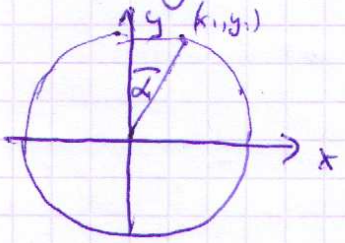
The index calculus attacks are the reason why q has to be chosen to have >3000 bits to achieve 128 bit security while the group can be as small as 2^{256}.

Elliptic curves

For suitably chosen elliptic curves (e.g. over prime fields, of prime order) no subexponential attacks are known, i.e. Pollard rho (with some tweaks and in parallel) is ~~the~~ the best attack. This means much smaller fields can be used which makes arithmetic faster and helps with implementations on small devices.

Let's start with an easier curve: the circle of radius 1 $x^2+y^2=1$. Points are (x_1, y_1) satisfying $x_1^2+y_1^2=1$



Define addition of 2 points as additions of the angles towards the y-axis. The coordinates then can be expressed using trigonometric functions in α: $x_1=\sin\alpha_1$, $y_1=\cos\alpha_1$; $x_2=\sin\alpha_2$, $y_2=\cos\alpha_2$

The trigonometric functions satisfy $\sin(\alpha_1+\alpha_2)=\sin\alpha_1\cos\alpha_2+\sin\alpha_2\cos\alpha_1$, $\cos(\alpha_1+\alpha_2)=\cos\alpha_1\cos\alpha_2-\sin\alpha_1\sin\alpha_2$. This means that the sum can be expressed as $P_1+P_2=P_3=(x_3,y_3)$ with:

$x_3=x_1y_2+x_2y_1$, $y_3=y_1y_2-x_1x_2$.

Can think of this addition as addition on the clock. We know 2:00 + 6:00 = 8:00, 3:00 + 6:00 = 9:00 etc. Try this with coordinates

$(\sqrt{3}/2, 1/2)+(0,-1)=(\sqrt{3}/2\cdot(-1)+1/2\cdot 0,\ 1/2\cdot(-1)-\sqrt{3}/2\cdot 0)=(-\sqrt{3}/2,-1/2)$

$(1,0)+(0,-1)=(1\cdot(-1)+0\cdot 0,\ 0\cdot(-1)-1\cdot 0)=(-1,0)$.

But there are points that do not correspond to times, e.g. $(3/5,4/5)=P$ is on the circle: $(3/5)^2+(4/5)^2=(9+16)/25=1$.

Can compute $(3/5,4/5)+(3/5,4/5)=(3/5\cdot 4/5+3/5\cdot 4/5,\ 4/5\cdot 4/5-3/5\cdot 3/5)$

$=(24/25, 7/25)=2P$

$3P = (24/25, 7/25) + (3/5, 4/5) = (\frac{24}{25}\cdot\frac{4}{5} + \frac{3}{5}\cdot\frac{7}{25}, \frac{7}{25}\cdot\frac{4}{5} - \frac{24}{25}\cdot\frac{3}{5})$

$= (117/125, -44/125)$

$4P = (336/625, -527/625)$ . . . .

Looks like the denominators in nP are 5^n – so not good as a DL system.

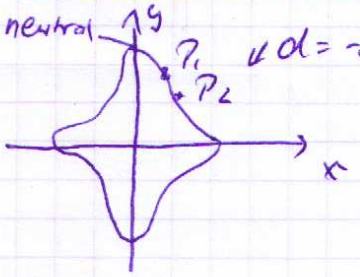
Btw. the points do form a group:

neutral el. is $(0,1)$, $-(x,y) = (x,-y)$, and (x_3,y_3) is on the curve.

Associativity and commutativity follow from the trigonometric formulas – $(\alpha_1+\alpha_2)+\alpha_3 = \alpha_1+(\alpha_2+\alpha_3)$ and $\alpha_1+\alpha_2 = \alpha_2+\alpha_1$.

Could get a DL system by doing computations modulo a large prime – but the security of this system is basically that of $\mathbb{F}_{p^2}^*$ – so circles don't help much.

Edwards curve $x^2+y^2 = 1+dx^2y^2$ for some $d \notin \{0,1\}$



$d = -30$

Define addition of $P_1 = (x_1,y_1)$ and $P_2 = (x_2,y_2)$ as $P_1+P_2 = P_3 = (x_3,y_3)$

$$x_3 = \frac{x_1y_2 + x_2y_1}{1+dx_1y_1x_2y_2} \qquad y_3 = \frac{y_1y_2 - x_1x_2}{1-dx_1y_1x_2y_2}$$

numerators are same as on circle; for $d=0$ the formulas match – but $d=0$ is excluded in the definition to avoid having the circle show up.

Can show (messy computation) that P_3 is on the curve, that the operation is associative and commutative. (easy) $(0,1)$ is the neutral element (same as on circle) and $-(x,y) = (x,-y)$.

Special attention: is the division defined? It depends. Take p a prime and d a non-square in $\mathbb{F}_p$ (i.e. for $\mathbb{F}_p^* = \langle g \rangle$ have that $d = g^k$ with k odd).

then the denominators are never $=0$: Let P_1, P_2 be on the curve, i.e. $x_i^2+y_i^2 = 1+d x_i^2 y_i^2$. Denominator $=0$ means $d x_1 y_1 x_2 y_2 = \pm 1$ for x_3 or y_3. Let's do some calculations, starting from a square $\cdot d$; so this is a non square:

$$d x_1^2 y_1^2 (x_2+y_2)^2 = d x_1^2 y_1^2 (\underbrace{x_2^2+y_2^2}_{\text{curve eq.}} + 2x_2y_2)$$

$$= d x_1^2 y_1^2 (1 + d x_2^2 y_2^2 + 2 x_2 y_2) \overset{\text{sort}}{=} \underbrace{d^2 x_1^2 y_1^2 x_2^2 y_2^2}_{=(\pm 1)^2 = 1} + d x_1^2 y_1^2 + \underbrace{(d x_1 y_1 x_2 y_2)}_{=\pm 1} x_1 y_1$$

$$= \overbrace{1 + d x_1^2 y_1^2}^{\text{curve eq.}} \pm 2 x_1 y_1 = (x_1^2 + y_1^2 \pm 2 x_1 y_1) = (x_1 \pm y_1)^2$$

but the result is a square. This is not a problem if both side $=0$ — but:

if $x_2+y_2 \neq 0$ have $d = \left(\frac{x_1 \pm y_1}{x_1 y_1 (x_2+y_2)}\right)^2$ [formulas work almost always also for $d=0$]

if $x_2 - y_2 \neq 0$ change sign on y_2 everywhere and get $d = \left(\frac{x_1 \mp y_1}{x_1 y_1 (x_2 - y_2)}\right)^2$

if $x_2+y_2=0$ and $x_2-y_2=0$ then $x_2=0$ and $y_2=0$ — but $(0,0)$ does not satisfy the curve equation $x^2+y^2 = 1+dx^2y^2$.

Can make computation faster: use double-and-add method (same as square-and-multiply) and use special formulas for ~~addition~~ doubling: If $P_1 = P_2$ have

$$x_3 = \frac{x_1 y_1 + x_1 y_1}{1 + d x_1 y_1 x_1 y_2} = \frac{2 x_1 y_1}{\underbrace{1 + d x_1^2 y_1^2}_{\text{curve equation}}} = \frac{2 x_1 y_1}{x_1^2 + y_1^2}$$

$$y_3 = \frac{y_1 y_1 - x_1 x_1}{1 - d x_1 y_1 x_1 y_1} = \frac{y_1^2 - x_1^2}{1 - d x_1^2 y_1^2} = \frac{y_1^2 - x_1^2}{2 - x_1^2 - y_1^2}$$

These formulas have smaller degree (2 instead of 4) which makes it faster to compute them.

Note on efficiency: Inversions are very expensive, ~~math~~ cost several multiplications (extreme case is to use little Fermat and to compute $a^{-1} = a^{p-2}$.)

Idea: ~~dely~~ delay the divisions, & work with fractions instead — this works over $\mathbb{Q}$, so why not over $\overline{\mathbb{F}}_p$. Simplify things by insisting on same denominator for x and y; put $x_i = \frac{X_i}{Z_i}$, $y = \frac{Y_i}{Z_i}$ then

$$x_3 = \frac{\frac{X_1}{Z_1}\cdot\frac{Y_2}{Z_2} + \frac{X_2}{Z_2}\frac{Y_1}{Z_1}}{1 + d\frac{X_1}{Z_1}\frac{Y_1}{Z_1}\frac{X_2}{Z_2}\frac{Y_2}{Z_2}} = \frac{Z_1Z_2(X_1Y_2 + X_2Y_1)}{(Z_1Z_2)^2 + dX_1Y_1X_2Y_2}$$

$$Y_3 = \frac{\frac{Y_1}{Z_1}\frac{Y_2}{Z_2} - \frac{X_1}{Z_1}\frac{X_2}{Z_2}}{1 - d\frac{X_1Y_1X_2Y_2}{Z_1Z_1Z_2Z_2}} = \frac{Z_1Z_2(Y_1Y_2 - X_1X_2)}{(Z_1Z_2)^2 - dX_1Y_1X_2Y_2}$$

bring on common denominator and get

$$X_3 = Z_1Z_2(X_1Y_2 + X_2Y_1)\cdot((Z_1Z_2)^2 - dX_1Y_1X_2Y_2)$$

$$Y_3 = Z_1Z_2(Y_1Y_2 - X_1X_2)\cdot((Z_1Z_2)^2 + dX_1Y_1X_2Y_2)$$

$$Z_3 = ((Z_1Z_2)^2 + dX_1Y_1X_2Y_2)((Z_1Z_2)^2 - dX_1Y_1X_2Y_2)$$

this can be computed in 10 mults, 1 sqr and 1 mult by d as follows:

$A = Z_1Z_2$, $B = A^2$, $C = X_1X_2$, $D = Y_1Y_2$

$E = d\cdot C\cdot D$, $F = B - E$, $G = B + E$

$X_3 = A\cdot F\cdot((X_1+Y_1)(X_2+Y_2) - C - D)$

$Y_3 = A\cdot G\cdot(D - C)$, $Z_3 = F\cdot G$.

Only surprising part is in X_3

$((X_1+Y_1)\cdot(X_2+Y_2) - C - D) = \underline{X_1X_2} + X_1Y_2 + Y_1X_2 + \underline{Y_1Y_2} - \underline{X_1X_2} - \underline{Y_1Y_2}$

$= X_1Y_2 + X_2Y_1$, which otherwise would have taken 2 M.

Doubling (could also be homework)

$$x_3 = \frac{2\frac{x_1}{z_1}\frac{y_1}{z_1}}{\frac{x_1^2}{z_1^2}+\frac{y_1^2}{z_1^2}} = \frac{2x_1y_1}{x_1^2+y_1^2}, \quad y_3 = \frac{\frac{y_1^2}{z_1^2}-\frac{x_1^2}{z_1^2}}{2-\frac{x_1^2}{z_1^2}-\frac{y_1^2}{z_1^2}} = \frac{y_1^2-x_1^2}{2z_1^2-x_1^2-y_1^2}$$

$B=(x_1+y_1)^2$, $C_1=x_1^2$, $D=y_1^2$, $E=C+D$, $H=z_1^2$

$J=E-2H$, $x_3=(B-E)J$, $y_3=E\cdot(C-D)$, $z_3=E\cdot J$.

needs $4S+3M$ (much faster than general addition).

Can make scalar multiplication faster by precomputing some multiples of P and/or by using a signed representation: $15P = 8P+4P+2P+P = 16P-P$; this is good for curves because $-(x,y)=(x,-y)$ is easy to compute; in finite fields the corresponding operation is division which does not pay off.

Slight generalizations: twisted Edwards curves: $ax^2+y^2=1+dx^2y^2$ with $a,d\neq 0$, $a\neq d$. Particularly fast formulas for $a=-1$. In general:

$$x_3 = \frac{x_1y_2+x_2y_1}{1+dx_1y_1x_2y_2}, \quad y_3 = \frac{y_1y_2-ax_1x_2}{1-dx_1y_1x_2y_2}$$

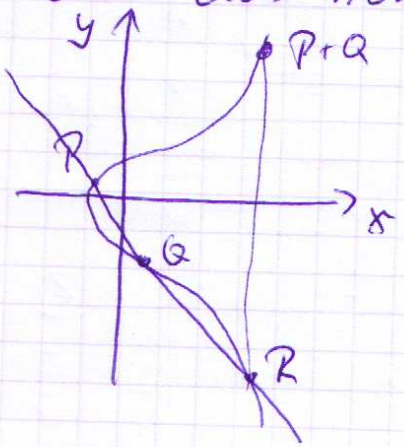
Elliptic curves: Edwards curves have easy to grasp arithmetic but there are more curves. Most common representation is the Weierstrass form

$$E: y^2+a_1xy+a_3y = x^3+a_2x^2+a_4x+a_6 \mid a_i\in k$$

If $\mathrm{char}(k)\neq 2$ and 3 one can simplify this equation to $y^2=x^3+a_4x+a_6$ with $4a^3+27b^2\neq 0$

For this curve the points are tuples $(x_1,y_1)=P_1$ satisfying the curve equation together with a point P_∞

Which does not appear on the pictures:



To add 2 points P and Q draw a line through them, this line intersects the curve in exactly 1 one more point R (see below for reasons), compute the mirror image of R w.r.t. the x-axis to find a point, call that point P+Q.

This addition law is commutative, seems to find a point on the curve – but does it give a group?

Problem: adding $(x,y)+(x,-y)$ gives a vertical line in the first step, there is no third point on it! Here comes P_∞: this point is defined to lie on every vertical line. Also define that if $R=P_\infty$ then the mirror image is also P_∞. With these conventions get that P_∞ is the neutral element: A line through $(x,y)=P$ and P_∞ is vertical and finds $(x,-y)$ as 3rd point, mirroring that gives again $(x,y)=P$. This also shows that $-P=(x,-y)$ because the line through (x,y) and $(x,-y)$ is vertical, thus has P_∞ as 3rd point; by convention it is its own mirror image.

Associativity could be shown but is a mess.

Still some problems: What if $P=(x,0)$? Then $-P=P$, so $\operatorname{ord}(P)=2$. What if $P=Q$? Then the connecting line changes to the tangent to the curve in P.

Now do the whole thing in formulas:

Slope of line through $P_1 = (x_P, y_P)$ and $Q = (x_Q, y_Q)$ is $\lambda = \frac{y_P - y_Q}{x_P - x_Q}$ for $x_1 \neq x_2$, for doubling get

$\lambda = \frac{3x_P^2 + a}{2y_P}$ by implicit differentiation.

R lies on $y^2 = x^3 + ax + b$ and on $y = \lambda x + \mu$, equate y^2 values to get

$$x^3 + ax + b = (\lambda x + \mu)^2$$

For this line there are 3 intersection points: P, Q, and R, so their x-coordinates x_P, x_Q, and x_R are the roots of this degree-3 polynomial.

$$x^3 - \lambda^2 x^2 + (-2\lambda\mu + a)x + b - \mu^2 = (x - x_P)(x - x_Q)\cdot(x - x_R)$$
$$= x^3 - (x_P + x_Q + x_R)x^2 + \cdots$$

Comparing the coefficients of x^2 can solve for x_R:
$x_R = \lambda^2 - x_P - x_Q$. Since P is on the line get
$y_P = \lambda \cdot x_P + \mu \Rightarrow \mu = y_P - \lambda x_P$ and so
$y_R = \lambda \cdot x_R + \mu = \lambda(x_R - x_P) + \mu_P$. The coordinates of $P+Q$ are x_R and $-y_R$, thus

$P+Q = (\lambda^2 - x_P - x_Q,\ \lambda(x_P - x_{R+Q}) - y_P)$

These formulas require more checks than those for Edwards curves, eg test $P = Q$?, $P = -Q$?, $P = P_\infty$? $Q = P_\infty$?

Relation? Can transform between the curves.

First look at one more curve shape

Montgomery curves

$By^2 = u^3 + Au^2 + u$ for $A \neq 0, 2$.

This curve has almost the same addition law as the Weierstrass curve above.

Put $\lambda = \frac{v_1 - v_2}{u_1 - u_2}$ for general odd and

$\lambda = \frac{3u^2 + 2Au + 1}{2Bv}$ for doubling and

$u_{P+Q} = B \cdot \lambda^2 - A - u_1 - u_2$

$v_{P+Q} = \lambda(u_P - u_{P+Q}) - v_P$

Define map from twisted Edwards curve $ax^2 + y^2 = 1 + dx^2y^2$ to Montgomery curve $Bv^2 = u^3 + Au^2 + u$:

define $A = 2(a+d)/(a-d)$, $B = 4/(a-d)$

$u = \frac{1+y}{1-y}$, $v = \frac{1+y}{x(1-y)} = \frac{u}{x}$, then (u, v) is on the Montgomery curve. Other direction $x = \frac{u}{v}$, $y = \frac{u-1}{u+1}$.

Can move from Montgomery curve to Weierstrass curve, but not always better note that $(0,0)$ is a point on an Edwards curve and that it has order 4: $2(0,0) = 2 \cdot (2 \cdot 1 \cdot 0/(1+0), \frac{0^2 - 1^2}{2 - 1^2 - 0^2}) = 2(0,-1) = (0,1)$.

but not every curve of the form $y^2 = x^3 + ax + b$ has a point of order 4.

Crypto: computations on Edwards curve are faster than on Weierstrass curve, so use Edwards curves for implementations — but Weierstrass curves are far more studied, so the transfer acts as a security argument that these new curves (from 2007) don't weaken the security.

Many more formulas for addition and doubling on elliptic curves in various shapes can be found at hyperelliptic.org/EFD.