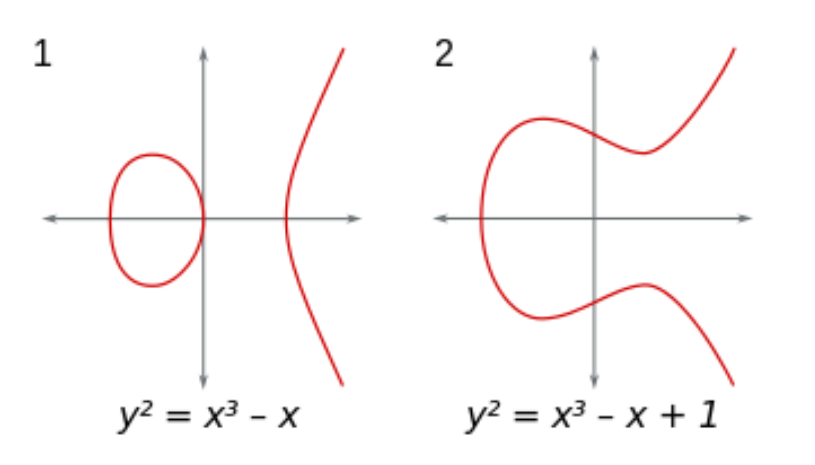
我们根据初始形式对 配方有

令 ， 分别对应 的系数，有

对于右侧

令 ，即

如果 Weierstrass 方程是光滑的，即判别式 ，其几何图形如图下



除了 Weierstrass 形式，椭圆曲线还有其他表示形式：下面给出两种表示形式，这两种形式将在ECM算法中使用

Montgomery 形式 的椭圆曲线

Montgomery 形式 的齐次椭圆曲线

下面是将 转化为 的步骤

假设我们有一条 Weierstrass形式 的椭圆曲线 ，我们可以通过适当的变量变换将其转换为 Montgomery形式。  
设 和 ，其中 是一个常数。代入 Weierstrass形式，我们得到：

展开并整理得

为了使其成为 Montgomery形式，我们需要使 的二次项系数为 且常数项为 。选择合适的 使得

通过求解这些方程，可以找到合适的 ，从而将 Weierstrass 形式 转换为 Montgomery 形式。

对于 转化为 ，只需要令 即可

**在 ECM 算法中我们使用 Montgomery 形式 的齐次椭圆曲线，在 2.2.2 中我们会讲述选择 Montgomery 形式 的齐次椭圆曲线而不是 Weierstrass形式 的椭圆曲线的原因。**

##### 椭圆曲线的加法

下面我们定义椭圆曲线上的加法：对于椭圆曲线上的任意两点 ，我们记其连线交椭圆曲线于点 ，取 关于 轴的对称点为 ，此时定义 。可以看到的是当 时，与曲线显然没有交点，故我们定义无穷远点 。**可以证明的是引入点 后，椭圆曲线对加法满足封闭性，进一步我们可以证明椭圆曲线是一个群**。

对无穷远点的理解需要注意下面几点：

* **一条直线上的无穷远点只有一个，因为过定点与已知直线平行的直线只能有一条，而两条直线的交点只有一个**。
* 一组相互平行的直线有公共的无穷远点。
* 一个平面上全体无穷远点构成一条无穷远直线。

对于上述加法的定义，我们给出不同形式椭圆曲线加法的代数形式：

**Weierstrass的点加法**

**Montgomery齐次形式的加法**

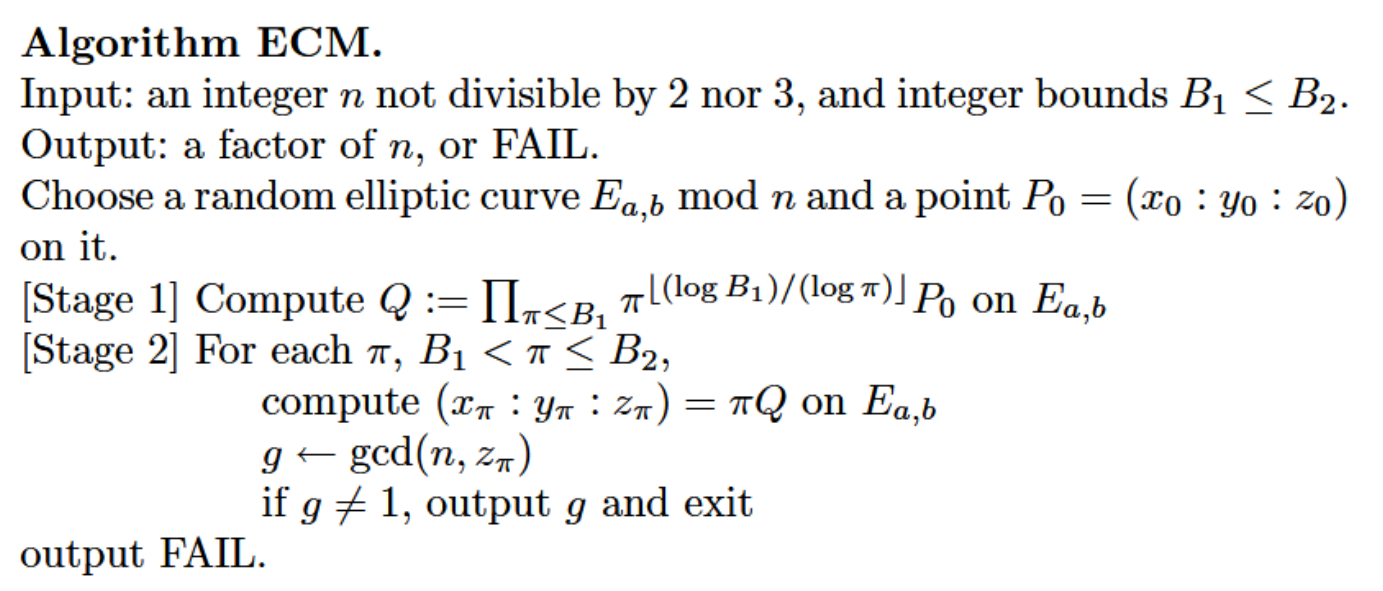
~~这里没有给出 的表达式是因为 可以由 确定，并且后续算法中我们并没有使用 ，故这里不予考虑。~~

可以看到的是Weierstrass的点加法需要计算 ，在 上我们需要 **不断计算模逆元操作**，而该操作的时间复杂度是 （基础运算为模乘，故基础运算的时间复杂度即为 ）；但对于Montgomery齐次形式的加法我们仅需要 **做常数次模乘** 即可。

此外我们假设 为椭圆曲线上任意一点， 为任意一正整数，我们定义

### ECM

下面我们首先给出 ECM 的伪代码和流程，后续将详细介绍其原理和细节处理。



不失一般性，这里我们考虑 ，其中 均为素数。

#### 算法流程

**第一阶段**

* 选择界限 和 。 是第一阶段的界限， 是第二阶段的界限。这些界限决定了算法在每个阶段中考虑的素数范围。
* 列出所有小于等于 的素数 ，对于每个素数我们做下面操作
* 计算
* 初始化
* 对 于从 到 ，执行
* 对于 Weierstrass形式 的椭圆曲线，如果出现除法失败(意味着不存在模逆)，那么意味着可能出现一个与 不互素的数，做 即可；对于 Montgomery形式 的椭圆曲线，如果出现 ，那么成功分解

**第二阶段**

* 列出所有 之间的素数 ，对于每个素数我们做下面操作
* 执行
* 检测(和第一阶段一样)

注：对于选择素数计算方幂，我们是为了得到光滑数，这也可以用阶乘计算。阶乘的检测原理运用的便是寻找 使得椭圆曲线 在模 下的阶整除 然后得到无穷远点进行计算

#### 算法原理

**前文提及了在实数域上 是一个群，不难证明的是对于任意一个大于 的素数， 在有限域 上仍然是一个群。**

对于一个未知分解的 ，不失一般性，我们假设 ，其中 均为大于 的素数。这里由于未知问分解，我们只能构建椭圆曲线在 上，但是**本质上我们可以看作分别在 上操作**。

这里我们随机选取 ，假设 在 上的阶分别为

我们在曲线上随机选择一个点 ，并且假设我们找到了一个 ，使得

记 ，根据拉格朗日定理，**在域 上**

在倍增法计算 的最后一步，假设我们计算的两点分别为 ，有

根据前面的定义必然有

由于实际是在 上进行操作，我们可能得到 ，至此我们成功分解

当然事实上我们还存在如下几个问题仍未解决：

* 如何计算
* 为什么只需要满足 （或者 ）中存在一个数较为光滑即可分解

##### 计算 K

我们记点乘法即为若干个点相加。

如果 或者 是 光滑的，类似 p-1方法我们可以取 ，这意味着我们要计算 次点乘法

**计算流程**

* 初始化
* 对于 从 到 ，执行

值得一提的是，我们执行点乘法的时间复杂度为 ，故总的时间复杂度为

但我们有一种更好的策略，我们可以选择合适的因子基，使得 。事实上，在我们上述的算法流程中，对于素数 ，我们知道当其比较小时很有可能有 ，故我们取 ；而其比较大时 的概率是极低的，我们取 即可。因此对上述算法我们分为了两个阶段。相较于上面类似p-1的做法，我们只需要计算 次基础操作。每个基础操作对于大素数时间复杂度与p-1方法一致，对于小素数则需要额外的开销（ 次点乘法）

如果按照上述算法的 的取值和 ，可以计算出点乘法次数为

按照p-1算法即为

**对于 的取值， 我们将在优化中给出其选取参考**。

##### 光滑性

根据 Hasse’s theorem，对于群 ，其中 为一个大于 的素数，有

即

相较于p-1方法，可以看到我们的限制条件由 光滑弱化至 存在某个数光滑

### 优化

#### Suyama’s Parametrizatio

待优化

* 选择参数 和   
  首先，选择两个整数 和 。通常，可以选择较小的整数，例如 和 。
* 生成椭圆曲线  
  使用参数 和 生成椭圆曲线 的系数 和
* 生成初始点  
  使用参数u和v生成初始点
* 使用生成的椭圆曲线和初始点进行ECM算法  
  将生成的椭圆曲线 和初始点 用于ECM算法的第一阶段和第二阶段，进行点倍增和点加法操作，尝试找到整数的因  
  子。

#### PRAC算法

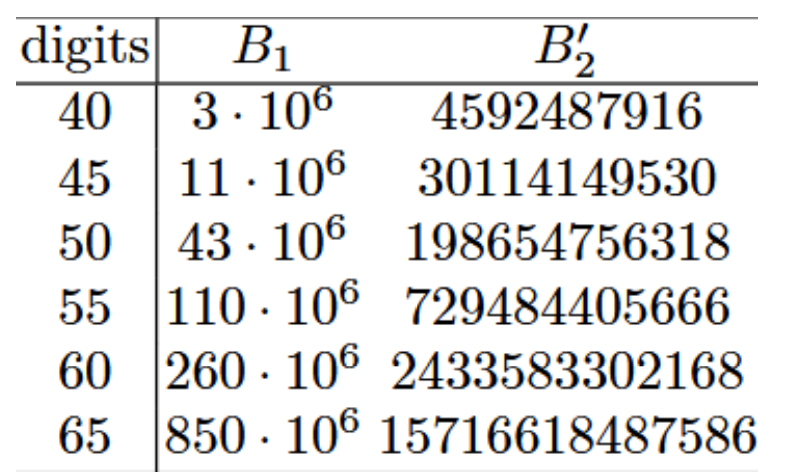
待优化

PRAC算法的核心思想是通过预计算和重组加法链来优化标量乘法。具体来说：

* 预计算：预先计算一些关键点，以便在后续的标量乘法中快速使用。
* 重组加法链：通过重组加法链，将标量k表示为多个较小整数的和，从而减少所需的点加法和点倍增操作次数。
* 标量乘法：按照重组后的加法链，逐步进行点加法和点倍增操作，直到计算出。如使用二进制方法进行与处理和计算

#### B1,B2 的经验选择

待优化



#### a,b 的选取

##### 结论1

设素数 ，试证明定义在有限域 上的形如 的椭圆曲线的个数为 。

**证明**

由于椭圆曲线是在 上，我们仅考虑

对于任意 ，有

那么根据 ， 的取值有 种可能

其中对于满足 的二元组 ，不满足椭圆曲线的定义(判别式为0)，我们应当舍去。下面考察满足 的二元组 的数量。

考察下面两个集合

由于素数 与 互素，那么在 上有

这里我们记 ，可以知道

那么原问题转化为形如 的解的个数

可以知道

当 时，

当 时，满足 的 共 种，并且对于这样的 ， 有且仅有 种取值

则满足 的二元组 的数量有

故有限域 上的形如 的椭圆曲线的个数为

可以看到在 F\_p 上椭圆曲线的个数要远大于其阶的取值范围，故一个关键是如何选择合适的 a,b，使得选择的椭圆曲线的阶不重合

### 评价

#### 优点

* 适用于中等大小的素因子：ECM算法在寻找中等大小的素因子（通常是20到50位的素数）时非常有效。对于这种规模的因子ECM通常比其他算法（如Pollard'srho算法或二次筛法）更快。
* 并行化能力强：ECM算法可以很容易地并行化，可以在多台计算机或多核处理器上同时运行多个实例，从而显著提高分解速度。

#### 不足

* 不适合大素因子：ECM算法在寻找非常大的素因子（超过50位）时效果不佳。对于这种规模的因子，其他算法（如数域筛法)可能更为有效。
* 概率性算法：ECM算法是一种概率性算法，即使选择了合适的参数，也不能保证一定能找到因子。有时需要多次运行才能成功。