Beyond *Varsity Math:* The red-and-blue-balls puzzle

An odds inversion problem

The red-and-blue balls puzzle, and much more

1 Introduction

> “From a bag containing red and blue balls, two are removed at random. The chances are 50-50 that they will differ in color.
> *What were the possible numbers of balls initially in the bag?*”

This problem appeared in the National Museum of Mathematics *Varsity Math Week* puzzle number 117. Here we solve that problem and then explore the solutions for other values of the odds.

Hence it is a problem in inverting a probability to find the input numbers that will produce it.

The style of exposition in this document is exploratory, rather than simply presenting the solution methods. I want to let the reader share in a way in the enjoyment I had, gradually unfolding the secrets of this problem. But the exposition does not follow my own track in solving the problem. I went down many blind alleys, or made discoveries that rendered earlier approaches obsolete. I have organized the topics in a logical sequence, omitting the false starts and digressions.

Notes

By Robert K. Moniot, Fordham University, 2020.

Disclaimer: this work has not been peer reviewed. There may be errors here and there. I welcome feedback pointing out any errors. This version dated 1 May, 2020.

1.1 Acknowledgements and references

The sources I have relied on for this analysis are:

- Wikipedia pages on Continued Fractions and Pell's Equation.
- Eric W. Weisstein, "Pell Equation." From MathWorld--A Wolfram Web Resource. http://mathworld.wolfram.com/PellEquation.html
- Dario Alejandro Alpern, "Methods to solve $A x^2 + B x y + C y^2 + D x + E y + F == 0$", web page https://www.alpertron.com.ar/METHODS.HTM
  A good site for beginners to start. He credits Iain Davidson for much of his material.
- Hua Loo Keng, *Introduction to Number Theory,* translated from the Chinese by Peter Shiu, Springer-Verlag, 1982, chapters 10 and 11.
- Trygve Nagell, *Introduction to Number Theory, 2nd Ed.,* Chelsea Publishing Co., New York (1964).
- Dick Hess, of the National Museum of Mathematics, private communication (July, 2018).

The puzzle (authored by Dick Hess) originated from the National Museum of Mathematics in New York City, in *Varsity Math Week 117,* available at https://momath.org/home/varsity-math/varsity-math-week-117/. I would like to thank MoMath for creating this puzzle.

The main results of this exploration are in an article accepted for publication in the *American Mathematical Monthly,* "Solution of an Odds Inversion Problem," by Robert K. Moniot, 2020 (to appear).

1.2 To the reader

The aim of this document is to share with you something of the enjoyment I had working out the solution of the problem and exploring various lines of inquiry that I encountered along the way. Hence, while the document mainly marches along on an orderly route to solving the different cases, there are many digressions and excursions. Where these are particularly arcane and not essential to the main task, I so note. You should be prepared for a leisurely ramble: the exposition is not designed to be read quickly. It is my hope that I have provided enough support in the form of proofs and examples so that the reader can follow along without too much effort. It is assumed that the reader is mathematically inclined and is comfortable with algebra, calculus, and elementary number theory. Some calculus is used in places when exploring trends and limits. I hope you have as much fun with this as I did.

The reader who is unfamiliar with Mathematica may have some difficulty following the steps in this document that use Mathematica expressions. Documentation of the Mathematica language and functions can be found at https://reference.wolfram.com/language/.

2 The *Varsity Math* problem

The problem posed in the *Varsity Math* puzzle can be solved with high-school level algebra. You might suppose at first that to have 50-50 odds of drawing balls of different colors, there should be an equal number of each color in the bag. But this is not the case. Suppose you draw a red ball initially. Now the bag contains more blue balls than red ones, so the odds of drawing a blue ball are higher than 50%. Let's solve the problem to see what the correct answer is.

Let x, y be the numbers of red and blue balls, respectively. The probability of picking a red and then a blue ball from the bag is

In[53]:= probredblue[{x_, y_}] := $\frac{x}{x+y}\,\frac{y}{x+y-1}$

The odds of picking a blue and then a red ball are

In[54]:= probbluered[{x_, y_}] := $\frac{y}{x+y}\,\frac{x}{x+y-1}$

Obviously these are equal. The probability of picking different-colored balls is their sum:

```
In[55]:= probdifferent[{x_, y_}] := probredblue[{x, y}] + probbluered[{x, y}]
```

So in order for the odds of picking different-colored balls to be 50-50, i.e. a probability of 1/2, we need to solve

In[56]:= probdifferent[{x, y}] == $\frac{1}{2}$

Out[56]= $\frac{2\,x\,y}{(-1+x+y)\,(x+y)} == \frac{1}{2}$

Cross multiplying and expanding terms, we obtain

```
In[57]:= Simplify[Expand[(x + y) (-1 + x + y) == 4 x y]]
```

Out[57]= $x^2 + y^2 == x + y + 2\,x\,y$

This rearranges to

$$x^2 - 2\,x\,y + y^2 == x + y$$
$$(x - y)^2 == x + y == t$$

where t is the total number of balls. This shows that t must be a square. Let $t == v^2$. Then $x - y == \pm v$ while $x + y == v^2$. Taking the minus sign so that $x \leq y$, we have $y - x == v$ and $y + x == v^2$. Adding these, we obtain

$$2\,y == v^2 + v, \quad y == \frac{v^2 + v}{2} == \frac{v\,(v+1)}{2}$$

Subtracting them instead,

$$2\,x == v^2 - v, \quad x == \frac{v^2 - v}{2} == \frac{v\,(v-1)}{2}$$

Observe that x and y are successive triangular numbers. The sum of successive triangular numbers is always a square. Pretty cool.

Here is a table of the first 10 solutions.

In[∘]:=
```
TableForm[Table[{1/2 (v^2 - v), 1/2 (v^2 + v), v^2}, {v, 10}],
 TableHeadings -> {None, {x, y, t}}]
```

Out[∘]//TableForm=

x	y	t
0	1	1
1	3	4
3	6	9
6	10	16
10	15	25
15	21	36
21	28	49
28	36	64
36	45	81
45	55	100

The first row formally satisfies the equation but is not an acceptable solution, since there are not 2 balls to draw out. (Including $v == 0$ would yield another unacceptable solution (0, 0). Negative values of v generate the same table but with x and y swapped.)

3 Generalizing the problem

3.1 Questions to be answered

Having solved the *Varsity Math* problem, an obvious followup question is to ask what we can say about other odds than 50-50. This generalized problem turns out to be solvable, with a rich variety of different cases requiring different solution methods. We seek to answer the questions:

- For a given probability ratio, is there a solution to the problem? That is, does there exist a pair of numbers of red and blue balls that give that probability?
- If there is a solution, is the number of solutions finite or infinite?
- For instances with a finite number of solutions, can we list them all?
- For instances with an infinite number of solutions, can we obtain a formula or recurrence that will generate as many solutions as desired?
- Can we show that the methods used are capable of finding all solutions, not missing any that exist?

It turns out that all of these questions can be answered satisfactorily.

3.2 Main results

The main results I obtained are:

- For any probability ratio greater than 50%, the number of solutions is either none or finite, and an algorithm exists to determine all the solutions that exist or show that there are none.
- There are only a few probability ratios greater than 55% that have solutions, and only one ratio greater than 70% that has one (namely 100%, which has only one solution).
- The probability ratio of 50% is special, and has the infinite set of solutions obtained already. The numbers of red and blue balls are any successive triangular numbers, whose sum is a square.
- For any probability ratio less than 50%, except for a minority (those for which the probability equation factors), the number of solutions is infinite, and an algorithm exists to generate as many solutions as desired. The algorithm can find all solutions (up to any desired limit).
- For the minority of ratios less than 50% for which the probability equation factors, the number of solutions is either none or finite, and an algorithm exists to determine all the solutions that exist or show that there are none.
- The algorithms mentioned have finite running time, although for certain probability ratios the running time may be long. In particular, probability ratios very slightly over or under 50% are problematic. However, for any probability that can be expressed as a ratio of moderate-sized integers (e.g. up to 3 digits), the problem can be solved quickly (using a computer). Also, for those probability ratios below 50% having an infinite number of solutions, if the numerator is prime or equal to 4, a relatively simple and quick method can find all solutions.
- Interestingly, if the bag initially contains an equal number of red and blue balls, removing one ball does not change the odds of subsequently drawing two balls of different colors.

4 Preliminaries

4.1 Notation

As before, let x, y be the numbers of red and blue balls, respectively. The probability of drawing balls of different colors is obviously a rational number. Denote it by p/q where p, q are relatively prime integers. The basic problem then is, given a ratio p/q, find integer x, y values which produce that probability of drawing two balls of different colors. The equation is

$$\frac{2\,x\,y}{(x+y)\,(x+y-1)} == \frac{p}{q} \tag{1}$$

Put the equation into a function. Note: here and in other functions of solutions of the equation, the argument is a list, rather than separate variables (i.e. `f[{x,y}]` not `f[x,y]`). This allows the functions to be applied to a list of solutions easily using `/@` notation.

```
In[58]:= probequation[{x_, y_}] := (2 x y)/((x + y) (x + y - 1)) == p/q
```

4.2 Setting up the problem

4.2.1 The Diophantine equation

Put the probability equation into the form of a Diophantine equation, i.e. a polynomial equation involving only integers.

In[59]:=
```
MultiplySides[probequation[{x, y}],
 q (x + y - 1) (x + y), Assumptions → q (x + y - 1) (x + y) ≠ 0]
```

Out[59]= $2\,q\,x\,y == p\,(-1 + x + y)\,(x + y)$

In[60]:=
```
SubtractSides[%, 2 q x y]
```

Out[60]= $0 == -2\,q\,x\,y + p\,(-1 + x + y)\,(x + y)$

In[61]:=
```
Collect[Expand[%], {x^2, x y, y^2, x, y}]
```

Out[61]= $0 == -p\,x + p\,x^2 - p\,y + (2\,p - 2\,q)\,x\,y + p\,y^2$

Rearranging,

$$p\,x^2 - 2\,(q - p)\,x\,y + p\,y^2 - p\,x - p\,y == 0 \tag{2}$$

Put this into a formula for convenience later. The function returns true if the argument is a solution. The values of p and q are left as parameters rather than arguments.

In[62]:=
```
xyeqn[{x_, y_}] := Evaluate[%]
```

Show that it works on one of the Varsity Math puzzle solutions.

In[63]:=
```
xyeqn[{3, 6}] /. {p → 1, q → 2}
```

Out[63]= True

The values of p, q are givens and x, y are to be solved for. Since p/q is a probability, we require

$$0 \le \frac{p}{q} \le 1 \tag{3}$$

and p/q in lowest terms, i.e. gcd(p, q) == 1.

4.2.2 Formal and admissible solutions

We will call any pair of integers (x, y) satisfying Equation (2) a ***formal*** solution. Since the number of balls of either color cannot be negative and there need to be at least two balls in the bag in order to be able to draw two out, not all formal solutions are ***admissible*** as solutions to the original problem. For a solution to be admissible it must satisfy

$$x \ge 0, \quad y \ge 0, \quad x + y \ge 2 \tag{4}$$

The general quadratic Diophantine equation $a\,x^2 + b\,x\,y + c\,y^2 + d\,x + e\,y + f == 0$ is a fully solved problem, i.e. there are methods for finding all the solutions that exist, though the solution method is complicated in many instances. (See Alpern for details.) Solutions are not, in general, guaranteed to exist for arbitrary values of the coefficients. Our equation has a very special form, in that all but one of the coefficients are the same in magnitude, with opposite signs for the linear and squared terms of each variable. This structure turns out to simplify the solution greatly relative to the general case, as we shall see shortly.

4.2.3 Trivial solutions

The lack of a constant term in Equation (2) leads to the existence of three trivial solutions that always satisfy the equation for any values of p and q:

$$\begin{aligned} x &= 0, \quad y = 0 \\ x &= 0, \quad y = 1 \\ x &= 1, \quad y = 0 \end{aligned} \tag{5}$$

These are not admissible solutions, but they will prove useful in finding admissible solutions. When plugged into Equation (1), they yield $p/q == 0/0$. This is undefined, which is how they are able to satisfy Equation (2) for any values of p and q.

4.2.4 Symmetry

The equation is symmetric in x and y, so that if (x, y) is a solution, then (y, x) is also a solution. When listing solutions, we will usually list just one member of the pair, usually choosing $x \leq y$ to provide uniqueness. We will say two different solutions are ***distinct*** if they are not composed of the same two x, y values.

4.2.5 Change of variables

Equation (2) simplifies considerably if we change variables. As we did in solving the Varsity Math puzzle, let

$$t == y + x, \quad v == y - x$$

```
In[64]:= tvfromxy[{x_, y_}] := {y + x, y - x}
```

Define functions for mapping back.

```
In[65]:= Solve[{tvfromxy[{x, y}] == {t, v}}, {x, y}]
```

Out[65]= $\left\{\left\{x \to \frac{t - v}{2}, y \to \frac{t + v}{2}\right\}\right\}$

In[66]:= $\texttt{xyfromtv[\{t\_, v\_\}] :=} \left\{\frac{t - v}{2}, \frac{t + v}{2}\right\}$

Rewrite the probability equation in terms of these variables:

```
In[67]:= Simplify[xyeqn[{x, y}] /. Solve[{t == y + x, v == y - x}, {x, y}]]
```

Out[67]= $\left\{2\, p\, (-1 + t)\, t + q\, \left(-t^2 + v^2\right) == 0\right\}$

Collect terms

In[68]:=
```
Collect[Expand[%], {v^2, t, t^2}]
```

Out[68]= $\{-2\,p\,t + (2\,p - q)\,t^2 + q\,v^2 == 0\}$

Rewrite in nicer form.

$$(q - 2\,p)\,t^2 + 2\,p\,t - q\,v^2 == 0 \tag{6}$$

Put it into a function for later use.

In[69]:=
```
tveqn[{t_, v_}] := (q - 2 p) t^2 + 2 p t - q v^2 == 0
```

If $q - 2\,p == 0$ then $p/q == 1/2$, which is the Varsity Math puzzle, already solved. Otherwise, we can simplify (6) further by another change of variables, as follows.

If $q - 2\,p \neq 0$ we can eliminate the linear term by completing the square. First step: multiply LHS by $(q - 2\,p)$ to make the coefficient of t^2 square.

$$(q - 2\,p)^2\,t^2 + 2\,p\,(q - 2\,p)\,t - q\,(q - 2\,p)\,v^2 == 0$$

Add p^2 to both sides.

$$(q - 2\,p)^2\,t^2 + 2\,p\,(q - 2\,p)\,t + p^2 - q\,(q - 2\,p)\,v^2 == p^2$$

Observe that the first three terms are a square:

In[70]:=
```
Simplify[(q - 2 p)^2 t^2 + 2 p (q - 2 p) t + p^2]
```

Out[70]= $(p - 2\,p\,t + q\,t)^2$

Let

$$u == p - 2\,p\,t + q\,t == (q - 2\,p)\,t + p \tag{7}$$

Then the equation is

$$u^2 - q\,(q - 2\,p)\,v^2 == p^2 \tag{8}$$

If we let

$$D == q\,(q - 2\,p) \tag{9}$$

and

$$f == p^2 \tag{10}$$

we can write Equation (8) as

$$u^2 - D\,v^2 == f \tag{11}$$

This is a well-studied equation. If $D > 0$ is nonsquare and $f == 1$, it is called the Pell equation. Otherwise it is called a Pell-like equation. Weisstein gives a concise overview of the Pell and Pell-like equations. Hua (chapter 11) and Nagell (chapter VI) discuss the solution of Equation (11) in detail.

Define some functions that will be useful when solving.

In[71]:=
```
uveqn[{u_, v_}] := u^2 - q (q - 2 p) v^2 == p^2
```

In[72]:=
```
uvfromtv[{t_, v_}] := {(q - 2 p) t + p, v}
```

In[73]:=
```
Solve[uvfromtv[{t, v}] == {u, v}, t]
```

Out[73]= $\left\{\left\{t \to \frac{p-u}{2\,p-q}\right\}\right\}$

In[74]:=
```
tvfromuv[{u_, v_}] := {(p - u)/(2 p - q), v}
```

In[75]:=
```
tvfromuv[{u, v}]
```

Out[75]= $\left\{\frac{p-u}{2\,p-q},\ v\right\}$

Note that even if u is integer, t may not be integer, unless $u \equiv p \mod (q - 2\,p)$.

One last pair of functions to go directly from (u, v) to (x, y) and vice versa.

In[76]:=
```
xyfromuv[{u_, v_}] := xyfromtv[tvfromuv[{u, v}]]
```

In[77]:=
```
xyfromuv[{u, v}]
```

Out[77]= $\left\{\frac{1}{2}\left(\frac{p-u}{2\,p-q} - v\right),\ \frac{1}{2}\left(\frac{p-u}{2\,p-q} + v\right)\right\}$

In[78]:=
```
uvfromxy[{x_, y_}] := uvfromtv[tvfromxy[{x, y}]]
```

In[79]:=
```
uvfromxy[{x, y}]
```

Out[79]= $\{p + (-2\,p + q)\,(x + y),\ -x + y\}$

4.2.6 Existence and completeness of solutions

Equation (8) is not guaranteed to have integer solutions beyond the trivial ones for all values of p, q. In the later sections, the existence of solutions will be examined for the different cases. For some values of p/q no admissible solutions exist; for others a finite number exist; and for others the number of solutions is infinite. The integer solutions (u, v) of Equation (8) that one may find need to be transformed back to (t, v) which in turn need to be transformed back to (x, y). From given u, v it is not guaranteed that t will be integer, and if t and v are integer it is not guaranteed that x and y will be.

Because integer (x, y) map to integer (u, v), if Equation (2) has integer solutions, they will correspond to integer solutions of Equation (8). So solving the latter will turn up any solutions to the former that exist. Below, methods will be developed to solve Equation (8). The methods are capable of finding all solutions.

4.2.7 Mathematica

Mathematica is able to solve Equation (2) for admissible solutions or show that none exist. Examples of using Mathematica to solve the problem are given in Section 13 at the end of this document. Our aim, however, is not simply to obtain solutions but to see how they are obtained, and to obtain insight about when they exist and how many exist. So except for Section 13 , Mathematica is used here only to perform tedious algebra or lengthy numerical calculations. There are a lot of both of those here, so I am very grateful to Mathematica for its help.

5 Exploring the problem

5.1 The character of the equation

Equation (2) is a conic, so it can be an ellipse, a parabola, or a hyperbola.

The transformations from (x, y) to (t, v) and (u, v) are linear, so they only rotate and scale the curve, preserving its basic character. It is easily seen from its form that Equation (11) is an ellipse if $D < 0$ and a hyperbola if $D > 0$, while if $D == 0$, which requires $p/q == 1/2$, Equation (6) reduces to $t - v^2 == 0$ (as we saw in Section 2), which is a parabola. In (u, v)-space, the curves have their axis along the u-axis. The parabola has its vertex at the origin, while the ellipse and the hyperbola are centered on the origin. Graphs are plotted below.

In (x, y)-space all the curves represented by Equation (2) are constrained to pass through the three trivial solutions $(x, y) == (0, 0), (0, 1)$, and $(1, 0)$. Since x and y are interchangeable, the curves are symmetric about the line $x == y$. Thus the curves are tilted so that they lie mainly in the first quadrant, and, for the hyperbola's other branch, also in the third quadrant. The curves cross the axes from the first quadrant into the second and fourth only for a short distance in between the trivial solution points. All integer solutions for the ellipse and parabola will be non-negative, while the hyperbola may have negative solutions from its third-quadrant branch. In all nontrivial integer solutions, x and y have the same sign.

The discriminant of Equation (2) is

```
In[80]:= Simplify[b^2 - 4 a c /. {a → p, b → 2 (p - q), c → p}]
Out[80]= 4 q (-2 p + q)
```

This is equal to $4 D$, which is also the discriminant of Equation (8). For purposes of determining sign or squareness we can ignore the factor of 4, so for simplicity we will call D the discriminant of the equation.

Since $q > 0$, the sign of D is that of $q - 2 p$. Therefore the classification of the conic is

$$\begin{array}{lll} q < 2 p & D < 0 & \text{ellipse} \\ q == 2 p & D == 0 & \text{parabola} \\ q > 2 p & D > 0 & \text{hyperbola} \end{array} \tag{12}$$

It is often convenient to work with the parameter

$$z == \frac{p}{q}$$

In terms of z Equation (8) becomes

In[81]:=
```
Simplify[uveqn[{u, v}] /. {p → z, q → 1}]
```

Out[81]= $u^2 + v^2 (-1 + 2 z) == z^2$

and the classification of the conic is

$$\begin{array}{ll} z > 1/2 & \text{ellipse} \\ z == 1/2 & \text{parabola} \\ z < 1/2 & \text{hyperbola} \end{array} \tag{13}$$

For our problem, we restrict to $0 \leq z \leq 1$ since it represents a probability, namely the probability of drawing balls of different colors.

Since the ellipse is a closed curve, the number of solutions possible for those cases is finite. For the parabola and hyperbola, the number of solutions may be (and in many cases is) infinite.

5.1.1 Plotting the equations

Note on efficiency of `Plot` in Mathematica: it runs rather slowly (~ 0.5s on my laptop for these examples) if the `Solve` function is simply placed inside the `Plot` function. One can use `Evaluate` to speed it up (~ 0.025 s, a factor of 20 better). However, using `Evaluate` around the `Solve` function is still slow, while using it around the whole y expression is fast but treats a two-valued curve as two curves in different colors. The only solution I have found that yields fast plotting with a single color is to use an intermediate variable.

Parabola

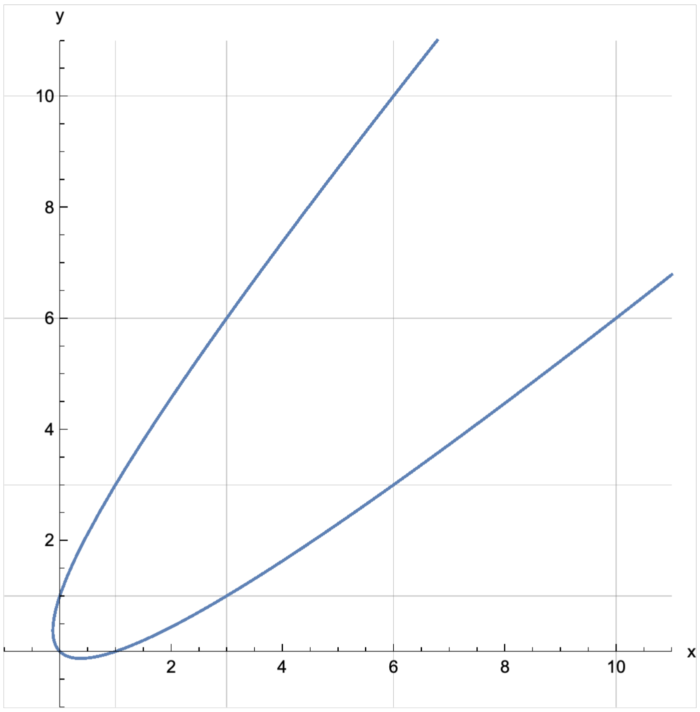
First plot this example in (x, y)-space. Force Mathematica to scale the axes the same so its shape is true. Add grid lines to show the solution points where they intersect the curve.

In[82]:=
```
yvaluesforplot = Solve[xyeqn[{x, y}] /. {p → 1, q → 2}, y]
```

Out[82]= $\left\{\left\{y \to \frac{1}{2}\left(1 + 2x - \sqrt{1 + 8x}\right)\right\}, \left\{y \to \frac{1}{2}\left(1 + 2x + \sqrt{1 + 8x}\right)\right\}\right\}$

In[83]:=
```
Plot[y /. yvaluesforplot, {x, -1, 11},
 PlotRange → {{-1, 11}, {-1, 11}}, AspectRatio → 1,
 GridLines → {{1, 3, 6, 10}, {1, 3, 6, 10}},
 AxesLabel → {"x", "y"}]
```

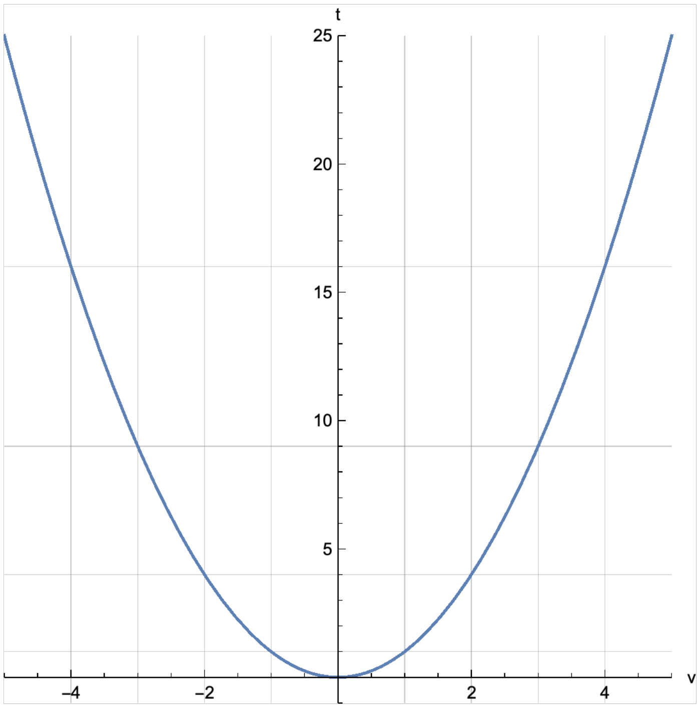
Out[83]=



Now plot it in (t, v)-space, where the equation is $t == v^2$:

In[84]:=
```
Plot[v^2, {v, -5, 5},
 PlotRange → {{-5, 5}, {-1, 25}}, AspectRatio → 1,
 GridLines → {{-4, -3, -2, -1, 1, 2, 3, 4}, {1, 4, 9, 16}},
 AxesLabel → {"v", "t"}]
```

Out[84]=

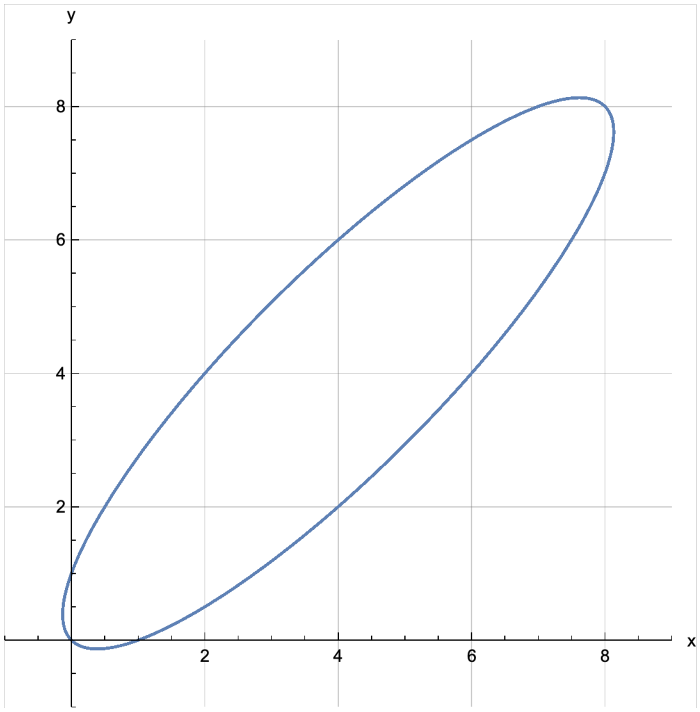


Ellipse

We use an example $p/q == 8/15$, solved later, that has several integer solutions besides the trivial ones, namely $(x, y) == (2, 4), (4, 6), (7, 8)$, and $(8, 8)$ and their symmetric partners. First, in (x, y)-space.

In[85]:=
```
yvaluesforplot = Solve[xyeqn[{x, y}] /. {p → 8, q → 15}, y];
Plot[y /. yvaluesforplot, {x, -1, 9},
 PlotRange → {{-1, 9}, {-1, 9}}, AspectRatio → 1,
 GridLines → {{2, 4, 6, 8}, {2, 4, 6, 8}},
 AxesLabel → {"x", "y"}]
```

Out[86]=



Now in (t, v)-space. First compute the grid points where solutions lie:

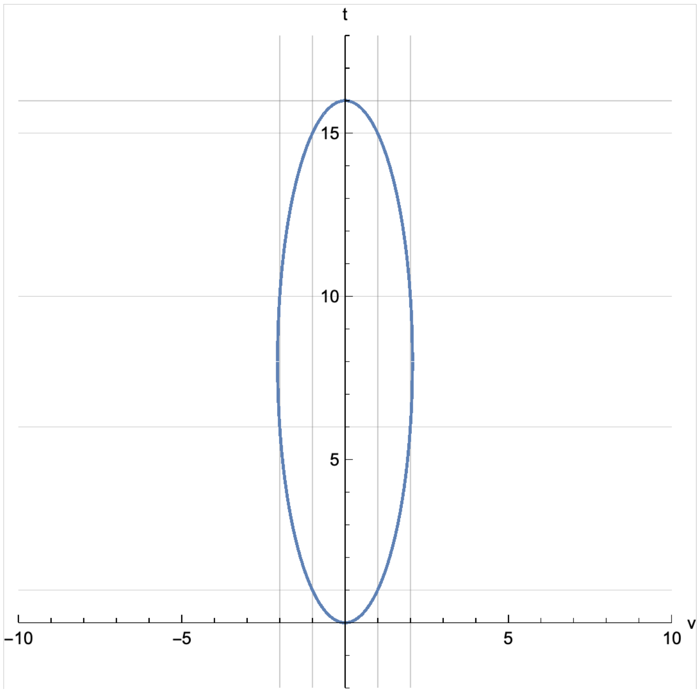
In[87]:=
```
tvsolns8o15 = tvfromxy /@ {{0, 0}, {0, 1}, {2, 4}, {4, 6}, {7, 8}, {8, 8}}
```

Out[87]=
```
{{0, 0}, {1, 1}, {6, 2}, {10, 2}, {15, 1}, {16, 0}}
```

The values of v can also be the negatives of these.

In[88]:=
```
tvaluesforplot = Solve[tveqn[{t, v}] /. {p → 8, q → 15}, t];
Plot[t /. tvaluesforplot, {v, -5 / 2, 5 / 2},
 PlotRange → {{-10, 10}, {-2, 18}}, AspectRatio → 1,
 GridLines → {{-2, -1, 1, 2}, {1, 6, 10, 15, 16}},
 AxesLabel → {"v", "t"}]
```

Out[89]=



The ellipse is tangent to the *x*-axis.

Now in (*u*, *v*)-space. Compute *u* values of solutions for grid lines.

In[90]:=
```
uvsolns8o15 = (uvfromtv /@ tvsolns8o15) /. {p → 8, q → 15}
```

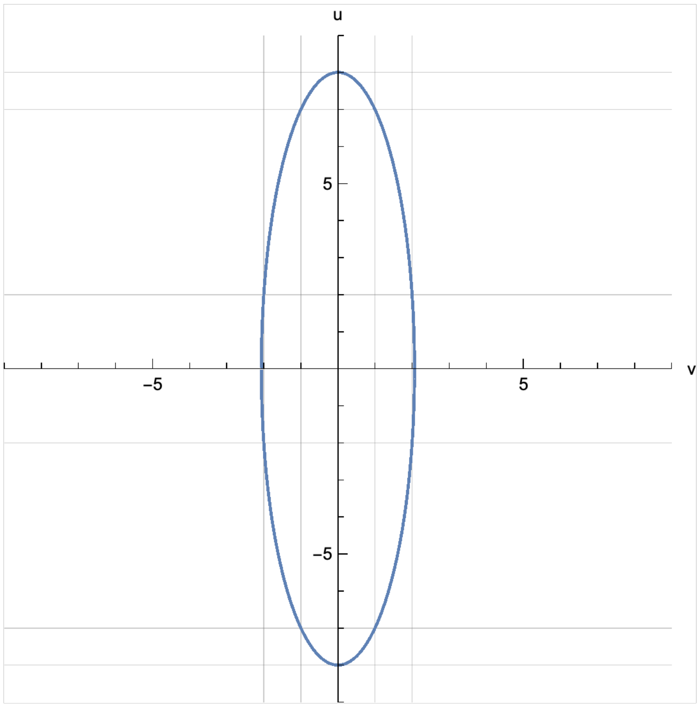
Out[90]= {{8, 0}, {7, 1}, {2, 2}, {-2, 2}, {-7, 1}, {-8, 0}}

The negatives of *v* are also allowable.

The plot has *v* on the horizontal axis and *t* vertically.

In[91]:=
```
uvaluesforplot = Solve[uveqn[{u, v}] /. {p → 8, q → 15}, u];
Plot[u /. uvaluesforplot, {v, -5 / 2, 5 / 2},
 PlotRange → {{-9, 9}, {-9, 9}}, AspectRatio → 1,
 GridLines → {{-2, -1, 1, 2}, {-8, -7, -2, 2, 7, 8}},
 AxesLabel → {"v", "u"}]
```

Out[92]=



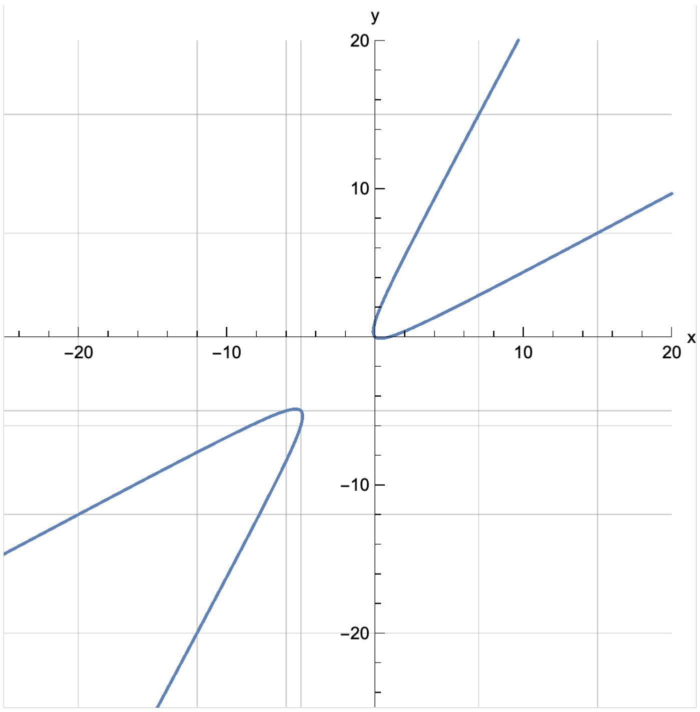
The ellipse is centered on the origin. It is worth noting that the ellipse turns upside-down in going from (t, v)-space to (u, v)-space: $t == 0$ corresponds to $u == 8$ and $t == 16$ to $u == -8$.

Hyperbola

For this example, for the sake of completeness we will include inadmissible negative solutions, with gridlines locating the integer points other than the trivial solutions. This example $p/q == 5/11$ has admissible solution $(x, y) == (7, 15)$ as well as negative integer solutions $(-5, -5)$, $(-5, -6)$, $(-12, -20)$ plus their symmetric partners. It also has many other solutions not shown. In (x, y)-space:

In[93]:=
```
yvaluesforplot = Solve[xyeqn[{x, y}] /. {p → 5, q → 11}, y];
Plot[y /. yvaluesforplot, {x, -25, 20},
 PlotRange → {{-25, 20}, {-25, 20}},
 AspectRatio → 1,
 GridLines → {{-20, -12, -6, -5, 7, 15}, {-20, -12, -6, -5, 7, 15}},
 AxesLabel → {"x", "y"}]
```

Out[94]=



Now plot it in (t, v)-space. Compute the gridlines. To avoid clutter we will only put the gridlines for the solutions away from the vertices. Calculate the gridlines for solutions.

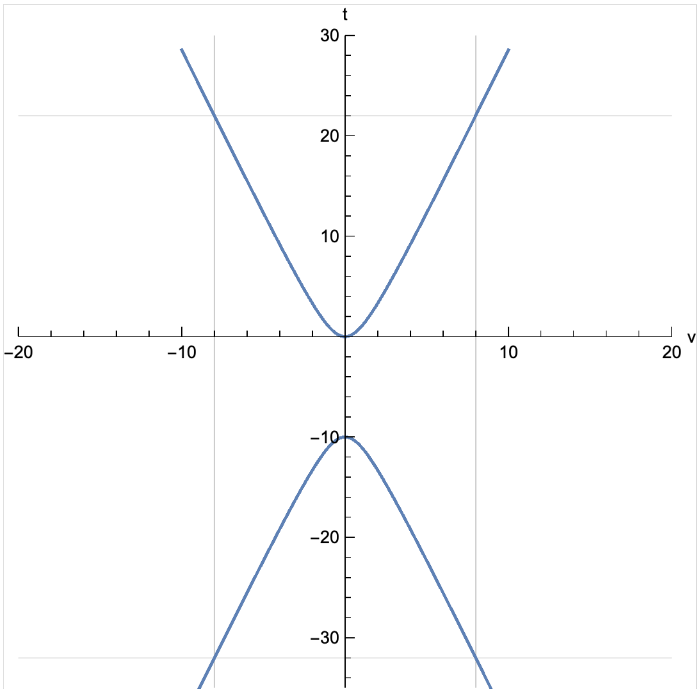
In[95]:=
```
tvsolns5o11 = tvfromxy /@ {{-12, -20}, {7, 15}}
```

Out[95]=
```
{{-32, -8}, {22, 8}}
```

The plot has v on the horizontal axis and u vertically.

In[96]:=
```
tvaluesforplot = Solve[tveqn[{t, v}] /. {p → 5, q → 11}, t];
Plot[t /. tvaluesforplot, {v, -10, 10},
 PlotRange → {{-20, 20}, {-35, 30}}, AspectRatio → 1,
 GridLines → {{-8, 8}, {-32, 22}},
 AxesLabel → {"v", "t"}]
```

Out[97]=



The upper branch is tangent to the *x*-axis. For this example, there are three integer solutions at each vertex.

Now plot it in (*u*, *v*)-space.

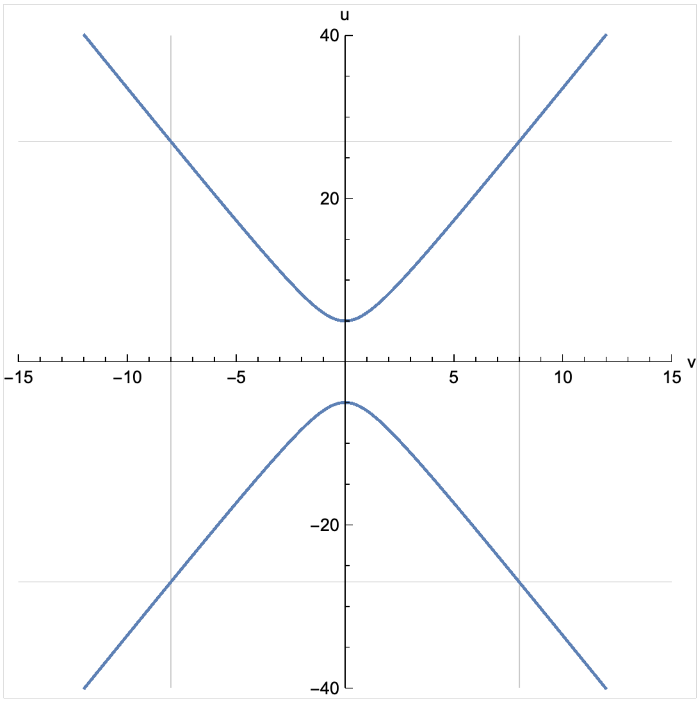
Calculate *u* values for gridlines.

In[98]:=
```
(uvfromtv /@ tvsolns5o11) /. {p → 5, q → 11}
```

Out[98]=
```
{{-27, -8}, {27, 8}}
```

In[99]:=
```
uvaluesforplot = Solve[uveqn[{u, v}] /. {p → 5, q → 11}, u];
Plot[u /. uvaluesforplot, {v, -15, 15},
 PlotRange → {{-15, 15}, {-40, 40}}, AspectRatio → 1,
 GridLines → {{-8, 8}, {-27, 27}},
 AxesLabel → {"v", "u"}]
```

Out[100]=



In these coordinates it is centered on the origin and aligned with the axes. There are three integer solutions at each vertex.

5.1.2 What values of *D* are possible?

Answer: any odd value, positive or negative, and any multiple of 8, positive or negative or 0. But no even values that are not multiples of 8. Proof follows. The definition of *D* is

$D == q\,(q - 2\,p)$

Case of *D* == 0

The value $D == 0$ occurs only when $p/q == 1/2$, since it requires $q - 2\,p == 0$.

Case of odd *D*

Supposing $D > 0$ is odd, set $D == 2\,m + 1$, $m \geq 0$. We can obtain this using

$$q == D == 2\,m + 1, \quad p == \frac{D - 1}{2} == m$$

Special case is $m == 0$, $p == 0$, $q == 1$, $D == 1$. This is a legitimate probability of 0. (All balls of same color, so that drawing balls of different colors is impossible.)

For $m > 0$ it is obvious that $p == m$ and $q == 2\,m + 1$ are relatively prime, so p/q is in lowest terms. So all odd values of $D > 0$ are possible. This implies all primes greater than 2 are possible values of D.

There may be other ways to obtain the same D. For instance, $D == 21$ results from the above with $m == 10$, which gives $p/q == 10/21$, but also from $p/q == 2/7$.

Supposing $D < 0$ is odd, set $D == -2\,m + 1$, $m > 0$. We can obtain this using

$$q == -D == 2\,m - 1,\quad p == \frac{-D + 1}{2} == m$$

For $m > 0$, $p == m$ and $q == 2\,m - 1$ are relatively prime, giving p/q in lowest terms. So all odd values of $D < 0$ are possible. Again, there may be other ways to obtain the same D.

Case of even *D*

Now consider D even. We can show that all positive multiples of 8 appear. We saw above that $D == 0$ results for $p/q == 1/2$. Zero is a multiple of 8. For $D > 0$, set $D == 8\,m$, $m > 0$. We can obtain this using

$$p == 2\,m - 1,\quad q == 4\,m \Rightarrow q\,(q - 2\,p) == 4\,m\,(4\,m - 4\,m + 2) == 8\,m == D$$

This clearly gives $p == 2\,m - 1$, $q == 4\,m$ relatively prime, so any positive multiple of 8 is obtainable.

For $D < 0$ even set $D == -8\,m$, $m > 0$. Use

$$p == 2\,m + 1,\quad q == 4\,m \Rightarrow q\,(q - 2\,p) == 4\,m\,(4\,m - 4\,m - 2) == -8\,m == D$$

Multiples of 8 also appear for $p == 1$ with any even q: put $q == 2\,m$, $m > 0$.

$$D == q\,(q - 2\,p) == 2\,m\,(2\,m - 2) == 4\,m\,(m - 1)$$

One of m, $m - 1$ is even, so D is a multiple of 8. In these cases only non-negative D appear.

Now prove the non-existence of even values of D that are not multiples of 8. For D to be even, q must be even since $q - 2\,p$ is the same parity as q. Put $q == 2\,m$.

$$D == q\,(q - 2\,p) == 2\,m\,(2\,m - 2\,p) == 4\,m\,(m - p)$$

Now, p must be odd to be relatively prime to q. So one of m, $m - p$ is even, making D a multiple of 8. QED.

There can be other ways to achieve the same D with different ratios. For instance, setting $m == 10$ in the $D < 0$ case gives $D == -80$, $p/q == 21/40$, but $D == -80$ also results from $p/q == 9/10$.

Alternative ways to achieve the same *D*

As we saw above, there can be more than one p/q ratio for a given D.

In general, q must be a divisor of D. Set $d == D/q$, then find p and see if it gives a valid probability ratio $0 \le p/q \le 1$.

In[101]:=
```
Solve[q - 2 p == d, p]
```

Out[101]= $\left\{\left\{p \to \frac{1}{2}\,(-d + q)\right\}\right\}$

In[102]:=
```
Reduce[0 ≤ (q - d)/2 ≤ q && q > 0, d]
```
Out[102]=
```
q > 0 && -q ≤ d ≤ q
```

A valid probability will result so long as $|D|/q \le q$ or $q \ge \sqrt{|D|}$. However, p/q obtained this way is not guaranteed to be in lowest terms. We need to require p, q relatively prime. Furthermore, for p to be integer requires q and d to be the same parity. These conditions are sufficiently mild that there are many instances where more than one ratio gives the same D.

5.2 Trivial solutions

Above, we noted that there are always three trivial solutions to Equation (2), namely $(x, y) == (0, 0)$, $(0, 1)$, and $(1, 0)$. Mapping these to (u, v) we obtain

In[103]:=
```
uvfromxy[{0, 0}]
```
Out[103]=
```
{p, 0}
```

In[104]:=
```
uvfromxy[{0, 1}]
```
Out[104]=
```
{-p + q, 1}
```

In[105]:=
```
uvfromxy[{1, 0}]
```
Out[105]=
```
{-p + q, -1}
```

So the corresponding (u, v) solutions are $(p, 0)$, $(q - p, 1)$, and $(q - p, -1)$.

5.2.1 Additional solutions that always exist if $q == 2\,p - 1$

But note that if (u, v) is a solution to Equation (8), then so also is $(\pm u, \pm v)$. This implies three more solutions obtained by changing the sign of u or v. (Changing the sign of 0 of course does not give a different result.) Mapping these back to (x, y), and verifying that they satisfy Equation (2):

In[106]:=
```
{x, y} = Simplify[xyfromuv[{-p, 0}]]
```
Out[106]=
$$\left\{\frac{p}{2\,p - q}, \frac{p}{2\,p - q}\right\}$$

In[107]:=
```
Simplify[xyeqn[{x, y}]]
```
Out[107]=
```
True
```

In[108]:=
```
{x, y} = Simplify[xyfromuv[{p - q, 1}]]
```
Out[108]=
$$\left\{\frac{-p + q}{2\,p - q}, \frac{p}{2\,p - q}\right\}$$

In[109]:=
```
Simplify[xyeqn[{x, y}]]
```
Out[109]=
```
True
```

In[110]:= `{x, y} = Simplify[xyfromuv[{p - q, -1}]]`

Out[110]= $\left\{\frac{p}{2\,p-q}, \frac{-p+q}{2\,p-q}\right\}$

In[111]:= `Simplify[xyeqn[{x, y}]]`

Out[111]= `True`

In[112]:= `Clear[x, y]`

These correspond to points on the curve that are symmetric partners of the trivial solutions, i.e. at the far end of the ellipse or on the other branch of the hyperbola. However, while in (u, v)-space these solutions are always integer, only in special cases are x, y integer.

These solutions will clearly be integer if $|\,2\,p-q\,| == 1$, i.e. $q == 2\,p \pm 1$. This is an if and only if, for suppose that there is some $k == |2\,p-q| > 1$ that divides p, then k also divides q, contradicting p and q relatively prime. For these integer solutions, the numerators simplify:

In[113]:= `Simplify[-p + q /. {q → 2 p - 1}]`

Out[113]= $-1+p$

In[114]:= `Simplify[-p + q /. {q → 2 p + 1}]`

Out[114]= $1+p$

These three additional solutions will be positive only if $2\,p-q>0$, i.e. $p/q>1/2$. Hence they are admissible only for $q == 2\,p-1$, and negative for $q == 2\,p+1$.

So we already have one result: for ratios of form $\frac{p}{2\,p-1}$, there always exist three solutions $(x, y) == (p, p)$, $(p-1, p)$, and $(p, p-1)$. If $p>1$ these are all admissible.

For hyperbolic ratios of the form $\frac{p}{2\,p+1}$, there are integer solutions $(x, y) == (-p, -p)$, $(-p-1, -p)$, and $(-p, -p-1)$. These are not admissible since they are negative. The examples plotted in Section 5.1.1 are of these forms.

5.3 Reverse search

An initial approach I took was to reverse the problem by computing p/q ratios resulting from the probability equation (1) for all values of x and y below 1000 using a simple program written in Python. Usefully, Python has a `Fraction` class in the `fraction` package, which automatically reduces fractions to lowest terms. (It also has unlimited precision integer arithmetic.) I sorted the resulting list by p/q ratio, displaying all the distinct (x, y) pairs (omitting the symmetric partners) for each ratio.

Here is the Python3 code for the search.

```
# Performs reverse search for red-blue balls puzzle.
# Calculates probability of balls being different colors for all (x,y)
pairs
# with 1 <= x <= y < nmax.
```

```
import sys
from fractions import Fraction

# optional first argument is range of search, default 100
if len(sys.argv) > 1:
    nmax = int(sys.argv[1])
else:
    nmax = 100

result_list = {}

for y in range(1,nmax): # range does not include second value
    for x in range(1,y+1):
        n = y+x
        odds_differ = Fraction(2*y*x,n*(n-1))
        if odds_differ in result_list:
            result_list[odds_differ] += [ [x,y,n] ]
        else:
            result_list[odds_differ] = [ [x,y,n] ]

print("nmax=",nmax)
print("odds","[R,B,N]")
for f in sorted(result_list):
    print(f,result_list[f])
```

I ran this search with nmax set to 1000, i.e. x, $y \leq 999$. (It took only about 47 seconds on my laptop.) The number of (x, y) pairs included in the search was

In[115]:= `n (n + 1) / 2 /. n → 999`

Out[115]= 499 500

The output contained 494 396 lines, indicating that most of the ratios generated had only one solution within the range of the search.

Inspection of the results showed some interesting patterns, suggesting avenues to explore. These will be mentioned in later sections.

5.3.1 Selected results of reverse search

Here are some results of the reverse search, selecting ratios with p of one digit and q of one or two digits. Each line has the ratio p/q followed by a list of solutions in the form of triplets $\{x, y, x+y\}$. Only distinct solutions with $x \leq y$ are listed. They are in order of ascending p/q. The ratios from 1/98 through 2/41 are suppressed to save space, since they follow the same pattern as 1/20 through 1/17.

1/99 {{1, 197, 198}}

[108 lines omitted]

1/20 {{1, 39, 40}}

2/39 {{1, 38, 39}}

1/19 {{1, 37, 38}}

2/37 {{1, 36, 37}}
1/18 {{1, 35, 36}}
2/35 {{1, 34, 35}}
1/17 {{1, 33, 34}}
2/33 {{1, 32, 33}, {32, 992, 1024}}
1/16 {{1, 31, 32}, {31, 930, 961}}
2/31 {{1, 30, 31}, {30, 870, 900}}
1/15 {{1, 29, 30}, {29, 812, 841}}
2/29 {{1, 28, 29}, {28, 756, 784}}
1/14 {{1, 27, 28}, {27, 702, 729}}
2/27 {{1, 26, 27}, {26, 650, 676}}
1/13 {{1, 25, 26}, {25, 600, 625}}
2/25 {{1, 24, 25}, {24, 552, 576}}
1/12 {{1, 23, 24}, {23, 506, 529}}
2/23 {{1, 22, 23}, {22, 462, 484}}
1/11 {{1, 21, 22}, {21, 420, 441}}
2/21 {{1, 20, 21}, {20, 380, 400}}
1/10 {{1, 19, 20}, {19, 342, 361}}
9/88 {{38, 666, 704}}
2/19 {{1, 18, 19}, {18, 306, 324}}
1/9 {{1, 17, 18}, {17, 272, 289}}
2/17 {{1, 16, 17}, {16, 240, 256}}
9/76 {{6, 90, 96}}
5/42 {{4, 60, 64}, {60, 885, 945}}
3/25 {{40, 585, 625}}
5/41 {{24, 345, 369}}
1/8 {{1, 15, 16}, {15, 210, 225}}
7/55 {{3, 42, 45}, {42, 574, 616}}
9/70 {{58, 783, 841}}
2/15 {{1, 14, 15}, {14, 182, 196}}
7/51 {{10, 126, 136}}
5/36 {{6, 75, 81}, {75, 925, 1000}}
1/7 {{1, 13, 14}, {13, 156, 169}}
8/55 {{39, 456, 495}}
2/13 {{1, 12, 13}, {12, 132, 144}}
7/44 {{46, 483, 529}}
1/6 {{1, 11, 12}, {11, 110, 121}}
5/29 {{22, 210, 232}}
7/40 {{93, 868, 961}}
2/11 {{1, 10, 11}, {10, 90, 100}, {90, 801, 891}}
9/49 {{5, 45, 50}, {45, 396, 441}}

9/47 {{15, 126, 141}}

5/26 {{21, 175, 196}}

1/5 {{1, 9, 10}, {9, 72, 81}, {72, 568, 640}}

8/39 {{3, 24, 27}, {24, 184, 208}}

5/24 {{34, 255, 289}}

9/41 {{5, 36, 41}, {36, 252, 288}}

2/9 {{1, 8, 9}, {8, 56, 64}, {56, 385, 441}}

9/40 {{124, 837, 961}}

7/31 {{4, 28, 32}, {28, 189, 217}}

5/22 {{69, 460, 529}}

8/35 {{115, 760, 875}}

7/30 {{2, 14, 16}, {14, 91, 105}, {91, 585, 676}}

5/21 {{116, 725, 841}}

1/4 {{1, 7, 8}, {7, 42, 49}, {42, 246, 288}}

9/35 {{3, 18, 21}, {18, 102, 120}}

6/23 {{14, 78, 92}, {78, 429, 507}}

9/34 {{40, 216, 256}}

4/15 {{57, 304, 361}}

7/26 {{100, 525, 625}}

7/25 {{4, 21, 25}, {21, 105, 126}, {105, 520, 625}}

2/7 {{1, 6, 7}, {6, 30, 36}, {30, 145, 175}, {145, 696, 841}}

7/24 {{51, 238, 289}}

5/17 {{3, 15, 18}, {15, 70, 85}, {70, 322, 392}}

3/10 {{22, 99, 121}, {99, 441, 540}}

5/16 {{186, 775, 961}}

9/28 {{11, 45, 56}, {45, 180, 225}, {180, 716, 896}}

1/3 {{1, 5, 6}, {5, 20, 25}, {20, 76, 96}, {76, 285, 361}}

9/26 {{46, 162, 208}, {162, 567, 729}}

8/23 {{10, 36, 46}, {36, 126, 162}}

7/20 {{18, 63, 81}, {63, 217, 280}, {217, 744, 961}}

6/17 {{280, 945, 1225}}

5/14 {{39, 130, 169}, {130, 430, 560}}

9/25 {{188, 612, 800}}

4/11 {{105, 336, 441}}

8/21 {{5, 16, 21}, {16, 48, 64}, {48, 141, 189}}

5/13 {{3, 10, 13}, {10, 30, 40}, {30, 87, 117}, {189, 540, 729}}

7/18 {{2, 7, 9}, {7, 21, 28}, {21, 60, 81}, {95, 266, 361}, {266, 742, 1008}}

9/23 {{6, 18, 24}, {18, 51, 69}, {196, 540, 736}}

2/5 {{1, 4, 5}, {4, 12, 16}, {12, 33, 45}, {33, 88, 121}, {88, 232, 320}, {232, 609, 841}}

9/22 {{3, 9, 12}, {9, 24, 33}, {50, 126, 176}, {126, 315, 441}, {315, 785, 1100}}

5/12 {{14, 35, 49}, {35, 85, 120}, {85, 204, 289}}

8/19 {{51, 120, 171}, {120, 280, 400}, {280, 651, 931}}
3/7 {{2, 6, 8}, {6, 15, 21}, {15, 35, 50}, {52, 117, 169}, {117, 261, 378}, {261, 580, 841}}
7/16 {{310, 651, 961}}
9/20 {{13, 27, 40}, {27, 54, 81}, {54, 106, 160}}
5/11 {{7, 15, 22}, {15, 30, 45}, {30, 58, 88}, {184, 345, 529}, {345, 645, 990}}
6/13 {{4, 9, 13}, {9, 18, 27}, {18, 34, 52}, {70, 126, 196}, {126, 225, 351}, {225, 400, 625}}
7/15 {{3, 7, 10}, {7, 14, 21}, {14, 26, 40}, {44, 77, 121}, {77, 133, 210}, {133, 228, 361}, {370, 630, 1000}}
8/17 {{6, 12, 18}, {12, 22, 34}, {57, 96, 153}, {96, 160, 256}, {160, 265, 425}}
9/19 {{21, 36, 57}, {36, 60, 96}, {365, 585, 950}, {585, 936, 1521}}
1/2 {{1, 3, 4}, {3, 6, 9}, {6, 10, 16}, {10, 15, 25}, {15, 21, 36}, {21, 28, 49}, {28, 36, 64}, {36, 45, 81}, {45, 55, 100}, {55, 66, 121}, {66, 78, 144}, {78, 91, 169}, {91, 105, 196}, {105, 120, 225}, {120, 136, 256}, {136, 153, 289}, {153, 171, 324}, {171, 190, 361}, {190, 210, 400}, {210, 231, 441}, {231, 253, 484}, {253, 276, 529}, {276, 300, 576}, {300, 325, 625}, {325, 351, 676}, {351, 378, 729}, {378, 406, 784}, {406, 435, 841}, {435, 465, 900}, {465, 496, 961}, {496, 528, 1024}, {528, 561, 1089}, {561, 595, 1156}, {595, 630, 1225}, {630, 666, 1296}, {666, 703, 1369}, {703, 741, 1444}, {741, 780, 1521}, {780, 820, 1600}, {820, 861, 1681}, {861, 903, 1764}, {903, 946, 1849}, {946, 990, 1936}}
9/17 {{8, 9, 17}, {9, 9, 18}}
8/15 {{2, 4, 6}, {4, 6, 10}, {7, 8, 15}, {8, 8, 16}}
7/13 {{6, 7, 13}, {7, 7, 14}}
6/11 {{5, 6, 11}, {6, 6, 12}}
5/9 {{4, 5, 9}, {5, 5, 10}}
4/7 {{3, 4, 7}, {4, 4, 8}}
3/5 {{2, 3, 5}, {3, 3, 6}}
2/3 {{1, 2, 3}, {2, 2, 4}}
1/1 [[1, 1, 2]]

Observations

It is notable that the list has many more ratios less than 1/2 than greater than 1/2. Many of the ratios less than 1/2 have only one solution listed, but it is shown in Section 12.2 that all of them have more solutions (infinitely many in fact, since all have nonsquare D), but those lie beyond the limit of $x, y < 1000$ used in the search. Of the ratios where more than one solution appears, it is notable that very often the y value of one solution is the same as the x value of the next. The case $p/q == 1/2$ shows this pattern continuously for all solutions, which makes sense since we saw that the solutions are pairs of successive triangular numbers. For many other ratios, however, each sequence of solutions related in this way has at most 3 members.

For ratios between 0 and 1/2, the number of solutions listed for each ratio tends to increase the closer the ratio is to 1/2. Since each ratio in this range actually has an infinite number of solutions, this trend is simply the result of the fact that more of them are within the range of the search. Thus the solutions for a given ratio tend to increase in size more slowly, for the ratios closer to 1/2.

Some other observations: if $p == 1$ then the total number of balls in the smallest solution is $2q$, while if $p == 2$ the total is q. In both cases $x == 1$. For ratios greater than 1/2, all the ratios in these results have a solution in which $x == y == p$, along with another in which $x == p - 1$, $y == p$. We shall see that this only holds for ratios of the form $p/(2p - 1)$, which is the case for all of these.

In later sections of this notebook, we will find the explanations for these observations.

5.4 The "recycling" recurrence

One pattern I noticed in the reverse search results was that for many p/q ratios, there were series of solutions (x, y) in which the larger value (y) in one solution became the smaller value (x) in another, larger solution. For some ratios these series continued indefinitely up to the limit of the search; but for many ratios there were only a few members of any such series, typically three although sometimes only two. (I have found none with four or more admissible solutions in a series that did not continue indefinitely.) With this hint, we can solve for a formula that, given one (x, y) pair, generates another (y, z) pair with $x \neq z$, having the same p/q value.

```
In[116]:= Solve[{probdifferent[{x, y}] == probdifferent[{y, z}], x ≠ z}, z]
```

Out[116]= $\left\{\left\{z \to \frac{(-1 + y)\, y}{x}\right\}\right\}$

The formula preserves the probability ratio p/q, although it is not guaranteed to yield integer values. Since the sequence re-uses a number from one solution in the next, I call this the *"recycling recurrence"* to distinguish it from another recurrence that will be discussed in a later section.

In order for this formula to yield increasingly large solutions, i.e. to go from (x, y) with $x < y$ to a larger solution (y, z) with $y < z$, we need to have

$$z == \frac{y\,(y - 1)}{x} > y$$

which requires $y - 1 > x$, i.e. $y > x + 1$.

The trivial solutions are problematic for the recycling recurrence. Only one of them even gives a defined result: $(1, 0) \to (0, 0)$ but $(0, 0)$ or $(0, 1)$ yields $z == 0/0$.

Define a function to calculate the recycling recurrence neighbor to (x, y).

In[117]:= `recycle[{x_, y_}] :=` $\left\{y, \frac{y\,(y - 1)}{x}\right\}$

If $x < y$ this yields the next larger member of the series (except for some special cases discussed below). If $x \geq y$ then it runs in reverse, yielding the previous smaller member of the series.

5.4.1 Examples

The *Varsity Math* case follows this recurrence.

In[118]:=
```
TableForm[RecurrenceTable[{{x[i + 1], y[i + 1]} == recycle[{x[i], y[i]}],
    x[1] == 1, y[1] == 3},
   {x[i], y[i]}, {i, 9}]]
```

Out[118]//TableForm=

1	3
3	6
6	10
10	15
15	21
21	28
28	36
36	45
45	55

Here is another example in which the recurrence yields integer values indefinitely:

In[119]:=
```
xypairs = RecurrenceTable[{{x[i + 1], y[i + 1]} == recycle[{x[i], y[i]}],
    x[1] == 1, y[1] == 6},
   {x[i], y[i]}, {i, 10}];
```

In[120]:=
```
TableForm[xypairs]
```

Out[120]//TableForm=

1	6
6	30
30	145
145	696
696	3336
3336	15 985
15 985	76 590
76 590	366 966
366 966	1 758 241
1 758 241	8 424 240

Verify that these all have the same probability of different colors.

In[121]:=
```
DeleteDuplicates[probdifferent /@ xypairs]
```

Out[121]= $\left\{\frac{2}{7}\right\}$

Here is an example where the recurrence yields just 3 integer pairs before producing inadmissible values.

In[122]:=
```
probdifferent[{2, 5}]
```

Out[122]= $\frac{10}{21}$

In[123]:=
```
TableForm[RecurrenceTable[{{x[i + 1], y[i + 1]} == recycle[{x[i], y[i]}],
    x[1] == 2, y[1] == 5},
   {x[i], y[i]}, {i, 4}]]
```

Out[123]//TableForm=

2	5
5	10
10	18
18	$\frac{153}{5}$

Running it in reverse also gives a fractional value for the preceding pair.

In[124]:=
```
recycle[{5, 2}]
```

Out[124]= $\left\{2, \frac{2}{5}\right\}$

So the related solutions form a triplet. We shall see there is a reason why many solutions occur in triplets. However, I don't know why there are never more than three admissible solutions in a sequence except when the recurrence continues infinitely.

5.4.2 Elliptical case

The recycling recurrence has to stop progressing to larger values for the elliptical case, since the ellipse has a finite range. In Section 5.2.1 we showed that for ratios of the form $p/(2p-1)$ there are solutions $(p-1, p)$, (p, p), and $(p, p-1)$ which form a recycling triplet. They do not continually advance to larger values, and indeed, these fail to satisfy the requirement $y > x - 1$. Here is what happens for the ratio $p/q = 4/7$:

In[125]:=
```
TableForm[
 RecurrenceTable[{{x[i + 1], y[i + 1]} == recycle[{x[i], y[i]}],
    x[1] == 4, y[1] == 5},
   {x[i], y[i]}, {i, 4}]]
```

Out[125]//TableForm=

4	5
5	5
5	4
4	$\frac{12}{5}$

In general, the recycling recurrence around the vertex of the ellipse runs as follows:

$(x - 1, x) \to (x, x) \to (x, x - 1)$

Running the recurrence further may or may not yield an integer. Here is an example where it does yield integers (which are trivial solutions) for a couple more steps:

In[126]:=
```
RecurrenceTable[{{x[i + 1], y[i + 1]} == recycle[{x[i], y[i]}],
  x[1] == 1, y[1] == 2},
 {x[i], y[i]}, {i, 5}]
```

Out[126]=
```
{{1, 2}, {2, 2}, {2, 1}, {1, 0}, {0, 0}}
```

In[127]:=
```
probdifferent[{2, 2}]
```

Out[127]= $\frac{2}{3}$

5.4.3 When removing one ball does not change the odds

An interesting observation is that if $x == y$, the recycling recurrence gives as the next value $(x,\ x-1)$.

In[128]:=
```
recycle[x, x]
```

Out[128]=
```
recycle[x, x]
```

That is, assuming $x \geq 2$ so that the result is admissible, the probability ratio p/q for $(x,\ x)$ is the same as for $(x-1,\ x)$ and $(x,\ x-1)$. In words: if the bag initially contains equal numbers of red and blue balls, removing one ball does not change the odds that the next two balls drawn will be different colors. That's pretty neat. For example:

In[129]:=
```
probdifferent[{8, 8}]
```

Out[129]= $\frac{8}{15}$

In[130]:=
```
probdifferent[{7, 8}]
```

Out[130]= $\frac{8}{15}$

5.4.4 Function to generate recycling series

The recycling recurrence gives us a way, once one solution to (2) has been found for a given p/q, to generate other solutions. It is not guaranteed to yield integer solutions, however. As discussed above, usually these series are limited to at most three solutions.

Let us define a function to find neighbors to a given solution via the recycling recurrence.

Argument `xylist` is a list of $\{x,\ y\}$ pairs whose recycling neighbors are to be found. The function returns the original list plus any companions up to two steps away. Only non-negative integer solutions are returned. (The trivial solutions may be returned.) Optional argument `tableform` is `True` to select tabular output (default) or `False` to return in list form (useful when feeding result to another function).

In[131]:=
```
recycleSolutions[xylist_, tableform_: True] :=
 Module[{xyfwd1, xyfwd2, xyrev1, xyrev2, xyall, xlist, ylist},
  xyfwd1 = Table[If[x ≠ 0, {y, y (y - 1) / x}, {-9, -9}] /.
      {x -> xylist[[i]][[1]], y → xylist[[i]][[2]]}
    , {i, Length[xylist]}];
  xyrev1 = Table[If[y ≠ 0, {x (x - 1) / y, x}, {-9, -9}] /.
      {x -> xylist[[i]][[1]], y → xylist[[i]][[2]]}
    , {i, Length[xylist]}];
  xyfwd2 = Table[If[x ≠ 0, {y, y (y - 1) / x}, {-9, -9}] /.
      {x -> xyfwd1[[i]][[1]], y → xyfwd1[[i]][[2]]}
    , {i, Length[xyfwd1]}];
  xyrev2 = Table[If[y ≠ 0, {x (x - 1) / y, x}, {-9, -9}] /.
      {x -> xyrev1[[i]][[1]], y → xyrev1[[i]][[2]]}
    , {i, Length[xylist]}];
  xyall = DeleteDuplicates[Join[xylist, xyfwd1, xyrev1, xyfwd2, xyrev2]];
  xlist = Table[xyall[[i]][[1]], {i, Length[xyall]}];
  ylist = Table[xyall[[i]][[2]], {i, Length[xyall]}];
  ylist = Extract[ylist, Position[xlist, _Integer?NonNegative]];
  xlist = Extract[xlist, Position[xlist, _Integer?NonNegative]];
  xlist = Extract[xlist, Position[ylist, _Integer?NonNegative]];
  ylist = Extract[ylist, Position[ylist, _Integer?NonNegative]];
  If[tableform,
   TableForm[
    Sort[Table[{xlist[[i]], ylist[[i]]}, {i, Length[xlist]}]]
    , TableHeadings → {None, {"x", "y"}}]
   , Sort[Table[{xlist[[i]], ylist[[i]]}, {i, Length[xlist]}]]
  ]
 ]
```

Exercise it on a couple of examples.

A hyperbolic case with only three integer solutions in a series.

In[132]:=
```
recycleSolutions[{{5, 10}}]
```

Out[132]//TableForm=

x	y
2	5
5	10
10	18

The *Varsity Math* case:

In[133]:=
```
recycleSolutions[{{6, 10}}, False]
```

Out[133]= {{1, 3}, {3, 6}, {6, 10}, {10, 15}, {15, 21}}

An elliptical case that hits all integers in its range, including the trivial solutions.

In[134]:=
```
recycleSolutions[{{1, 2}, {2, 2}}, False]
```
Out[134]=
```
{{0, 0}, {0, 1}, {1, 0}, {1, 2}, {2, 1}, {2, 2}}
```

6 Special cases

There are a few special cases that can be solved with simple algebra.

6.1 Special case: $p/q == 0$

If $p/q == 0$, it means that the probability of drawing balls of different colors is zero. This is obviously the case if and only if all the balls are the same color. So there are two infinite families of admissible solutions:

$x == 0, \quad y \in \mathbb{Z}, \quad y \geq 2$

$y == 0, \quad x \in \mathbb{Z}, \quad x \geq 2$

Solving Equation (2) formally for this case, set $p == 0$. In lowest terms, $q == 1$. The equation reduces to

In[135]:=
```
Simplify[xyeqn[{x, y}] /. {p → 0, q → 1}]
```
Out[135]=
```
x y == 0
```

Clearly this requires either $x == 0$ or $y == 0$. Requiring admissibility yields the above solutions.

In[136]:=
```
Reduce[{xyeqn[{x, y}] /. {p → 0, q → 1}, x ≥ 0, y ≥ 0, x + y ≥ 2}]
```
Out[136]=
```
(x ≥ 2 && y == 0) || (y ≥ 2 && x == 0)
```

6.2 Special case: $p/q == 1$

If $p/q == 1$, it means the likelihood of drawing balls of different colors is certainty. Obviously there is just one admissible solution, in which the bag contains two balls of different colors:

$x == 1, \quad y == 1$

Solving Equation (2) formally for this case, set $p == 1$, $q == 1$. The equation becomes

In[137]:=
```
Simplify[xyeqn[{x, y}] /. {p → 1, q → 1}]
```
Out[137]= $x^2 + y^2 == x + y$

$x^2 + y^2 - x - y == 0$

$x\,(x - 1) == -y\,(y - 1)$

Since we require $x \geq 0$ and $y \geq 0$, there will be an irresolvable sign conflict unless $x(x - 1) == y(y - 1) == 0$. The solutions (0, 0), (0, 1), and (1, 0) are not admissible, leaving only (1, 1). Mathematica agrees:

In[138]:=
```
Reduce[{xyeqn[{x, y}] /. {p → 1, q → 1}, x ≥ 0, y ≥ 0, x + y ≥ 2}]
```
Out[138]=
```
y == 1 && x == 1
```

6.3 Special case: p == 1 or 2

Suppose $p == 1$. Then the probability of drawing different colors is

In[139]:=
```
probequation[{x, y}] /. p → 1
```

Out[139]= $\frac{2\,x\,y}{(-1+x+y)\,(x+y)} == \frac{1}{q}$

If $p == 2$, it is

In[140]:=
```
probequation[{x, y}] /. p → 2
```

Out[140]= $\frac{2\,x\,y}{(-1+x+y)\,(x+y)} == \frac{2}{q}$

These can be unified by writing

$$(x+y)\,(x+y-1) == a\,x\,y \qquad (14)$$

where $a == 2\,q$ if $p == 1$, and $a == q$ if $p == 2$. Since $a > 1$ this equation always has an admissible solution with $x == 1$:

In[141]:=
```
Solve[(x + y) (x + y - 1) == a x y /. x → 1, y]
```

Out[141]= {{y → 0}, {y → -1 + a}}

In[142]:=
```
{{y → 0}, {y → -1 + a}}
```

Out[142]= {{y → 0}, {y → -1 + a}}

So there is always the solution

$$\begin{array}{ll} (1,\ 2\,q-1), & p == 1 \\ (1,\ q-1), & p == 2 \end{array}$$

The recycling recurrence will generate additional solutions. We can show that it will always yield integer values. This requires that x divide $y(y-1)$. From Equation (14), and the fact that a is an integer, x must divide $(x+y)\,(x+y-1)$. Using the fact that a divides $b\,c$ if and only if it divides $(b+a)\,(c+a)$, we have the result that x divides $y(y-1)$ and so $y(y-1)/x$ is always integer. However, this does not imply that the recycling recurrence will yield an infinite number of solutions. For $a == 2$ or 3, $y == a-1$ is 1 or 2, violating the requirement we saw earlier that $y > x+1$ for the recurrence to advance to larger values. These are elliptical cases: $a == 2$ corresponds to $p/q = 1/1$, while $a == 3$ corresponds to $p/q == 2/3$. For $a == 4$, $p/q == 1/2$, and for $a > 4$, $p/q < 1/2$. For those cases, which are in the parabolic and the hyperbolic regime respectively, the sequence of solutions is infinite.

Here is an example, for $p/q == 2/7$, $a == 7$.

```
In[143]:= TableForm[RecurrenceTable[{{x[i + 1], y[i + 1]} == recycle[{x[i], y[i]}],
    x[1] == 1, y[1] == 6},
   {x[i], y[i]}, {i, 5}], TableHeadings → {None, {"x", "y"}}]
```

Out[143]//TableForm=

x	y
1	6
6	30
30	145
145	696
696	3336

Comment

We have not shown that there are no other solutions besides those generated by the recycling recurrence from the initial solution. That is in fact the case, as is proved in Section 12.4.

7 Parabolic case: $p/q == 1/2$

The parabolic case $p/q == 1/2$ is the *Varsity Math* problem. We have already solved it in Section 2. The number of solutions is infinite. If we impose uniqueness by requiring $x \leq y$, every solution is of the form

$$x == \frac{v\,(v-1)}{2},\quad y == \frac{v\,(v+1)}{2},\quad t == v^2,\quad v == 2,\,3,\,\ldots$$

i.e. pairs of successive triangular numbers making the total number of balls a square.

In terms of the derivation in Section 4.2.5, we need to use Equation (6), the equation in t and v, not (8), the equation in u and v, because the transformation to u, v breaks down when $q == 2\,p$. Recall Equation (6) is

```
In[144]:= tveqn[{t, v}]
```

Out[144]= $2\,p\,t + (-2\,p + q)\,t^2 - q\,v^2 == 0$

Here this reduces to

```
In[145]:= Simplify[tveqn[{t, v}] /. {p → 1, q → 2}]
```

Out[145]= $t == v^2$

as found before.

8 Elliptical case: $p/q > 1/2$

8.1 Method of bracketed direct search

I have not come across a method for solving the elliptical case other than direct search, i.e. trying all viable possibilities. This is feasible in principle since the number of possible solutions is finite, although it may become expensive for values of p/q close to 1/2.

The search can be confined to the limits of the minimum and maximum values of x reached by the ellipse, testing the integral values of x to see if they yield integral values of y in Equation (2). Alternatively, one can search within the range of v and u values in Equation (8). This turns out to be a much better way to go. We can rewrite Equation (8) in terms of $z == p/q$ as

In[146]:=
```
Simplify[uveqn[{u, v}] /. {p → z, q → 1}]
```

Out[146]= $u^2 + v^2 \, (-1 + 2\,z) == z^2$

For $z > 1/2$ the coefficient of v is positive. The maximum value of v occurs when $u == 0$:

In[147]:=
```
maxvsolns = Solve[uveqn[{0, v}] /. {p → z, q → 1}, v]
```

Out[147]= $\left\{\left\{v \to -\frac{z}{\sqrt{-1+2\,z}}\right\}, \left\{v \to \frac{z}{\sqrt{-1+2\,z}}\right\}\right\}$

Since changing the sign of v simply reverses the order of x, y in any solution, we can restrict to $v \geq 0$.

In[148]:=
```
maxv[z_] := Evaluate[Part[maxvsolns, 2, 1, 2]]
```

In[149]:=
```
maxv[z]
```

Out[149]= $\frac{z}{\sqrt{-1+2\,z}}$

For $z == 1$,

In[150]:=
```
maxv[1]
```

Out[150]= 1

Let's examine this bound, to see how it grows as the probability ratio approaches 1/2 from above. Let ϵ be the amount by which the ratio differs from 1/2.

In[151]:= $\texttt{Simplify}\left[\texttt{maxv[z] /. }\left\{z \to \frac{1}{2} + \epsilon\right\}\right]$

Out[151]= $\frac{1 + 2\,\epsilon}{2\,\sqrt{2}\,\sqrt{\epsilon}}$

As $\epsilon \to 0$, the $2\,\epsilon$ in the numerator can be neglected and the bound becomes approximately

$$|v| \leq \frac{1}{2\,\sqrt{2}\,\sqrt{\epsilon}}$$

which grows only as $\epsilon^{-1/2}$. For instance, if $\epsilon == 10^{-6}$, the maximum value of v is only 353.

In[152]:= $\texttt{Floor}\left[\texttt{maxv}\left[\frac{1}{2} + 10^{-6}\right]\right]$

Out[152]= 353

For $\epsilon == 10^{-12}$, a factor of a million smaller, the maximum value of v is only about a factor of a thousand larger, still not beyond reach of a search by computer in a reasonable time.

In[153]:=
```
Floor[maxv[1/2 + 10^-12]]
```

Out[153]= 353 553

Comparison to direct search in (*x*, *y*)-space

The direct search in (*x*, *y*)-space, using Equation (2), is much less favorable. Solve for the endpoint as a function of probability ratio *z*. (The maximum *x* reached by the ellipse is slightly larger, but that does not matter for our purposes here.)

In[154]:=
```
Solve[probdifferent[{x, x}] == z, x]
```

Out[154]= $\left\{\left\{x \to \frac{z}{-1+2\,z}\right\}\right\}$

Now determine the asymptotic behavior as *z* approaches 1/2.

In[155]:=
```
Simplify[z/(2 z - 1) /. z -> 1/2 + ϵ]
```

Out[155]= $\frac{1}{4}\left(2+\frac{1}{\epsilon}\right)$

For small values of ϵ, the endpoint is asymptotically

$$x == \frac{1}{4\,\epsilon}$$

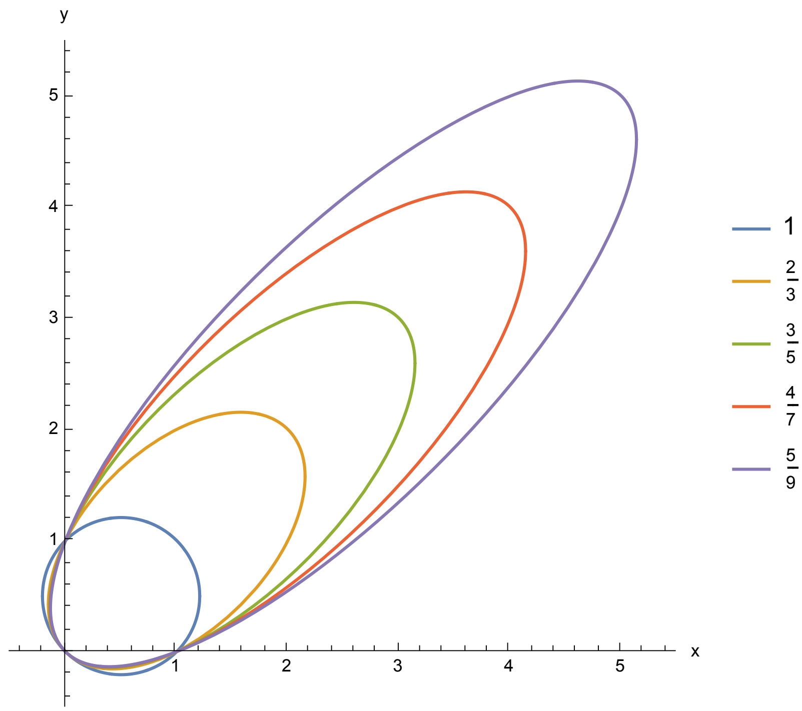
For $\epsilon == 10^{-6}$, this is 250 000, 700 times larger than the maximum *v*. Looking at the graphs of the ellipse in (*x*, *y*)-space vs. (*u*, *v*)-space, it is clear why: the former is tilted to have its axis along *x* == *y*, so one needs to search along its whole length, whereas for *u*, *v* it is aligned with the axes and is much narrower in the *v*-direction than in the *u*-direction. This is shown with plots in the next section.

8.1.1 Plotting the elliptical case vs. *z*

Here we plot a family of ellipses as the *p*/*q* ratio varies. Use ratios of the form *p*/(2 *p* – 1) from 1/1 to 5/9.

```
In[156]:= yvaluesforplot = Table[{y /. Solve[xyeqn[{x, y}] /. q → 2 p - 1, y]}, {p, 5}];
Plot[yvaluesforplot,
 {x, -1 / 2, 11 / 2},
 PlotRange → {{-1 / 2, 11 / 2}, {-1 / 2, 11 / 2}}, AspectRatio → 1,
 PlotLegends → Table[p / (2 p - 1), {p, 5}], AxesLabel → {"x", "y"}]
```

Out[157]=

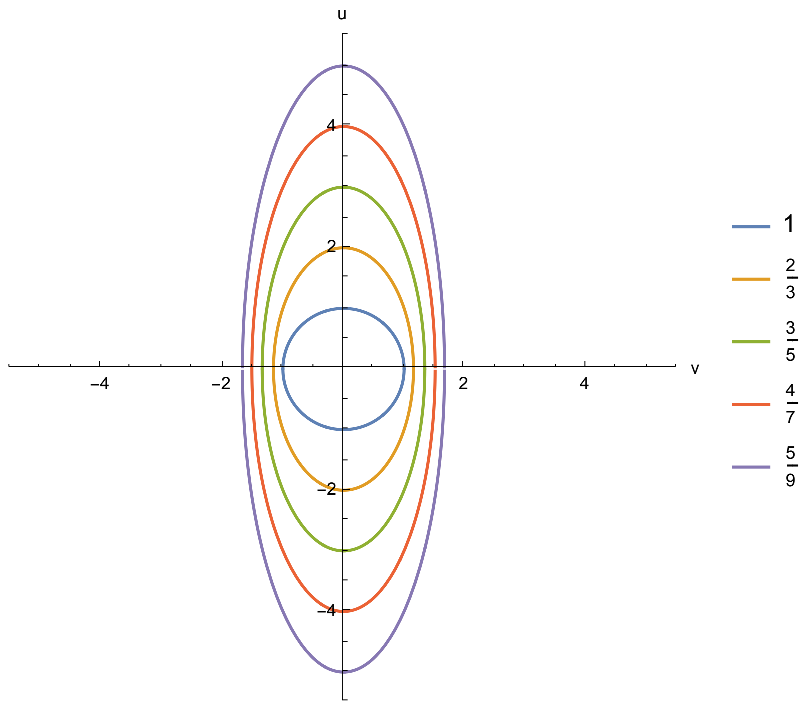


For p/q == 1, the ellipse is a circle. As p/q decreases toward 1/2, the ellipses elongate, approaching the parabola at p/q == 1/2.

Here is the same set of cases plotted in (u, v)-space. From Equation (8), the maximum value of u is p, while the maximum value of v is $\frac{p}{\sqrt{q(2p-q)}}$.

In[158]:=
```
uvaluesforplot = Table[u /. Solve[uveqn[{u, v}] /. q → 2 p - 1, u], {p, 5}];
Plot[uvaluesforplot,
 {v, -2, 2},
 PlotRange → {{-11 / 2, 11 / 2}, {-11 / 2, 11 / 2}}, AspectRatio → 1,
 PlotLegends → Table[p / (2 p - 1), {p, 5}], AxesLabel → {"v", "u"}]
```

Out[159]=



One can see that *v* grows more slowly than *u*.

8.2 Function to solve elliptic case by bracketed search

Let us put the bracketed search method into a function. Argument $z == p/q$. Optional argument `table-form` is `True` to format results in a table (default), `False` to produce a list. Include an `If` to proceed only for ratios in the elliptical range. It's written independent of any local function definitions so it can be copied into another notebook and still work.

Outline of the function:

- Compute range of *v* to search, using formula found above: $0 \leq v \leq v_{max}$.
- Compute positive *u* values corresponding to each integer value of *v* in range.
- Sift out integer values of *u* and pair with corresponding *v* values. Note that *v* values are offset from `Position` values by 1 since *v* starts at 0 while `Position` starts at 1. Augment list with $-u$ values since they are also solutions and give distinct (x, y).
- Convert (u, v) solutions to (t, v).

- Convert (t, v) solutions to (x, y).
- Sift out positive integer values of x and y to yield only admissible solutions. (The trivial solutions are not output.)

In[160]:=
```
solveEllipticalBySearch[z_, tableform_: True] := Module[
  {p, q, maxv, ufromv, uvaluesall, hits,
   uvalues, vvalues, tvalues, xvalues, yvalues, xyvalues},
  p = Numerator[z];
  q = Denominator[z];
  If[1 / 2 < z ≤ 1, (* make sure this is an elliptical case *)
   maxv = Floor[z / Sqrt[2 z - 1]];
   ufromv = Sqrt[p^2 - q (2 p - q) v^2];
   uvaluesall = Table[ufromv, {v, 0, maxv}];
   hits = Position[uvaluesall, _Integer, {1}];
   uvalues = Join[Extract[uvaluesall, hits], -Extract[uvaluesall, hits]];
   vvalues = Flatten[Join[hits - 1, hits - 1]];
   tvalues = (u - p) / (q - 2 p) /. u → uvalues;
   xvalues =
    Table[(t - v) / 2 /. {t → tvalues[[i]], v → vvalues[[i]]}, {i, Length[vvalues]}];
   yvalues = Table[(t + v) / 2 /. {t → tvalues[[i]], v → vvalues[[i]]},
     {i, Length[vvalues]}];
   (* extract the admissible solutions *)
   yvalues = Extract[yvalues, Position[xvalues, _Integer?Positive, {1}]];
   xvalues = Extract[xvalues, Position[xvalues, _Integer?Positive, {1}]];
   xyvalues = Sort[Table[{xvalues[[i]], yvalues[[i]]}, {i, Length[xvalues]}]];
   If[tableform,
    TableForm[xyvalues, TableHeadings → {None, {x, y}}]
    , (* else *)
    xyvalues]
   , (* Else *) Print["argument out of range"]]
 ]
```

8.2.1 Examples

Exercise it on some examples.

In[161]:=
```
solveEllipticalBySearch[1]
```

Out[161]//TableForm=

x	y
1	1

Here is a ratio of form $p/(2p-1)$ and hence has vertex solutions (see Section 5.2.1). It also has other solutions.

In[162]:=
```
solveEllipticalBySearch[200 / 399]
```

Out[162]//TableForm=

x	y
25	32
90	100
100	110
168	175
199	200
200	200

In[163]:=
```
solveEllipticalBySearch[3 / 5, False]
```

Out[163]=
```
{{2, 3}, {3, 3}}
```

Here is one that has no admissible solution.

In[164]:=
```
solveEllipticalBySearch[4 / 5]
```

Out[164]//TableForm=
```
{}
```

This ratio is not elliptical.

In[165]:=
```
solveEllipticalBySearch[1 / 9]
```

```
argument out of range
```

8.2.2 Example of a very elongated ellipse

Seeking a challenge, I thought it would be interesting to see how the bracketed search method fares on an elliptical case for which the ellipse is very large.

Among the results of the reverse search, which found p/q ratios for all pairs of numbers x, $y < 1000$, I found the ratio closest to and larger than 1/2. It has a solution $(x, y) == (947, 991)$. It may have larger solutions not found by the search. Here is the p/q ratio:

In[166]:=
```
zelongated = probdifferent[{947, 991}]
```

Out[166]=
$$\frac{938\,477}{1\,876\,953}$$

Here it is numerically.

In[167]:=
```
zelongated // N
```

Out[167]=
```
0.5
```

We know it is not exactly 1/2 so let’s see by how much it differs

In[168]:=
```
N[zelongated, 10]
```

Out[168]=
```
0.5000002664
```

Here is the difference.

In[169]:= `zelongated - 1/2 // N`

Out[169]= 2.66389×10^{-7}

Check to see whether this ratio is in the form $q == 2\,p - 1$ to have an endpoint solution.

In[170]:= `(2 p - q) /. {p → Numerator[zelongated], q → Denominator[zelongated]}`

Out[170]= 1

So it is indeed by happy chance in the required form. (Actually, Section 8.5.3 shows it is not by chance.) There will be a solution (x, x) with $x == p == 938\,477$ and companion solutions $(x - 1, x)$ and $(x, x - 1)$. Given the solution found by the reverse search, by symmetry we know there will also be solutions diametrically opposite to it:

In[171]:= `{Numerator[zelongated] - 991, Numerator[zelongated] - 947}`

Out[171]= {937 486, 937 530}

There will also of course be the counterparts obtained by swapping x and y.

However, we don't know if there are any more solutions besides these, since the reverse search was not exhaustive for this ratio. There could be other solutions somewhere above 1000. Let's assess how much searching will be needed.

In[172]:= `Floor[maxv[zelongated]]`

Out[172]= 685

That is not bad at all. Here we go, using our spiffy new function. Let's time it.

In[173]:= `Timing[solveEllipticalBySearch[zelongated]]`

Out[173]= {0.064212,

x	y
947	991
937 486	937 530
938 476	938 477
938 477	938 477

}

Far less than a second to find the solutions on my laptop, so this is not a challenging case after all. Clearly ratios much closer to 1/2 could be solved. Disappointingly, the search did not turn up any solutions that we had not already predicted based on the form of p/q and the one solution found by the reverse search.

8.3 Exhaustive enumeration of solutions

For elliptical cases where the maximum x is not very large, there can only be a few solutions at most. The ellipse grows smaller as p/q grows from 1/2 toward 1. Thus for p/q values above a certain value, the solutions will be for x and y within a very limited range. Only a limited number of p/q values can result. This allows all possible probability ratios above that chosen value having solutions to be enumerated, and all other ratios in that range have no solution. Let us choose a maximum x and y of 5 as a very manageable value. The value of z above which the endpoint $x \le 5$ (so that floor(x) ≤ 5) is given by

In[174]:=
```
zrangeforxle5 = Reduce[z/(2 z - 1) ≤ 5 && 1/2 < z ≤ 1, z]
```

Out[174]= $\frac{5}{9} \le z \le 1$

In[175]:=
```
minzforxle5 = zrangeforxle5[[1]]
```

Out[175]= $\frac{5}{9}$

Here is the lower limit numerically.

In[176]:=
```
minzforxle5 // N
```

Out[176]= 0.555556

So all solutions for ratios $\frac{5}{9} \le \frac{p}{q} \le 1$ must have $x,\ y \le 5$. Here is the table of ratios in this range that have solutions. For uniqueness we impose $x \le y$. (Mathematica note: since the table is triangular, grouping is needed to arrange it in the right format.)

In[177]:=
```
TableForm[Table[{{x, y, probdifferent[{x, y}]}}, {y, 1, 5}, {x, 1, y}]]
```

Out[177]//TableForm=

1 1 1				
1 2 $\frac{2}{3}$	2 2 $\frac{2}{3}$			
1 3 $\frac{1}{2}$	2 3 $\frac{3}{5}$	3 3 $\frac{3}{5}$		
1 4 $\frac{2}{5}$	2 4 $\frac{8}{15}$	3 4 $\frac{4}{7}$	4 4 $\frac{4}{7}$	
1 5 $\frac{1}{3}$	2 5 $\frac{10}{21}$	3 5 $\frac{15}{28}$	4 5 $\frac{5}{9}$	5 5 $\frac{5}{9}$

Some ratios in the table are smaller than the z threshold, and some are outside of the elliptical range (having $z \le 1/2$). Let’s arrange them in ascending order.

In[178]:=
```
DeleteDuplicates[Sort[Flatten[Table[probdifferent[{x, y}], {x, 1, 5}, {y, 1, x}]]]]
```

Out[178]= $\left\{\frac{1}{3}, \frac{2}{5}, \frac{10}{21}, \frac{1}{2}, \frac{8}{15}, \frac{15}{28}, \frac{5}{9}, \frac{4}{7}, \frac{3}{5}, \frac{2}{3}, 1\right\}$

Only five ratios, namely $\frac{5}{9}$, $\frac{4}{7}$, $\frac{3}{5}$, $\frac{2}{3}$, and 1, are in the range greater than or equal to the threshold. All other probability ratios in that range have no solutions. The solutions in the table are the only ones for these five ratios. Note that all of them correspond to solutions of the form (x, x) and, except for $\frac{p}{q}$ == 1, also (x – 1, x) and (suppressed in the table) (x, x – 1). So there are 3 solutions each for x == 2, 3, 4, 5 plus 1 solution for x == 1, a total of 13 solutions in this range of p/q values.

This range represents nearly half (more exactly, about 45%) of all admissible p/q ratios.

In[179]:=
```
(1 - minzforxle5) // N
```

Out[179]= 0.444444

8.3.1 Refining the bound

The ratio 5/9 is a conservative bound on the ratios that can have solutions with x, $y \leq 5$. One can actually use the smaller ratio for the far endpoint (6, 6) as a bound, as a strict inequality. In general, suppose we wish to define a bound on p/q such that solutions will be x, $y \leq x_{max}$ for some chosen x_{max}. The probability ratio giving an endpoint at (x, x) (not necessarily integer) is

In[180]:=
```
Solve[z/(2 z - 1) == x, z]
```

Out[180]= $\left\{\left\{z \to \frac{x}{-1 + 2x}\right\}\right\}$

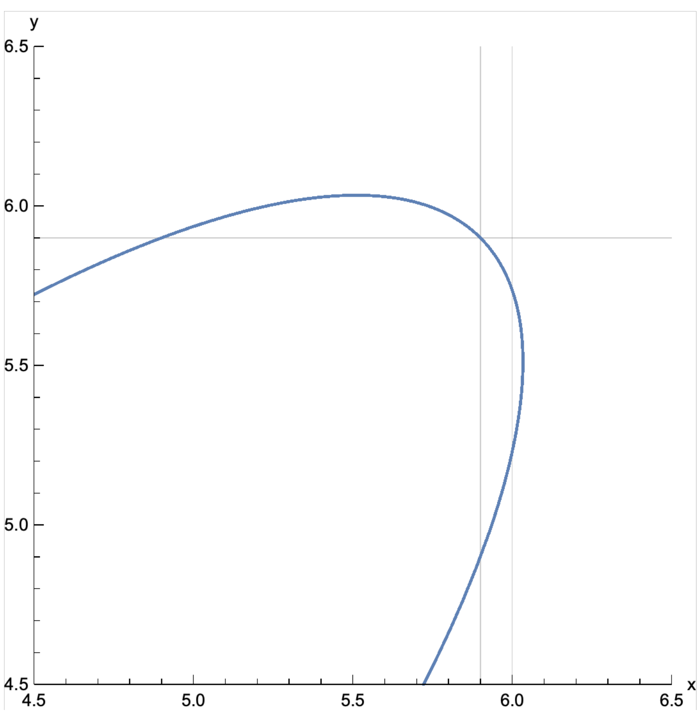
It may seem that there could be a solution with a larger x on this ellipse, in the short arc between (x, x) and $(x, x - 1)$. Here is a plot showing the case where the vertex is at (5.9, 5.9).

In[181]:=
```
probdifferent[{59 / 10, 59 / 10}]
```

Out[181]= $\frac{59}{108}$

In[182]:=
```
yvaluesforplot = Solve[xyeqn[{x, y}] /. {p → 59, q → 108}, y];
Plot[y /. yvaluesforplot, {x, 4.5, 6.5},
 PlotRange → {{4.5, 6.5}, {4.5, 6.5}}, AspectRatio → 1,
 GridLines → {{5.9, 6}, {5.9}},
 AxesLabel → {"x", "y"}]
```

Out[183]=



One can see that x values larger than the vertex value of 5.9 (left gridline) appear in the arc segment at the right. These include the integer value 6 (right gridline).

However, it turns out that no matter what the vertex coordinate is, this arc can never hit integer values for both x and y. A formal proof is in Section 12.1. The reason is basically that it is too close to the line $x == y$ and hence there is no room for $x - y$ to be other than 0 or 1, which correspond to (x, x) and $(x, x - 1)$ themselves, contradicting the hypothesis that the solution is strictly between these points. Hence we can use $x_{max} + 1$ as the vertex value in calculating the lower bound on z for $x \le x_{max}$.

For the case $x, y \le 5$ therefore we take probability ratio for the endpoint (6, 6).

In[184]:=
```
x/(2 x - 1) /. x → 6
```

Out[184]= $\frac{6}{11}$

Hence we can say that any ratio $z > 6/11$ cannot have solutions with $x, y > 5$.

In[185]:=
```
6 / 11 // N
```

Out[185]=
```
0.545455
```

Revisiting the exhaustive search on $x, y \le 5$, we can see where this new bound fits in the sequence of ratios that have solutions.

In[186]:=
```
DeleteDuplicates[
 Sort[Flatten[Join[Table[probdifferent[{x, y}], {x, 1, 5}, {y, 1, x}], {6 / 11}]]]]
```

Out[186]= $\left\{\frac{1}{3}, \frac{2}{5}, \frac{10}{21}, \frac{1}{2}, \frac{8}{15}, \frac{15}{28}, \frac{6}{11}, \frac{5}{9}, \frac{4}{7}, \frac{3}{5}, \frac{2}{3}, 1\right\}$

It is adjacent to the old bound of 5/9. The list of ratios having solutions in this range does not change, but additional ratios between the new bound and the old bound are now excluded.

8.3.2 Enumeration of elliptical solutions with $x, y < 999$

The reverse search described in Section 5.3 amounts to doing an exhaustive enumeration for values of $x, y < 1000$. Thus for ratios above a certain threshold it enumerates all solutions that exist. Here is the probability value corresponding to an endpoint of 1000:

In[187]:=
```
minzforapex1000 = probdifferent[{1000, 1000}]
```

Out[187]= $\frac{1000}{1999}$

Here it is numerically.

In[188]:=
```
minzforapex1000 // N
```

Out[188]=
```
0.50025
```

The probability ratio for an endpoint of 999 is:

In[189]:=
```
minzforapex999 = probdifferent[{999, 999}]
```
Out[189]= $\frac{999}{1997}$

This endpoint is within the range of the search, and the ratio 999/1997 appears as ratio number 477 704 of the output. I searched above it for the smallest ratio greater than the ratio 1000/1999 for vertex 1000. This appears as ratio number 477 683. The ratio is 115 993/231 870. It has one distinct solution, (579, 601). The ratio for vertex 1000 itself does not appear in the reverse search results since it has no solutions other than the vertex solutions, which are beyond the range of the search.

The last line of output, for probability ratio 1/1, is number 494 396. Therefore the number of ratios greater than 1000/1999 having solutions is:

In[190]:=
```
494 396 - 477 683 + 1
```
Out[190]= 16 714

8.4 Placing bounds on *p* and *q*

This section is heavy going, and the result is not used in later sections, so it can be skipped unless you are interested in the details.

Not just any ratio in the elliptical regime is possible. There are bounds on *p* and *q* that get smaller as the ratio approaches 1. This is because they arise from the probability formula, so the bounds on *x* and *y* place bounds on *p* and *q*.

The result is

$$p \le \left(\frac{z}{2z-1}\right)^2$$

$$q \le \frac{z}{(2z-1)^2}$$

where $z == p/q$. The derivation of these formulas follows.

Start with the probability formula, Equation (1):

$$\frac{p}{q} == \frac{2xy}{(x+y)(x+y-1)}$$

If there is cancellation of common factors, then numerator and denominator may be smaller than these, but they cannot be larger. The factor 2 always divides the denominator. Thus we have

$$p \le x\,y$$

$$q \le (x+y)(x+y-1)/2$$

Now, a bound can be calculated using $x \le x_{max}$ and $y \le x_{max}$ where x_{max} is the rightmost *x* value. However, this bound is loose since when $x == x_{max}$, then $y < x$. In fact, the products on the RHS of the inequalities above are maximized at the vertex. Proof: start with *y* as a function of *x*, calculated in the previous section, repeated here:

In[191]:=
```
yvsx = Solve[probdifferent[{x, y}] == z, y]
```

Out[191]= $\left\{\left\{y \to \frac{2x + z - 2xz - \sqrt{4x^2 + 4xz - 8x^2 z + z^2}}{2z}\right\}, \left\{y \to \frac{2x + z - 2xz + \sqrt{4x^2 + 4xz - 8x^2 z + z^2}}{2z}\right\}\right\}$

Extract the positive branch to get $y \geq x$.

In[192]:=
```
Part[yvsx, 2, 1]
```

Out[192]= $y \to \frac{2x + z - 2xz + \sqrt{4x^2 + 4xz - 8x^2 z + z^2}}{2z}$

Use this to get the product $x\,y$ in terms of x alone.

In[193]:=
```
xyvsx = Simplify[x y /. Part[yvsx, 2, 1]]
```

Out[193]= $\frac{x\left(-2x(-1+z) + z + \sqrt{x^2(4-8z) + 4xz + z^2}\right)}{2z}$

Find the extrema of this expression by solving for where the derivative is 0. This is where Mathematica shines.

In[194]:=
```
dxydxzeros = Solve[D[xyvsx, x] == 0, x]
```

Out[194]= $\left\{\{x \to 0\}, \left\{x \to \frac{z}{-1+2z}\right\}, \left\{x \to \frac{1}{4}\left(1 - \sqrt{1+2z}\right)\right\}, \left\{x \to \frac{1}{4}\left(1 + \sqrt{1+2z}\right)\right\}\right\}$

The first solution is the vertex at the origin, the second is the far vertex. The third and fourth are near the origin. Let's explore graphically. Here they are numerically for a $p/(2p-1)$ instance where $x_e == 13$.

In[195]:=
```
% /. z → 13 / 25 // N
```

Out[195]= {{x → 0.}, {x → 13.}, {x → -0.107071}, {x → 0.607071}}

These are not at the extrema of x. Find the extrema by setting the term inside the radical for y vs. x to 0, so the positive and negative branches coincide.

In[196]:=
```
Part[yvsx, 1, 1, 2, 3, 4, 2, 1]
```

Out[196]= $4x^2 + 4xz - 8x^2 z + z^2$

In[197]:=
```
xmaxvsz = Solve[Part[yvsx, 1, 1, 2, 3, 4, 2, 1] == 0, x]
```

Out[197]= $\left\{\left\{x \to \frac{z - \sqrt{2}\,z^{3/2}}{2(-1+2z)}\right\}, \left\{x \to \frac{z + \sqrt{2}\,z^{3/2}}{2(-1+2z)}\right\}\right\}$

Plug in the example:

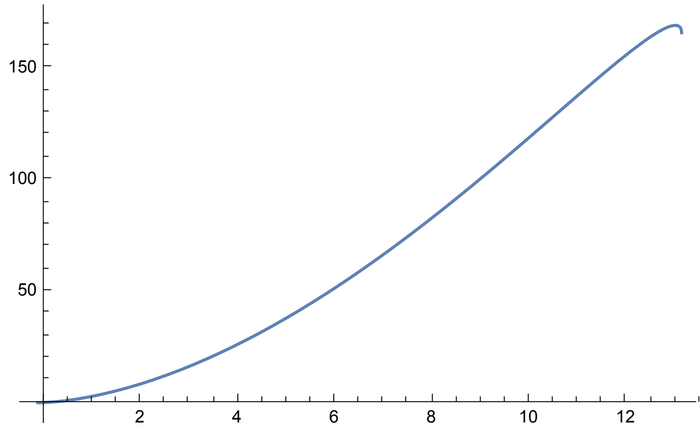
In[198]:=
```
xmaxvsz /. z → 13 / 25 // N
```

Out[198]= {{x → -0.128725}, {x → 13.1287}}

Plot $x\,y$ for this instance from end to end.

In[199]:= Plot[xyvsx /. z → $\frac{13}{25}$, {x, -0.2, 13.2}]

Out[199]=



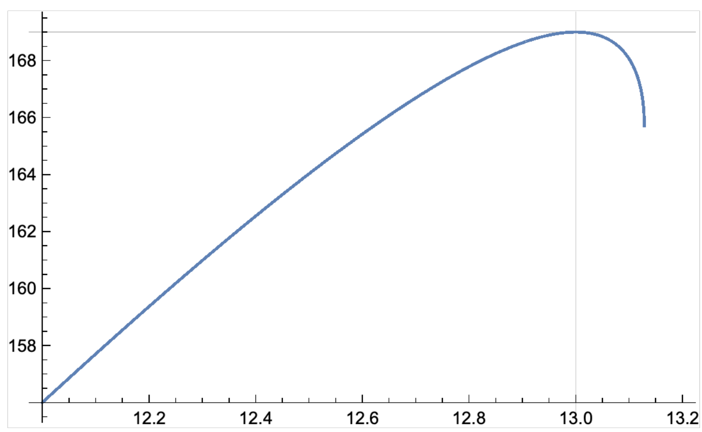
Look more closely at *x y* near the vertex:

In[200]:= xyvsx /. {x → 13, z → 13 / 25}

Out[200]= 169

In[201]:= Plot[xyvsx /. z → $\frac{13}{25}$, {x, 12, 13.2}, GridLines → {{13}, {169}}]

Out[201]=

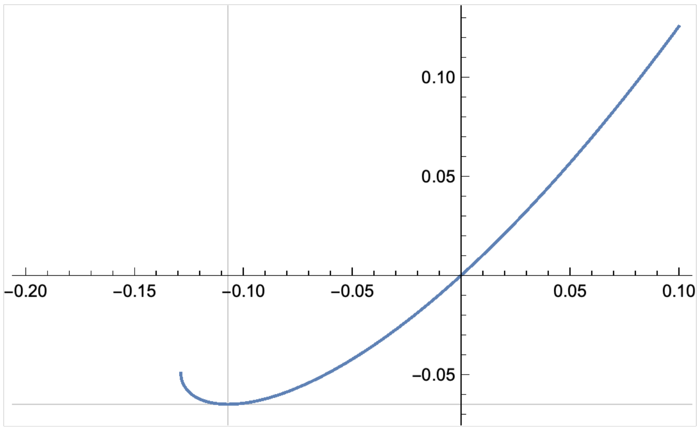


The vertex solution is indeed a maximum of *x y*.

Examine the first and third solutions.

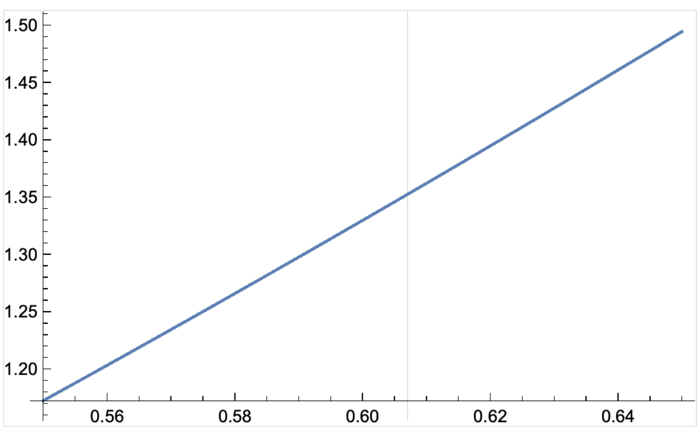
In[202]:=
```
Plot[xyvsx /. z → 13/25, {x, -0.2, 0.1},
 GridLines →
  {{1/4 (1 - Sqrt[1 + 2 z]) /. z → 13 / 25}, {(xyvsx /. x -> 1/4 (1 - Sqrt[1 + 2 z])) /. z → 13 / 25}}]
```

Out[202]=



In[203]:=
```
Plot[xyvsx /. z → 13/25, {x, 0.55, 0.65},
 GridLines → {{1/4 (1 + Sqrt[1 + 2 z]) /. z → 13 / 25}, None}]
```

Out[203]=



At the origin $x\, y$ does not have slope 0, nor at the small positive solution near 0.61. Evidently these are artifacts introduced when solving. The true min is at the small negative value, and the true max is at the far vertex.

Verify that the slope is 0 at the second and third solutions by plugging the x values into the derivative.

In[204]:= `Simplify[D[xyvsx, x] /. dxydxzeros, Assumptions → 1/2 < z ≤ 1]`

Out[204]=
$$\left\{1,\ 0,\ \frac{1}{4\,z\,\sqrt{1+z-\sqrt{1+2\,z}}}\left(-2\,\sqrt{2}\,z^2+2\left(\sqrt{2}-\sqrt{2+4\,z}+\sqrt{1+z-\sqrt{1+2\,z}}-\sqrt{(1+2\,z)\left(1+z-\sqrt{1+2\,z}\right)}\right)+z\left(\sqrt{2}+\sqrt{2+4\,z}+2\,\sqrt{(1+2\,z)\left(1+z-\sqrt{1+2\,z}\right)}\right)\right),\ \frac{1}{4\,z\,\sqrt{1+z+\sqrt{1+2\,z}}}\left(-2\,\sqrt{2}\,z^2+z\left(\sqrt{2}-\sqrt{2+4\,z}-2\,\sqrt{(1+2\,z)\left(1+z+\sqrt{1+2\,z}\right)}\right)+2\left(\sqrt{2}+\sqrt{2+4\,z}+\sqrt{1+z+\sqrt{1+2\,z}}+\sqrt{(1+2\,z)\left(1+z+\sqrt{1+2\,z}\right)}\right)\right)\right\}$$

Interestingly, Mathematica seems unable to simplify the third one to zero, even though it evidently is.

In[205]:= `% /. z → 13 / 25 // N`

Out[205]= `{1., 0., 0., 3.24149}`

Even using `FullSimplify` does not succeed.

In[206]:= `FullSimplify[D[xyvsx, x] /. dxydxzeros, Assumptions → 1/2 < z ≤ 1]`

Out[206]=
$$\left\{1,\ 0,\ \frac{1}{4\,z\,\sqrt{1+z-\sqrt{1+2\,z}}}\left(2\left(\sqrt{2}-\sqrt{2+4\,z}+\sqrt{1+z-\sqrt{1+2\,z}}-\sqrt{(1+2\,z)\left(1+z-\sqrt{1+2\,z}\right)}\right)+z\left(\sqrt{2}-2\,\sqrt{2}\,z+\sqrt{2+4\,z}+2\,\sqrt{(1+2\,z)\left(1+z-\sqrt{1+2\,z}\right)}\right)\right),\ \frac{1}{4\,z\,\sqrt{1+z+\sqrt{1+2\,z}}}\left(z\left(\sqrt{2}-2\,\sqrt{2}\,z-\sqrt{2+4\,z}-2\,\sqrt{(1+2\,z)\left(1+z+\sqrt{1+2\,z}\right)}\right)+2\left(\sqrt{2}+\sqrt{2+4\,z}+\sqrt{1+z+\sqrt{1+2\,z}}+\sqrt{(1+2\,z)\left(1+z+\sqrt{1+2\,z}\right)}\right)\right)\right\}$$

However, this method succeeds:

In[207]:= `Table[Simplify[(D[xyvsx, x] /. dxydxzeros)[[i]] == 0, Assumptions → 1/2 < z ≤ 1], {i, 4}]`

Out[207]= `{False, True, True, False}`

Anyway, this tells us that we can use $x\,y \le x_e^2$ for the bound on p. Now turn to the bound on q.

In[208]:= `xpyxpym1vsx = Simplify[(x + y) (x + y - 1) /. Part[yvsx, 2, 1]]`

Out[208]= $$\frac{x\left(-2\,x\,(-1+z)+z+\sqrt{x^2\,(4-8\,z)+4\,x\,z+z^2}\right)}{z^2}$$

In[209]:= `Solve[D[xpyxpym1vsx, x] == 0, x]`

Out[209]= $$\left\{\{x\to 0\},\left\{x\to\frac{z}{-1+2\,z}\right\},\left\{x\to\frac{1}{4}\left(1-\sqrt{1+2\,z}\right)\right\},\left\{x\to\frac{1}{4}\left(1+\sqrt{1+2\,z}\right)\right\}\right\}$$

The same four points.

In[210]:= `Table[Simplify[(D[xpyxpym1vsx, x] /. dxydxzeros)[[i]] == 0, Assumptions -> 1/2 < z <= 1], {i, 4}]`

Out[210]= `{False, True, True, False}`

So once again we can use the vertex values as the bound on q. Thus we can set

$$x == y == x_e == \frac{p}{2\,p-q} == \frac{z}{2\,z-1}$$

in the expressions on the RHS of the inequalities for p and q. Hence

$$p \le \left(\frac{z}{2\,z-1}\right)^2$$

$$q \le \frac{1}{2}\left(\frac{2\,z}{2\,z-1}\right)\left(\frac{2\,z}{2\,z-1}-1\right) == \left(\frac{z}{2\,z-1}\right)\left(\frac{1}{2\,z-1}\right) == \frac{z}{(2\,z-1)^2}$$

In[211]:= `pboundvsz[z_] := z^2/(2 z - 1)^2`

In[212]:= `qboundvsz[z_] := z/(2 z - 1)^2`

Let’s look at some values.

In[213]:=
```
TableForm[Table[{z, Floor[pboundvsz[z]], Floor[qboundvsz[z]]} /. z → i / 25,
  {i, 13, 25}], TableHeadings → {None, {"z", "pmax", "qmax"}}]
```

Out[213]//TableForm=

z	p_{max}	q_{max}
$\frac{13}{25}$	169	325
$\frac{14}{25}$	21	38
$\frac{3}{5}$	9	15
$\frac{16}{25}$	5	8
$\frac{17}{25}$	3	5
$\frac{18}{25}$	2	3
$\frac{19}{25}$	2	2
$\frac{4}{5}$	1	2
$\frac{21}{25}$	1	1
$\frac{22}{25}$	1	1
$\frac{23}{25}$	1	1
$\frac{24}{25}$	1	1
1	1	1

The reverse search found a ratio with largish $p == 99$, $q == 190$ compared to its neighbors, with solution $(x, y) == (9, 11)$.

In[214]:=
```
Floor[{pboundvsz[z], qboundvsz[z]} /. z → 99 / 190]
```

Out[214]= {153, 293}

It is comfortably within the bounds. Here is one with larger p, q, for $(x, y) == (41, 43)$.

In[215]:=
```
Floor[{pboundvsz[z], qboundvsz[z]} /. z → 1763 / 3486]
```

Out[215]= {1942, 3841}

This too is within the bounds.

We can rule out the following ratio, randomly chosen to be far from 1/2 so the bounds are small:

In[216]:=
```
Floor[{pboundvsz[z], qboundvsz[z]} /. z → 98 058 / 176 501]
```

Out[216]= {24, 44}

8.5 Vertex solutions

In Section 5.2.1 we showed that for ratios p/q where $q == 2p - 1$, which are elliptical, there are always three admissible solutions at the far vertex of the ellipse, at (p, p), $(p - 1, p)$, and $(p, p - 1)$.

Here is a table of probability ratios giving elliptical vertex solutions for $x \leq 10$.

In[217]:=
```
TableForm[Table[{x, probdifferent[{x, x}]}, {x, 10}],
 TableHeadings → {None, {x, p / q}}]
```

Out[217]//TableForm=

x	$\frac{p}{q}$
1	1
2	$\frac{2}{3}$
3	$\frac{3}{5}$
4	$\frac{4}{7}$
5	$\frac{5}{9}$
6	$\frac{6}{11}$
7	$\frac{7}{13}$
8	$\frac{8}{15}$
9	$\frac{9}{17}$
10	$\frac{10}{19}$

The existence of these solutions does not give any information about the possible existence of solutions at other points than around the vertex. For instance, 8/15 has non-vertex solutions.

In[218]:=
```
solveEllipticalBySearch[8 / 15]
```

Out[218]//TableForm=

x	y
2	4
4	6
7	8
8	8

8.5.1 Symmetry when vertex solutions exist

Consideration of the symmetry of the ellipse leads to another conclusion. If the far vertex coordinate $x_{max} == y_{max}$is an integer, then if there is a solution somewhere other than at the endpoint, there will be a companion solution at a point that is symmetrical about the center of the ellipse. I.e., if (x, y) is a solution, then

$$(x_{max} - x,\ y_{max} - y)$$

is diametrically opposite and is a solution. The example 8/15 shows this: the vertex is at (8, 8), and there are symmetrical solutions (2, 4) and (6, 4).

For other probability ratios that do not give integer vertex solutions, there may still be solutions in other parts of the ellipse, but they will not occur in symmetric pairs like these. An example is 15/28, for which there is one distinct solution (3, 5). The vertex is at

In[219]:= $\frac{p}{2p-q}$ /. {p → 15, q → 28}

Out[219]= $\frac{15}{2}$

8.5.2 Midsection solutions

While perusing the results of the reverse search, I noticed that there were a number of elliptical cases having vertex solutions in which another solution existed with x or $y == p/2$. Investigating, I found these corresponded to $u^2 == v^2$ in Equation (8). We can determine the conditions for this to occur:

In[220]:= `uveqn[{u, u}]`

Out[220]= $u^2 - q\,(-2\,p + q)\,u^2 == p^2$

or

$$u^2 == \frac{p^2}{1 - q\,(q - 2\,p)}$$

In order for this solution to be admissible, the denominator must be a square. (It is always negative in the hyperbolic case, so this situation can only occur for the elliptical case.) If the probability ratio has vertex solutions, i.e. $q == 2\,p - 1$, then

$$u^2 == \frac{p^2}{1 + (2\,p - 1)} == \frac{p}{2}$$

which requires $p/2$ to be square. Set $p == 2\,k^2$, giving

$$u^2 == v^2 == k^2$$

i.e. $u == \pm k$, $v == \pm k$ (for any combination of signs). As usual we limit ourselves to $v \geq 0$ to yield $x \leq y$ for distinctness. Then converting to x, y we obtain for $u == k$:

In[221]:= Simplify[xyfromuv[{k, k}] /. {p → 2 k^2, q → 4 k^2 - 1}]

Out[221]= $\{(-1 + k)\,k,\; k^2\}$

and for $u == -k$:

In[222]:= Simplify[xyfromuv[{-k, k}] /. {p → 2 k^2, q → 4 k^2 - 1}]

Out[222]= $\{k^2,\; k\,(1 + k)\}$

We can unify these in a single elegant formula if we don't require $x \leq y$:

$$(x,\, y) == \left(k^2,\; k^2 \pm k\right)$$

Since $k^2 == p/2$, and the endpoint of the ellipse is $(p,\, p)$, these solutions are around the midsection of the ellipse. The example 8/15 seen earlier is of this class, with $k == 2$. Another example, for $k == 3$, is

In[223]:=
```
solveEllipticalBySearch[18 / 35]
```

Out[223]//TableForm=

x	y
6	9
9	12
17	18
18	18

8.5.3 Near-triangular solutions

The reader who would like to move along can skip this section. I can't resist including this discovery since it is pretty cool, but fairly far into the weeds.

The first 7 solutions found in the reverse search following $p/q == 1/2$, i.e. the smallest elliptical ratios having solutions within range of the search, show an intriguing pattern. The solutions obey the rule that the smaller value of the solution for one ratio reappears in the solution for the next ratio as the larger value. This is not the recycling recurrence, since they are for different p/q. These solutions turn out to be related to the triangular numbers that characterize the $p/q == 1/2$ solutions.

I was able to follow the sequence further down the list, with other cases interspersed. Here are the first several ratios obeying that pattern. I reverse the order of the ratios to be largest to smallest so that the solutions will be in order of increasing size.

In[224]:=
$$\text{topellipticals} = \text{Reverse}\left[\left\{\frac{938477}{1876953}, \frac{856088}{1712175}, \frac{779248}{1558495}, \frac{707702}{1415403}, \frac{641201}{1282401}, \frac{579502}{1159003}, \frac{522368}{1044735}, \frac{469568}{939135}, \frac{420877}{841753}, \frac{376076}{752151}, \frac{334952}{669903}, \frac{297298}{594595}, \frac{262913}{525825}\right\}\right]$$

Out[224]=
$$\left\{\frac{262913}{525825}, \frac{297298}{594595}, \frac{334952}{669903}, \frac{376076}{752151}, \frac{420877}{841753}, \frac{469568}{939135}, \frac{522368}{1044735}, \frac{579502}{1159003}, \frac{641201}{1282401}, \frac{707702}{1415403}, \frac{779248}{1558495}, \frac{856088}{1712175}, \frac{938477}{1876953}\right\}$$

These are all of the form $p/(2p-1)$, hence have vertex solutions at (p, p) and $(p-1, p)$.

In[225]:=
```
DeleteDuplicates[
 2 p - q /. {p → Numerator /@ topellipticals, q → Denominator /@ topellipticals}]
```

Out[225]= {1}

Here are the complete solutions. The reverse search only found those in the range of 999.

In[226]:=
```
solntopellipticals = Table[solveEllipticalBySearch[topellipticals[[i]], False],
   {i, Length[topellipticals]}];
```

In[227]:=
```
TableForm[solntopellipticals]
```

Out[227]//TableForm=

497	262 384	262 912	262 913		
529	262 416	262 913	262 913		
529	296 736	297 297	297 298		
562	296 769	297 298	297 298		
562	334 356	334 951	334 952		
596	334 390	334 952	334 952		
596	375 445	376 075	376 076		
631	375 480	376 076	376 076		
631	420 210	420 876	420 877		
667	420 246	420 877	420 877		
667	468 864	469 567	469 568		
704	468 901	469 568	469 568		
704	139 072	382 844	521 626	522 367	522 368
742	139 524	383 296	521 664	522 368	522 368
742	578 721	579 501	579 502		
781	578 760	579 502	579 502		
781	640 380	641 200	641 201		
821	640 420	641 201	641 201		
821	706 840	707 701	707 702		
862	706 881	707 702	707 702		
862	778 344	779 247	779 248		
904	778 386	779 248	779 248		
904	855 141	856 087	856 088		
947	855 184	856 088	856 088		
947	937 486	938 476	938 477		
991	937 530	938 477	938 477		

Format of table: each solution is a stack with x above y, solutions for one ratio go right to left.

Now get a list of the small solutions.

In[228]:=
```
smallsolntopellipticals =
 Table[solntopellipticals[[i]][[1]], {i, Length[solntopellipticals]}]
```

Out[228]=
```
{{497, 529}, {529, 562}, {562, 596}, {596, 631}, {631, 667}, {667, 704}, {704, 742},
 {742, 781}, {781, 821}, {821, 862}, {862, 904}, {904, 947}, {947, 991}}
```

The big non-vertex solutions are symmetrically placed with respect to the center of the ellipse, i.e. $p - x$ and $p - y$ where (x, y) is a small solution.

In[229]:=
```
bigsolntopellipticals =
 Table[solntopellipticals[[i]][[Length[solntopellipticals[[i]]] - 2]],
  {i, Length[solntopellipticals]}]
```

Out[229]=
```
{{262 384, 262 416}, {296 736, 296 769}, {334 356, 334 390},
 {375 445, 375 480}, {420 210, 420 246}, {468 864, 468 901},
 {521 626, 521 664}, {578 721, 578 760}, {640 380, 640 420},
 {706 840, 706 881}, {778 344, 778 386}, {855 141, 855 184}, {937 486, 937 530}}
```

In[230]:=
```
Numerator /@ topellipticals - bigsolntopellipticals
```
Out[230]=
```
{{529, 497}, {562, 529}, {596, 562}, {631, 596}, {667, 631}, {704, 667}, {742, 704},
 {781, 742}, {821, 781}, {862, 821}, {904, 862}, {947, 904}, {991, 947}}
```

These are the same as the small solutions with (x, y) swapped.

The small solutions are close to the square roots of the p values.

In[231]:=
```
Round[(Sqrt /@ Numerator /@ topellipticals)]
```
Out[231]=
```
{513, 545, 579, 613, 649, 685, 723, 761, 801, 841, 883, 925, 969}
```

Look at the v values for these small solutions.

In[232]:=
```
smallsolnvvalues =
 Table[smallsolntopellipticals[[i]][[2]] - smallsolntopellipticals[[i]][[1]],
  {i, Length[smallsolntopellipticals]}]
```
Out[232]=
```
{32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44}
```

Holy cow! These are successive integers. Successive triangular numbers differ by successive integers. The x, y are not exactly successive triangular numbers. They are offset by 1 from triangular. For instance the 43rd triangular number is

In[233]:=
$$\frac{44 \times 43}{2}$$
Out[233]=
```
946
```

The solution value is 947. Verify that this offset holds for all the solutions. Generate the series of triangular numbers +1 to reproduce the solution set.

In[234]:=
$$\text{Table}\left[\left\{\frac{n\,(n-1)}{2}+1,\ \frac{n\,(n+1)}{2}+1\right\},\ \{n,\ 32,\ 44\}\right]$$
Out[234]=
```
{{497, 529}, {529, 562}, {562, 596}, {596, 631}, {631, 667}, {667, 704}, {704, 742},
 {742, 781}, {781, 821}, {821, 862}, {862, 904}, {904, 947}, {947, 991}}
```

Check that they all agree.

In[235]:=
```
smallsolntopellipticals == %
```
Out[235]=
```
True
```

The triangular numbers themselves appear in the $p/q = 1/2$ solutions, of course.

So it turns out that if x and y are 1 + successive triangular numbers, then their p/q ratio has both vertex solutions and the other solutions near the square root of the vertex and symmetrically opposite. Get expressions for p and q, knowing the factor 2 in the numerator of the probability formula always divides the denominator.

In[236]:=
```
pfortriangplus1 = Factor[Simplify[x y /. {x → 1 + v (v - 1) / 2, y → 1 + v (v + 1) / 2}]]
```
Out[236]=
$$\frac{1}{4}\left(2 - v + v^2\right)\left(2 + v + v^2\right)$$

Here v can be any positive integer.

In[237]:=
```
qfortriangplus1 =
  Simplify[(x + y) (x + y - 1) / 2 /. {x → 1 + v (v - 1) / 2, y → 1 + v (v + 1) / 2}]
```

Out[237]= $\frac{1}{2}\left(1+v^2\right)\left(2+v^2\right)$

Show that these formulas give $2\,p - q == 1$. This implies they are relatively prime.

In[238]:=
```
Simplify[2 pfortriangplus1 - qfortriangplus1]
```

Out[238]= 1

We can put this into a simpler form if we take advantage of the appearance of $2 + v^2$ in both expressions. Let $w == 2 + v^2$. Then

$$p == \frac{1}{4}(w-v)(w+v) == \frac{1}{4}\left(w^2-v^2\right) == \frac{1}{4}\left(w^2-w+2\right)$$

$$q == \frac{1}{2}w\left(1+v^2\right) == \frac{1}{2}w(w-1)$$

$$\frac{p}{q} == \frac{w^2-w+2}{2\,w\,(w-1)} == \frac{w\,(w-1)+2}{2\,w\,(w-1)} == \frac{1}{2}+\frac{1}{w\,(w-1)}$$

This is simpler, but w can only take on special values.

In[239]:=
```
Table[2 + v^2, {v, 1, 10}]
```

Out[239]= {3, 6, 11, 18, 27, 38, 51, 66, 83, 102}

Here are the first 10 ratios obeying this pattern.

In[240]:=
```
triangplus1ratios = Table[1/2 + 1/(w (w - 1)) /. w → 2 + v^2, {v, 1, 10}]
```

Out[240]= $\left\{\frac{2}{3}, \frac{8}{15}, \frac{28}{55}, \frac{77}{153}, \frac{176}{351}, \frac{352}{703}, \frac{638}{1275}, \frac{1073}{2145}, \frac{1702}{3403}, \frac{2576}{5151}\right\}$

In[241]:=
```
Table[{triangplus1ratios[[i]],
  solveEllipticalBySearch[triangplus1ratios[[i]]]}, {i, 10}]
```

Out[241]=

$$\left\{\left\{\frac{2}{3}, \begin{array}{cc} x & y \\ \hline 1 & 2 \\ 2 & 2 \end{array}\right\}, \left\{\frac{8}{15}, \begin{array}{cc} x & y \\ \hline 2 & 4 \\ 4 & 6 \\ 7 & 8 \\ 8 & 8 \end{array}\right\}, \left\{\frac{28}{55}, \begin{array}{cc} x & y \\ \hline 4 & 7 \\ 21 & 24 \\ 27 & 28 \\ 28 & 28 \end{array}\right\}, \left\{\frac{77}{153}, \begin{array}{cc} x & y \\ \hline 7 & 11 \\ 66 & 70 \\ 76 & 77 \\ 77 & 77 \end{array}\right\},\right.$$

$$\left\{\frac{176}{351}, \begin{array}{cc} x & y \\ \hline 11 & 16 \\ 160 & 165 \\ 175 & 176 \\ 176 & 176 \end{array}\right\}, \left\{\frac{352}{703}, \begin{array}{cc} x & y \\ \hline 16 & 22 \\ 330 & 336 \\ 351 & 352 \\ 352 & 352 \end{array}\right\}, \left\{\frac{638}{1275}, \begin{array}{cc} x & y \\ \hline 22 & 29 \\ 609 & 616 \\ 637 & 638 \\ 638 & 638 \end{array}\right\},$$

$$\left.\left\{\frac{1073}{2145}, \begin{array}{cc} x & y \\ \hline 29 & 37 \\ 1036 & 1044 \\ 1072 & 1073 \\ 1073 & 1073 \end{array}\right\}, \left\{\frac{1702}{3403}, \begin{array}{cc} x & y \\ \hline 37 & 46 \\ 1656 & 1665 \\ 1701 & 1702 \\ 1702 & 1702 \end{array}\right\}, \left\{\frac{2576}{5151}, \begin{array}{cc} x & y \\ \hline 46 & 56 \\ 2520 & 2530 \\ 2575 & 2576 \\ 2576 & 2576 \end{array}\right\}\right\}$$

The $v == 1$, $p/q == 2/3$ is a special case, where the vertex solution (1, 2) coincides with the triangular+1 solution: it is of the form $(0+1, 1+1)$. Mathematically 0 and 1 are triangular, of form $n(n+1)/2$ for $n == 0, 1$ respectively, so this is formally of the form 1 + triangular. The rest indeed have a non-vertex solution that is 1 + successive triangular numbers, its centro-symmetric counterpart, and the vertex solutions. They may also have other solutions, as the instance for $v == 38$ in the table of the originally discovered instances above shows.

Other families of solutions

The results of the reverse search showed other series of solutions with recurring values, similar to the triangular +1 family. These turned out to be also related to the triangular numbers, but offset by 2 or more. In these cases, the formulas for p and q designed to give solutions consisting of successive triangular $+k$ numbers have $2p - q == k$ and therefore are not guaranteed to be in lowest terms or to yield vertex solutions if $k > 1$. Without a vertex solution, the centro-symmetric counterpart of the small solution disappears, since it requires the vertex point to be integer. I did not pursue this line of inquiry further. Probably many other special cases are waiting to be discovered.

9 Hyperbolic case: $p/q < 1/2$

Solution of the hyperbolic case is considerably more complicated than for the elliptical or parabolic cases. There is a special category of cases with $p == 1$ or 2, treated in Section 6.3, that are easily solved. There is also a subcase that is relatively easy to deal with, when the probability equation factors. This occurs when D is square. It is treated in Section 10. When D is nonsquare, the problem is related to the Pell Equation, and solution requires some advanced mathematics. This case is treated in Section 11.

9.1 Sign of *u*, *v* and admissibility of solutions

Because u and v appear in Equation (8) squared, either sign of either variable will satisfy the equation. But for the hyperbolic case, negative u values will not yield admissible x, y. Recall from the derivation of Equation (8) in Section 4.2.5 that u is related to the total number of balls t by

$$t == \frac{u - p}{q - 2\,p}$$

For the hyperbolic case, $q > 2\,p$ so the denominator is positive. If $u < 0$, then, $t < 0$, which is inadmissible.

Negative v values give admissible solutions, but not distinct from the ones for positive v. Changing the sign of v corresponds to simply swapping x, y. We generally make solutions distinct by requiring $x \le y$. So to obtain only distinct solutions we can omit negative v.

Therefore in pursuing distinct admissible solutions x, y for the hyperbolic case, we need only make use of solutions of (8) with $u > 0$, $v > 0$.

9.2 Growth rate of solutions for small *p*/*q*

In Section 5.3.1 it was noted that the reverse search results tend to show few solutions for small p/q ratios, implying that the solutions in a series for a given p/q must grow in size rapidly so that they soon exceed the limit of the search. We can understand this qualitatively as follows. Small p/q ratios require a large inequality between x and y. Let $x == \beta\,y$ where β is small compared to 1. Then the probability of different colors is

In[242]:=
```
Simplify[probdifferent[{β y, y}]]
```

Out[242]=
$$\frac{2\,y\,\beta}{(1+\beta)\,(-1+y+y\,\beta)}$$

Rewrite this as

$$\frac{2\,\beta}{(1+\beta)\,(1+\beta-1/y)}$$

If y is large and β small, this will be approximately $2\,\beta$. It approaches $2\,\beta$ as β gets smaller. The next solution given by the recycling recurrence is

In[243]:=
```
Simplify[recycle[{β y, y}]]
```

Out[243]=
$$\left\{y,\ \frac{-1+y}{\beta}\right\}$$

Thus the next solution is approximately a factor $1/\beta$ larger. The smaller $\beta \simeq \frac{1}{2}\frac{p}{q}$ is, the faster the solutions grow. Of course, the recycling recurrence does not continue giving integer results. But solutions often come in recycling triplets. So this result indicates that rapid growth should be expected.

Ratio x/y for small p/q

Before leaving this section, let's use the result to see how the ratio $\beta == x/y$ behaves as a function of $z == p/q$ for small z. Put $\epsilon == 1/y$, which will be small and tend to zero for larger solutions.

In[1108]:=
$$\texttt{betavsz = Simplify}\left[\texttt{Solve}\left[z == \frac{2\beta}{(1+\beta)\,(1+\beta-\epsilon)}, \beta\right]\right]$$

Out[1108]=
$$\left\{\left\{\beta \to \frac{2+z\,(-2+\epsilon)-\sqrt{4+4\,z\,(-2+\epsilon)+z^2\,\epsilon^2}}{2\,z}\right\}, \left\{\beta \to \frac{2+z\,(-2+\epsilon)+\sqrt{4+4\,z\,(-2+\epsilon)+z^2\,\epsilon^2}}{2\,z}\right\}\right\}$$

Expand in Taylor series. Linear term suffices.

In[1111]:=
```
Series[betavsz[[1]][[1]][[2]], {z, 0, 1}]
```

Out[1111]=
$$\frac{1}{2}\,(1-\epsilon)\,z+O[z]^2$$

Thus $\beta \to z/2$ as $\epsilon \to 0$. This is reasonable, since for $z == 0$, $x == 0$ with y finite. Thus β must get smaller as z gets smaller. We conclude that y values will need to be large even for the smallest solutions for small z. For instance, for $z == 1/1000$, $\beta \simeq 1/2000$, requiring $y \gtrsim 2000$.

The second solution for β corresponds to $x > y$.

In[1112]:=
```
Series[betavsz[[2]][[1]][[2]], {z, 0, 1}]
```

Out[1112]=
$$\frac{2}{z}+(-2+\epsilon)+\frac{1}{2}\,(-1+\epsilon)\,z+O[z]^2$$

10 Hyperbolic case, $D > 0$ square

We now turn to the solution of the hyperbolic case, treating square D in this section, and nonsquare D in the following section.

If the discriminant $D == q(q-2\,p)$ is square, Equation (8) can be written in a factored form and solved directly. First we examine which ratios p/q yield D a square.

10.1 Ratios p/q giving D square

The main results of this section are the conditions for D to be square:

- If p is even, then q must be square, and also $q-2\,p$ must be square.
- If p is odd, then $2\,q$ must be square, and also $2\,(q-2\,p)$ must be square.

Alternatively, set $P/Q == p/2\,q$, reduced to lowest terms. Then the conditions are more simply expressed as

- Require Q and $Q-4\,P$ to both be square.

This question is most cleanly analyzed using half the probability of different colors,

$$\frac{P}{Q} == \frac{1}{2}\frac{p}{q}$$

If p is even, then $P == p/2$, $Q == q$ and $D == q(q-2p) == Q(Q-4P)$, whereas if p is odd, $P == p$, $Q == 2q$ and $D == \frac{Q}{2}\left(\frac{Q}{2}-2P\right) == \frac{1}{4}Q(Q-4P)$. So in either case we can analyze $Q(Q-4P)$ to determine if D is square.

In order for $Q(Q-4P)$ to be square, it is necessary that Q itself be square. Proof: suppose Q nonsquare, then it contains a prime factor k to an odd power. To make $Q(Q-4P)$ square, then k must also be a prime factor to an odd power of $Q-4P$. But this requires k to divide $4P$. It cannot divide P since $\gcd(Q, P) == 1$. So k must divide 4. The only possibility is $k == 2$. But 4 is an even power of 2. QED.

Then since Q is square, it is also necessary for $Q-4P$ to be square, say m^2, $Q == 4P + m^2$. Setting $Q == l^2$, then $l^2 - m^2 == 4P$. The difference of two squares is a multiple of 4 iff they are both even or both odd. We can list the first several p/q ratios that meet these requirements. In the table we convert P/Q back to p/q simply using a factor of 1/2 and let Mathematica reduce to lowest terms.

```
In[244]:= ratiosforsquareD = DeleteDuplicates[Sort[Flatten[Join[
        Table[(l^2 - m^2)/2 / l^2, {l, 5, 9, 2}, {m, 1, l - 2, 2}] (* odd l, m *),
    Table[(l^2 - m^2)/2 / l^2, {l, 4, 10, 2}, {m, 2, l - 2, 2}] (* even l, m *)]]]]
```

Out[244]= $\left\{\frac{9}{50}, \frac{16}{81}, \frac{7}{32}, \frac{12}{49}, \frac{5}{18}, \frac{8}{25}, \frac{28}{81}, \frac{3}{8}, \frac{20}{49}, \frac{21}{50}, \frac{4}{9}, \frac{15}{32}, \frac{12}{25}, \frac{24}{49}, \frac{40}{81}\right\}$

We needed `DeleteDuplicates` because in some cases reduction to lowest terms yields odd q in the even q table. The reverse search to 999 found solutions for 21/50, 12/25, 24/49, and 40/81 but not the others. Below, we will show that the other ratios in this list in fact have no admissible solutions.

The special case $p == 1$ or 2 is never square D

In Section 6.3 we saw that there is a simple solution when $p == 1$ or 2, which correspond to $P == 1$ in the half-probability convention. For $Q > 4$ these all are in the hyperbolic regime and there is an infinite number of solutions for each, beginning with $(1, Q-1)$ and continuing via the recycling recurrence. In Section 10.2 we will show that if D is square, then the number of solutions is finite. Therefore the case $P == 1$ with $Q > 4$ must never have square D. We can see that $Q(Q-4P) == Q(Q-4)$ is never square: that would require both Q and $Q-4$ to be square. This would imply existence of a Pythagorean triple involving 2, which does not exist.

10.2 Method of factorization

Recall Equation (11)

$$u^2 - D v^2 == f$$

where

$$D == q\ (q-2\,p),\quad f == p^2$$

If D is a square, the equation factors as

$$\left(u-\sqrt{D}\ v\right)\ \left(u+\sqrt{D}\ v\right) == f$$

where $\sqrt{D}$ is integer by assumption. The two parenthesized factors on the LHS must equate to divisors (positive or negative) of the RHS, f. So the procedure is to list all the divisors of f, call them d_1, d_2, ... d_k, and then equate the first factor to d_i and the second factor to f/d_i. This yields a pair of linear equations to be solved for u and v. Thus

In[245]:=
```
Simplify[Solve[{u - √D v == dᵢ, u + √D v == f / dᵢ}, {u, v}]]
```

Out[245]=
$$\left\{\left\{u\to\frac{f+d_i^2}{2\,d_i},\ v\to\frac{f-d_i^2}{2\,\sqrt{D}\ d_i}\right\}\right\}$$

Negative divisors correspond simply to the opposite sign of u or v, so only positive divisors need to be tested. Exchanging d for f/d yields the same u and $-v$, so we only need to test divisors $d\le\sqrt{f}$, i.e. half the complete list of divisors.

Keep in mind that all quantities in these equations are integer. The solutions u and v are not guaranteed to be integer but are always rational. Discard those that are fractional or that yield inadmissible solutions for x and y.

10.3 Existence and completeness of solutions

It is evident that this method yields all solutions that exist. The number of possible solutions is finite and it tests them all. There can be at most as many solutions as positive and negative divisors of p^2. It is possible for a specific instance that there will be no admissible solutions. The case $p/q == 4/9$ is an example with no admissible solutions.

10.4 Bound on magnitude of solutions

We can place a bound on the magnitude of the solutions for the hyperbolic square-D case.

$$|u| == \left|\frac{f+d^2}{2\,d}\right| == \frac{1}{2}\left(\frac{f}{|d|}+|d|\right)$$

The maximum occurs for the extrema $d == 1$ or $d == f$, where $u == \frac{1}{2}(f+1)$. Converting to t for positive u (so the bound is on admissible solutions):

In[246]:=
```
Simplify[tvfromuv[{1/2 (f + 1) /. f → p^2, v}][[1]]]
```

Out[246]=
$$-\frac{(-1+p)^2}{4p-2q}$$

So the bound is

$$t \le \frac{(p-1)^2}{2(q-2p)}$$

In[247]:=
```
tboundSquareD[z_] := (Numerator[z] - 1)^2 / (2 (Denominator[z] - 2 Numerator[z]))
```

All the ratios listed above giving square D have a bound within the reverse search limit of $t = 2 \times 999$.

In[248]:=
```
Floor[tboundSquareD /@ ratiosforsquareD]
```

Out[248]=
```
{1, 2, 1, 2, 1, 2, 14, 1, 20, 25, 4, 49, 60, 264, 760}
```

Therefore, if the reverse search did not find a solution for one of these ratios, the ratio has no admissible solution. For several of these ratios, the explanation for the lack of solutions is obvious: if $t \le 2$ there simply isn't room for nontrivial solutions.

10.5 Function to solve hyperbolic square *D* case

This function takes the probability ratio $z == p/q$ as argument. It includes an optional argument `tableform` that is `True` (default) if the results should be formatted as a table, if `False` then the output is a list. The trivial solutions and negative solutions are excluded.

Outline of function:

- Make sure $D > 1$ and square. ($D == 1$ is square but is the degenerate case $p == 0$ treated in Section 6.1.)
- Get list of positive divisors of f that are $\le f/2$.
- For each divisor, calculate u and v by the factoring method. These are rational, not necessarily integer.
- Sift for integer values of u and v.
- Convert the list of (u, v) values to (t, v) values. Sift again for integer values of t.
- Convert list of (t, v) values to (x, y) values. Sift again for integer values.

In[249]:=
```
solveHyperbolicDsquare[z_, tableform_: True] := Module[{p, q, D, f, dlist,
   ucandidates, vcandidates, hits, uvalues,
   vvalues, tvaluesall, tvalues, xvalues, yvalues, xyvalues},
  p = Numerator[z]; q = Denominator[z];
  D = q (q - 2 p); f = p^2;
  If[D > 1 && Element[Sqrt[D], Integers],
   dlist = Divisors[f];
   dlist = dlist[[1 ;; Ceiling[Length[dlist] / 2]]];
   ucandidates = Table[(f + d^2)/(2 d), {d, dlist}];
   vcandidates = Table[(f - d^2)/(2 Sqrt[D] d), {d, dlist}];
   (* first pass: pick out positions where u is integer *)
   hits = Position[ucandidates, _Integer];
   uvalues = Extract[ucandidates, hits];
   vvalues = Extract[vcandidates, hits];
   (* second pass: pick out positions where v is integer.  It is always ≥0 *)
   hits = Position[vvalues, _Integer];
   uvalues = Extract[uvalues, hits];
   vvalues = Extract[vvalues, hits];
   tvaluesall = (u - p) / (q - 2 p) /. u → uvalues;
   hits = Position[tvaluesall, t_ /; t > 1 && IntegerQ[t]];
   (* require admissible *)
   tvalues = Extract[tvaluesall, hits];
   vvalues = Extract[vvalues, hits];
   xvalues = (tvalues - vvalues) / 2;
   yvalues = (tvalues + vvalues) / 2;
   (* x,y may still be fractional if t, v not same parity *)
   yvalues = Extract[yvalues, Position[xvalues, _Integer, {1}]];
   xvalues = Extract[xvalues, Position[xvalues, _Integer, {1}]];
   xyvalues = Sort[Table[{xvalues[[i]], yvalues[[i]]}, {i, Length[xvalues]}]];
   If[tableform,
    TableForm[xyvalues, TableHeadings → {None, {x, y}}],
    xyvalues]
   ,
   (* else *) Print["D=", D, " not OK"]]
 ]
```

Exercise the discriminant test by giving an elliptical ratio and a hyperbolic ratio with nonsquare *D*.

In[250]:=
```
solveHyperbolicDsquare[2 / 3]
```

```
D=-3 not OK
```

In[251]:= `solveHyperbolicDsquare[1 / 3]`

```
D=3 not OK
```

Also test degenerate case $p == 0$. Although $D == 1$ is square, the method does not work for this case.

In[252]:= `solveHyperbolicDsquare[0]`

```
D=1 not OK
```

10.5.1 Examples: ratios with smallest p, q giving square D

We can apply the function to the list of ratios found earlier, giving square D calculated above. The table lists the ratio and the set of solutions x, y it has, if any.

In[253]:= `TableForm[Table[{z, solveHyperbolicDsquare[z, False]}, {z, ratiosforsquareD}]]`

Out[253]//TableForm=

$\frac{9}{50}$	
$\frac{16}{81}$	
$\frac{7}{32}$	
$\frac{12}{49}$	
$\frac{5}{18}$	
$\frac{8}{25}$	
$\frac{28}{81}$	
$\frac{3}{8}$	
$\frac{20}{49}$	
$\frac{21}{50}$	7 18
$\frac{4}{9}$	
$\frac{15}{32}$	
$\frac{12}{25}$	9 16
$\frac{24}{49}$	20 30
$\frac{40}{81}$	70 92

The results agree with what was found before in identifying the four ratios from this list that have solutions. Each has only one solution.

10.5.2 Example with larger q, p / q close to 1/2

This section can be skipped without loss of continuity.

We may suppose that as p / q approaches 1/2, and p and q get larger, the number of solutions may increase. Just for fun, let us try a value close to 1/2 with $q == 99^2$ and $q - 2\,p == 1$ so that D is square. We use the formula for finding p / q for square D developed in Section 10.1:

In[254]:=
```
zforD99sq = (l^2 - m^2)/(2 l^2) /. {l → 99, m → 1}
```

Out[254]= $\frac{4900}{9801}$

The reverse search did not list this ratio among its results. The bound on $t == x + y$ is

In[255]:=
```
Floor[(p - 1)^2/(2 (q - 2 p)) /. {p → Numerator[zforD99sq], q → Denominator[zforD99sq]}]
```

Out[255]= 12 000 100

So solutions can exist outside the range of the reverse search.

Discriminant is 99^2 by construction.

In[256]:=
```
D99sq = q (q - 2 p) /. {p → Numerator[zforD99sq], q → Denominator[zforD99sq]}
```

Out[256]= 9801

RHS f of Equation (11):

In[257]:=
```
fforD99sq = p^2 /. {p → Numerator[zforD99sq], q → Denominator[zforD99sq]}
```

Out[257]= 24 010 000

This has a lot of divisors.

In[258]:=
```
Length[Divisors[fforD99sq]]
```

Out[258]= 125

But most of the divisors do not yield admissible solutions. Here are the ones that exist.

In[259]:=
```
solveHyperbolicDsquare[zforD99sq]
```

Out[259]//TableForm=

x	y
9560	9800
9800	10 045
116 400	118 825

So there are a few admissible solutions, not a lot. This ratio did not appear in the reverse search results since all the solutions are above the search limit of 999.

11 Hyperbolic case, $D > 0$ nonsquare

When $D > 0$ nonsquare, the solution involves the Pell equation. There are various methods for solving the Pell equation. We will use continued fractions.

11.1 Continued fractions

For the reader who is unfamiliar with the theory of continued fractions (as I was when I started this project), here is some basic background. Hua (1982), Chapter 10, has a full discussion and proofs of the claims stated in this section. Continued fractions are fractions of the form

$$x == a_0 + \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{\cdots}}}$$

where the a_i are integers. Continued fractions are usually written in the compact form

$$x == [a_0, a_1, a_2, \ldots]$$

For positive values of x all of the a_i are positive. There are two sign conventions for negative x: the one that seems to be more common among mathematicians is that a_0 carries the sign of x and the rest of the coefficients are positive. Mathematica has a function to compute continued fractions; for negative arguments the coefficients are simply the negatives of those generated for the absolute value of x. For our problem we will always be working with positive values so this difference does not matter.

If x is a rational number, then the continued fraction will terminate at some point. For irrational numbers, it continues without end. An important fact is that if x is a quadratic irrational, i.e. the irrational root of a quadratic equation with rational coefficients, the series repeats.

The continued fraction representation of any real number x can be calculated by the following algorithm. Start the recurrence with:

$$\begin{aligned} \alpha_0 &== x \\ a_0 &== \lfloor \alpha_0 \rfloor \end{aligned} \tag{15}$$

where $\lfloor \alpha_0 \rfloor$ denotes the floor of α_0. Then for $n > 0$, and as long as $\alpha_{n-1} - a_{n-1} \neq 0$,

$$\begin{aligned} \alpha_n &== \frac{1}{\alpha_{n-1} - a_{n-1}} \\ a_n &== \lfloor \alpha_n \rfloor \end{aligned} \tag{16}$$

11.2 Convergents

The *convergents* of a continued fraction are the rational numbers obtained by terminating the continued fraction at any point. These are the best approximations to the exact value of the full continued fraction obtainable with a denominator of a given size. They can be generated by a recurrence alongside the continued fraction expansion. Let h_n and k_n be the numerator and denominator, respectively of the n^{th} convergent. The recurrence begins with

$$\begin{aligned} &h_0 = a_0, \quad h_1 = a_1\, a_0 + 1 \\ &k_0 = 1, \quad k_1 = a_1 \end{aligned} \tag{17}$$

Then for $n > 1$,

$$\begin{aligned} h_n &= a_n\, h_{n-1} + h_{n-2} \\ k_n &= a_n\, k_{n-1} + k_{n-2} \end{aligned} \tag{18}$$

(The recurrence can be started at $n = 0$ by defining $h_{-2} = 0$, $h_{-1} = 1$, $k_{-2} = 1$, $k_{-1} = 0$.)

Of course, Mathematica has functions to compute continued fractions and convergents. We will work the recurrence in an example to demonstrate how it goes, but afterward we will let Mathematica do it.

11.3 Conversion to Pell equation

Because the RHS of Equation (8) is p^2, a square, the equation can always be divided through by it to yield the Pell Equation:

$$\left(\frac{u}{p}\right)^2 - q\,(q-2\,p)\,\left(\frac{v}{p}\right)^2 == 1 \tag{19}$$

Let $r == u/p$ and $s == v/p$, and set $D == q(q-2\,p)$:

$$r^2 - D\,s^2 == 1 \tag{20}$$

Once Equation (20) is solved, set $u == p\,r$ and $v == p\,s$ to obtain a solution to Equation (8).

I will call this method of solution the “Pell method” to distinguish it from other solution methods to be developed later.

Put the equation into a formula for later use.

```
In[260]:= pelleqn[{r_, s_}, D_] := r^2 - D s^2 == 1
```

11.4 Solution of Pell equation

The Pell equation $r^2 - D\,s^2 == 1$ always has solutions in integers if $D > 0$ is nonsquare. The smallest solution with both $u > 0$ and $v > 0$, which I will call the **base** solution, can be found by setting r/s equal to the convergent of $\sqrt{D}$ just before the end of the first repeat cycle if the repeat length is even, or the end of the second cycle if the repeat length is odd. The fact that this works is plausible since the convergents r/s are the best rational approximations to $\sqrt{D}$, and it can be shown that the one just before the end of the repeat cycle is especially close. (The need for a second cycle in case the repeat length is odd is due to the errors in the approximation alternating in sign, so the last convergent in the first cycle is negative, and gives a solution to the Pell equation with the RHS replaced by −1. A second cycle is needed to get a positive value.)

Additional solutions can be found by running the convergents calculation for additional cycles, using the convergent at the end of each repeat cycle (or every other if the repeat length is odd). However, it is preferable to use a recurrence. If (r_0, s_0) is the base solution, other solutions are given by

$$r + s\,\sqrt{D} == \pm\left(r_0 + s_0\,\sqrt{D}\right)^n, \quad n \in \mathbb{Z} \tag{21}$$

When the expression on the RHS is expanded, after collecting terms it is always in the form of $r + s\,\sqrt{D}$. Equating rational and irrational parts on each side of the equation gives the new solution. In this way, Equation (21) produces all solutions to the Pell Equation. We can turn this formula into a recurrence. From (21) it follows that for positive solutions

$$r_{n+1} + s_{n+1}\,\sqrt{D} == \left(r_0 + s_0\,\sqrt{D}\right)^{n+1} == \left(r_n + s_n\,\sqrt{D}\right)\left(r_0 + s_0\,\sqrt{D}\right)$$

In[261]:=
```
Collect[(r_n + s_n √D) (r_0 + s_0 √D), {1, √D}]
```

Out[261]= $r_0\, r_n + D\, s_0\, s_n + \sqrt{D}\, (r_n\, s_0 + r_0\, s_n)$

Equating rational and irrational parts yields the recurrence formula

$$\begin{aligned} r_{n+1} &== r_0\, r_n + D\, s_0\, s_n \\ s_{n+1} &== s_0\, r_n + r_0\, s_n \end{aligned} \tag{22}$$

I will call (22) the "Pell recurrence" to distinguish it from other recurrences discussed in this document.

We can verify that this recurrence yields a new solution to the Pell equation:

In[262]:=
```
Simplify[Expand[r^2 - D s^2 /. {r → r_0 r_n + D s_0 s_n, s → s_0 r_n + r_0 s_n}]]
```

Out[262]= $\left(r_0^2 - D\, s_0^2\right) \left(r_n^2 - D\, s_n^2\right)$

This equals $1 \cdot 1 == 1$.

From this it follows that the Pell Equation always has an infinite number of solutions.

Since

$$\left(r + s\sqrt{D}\right)^{-1} == \frac{r - s\sqrt{D}}{\left(r + s\sqrt{D}\right)\left(r - s\sqrt{D}\right)} == \frac{r - s\sqrt{D}}{r^2 - D\, s^2} == r - s\sqrt{D},$$

the recurrence can be run in reverse simply by negating s_0.

Define a formula for running the Pell recurrence. In the formula, $\{r, s\}$ is a solution, and $\{h, k\}$ is the base solution, treated as parameters along with D.

In[263]:=
```
nextPell[{r_, s_}] := {h r + D k s, k r + h s}
```

Trivial solutions

The Pell Equation always has a pair of trivial solutions, $(r, s) == (\pm 1, 0)$. These correspond to solutions of Equation (8) $(u, v) == (\pm p, 0)$ noted earlier (in Section 5.2). The solution $(u, v) == (p, 0)$ maps to $(x, y) == (0, 0)$. It is worth noting that applying the Pell recurrence to this trivial solution yields the first nontrivial solution:

In[264]:=
```
nextPell[{1, 0}]
```

Out[264]= {h, k}

11.4.1 Example: $p/q == 7/18$

We are now ready to solve a hyperbolic nonsquare D case.

This example, $p/q == 7/18$ was picked more or less at random from the results of the reverse search, which found the following 5 distinct admissible solutions in the range $x, y < 1000$:

(2, 7), (7, 21), (21, 60), (95, 266), (266, 742)

The continued fraction for $\sqrt{D}$ for this example has a short, even repeat length so we can easily work out the convergents by hand (with arithmetical help from Mathematica).

In[265]:=
```
D7o18 = q (q - 2 p) /. {p → 7, q → 18}
```

Out[265]= 72

This is nonsquare. We need to solve the Pell equation $r^2 - 72\,s^2 == 1$. We will do it manually for this example, as a way of getting familiar with continued fractions and convergents. We will even refrain from using the square root function. We know that 72 is between $8^2 == 64$ and $9^2 == 81$. So the calculation can be done without evaluating $\sqrt{72}$.

Finding the convergent to solve Pell equation for $p / q == 7 / 18$

You can skip to the next section if you are not interested in the gory details of the continued fraction calculation.

As we run the recurrence, we will put the α_n, a_n, h_n, k_n into lists. The indexing of these will be one higher than the subscripts in the recurrence formulas given earlier.

In[266]:= cfalpha7o18 = $\{\sqrt{72}\}$

Out[266]= $\{6\sqrt{2}\}$

Set $a_0 == \lfloor \sqrt{72} \rfloor$ which we know is 8.

In[267]:=
```
cfa7o18 = {8}
```

Out[267]= {8}

Calculate α_1.

In[268]:=
```
AppendTo[cfalpha7o18, Simplify[1 / (cfalpha7o18[[1]] - cfa7o18[[1]])]]
```

Out[268]= $\left\{6\sqrt{2},\ 1 + \frac{3}{2\sqrt{2}}\right\}$

Now we need to find the floor of α_1. Continuing to insist on not evaluating the square root, we observe $2\sqrt{2} == \sqrt{8} > 2$. So

$$1 + \frac{3}{2\sqrt{2}} < 1 + \frac{3}{2} == \frac{5}{2} == 2.5$$

So the floor of this is 2.

In[269]:=
```
AppendTo[cfa7o18, 2]
```

Out[269]= {8, 2}

Next up: $n == 2$, list index 3.

In[270]:=
```
AppendTo[cfalpha7o18, Simplify[1 / (cfalpha7o18[[2]] - cfa7o18[[2]])]]
```

Out[270]= $\left\{6\sqrt{2},\ 1 + \frac{3}{2\sqrt{2}},\ 8 + 6\sqrt{2}\right\}$

Evaluate $a_2 == \lfloor \alpha_2 \rfloor$.

$8 + 6\sqrt{2} == 8 + \sqrt{72} < 8 + 8 == 16$

So the floor is 16.

In[271]:= `AppendTo[cfa7o18, 16]`

Out[271]= {8, 2, 16}

Now for $n == 3$.

In[272]:= `AppendTo[cfalpha7o18, Simplify[1 / (cfalpha7o18[[3]] - cfa7o18[[3]])]]`

Out[272]= $\left\{6\sqrt{2},\ 1 + \frac{3}{2\sqrt{2}},\ 8 + 6\sqrt{2},\ 1 + \frac{3}{2\sqrt{2}}\right\}$

We observe that the coefficients are repeating.

In[273]:= `AppendTo[cfa7o18, 2]`

Out[273]= {8, 2, 16, 2}

Stop here. Confirm this with Mathematica, which by default lists the non-repeating terms and then the repeating terms in a nested list.

In[274]:= `ContinuedFraction[`$\sqrt{72}$`]`

Out[274]= {8, {2, 16}}

It is not coincidental that the last term before the repeat begins is $2\,a_0$. This always happens with quadratic irrationals.

Now compute the convergents. Since the solution we are seeking is associated with the continued fraction term just before the last term of the first repeat cycle, i.e. a_1 in this case, we don't need to compute many. Here is a reminder of the initialization:

$h_0 = a_0,\ \ h_1 = a_1\, a_0 + 1$

$k_0 = 1,\ \ k_1 = a_1$

Remember the Mathematica list index is $n + 1$ since n starts at 0.

In[275]:= `cvh7o18 = {cfa7o18[[1]], cfa7o18[[2]] × cfa7o18[[1]] + 1}`

Out[275]= {8, 17}

In[276]:= `cvk7o18 = {1, cfa7o18[[2]]}`

Out[276]= {1, 2}

Now a reminder of the recurrence.

$h_n = a_n\, h_{n-1} + h_{n-2}$

$k_n = a_n\, k_{n-1} + k_{n-2}$

In[277]:= `AppendTo[cvh7o18, cfa7o18[[3]] × cvh7o18[[2]] + cvh7o18[[1]]]`

Out[277]= {8, 17, 280}

In[278]:=
```
AppendTo[cvk7o18, cfa7o18[[3]] × cvk7o18[[2]] + cvk7o18[[1]]]
```
Out[278]=
```
{1, 2, 33}
```

Check with Mathematica:

In[279]:=
```
Convergents[√72]
```
Out[279]= $\left\{8, \frac{17}{2}, 6\sqrt{2}\right\}$

Mathematica lists the convergents up to the end of the first repeat, then gives the quadratic irrational of the representation of the argument.

The convergent we want is the last before the repeat, 17/2. This can also be found by expanding the continued fraction:

In[280]:= $8 + \frac{1}{2}$

Out[280]= $\frac{17}{2}$

Proceeding with the solution of $p/q == 7/18$ by Pell equation method

To solve the Pell equation, we use $h_1 == 17$, $k_1 == 2$ since a_2 is the last continued fraction coefficient but one of the first cycle before it begins repeating. Let's make sure:

In[281]:=
```
h^2 - D k^2 == 1 /. {D → 72, h → 17, k → 2}
```
Out[281]=
```
True
```

The solution to Equation (11) is found by multiplying these values by p.

In[282]:=
```
{u7o18, v7o18} = p {17, 2} /. p → 7
```
Out[282]=
```
{119, 14}
```

Verify that it satisfies Equation (11).

In[283]:=
```
uveqn[{u7o18, v7o18}] /. {p → 7, q → 18}
```
Out[283]=
```
True
```

Convert (u, v) to (x, y).

In[284]:=
```
{x7o18, y7o18} = xyfromuv[{u7o18, v7o18}] /. {p → 7, q → 18}
```
Out[284]=
```
{7, 21}
```

This is the second smallest solution found by the reverse search. Larger solutions can be found via the Pell recurrence. First we find the solutions (r, s) of the Pell equation itself, then convert those to (u, v) and then (x, y).

In[285]:=
```
Pellsolns = RecurrenceTable[
   {{r[n + 1], s[n + 1]} == nextPell[{r[n], s[n]}] /. {h → 17, k → 2, D → 72},
    r[1] == 17, s[1] == 2}, {r, s}, {n, 3}]
```

Out[285]=
```
{{17, 2}, {577, 68}, {19 601, 2310}}
```

In[286]:=
```
uvsolns7o18 = p Pellsolns /. p → 7
```

Out[286]=
```
{{119, 14}, {4039, 476}, {137 207, 16 170}}
```

In[287]:=
```
xysolns7o18 = (xyfromuv /@ uvsolns7o18) /. {p → 7, q → 18}
```

Out[287]=
```
{{7, 21}, {266, 742}, {9065, 25 235}}
```

We see here two of the solutions found by the reverse search, plus one more that was beyond its range. Just to make sure all is working as expected, verify that these each give the desired probability ratio.

In[288]:=
```
DeleteDuplicates[probdifferent /@ xysolns7o18]
```

Out[288]= $\left\{\frac{7}{18}\right\}$

We can find the other solutions found by the reverse search by running the recycling recurrence backward and forward.

In[289]:=
```
recycleSolutions[xysolns7o18]
```

Out[289]//TableForm=

x	y
2	7
7	21
21	60
95	266
266	742
742	2067
3256	9065
9065	25 235
25 235	70 246

Indeed, all 5 solutions found by the reverse search are here, plus another 4 that are outside its range of x, $y < 1000$.

The reverse search did not find any solutions that were missed by this combination of solving the Pell equation and then finding additional solutions with the recycling recurrence and the Pell recurrence.

11.5 Functions to solve the Pell equation

Here is a function to find the first nontrivial solution of the Pell equation $r^2 - D s^2 == 1$ where $D > 0$ is nonsquare. Argument is D. The function is written to be independent of other local functions so it can be used in another notebook. A second function finds the first n solutions by using the Pell recurrence.

Outline of function:

- Check $D > 0$ and nonsquare
- Calculate first cycle of continued fraction using the Mathematica function `ContinuedFraction`. The result is of form $\{a_0, a_1, \ldots, a_k, \{b_1, b_2, \ldots, b_r\}\}$ where the b_i are the repeating cycle.
  - If the repeat length r is even, then this result, flattened to remove the structure, is passed to the Mathematica function `Convergents` to produce the list of convergents. The r-th convergent gives the Pell solution.
  - If the repeat length r is odd, then the repeating cycle is appended to the result, which is flattened and passed to `Convergents`. The $2\,r$-th convergent gives the Pell solution.
- Return numerator and denominator of the chosen convergent. Since this is the base case, they are named (h, k).

In[290]:=
```
solvePell[D_] := Module[{cf, repeat, replen, hoverk} ,
  If[D > 0 && √D ∉ Rationals,
   cf = ContinuedFraction[√D];
   repeat = cf[[-1]]; (* repeat cycle is a sub-list as last element *)
   replen = Length[repeat];
   If[Mod[replen, 2] == 0,
    hoverk = Convergents[Flatten[cf]][[ ;; replen]], (* even replen *)
    hoverk = Convergents[Join[Flatten[cf], repeat]][[ ;; 2 replen]] (* odd *)
   ];
   {Numerator[hoverk[[-1]]], Denominator[hoverk[[-1]]]}
   (* return last h,k as solution *)
   , (* else *) Print["D=", D, " not OK"]
  ]
 ]
```

Define another function that generates a list of the first n solutions using the Pell recurrence. First argument is D. Default for optional second argument n is 1. Optional third argument `tableform` is `True` (default) to produce solutions in table form, `False` for a list of $\{r, s\}$ pairs. This function calls `solvePell` to get the base solution of the Pell equation.

In[291]:=
```
solvePellRecurrence[D_, n_: 1, tableform_: True] := Module[
  {hkPell, h, k, rsvalues} ,
  hkpell = solvePell[D]; (* get the base solution *)
  If[Length[hkpell] == 2, (* solution was found *)
   {h, k} = hkpell;
   rsvalues = RecurrenceTable[{
      r[i + 1] == h r[i] + D k s[i],
      s[i + 1] == k r[i] + h s[i],
      r[1] == h, s[1] == k},
     {r, s}, {i, n}];
   If[tableform,
    TableForm[rsvalues, TableHeadings → {None, {"r", "s"}}],
    rsvalues
   ]
  ]
 ]
```

Exercise the test on $D > 0$ and nonsquare. Here it is positive but square.

In[292]:=
```
solvePell[64]
```

```
D=64 not OK
```

Here it is negative.

In[293]:=
```
solvePell[-7]
```

```
D=-7 not OK
```

In[294]:=
```
solvePell[1]
```

```
D=1 not OK
```

Examples

Exercise the function on some examples. Here is the $D == 72$ from the previously worked-out $p/q == 7/18$.

In[295]:=
```
solvePell[72]
```

Out[295]=
```
{17, 2}
```

Generate the first 3 solutions.

In[296]:=
```
solvePellRecurrence[72, 3]
```

Out[296]//TableForm=

r	s
17	2
577	68
19 601	2310

Produce the result as a list.

In[297]:=
```
solvePellRecurrence[72, 3, False]
```

Out[297]=
```
{{17, 2}, {577, 68}, {19 601, 2310}}
```

Use these to find the x, y solutions for this ratio.

In[298]:=
```
recycleSolutions[xyfromuv /@ (p solvePellRecurrence[72, 3, False]) /. {p → 7, q → 18}]
```

Out[298]//TableForm=

x	y
2	7
7	21
21	60
95	266
266	742
742	2067
3256	9065
9065	25 235
25 235	70 246

Here is a famous example, where the repeat length (11) is long and odd. Fermat challenged contemporary mathematicians to solve this instance.

In[299]:=
```
solvePell[61]
```

Out[299]=
```
{1 766 319 049, 226 153 980}
```

11.6 A cautionary example: $p/q == 6/17$

At this point it may seem that the problem is completely solved: the Pell Equation is always solvable, and with the help of the recycling recurrence we can find some solutions it misses. However, we have not proved that the (u, v) solution obtained by solving the Pell Equation always maps to integer (x, y), and in fact it does not always. Furthermore, there can be solutions of Equation (2) that the method misses. Here is an example illustrating both of these issues.

Calculate D.

In[300]:=
```
D6o17 = q (q - 2 p) /. {p → 6, q → 17}
```

Out[300]=
```
85
```

Solve the Pell Equation and multiply the solution by p to obtain (u, v). Find the first 3 solutions.

In[301]:=
```
uvsolns6o17 = 6 solvePellRecurrence[D6o17, 3, False]
```

Out[301]=
```
{{1 714 614, 185 976}, {979 967 056 326, 106 292 351 088},
 {560 088 411 436 734 774, 60 750 117 755 947 368}}
```

My, my, those are big. This happens sometimes with the Pell Equation.

Convert these to (x, y).

In[302]:=
```
(xyfromuv /@ uvsolns6o17) /. {p → 6, q → 17}
```

Out[302]=
$$\left\{\left\{\frac{392\,364}{5}, \frac{1\,322\,244}{5}\right\}, \{44\,850\,530\,088, 151\,142\,881\,176\}, \left\{\frac{128\,168\,911\,328\,498\,964}{5}, \frac{431\,919\,500\,108\,235\,804}{5}\right\}\right\}$$

So the first and third solutions are fractional, not admissible. In Section 12.2 we show that the second Pell-method solution is always admissible, and the second obtained by applying the Pell recurrence to that, and so on *ad infinitum*.

Here are the nearby solutions found using the recycling recurrence.

In[303]:=
```
recycleSolutions[%]
```

Out[303]//TableForm=

x	y
13 309 062 481	44 850 530 088
44 850 530 088	151 142 881 176
151 142 881 176	509 340 034 225

There are also solutions missed by the Pell method. The reverse search turned up the solution $(x, y) == (280, 945)$.

In[304]:=
```
probdifferent[{280, 945}]
```

Out[304]=
$$\frac{6}{17}$$

We can find additional solutions by recycling.

In[305]:=
```
recycleSolutions[{{280, 945}}]
```

Out[305]//TableForm=

x	y
280	945
945	3186
3186	10 738

These solutions are much smaller than the smallest solution obtained using the Pell Equation. Later we will develop methods that find all solutions.

11.7 Function to solve hyperbolic case via Pell Equation

Now define a function to use the Pell method to solve Equation (2) for x, y. Include optional argument n to get solutions after the first, and optional `tableform` for formatting output as table vs. list. This function does not rely on any local function definitions except `solvePell` and `solvePellRecur`-rence.

Outline of the function:

- Solve Pell Equation for base solution $\{h, k\}$. The function will return null and issue an error message if D is negative or square. It will always return a list of length 2 if $D > 0$ is nonsquare, since the Pell Equation always has a solution.
- Run the Pell recurrence to generate the first n solutions, including $\{h, k\}$ as the first. Convert these to $\{u, v\}$ pairs by multiplying by p.
- Convert the $\{u, v\}$ pairs to $\{x, y\}$. These are not necessarily integer, though they are always non-negative.
- Sift out the integer solutions.

In[306]:=
```
solveHyperbolicByPell[z_, n_: 1, tableform_: True] := Module[
  {p, q, D, uvsolns, xysolns},
  p = Numerator[z];
  q = Denominator[z];
  D = q (q - 2 p);
  uvsolns = solvePellRecurrence[D, n, False];
  If[Length[uvsolns] > 0,
   uvsolns = p uvsolns;
   xysolns = Table[
     1/2 {((p - u)/(2 p - q) - v), ((p - u)/(2 p - q) + v)} /. {u → uvsolns[[i]][[1]], v → uvsolns[[i]][[2]]},
     {i, Length[uvsolns]}];
   xysolns = Cases[xysolns, {_Integer, _Integer}];
   If[tableform,
    TableForm[xysolns, TableHeadings → {None, {"x", "y"}}],
    xysolns]
  ]
 ]
```

Exercise the test for valid D.

In[307]:=
```
solveHyperbolicByPell[4 / 9]
```

```
D=9 not OK
```

In[308]:=
```
solveHyperbolicByPell[5 / 9]
```

```
D=-9 not OK
```

Examples

Exercise it on some examples. This one yields an admissible solution from each Pell generation.

In[309]:=
```
solveHyperbolicByPell[7 / 18, 3]
```

Out[309]//TableForm=

x	y
7	21
266	742
9065	25 235

In[310]:=
```
solveHyperbolicByPell[7 / 18, 3, False]
```

Out[310]=
```
{{7, 21}, {266, 742}, {9065, 25 235}}
```

The next example, as we saw above, yields fractional solutions from the first and third Pell solution. If we only ask for one solution, we get none. When we provide $n == 3$, we only get 1 solution.

In[311]:=
```
solveHyperbolicByPell[6 / 17]
```

Out[311]//TableForm=
```
{}
```

In[312]:=
```
solveHyperbolicByPell[6 / 17, 3]
```

Out[312]//TableForm=

x	y
44 850 530 088	151 142 881 176

For most ratios, the function skips the trivial solutions, but ratios with $p == 1$ are a special case.

In[313]:=
```
solveHyperbolicByPell[1 / 7, 3]
```

Out[313]//TableForm=

x	y
0	1
1	13
13	156

We will examine the case $p == 1$ in the context of the Pell Equation in Section 11.9. Therefore I regard this as a feature of the function, rather than a bug to fix. Note that (0, 1) does not correspond to the trivial solution of the Pell equation; that maps to (0, 0) which indeed is not included in the function results.

Here is an example that gives admissible solutions on each generation.

In[314]:=
```
solveHyperbolicByPell[2 / 7, 3]
```

Out[314]//TableForm=

x	y
6	30
696	3336
76 590	366 966

11.8 Solutions from the trivial solutions via Pell recurrence

From the Pell recurrence formula (21) and the fact that $(u, v) == (p\,r, p\,s)$, we can rewrite the formula so the recurrence works directly on (u, v). Let (u, v) be a solution of (11):

$u^2 - D\,v^2 == f \quad (11)$

and let (r, s) be a solution to the Pell equation (20):

$r^2 - D\,s^2 == 1 \quad (20)$

Then another solution (u', v') of Equation (11) is given by equating rational and irrational terms on each side of

$$u' + v'\sqrt{D} == \left(r + s\sqrt{D}\right)\left(u + v\sqrt{D}\right) \tag{23}$$

Show that this works:

```
In[315]:= Collect[(r + s √D) (u + v √D), {1, √D}]
```

$\text{Out[315]=}\ r\,u + D\,s\,v + \sqrt{D}\,(s\,u + r\,v)$

Identify $u' == r\,u + D\,s\,v$, $v' == s\,u + r\,v$. Show that they satisfy Equation (11).

```
In[316]:= Expand[(r u + D s v)^2 - D (s u + r v)^2]
```

$\text{Out[316]=}\ r^2\,u^2 - D\,s^2\,u^2 - D\,r^2\,v^2 + D^2\,s^2\,v^2$

```
In[317]:= Collect[%, {r^2, s^2}]
```

$\text{Out[317]=}\ r^2\,(u^2 - D\,v^2) + s^2\,(-D\,u^2 + D^2\,v^2)$

Using the fact that (u, v) satisfies Equation (11), and (r, s) satisfies Equation (20) this is

$r^2\,f + -D\,s^2\,f == (r^2 - D\,s^2)\,f == f$

Thus Equation (23) is valid even if (u, v) was not obtained from a Pell Equation solution.

We can write it in the form of a recurrence, where we set h, k to the base solution of (20).

$$\begin{aligned} u_{n+1} &== h\,u_n + D\,k\,v_n \\ v_{n+1} &== k\,u_n + h\,v_n \end{aligned} \tag{24}$$

If we apply this recurrence to the three trivial solutions, we obtain the Pell solutions and their recycling neighbors. For example, here is what we get applying this to 7/18. Recall (from Section 5.2) the three trivial solutions with $u > 0$ are $(u, v) == (p, 0)$, $(q - p, \pm 1)$. For $p == 7$ and $q == 18$ these are (7, 0) and (11, ±1).

```
In[318]:= uvsolns7o18 = Flatten[Table[RecurrenceTable[{
               r[i + 1] == h r[i] + D k s[i],
               s[i + 1] == k r[i] + h s[i],
               r[1] == uv[[1]],
               s[1] == uv[[2]]} /. {h → 17, k → 2, D → 72},
           {r, s}, {i, 3}], {uv, {{7, 0}, {11, -1}, {11, 1}}}], 1]

Out[318]= {{7, 0}, {119, 14}, {4039, 476}, {11, -1},
           {43, 5}, {1451, 171}, {11, 1}, {331, 39}, {11 243, 1325}}

In[319]:= xysolns7o18 = (xyfromuv /@ uvsolns7o18) /. {p → 7, q → 18}

Out[319]= {{0, 0}, {7, 21}, {266, 742}, {1, 0}, {2, 7}, {95, 266}, {0, 1}, {21, 60}, {742, 2067}}
```

Apart from the trivial solutions these are the same as generated by recycling the Pell-method solutions.

In[320]:=
```
Sort[%]
```

Out[320]=
```
{{0, 0}, {0, 1}, {1, 0}, {2, 7}, {7, 21}, {21, 60}, {95, 266}, {266, 742}, {742, 2067}}
```

In[321]:=
```
recycleSolutions[solveHyperbolicByPell[7 / 18, 2, False], False]
```

Out[321]=
```
{{2, 7}, {7, 21}, {21, 60}, {95, 266}, {266, 742}, {742, 2067}}
```

In Section 12.3 it is proved that this always happens.

11.9 Classes of solutions

We now turn to methods that will allow us to find all solutions, and to be certain that we have found them all. In this section we do not apply the restrictions $u > 0$, $v \geq 0$ that we have been using to assure admissible solutions to the puzzle. We need to consider all solutions of Equation (11) that can exist, and later winnow the admissible (x, y).

11.9.1 Solution classes and fundamental solutions

Nagell, section 58, presents Equation (23) and defines **classes** of solutions of Equation (11) according to whether they are related via that equation for some values of (r, s). Define a **fundamental solution** of a class as the member of the class (u_0, v_0) for which $v_0 \geq 0$ is smallest. This determines u_0 except for sign. In most cases, different signs of u_0 belong to different classes. However, if $v_0 == 0$, the class is called ambiguous, and both signs of u_0 belong to it. In this case, for definiteness take $u_0 > 0$ for the fundamental solution. (In our problem, $u == 0$ is never part of a solution of Equation (11), since f is square and D is nonsquare.)

When $p == 1$ Equation (11) is the Pell equation. The Pell recurrence gives all solutions, so for this case all solutions are in the same class. The fundamental solution is the trivial solution, (1, 0).

The three trivial solutions are always present, and except when $p == 1$ are members of different classes (see later in this section). Hence the Pell method generates all solutions in these classes. There may, however, be other classes.

We make one change to the definition of a fundamental solution. Nagell adopts the convention for the fundamental solution that $v_0 \geq 0$ and allows u_0 to be opposite in sign for the conjugate class. Since we are interested in positive x, y solutions and $u < 0$ yields negative ones, it is more convenient for us to keep $u_0 > 0$ and let the sign of v_0 vary for the conjugate class. Thus our definition of a fundamental solution of a class idss the solution for which $u_0 > 0$ and $| v_0 |$ is least.

If (r, s) is any solution of the Pell equation $r^2 - D s^2 == 1$ and (u_0, v_0) is the fundamental solution of a class $\mathbb{C}$, then all solutions of class $\mathbb{C}$ are given by

$$u + v \sqrt{D} == \left(u_0 + v_0 \sqrt{D}\right) \left(r + s \sqrt{D}\right) \tag{25}$$

as (r, s) ranges over all solutions of the Pell equation. Thus we can test whether two solutions belong to the same class by solving for r and s and requiring them to be integer.

In[322]:=
```
Solve[{u1 r + D v1 s == u2, u1 s + v1 r == v2}, {r, s}]
```

Out[322]=
$$\left\{\left\{r \to -\frac{-u_1 u_2 + D v_1 v_2}{u_1^2 - D v_1^2},\ s \to -\frac{u_2 v_1 - u_1 v_2}{u_1^2 - D v_1^2}\right\}\right\}$$

Rewrite to remove the initial minus signs.

$$\left\{r \to \frac{u_1 u_2 - D v_1 v_2}{u_1^2 - D v_1^2},\ s \to \frac{u_1 v_2 - u_2 v_1}{u_1^2 - D v_1^2}\right\} \tag{26}$$

The denominator is f. Therefore require the numerators to be divisible by f.

Define a function for the next few steps where we show the trivial solutions form three classes except when $p == 1$.

In[323]:=
```
sameClass[{u_, v_}, {U_, V_}, D_] :=
  (u U - D v V) / (u^2 - D v^2) ∈ Integers && (v U - u V) / (u^2 - D v^2) ∈ Integers
```

Verify that the trivial solutions belong to different classes. They are $(p, 0)$ and $(q - p, \pm 1)$.

Show $(p, 0)$ is in a different class from $(q - p, \pm 1)$

In[324]:=
```
sameClass[{p, 0}, {q - p, 1}, q (q - 2 p)]
```

Out[324]=
$$\frac{-p + q}{p} \in \mathbb{Z}\ \&\&\ \frac{1}{p} \in \mathbb{Z}$$

In[325]:=
```
sameClass[{p, 0}, {q - p, -1}, q (q - 2 p)]
```

Out[325]=
$$\frac{-p + q}{p} \in \mathbb{Z}\ \&\&\ \frac{1}{p} \in \mathbb{Z}$$

Since $\frac{1}{p} \notin \mathbb{Z}$ except if $p == 1$, these are in different classes.

Show $(q - p, 1)$ and $(q - p, -1)$ are in different classes.

In[326]:=
```
Simplify[sameClass[{q - p, 1}, {q - p, -1}, q (q - 2 p)]]
```

Out[326]=
$$\frac{p^2 - 4 p q + 2 q^2}{p^2} \in \mathbb{Z}\ \&\&\ \frac{2 (p - q)}{p^2} \in \mathbb{Z}$$

Clearly this requirement is false if $p > 1$, so again these are in different classes, except if $p == 1$.

In[327]:=
```
Simplify[sameClass[{q - p, 1}, {q - p, -1}, q (q - 2 p)] /. p → 1,
  Assumptions → {q ∈ Integers}]
```

Out[327]=
```
True
```

However, $(\pm p, 0)$ are in the same class since it is ambiguous.

In[328]:=
```
sameClass[{p, 0}, {-p, 0}, q (q - 2 p)]
```

Out[328]=
```
True
```

11.9.2 Completeness

The concept of classes allows us to define precisely what it means for a solution method to be complete for the hyperbolic case: it must find at least one solution from each class. From those, all the other solutions can be found through the recurrence (24), including running it backwards if the solution is not a fundamental solution.

11.9.3 Functions for testing if two solutions are in the same class

Define a more robust function for testing whether two (u, v) solutions belong to the same class. We can usually obtain D and f from the two solutions, so we will not make them arguments of the function.

```
In[329]:= Solve[u^2 - D v^2 == U^2 - D V^2, D]
```

$$\text{Out[329]=}\ \left\{\left\{D \to \frac{u^2 - U^2}{v^2 - V^2}\right\}\right\}$$

Then $f == u^2 - D\,v^2 == U^2 - D\,V^2$j.

If $v^2 == V^2$ this is undefined, but in that case the two solutions are either the same or negatives of each other, or else conjugates, and are respectively in the same class or in different classes. So there is no need to provide D or f as an argument.

```
In[330]:= uvSameClass[{u_, v_}, {U_, V_}] := Module[{D, f},
   If[v^2 ≠ V^2,
    D = (u^2 - U^2) / (v^2 - V^2);
    f = u^2 - D v^2;
    (u U - D v V) / f ∈ Integers && (v U - u V) / f ∈ Integers
    , (* else v^2==V^2 case *)
    {u, v} == {U, V} || {u, v} == -{U, V}
   ]
  ]
```

Test this on a few examples drawn from 7/18. The trivial solutions:

```
In[331]:= uvSameClass[{7, 0}, {7, 0}]

Out[331]= True

In[332]:= uvSameClass[{7, 0}, {-7, 0}]

Out[332]= True

In[333]:= uvSameClass[{11, 1}, {-11, -1}]

Out[333]= True

In[334]:= uvSameClass[{11, 1}, {11, -1}]

Out[334]= False
```

Use this function to identify classes of the previous example 7/18, $D == 72$. The list was generated with 3 successive Pell generations grouped together, so we should see sequences of 3 in the same class, for 3 classes, 9 solution pairs in each class.

In[335]:=
```
uvsolns7o18
```

Out[335]=
```
{{7, 0}, {119, 14}, {4039, 476}, {11, -1},
 {43, 5}, {1451, 171}, {11, 1}, {331, 39}, {11 243, 1325}}
```

In[336]:=
```
Position[Table[uvSameClass[uvsolns7o18[[i]], uvsolns7o18[[j]]],
   {i, Length[uvsolns7o18]}, {j, Length[uvsolns7o18]}], True]
```

Out[336]=
```
{{1, 1}, {1, 2}, {1, 3}, {2, 1}, {2, 2}, {2, 3}, {3, 1}, {3, 2}, {3, 3},
 {4, 4}, {4, 5}, {4, 6}, {5, 4}, {5, 5}, {5, 6}, {6, 4}, {6, 5}, {6, 6},
 {7, 7}, {7, 8}, {7, 9}, {8, 7}, {8, 8}, {8, 9}, {9, 7}, {9, 8}, {9, 9}}
```

Here are the {*u*, *v*} solution pairs that are in the same class with each other.

In[337]:=
```
Table[{uvsolns7o18[[pos[[1]]]], uvsolns7o18[[pos[[2]]]]}, {pos, %}]
```

Out[337]=
```
{{{7, 0}, {7, 0}}, {{7, 0}, {119, 14}}, {{7, 0}, {4039, 476}}, {{119, 14}, {7, 0}},
 {{119, 14}, {119, 14}}, {{119, 14}, {4039, 476}}, {{4039, 476}, {7, 0}},
 {{4039, 476}, {119, 14}}, {{4039, 476}, {4039, 476}}, {{11, -1}, {11, -1}},
 {{11, -1}, {43, 5}}, {{11, -1}, {1451, 171}}, {{43, 5}, {11, -1}},
 {{43, 5}, {43, 5}}, {{43, 5}, {1451, 171}}, {{1451, 171}, {11, -1}},
 {{1451, 171}, {43, 5}}, {{1451, 171}, {1451, 171}}, {{11, 1}, {11, 1}},
 {{11, 1}, {331, 39}}, {{11, 1}, {11 243, 1325}}, {{331, 39}, {11, 1}},
 {{331, 39}, {331, 39}}, {{331, 39}, {11 243, 1325}}, {{11 243, 1325}, {11, 1}},
 {{11 243, 1325}, {331, 39}}, {{11 243, 1325}, {11 243, 1325}}}
```

We can also define a function that determines whether two (x, y) solutions are in the same class. This does not need to be given D or p/q, since p/q is a function of (x, y) except for the trivial solutions. If one of the arguments is one of the trivial solutions, use the p/q of the other. If both arguments are trivial, then they are in different classes unless they are identical, except when $p == 1$, which the function does not know. In that case alert the user as best we can.

In[338]:=
```
xySameClass[{x_, y_}, {X_, Y_}] := Module[
  {z, Z, p, q, u, v, U, V, D, f},
  z = -1; Z = -1;
  If[(x + y) (x + y - 1) ≠ 0,
   z = 2 x y / ((x + y) (x + y - 1));
  ];
  If[(X + Y) (X + Y - 1) ≠ 0,
   Z = 2 X Y / ((X + Y) (X + Y - 1));
  ];
  If[Z < 0, Z = z]; (* If one solution is trivial, get p/q from the other *)
  If[z < 0, z = Z];
  If[z ≥ 0,
   If[z == Z,
    p = Numerator[z]; q = Denominator[z];
    D = q (q - 2 p); f = p^2;
    {u, v} = {p + (q - 2 p) (y + x), y - x};
    {U, V} = {p + (q - 2 p) (Y + X), Y - X};
    (u U - D v V) / f ∈ Integers && (v U - u V) / f ∈ Integers,
    Print["Solutions do not satisfy same equation"]; (* z ≠ Z *)
    False
   ],
   {x, y} == {X, Y} || p == 1
   (* trivial solutions are in diff classes except if p==1 *)
  ]
 ]
```

Run this on the x, y solutions corresponding to the u, v solutions for 7/18. To avoid redundancy, just show the lower triangle $j \leq i$.

In[339]:=
```
Position[Table[xySameClass[xysolns7o18[[i]], xysolns7o18[[j]]],
  {i, Length[xysolns7o18]}, {j, i}], True]
```

Out[339]=
```
{{1, 1}, {2, 1}, {2, 2}, {3, 1}, {3, 2}, {3, 3}, {4, 4}, {5, 4}, {5, 5},
 {6, 4}, {6, 5}, {6, 6}, {7, 7}, {8, 7}, {8, 8}, {9, 7}, {9, 8}, {9, 9}}
```

Here are the (x, y) solution pairs that are in the same classes with each other.

In[340]:=
```
Table[{xysolns7o18[[pos[[1]]]], xysolns7o18[[pos[[2]]]]}, {pos, %}]
```

Out[340]=
```
{{{0, 0}, {0, 0}}, {{7, 21}, {0, 0}}, {{7, 21}, {7, 21}}, {{266, 742}, {0, 0}},
 {{266, 742}, {7, 21}}, {{266, 742}, {266, 742}}, {{1, 0}, {1, 0}},
 {{2, 7}, {1, 0}}, {{2, 7}, {2, 7}}, {{95, 266}, {1, 0}}, {{95, 266}, {2, 7}},
 {{95, 266}, {95, 266}}, {{0, 1}, {0, 1}}, {{21, 60}, {0, 1}}, {{21, 60}, {21, 60}},
 {{742, 2067}, {0, 1}}, {{742, 2067}, {21, 60}}, {{742, 2067}, {742, 2067}}}
```

As expected, this gives the same classes as for the corresponding u, v. Successive triplets are recycling neighbors, which belong to different classes. To show this more clearly, here is the first nontrivial recycling triplet. They are in different classes.

In[341]:=
```
recycleSolutions[{{7, 21}}]
```

Out[341]//TableForm=

x	y
2	7
7	21
21	60

In[342]:=
```
xySameClass[{2, 7}, {7, 21}]
```

Out[342]= False

In[343]:=
```
xySameClass[{7, 21}, {21, 60}]
```

Out[343]= False

In[344]:=
```
xySameClass[{2, 7}, {21, 60}]
```

Out[344]= False

Here is a case where $p == 1$ so every solution is in the same class, including recycling neighbors and the trivial solutions.

In[345]:=
```
recycle[{1, 7}]
```

Out[345]= {7, 42}

In[346]:=
```
xySameClass[{1, 7}, {7, 42}]
```

Out[346]= True

In[347]:=
```
xySameClass[{1, 7}, {0, 1}]
```

Out[347]= True

In[348]:=
```
xySameClass[{1, 7}, {0, 0}]
```

Out[348]= True

In[349]:=
```
xySameClass[{1, 7}, {1, 0}]
```

Out[349]= True

The trivial solutions are not in the same class with all solutions if $p > 1$. Each Pell generation solution is in the same class as its predecessor.

In[350]:=
```
xySameClass[{2, 7}, {0, 1}]
```

Out[350]= False

In[351]:=
```
xySameClass[{2, 7}, {0, 0}]
```

Out[351]= False

In[352]:=
```
xySameClass[{2, 7}, {1, 0}]
```

Out[352]= True

In[353]:=
```
xySameClass[{2, 7}, {95, 266}]
```
Out[353]=
```
True
```

The trivial solutions are in different classes unless they are the same or $p == 1$.

In[354]:=
```
xySameClass[{0, 1}, {0, 1}]
```
Out[354]=
```
True
```

This one is true if and only if $p == 1$.

In[355]:=
```
xySameClass[{0, 1}, {0, 0}]
```
Out[355]=
```
p$19031 == 1
```

11.10 Solving hyperbolic case by direct search

Searching for solutions might seem impractical since the hyperbola extends to infinity, but it turns out that the fundamental solutions are bounded in size, making search viable in many cases.

11.10.1 Bounds on size of fundamental solution

Nagel section 58 provides bounds on the magnitude of the fundamental solution for any class of Equation (11). In my notation, Nagell's bounds are (using our convention $u_0 > 0$ and v_0 may be either sign):

$$|v_0| \le \frac{k}{\sqrt{2\,(h+1)}}\sqrt{f} == \frac{k\,p}{\sqrt{2\,(h+1)}} \tag{27}$$

$$0 < u \le \sqrt{\frac{h+1}{2}}\sqrt{f} == p\sqrt{\frac{h+1}{2}} \tag{28}$$

where (h, k) is the fundamental (smallest nontrivial) solution of the Pell equation $h^2 - D\,k^2 == 1$. A method of solution is therefore to search within this range to find all fundamental solutions. All solutions of each class are then found by using the recurrence (24).

Asymptotically, as h and k grow, $h \simeq k\sqrt{D}$ so both expressions are proportional to $p\sqrt{h}$ or $p\sqrt{k}$, which means the bounds grow much more slowly than h or k. (Note the similarity to the bounds on v for the elliptical case, Section 8.2, where v_{max} grows inversely as the square root of $\epsilon == p/q - 1/2$, enabling search to be practical on very elongated ellipses.)

Putting $h \simeq k\sqrt{D}$ and neglecting 1 relative to h, the bound is approximately

$$|v_0| \lesssim \frac{k\,p}{\sqrt{2\,k\,\sqrt{D}}} == p\sqrt{\frac{k}{2\,\sqrt{D}}}$$

Define a function to calculate the bound for a given ratio. Use the `Floor` function to produce an integer result.

In[356]:=
```
vboundNagell[z_] := Module[{p, q, D, hkpell, h, k},
  p = Numerator[z]; q = Denominator[z];
  D = q (q - 2 p);
  hkpell = solvePell[D]; (* get the base solution *)
  If[Length[hkpell] == 2,
   {h, k} = hkpell;
   Floor[k p / Sqrt[2 (h + 1)]]
  ]
 ]
```

11.10.2 Function to find fundamental solutions (u, v) by direct search

Since $|u| > |v|$ for our problem ($f > 0$), the search is best done by testing values of v, as in the elliptical case.

Define a function to search for integer solutions of Equation (11) within the range of v set by Nagell's bound on v that can be a fundamental solution of any class. This function uses `solvePell` to find (h, k) needed for the bound. It produces a list of (u, v) solutions that are the fundamental solutions for their respective classes. Larger solutions can be found using the Pell recurrence. These solutions can be turned into (x, y) solutions. We only need to test values of $v \geq 0$. The conjugate classes are obtained by reversing the sign of v.

Outline of the function:

- Solve Pell equation to get base solution $\{h, k\}$ needed to calculate bound v_{max}.
- Solve $u^2 - D v^2 == p^2$ for u, for each integer v, $0 \leq v \leq v_{max}$.
- Sift out the integer values of u and pair them with their corresponding v values.

In[357]:=
```
solveuvBySearch[z_] := Module[
  {p, q, D, hkpell, h, k, vmax, testvalues, vsolns, usolns},
  p = Numerator[z];
  q = Denominator[z];
  D = q (q - 2 p);
  hkpell = solvePell[D]; (* get the base solution *)
  If[Length[hkpell] == 2,
   {h, k} = hkpell;
   vmax = Floor[k p / Sqrt[2 (h + 1)]];
   testvalues = Table[Sqrt[p^2 + D v^2], {v, 0, vmax}];
   vsolns = Position[testvalues, _Integer, {1}];
   usolns = Extract[testvalues, vsolns];
   vsolns = Flatten[vsolns] - 1; (* adjust position 1..n to value 0..n-1 *)
   Table[{usolns[[i]], vsolns[[i]]}, {i, Length[vsolns]}]
  ]
 ]
```

Exercise the error handling if ratio is not hyperbolic with nonsquare *D*.

In[358]:=
```
solveuvBySearch[9 / 17]
```

```
D=-17 not OK
```

In[359]:=
```
solveuvBySearch[4 / 9]
```

```
D=9 not OK
```

Including the conjugates

The function returns only solutions with $v \geq 0$. It is sometimes useful to have the complete set of fundamental solutions, including the conjugates. This function takes a set of solutions and joins it with the conjugates, sorting and eliminating any duplicates that may have been introduced.

In[360]:=
```
uvIncludeConjugates[uvsolns_] :=
 Sort[DeleteDuplicates[Join[uvsolns, Table[{uv[[1]], -uv[[2]]}, {uv, uvsolns}]]]]
```

Example: $p / q == 7 / 18$

Applying this to the ratio 7/18 yields just the trivial solutions.

In[361]:=
```
uvsolnsbysearch7o18 = solveuvBySearch[7 / 18]
```

Out[361]=
```
{{7, 0}, {11, 1}}
```

The function suppresses negative v, but we know {11, −1} is another solution. So there are 3 classes in all. Generate the classes and convert to x, y.

In[362]:=
```
uvfundsolns7o18 = uvIncludeConjugates[uvsolnsbysearch7o18]
```
Out[362]=
```
{{7, 0}, {11, -1}, {11, 1}}
```
In[363]:=
```
(xyfromuv /@ uvfundsolns7o18) /. {p → 7, q → 18}
```
Out[363]=
```
{{0, 0}, {1, 0}, {0, 1}}
```

Applying the Pell recurrence to the fundamental solutions yields all of the solutions that exist. So in this case, the method of solution via the Pell equation plus recycling recurrence is complete.

Example: $p/q == 6/17$

The function does not warn if the bound on v_0 is large, which it can be in some cases. In Section 11.6 we saw that 6/17 has a large first Pell solution. Let's see if solving it by search is practical.

In[364]:=
```
{h6o17, k6o17} = solvePell[D6o17]
```
Out[364]=
```
{285 769, 30 996}
```
In[365]:=
```
vboundNagell[6 / 17]
```
Out[365]=
```
245
```

So this is actually not bad. The approximate proportionality to $\sqrt{k}$ rather than k saves it.

In[366]:=
```
Timing[uvsolnsbysearch6o17 = solveuvBySearch[6 / 17]]
```
Out[366]=
```
{0.003693, {{6, 0}, {11, 1}, {74, 8}, {249, 27}, {839, 91}}}
```

Including the conjugates, there are 9 classes.

Example: $p/q == 25/51$

Before writing a function to apply the search method to solve the hyperbolic case for x, y, it is instructive to work through an example. The ratio 25/51 turns out to have more than the 3 classes of the trivial solutions, and a relatively small bound on v for the search, so it is a good choice for this.

Calculate D.

In[367]:=
```
D25o51 = q (q - 2 p) /. {p → 25, q → 51}
```
Out[367]=
```
51
```

Find the base solution of the Pell Equation.

In[368]:=
```
{h25o51, k25o51} = solvePell[51]
```
Out[368]=
```
{50, 7}
```

Calculate the search bound for v.

In[369]:=
```
vboundNagell[25 / 51]
```
Out[369]=
```
17
```

Find the set of $\{u, v\}$ solutions.

In[370]:=
```
uvsolns25o51 = solveuvBySearch[25 / 51]
```
Out[370]=
```
{{25, 0}, {26, 1}, {110, 15}}
```

The first is an ambiguous class since $v == 0$. Changing the sign of v in the others gives solutions that are in different classes.

In[371]:=
```
uvfundsolns25o51 = uvIncludeConjugates[uvsolns25o51]
```
Out[371]=
```
{{25, 0}, {26, -1}, {26, 1}, {110, -15}, {110, 15}}
```

So there are 5 classes in all. The solutions {25, 0} and {26, ±1} are the trivial solutions. Calculate the $\{x, y\}$ solutions that these give.

In[372]:=
```
xyfundsolns25o51 = (xyfromuv /@ uvfundsolns25o51) /. {p → 25, q → 51}
```
Out[372]=
```
{{0, 0}, {1, 0}, {0, 1}, {50, 35}, {35, 50}}
```

Only 3 of these are distinct, with $x \le y$. Now find the second generation of $\{u, v\}$ solutions.

In[373]:=
```
uvgen2solns25o51 = Table[{h25o51 u + D25o51 k25o51 v, k25o51 u + h25o51 v} /.
    {u → uvfundsolns25o51[[i]][[1]], v → uvfundsolns25o51[[i]][[2]]},
   {i, Length[uvfundsolns25o51]}]
```
Out[373]=
```
{{1250, 175}, {943, 132}, {1657, 232}, {145, 20}, {10 855, 1520}}
```

In[374]:=
```
xygen2solns25o51 = (xyfromuv /@ uvgen2solns25o51) /. {p → 25, q → 51}
```
Out[374]=
```
{{525, 700}, {393, 525}, {700, 932}, {50, 70}, {4655, 6175}}
```

Note that at this stage all the solutions are distinct, because now all $v > 0$. But we would not have obtained them all if we had used only the positive v solutions in the recurrence.

Compare with the solution found using the Pell Equation method.

In[375]:=
```
solveHyperbolicByPell[25 / 51, 1, False]
```
Out[375]=
```
{{525, 700}}
```

This is the first of the 2nd-generation set of solutions, i.e. obtained from the fundamental solution $(x, y) == (0, 0)$.

Put all the solutions into table form.

In[376]:=
```
TableForm[Sort[Join[xyfundsolns25o51, xygen2solns25o51]],
 TableHeadings → {None, {"x", "y"}}]
```

Out[376]//TableForm=

x	y
0	0
0	1
1	0
35	50
50	35
50	70
393	525
525	700
700	932
4655	6175

The table has some solutions we would normally suppress, namely the three trivial solutions and (50, 35) which is not distinct from (35, 50).

11.10.3 Function to solve hyperbolic case for (x, y) by search

We write a function to find the fundamental solutions (u, v) by search, optionally use them to generate additional solutions via the Pell recurrence, and convert them to (x, y) solutions. Optional argument `iters` is the number of iterations of the Pell recurrence to use, counting the fundamental solutions as iteration 1. Note that the trivial solutions are always fundamental solutions of their classes. (For $p == 1$ only one of them is fundamental.) This means that to get the smallest nontrivial (x, y) from them requires `iters` at least 2.. Any other classes that may exist have fundamental solutions that give nontrivial solutions immediately, although these are not necessarily admissible, and may need a Pell iteration to get admissible solutions. For these reasons, the default value of the optional parameter `iters` is 3. This function uses `solvePell` but no other local functions.
Outline of function:

- Solve for fundamental solutions having $v \geq 0$ as in `solveuvBySearch`.
- Extend the list to include conjugate solutions by changing sign of v. First solution is always $(p, 0)$ which has no conjugate partner; the rest always have $v > 0$ and have conjugates obtained by negating v. (Keeping $u > 0$ ensures that the Pell recurrence goes to larger positive values.)
- Use the Pell recurrence (24) `iters` times to generate additional solutions. Note: `iters` == 1 means only the fundamental solutions result; `iters` == 2 means the second generation appear.
- Convert the $\{u, v\}$ solutions to $\{x, y\}$.
- Sift the $\{x, y\}$ solutions for distinct admissible solutions. Apply `Sort` to each solution in the list to make $x \leq y$ for distinctness according to our convention, because the fundamental solutions with $v < 0$ yield $x > y$. Then sort the whole list to put the solutions in order of increasing size. Remove the duplicates created by the sort of each solution.

In[377]:=
```
solveHyperbolicBySearch[z_, iters_: 3, tableform_: True] :=
 Module[{p, q, D, hkpell, h, k, vmax, testvalues,
   usolns, vsolns, uvsolns, xysolns},
  p = Numerator[z];
  q = Denominator[z];
  D = q (q - 2 p);
  hkpell = solvePell[D]; (* get the base solution *)
  If[Length[hkpell] == 2,
   {h, k} = hkpell;
   vmax = Floor[k p / Sqrt[2 (h + 1)]];
   (* search for u,v solutions *)
   testvalues = Table[Sqrt[p^2 + D v^2], {v, 0, vmax}];
   vsolns = Position[testvalues, _Integer, {1}];
   usolns = Extract[testvalues, vsolns];
   vsolns = Flatten[vsolns] - 1; (* adjust position 1..n to value 0..n-1 *)
   usolns = Join[usolns, usolns[[2 ;;]]];
   vsolns = Join[vsolns, -vsolns[[2 ;;]]];
   uvsolns = Flatten[Table[RecurrenceTable[{
        u[i + 1] == h u[i] + D k v[i],
        v[i + 1] == k u[i] + h v[i],
        u[1] == usolns[[n]],
        v[1] == vsolns[[n]]},
       {u, v}, {i, iters}],
      {n, Length[usolns]}], 1];
   xysolns = Table[
     1/2 {((p - u)/(2 p - q) - v), ((p - u)/(2 p - q) + v)} /. {u -> uvsolns[[i]][[1]], v -> uvsolns[[i]][[2]]},
     {i, Length[uvsolns]}];
   xysolns = DeleteDuplicates[
     Sort[Cases[Sort /@ xysolns, {_Integer?Positive, _Integer?Positive}]]];
   If[tableform,
    TableForm[xysolns, TableHeadings -> {None, {"x", "y"}}],
    xysolns]
  ]
 ]
```

Examples

Try it out on the example 7/18 solved earlier. Here it uses the default of 3 iterations of the Pell recurrence. Since this has only the classes of the trivial solutions, nontrivial solutions appear only after an iteration.

In[378]:=
```
solveHyperbolicBySearch[7 / 18]
```

Out[378]//TableForm=

x	y
2	7
7	21
21	60
95	266
266	742
742	2067

We now know that there are no other solutions than these and the ones produced by running the iteration more.

Using only one iteration, only the trivial solutions result, and are suppressed.

In[379]:=
```
solveHyperbolicBySearch[7 / 18, 1]
```

Out[379]//TableForm=

{}

Here is the example 25/51 worked out above, which has 2 classes besides the trivial solutions. Their fundamental solutions yield admissible (x, y).

In[380]:=
```
solveHyperbolicBySearch[25 / 51]
```

Out[380]//TableForm=

x	y
35	50
50	70
393	525
525	700
700	932
4655	6175
6175	8190
40 524	53 725
53 725	71 225
71 225	94 424
466 690	618 675

Now try it on 6/17, which has 9 classes of solutions, as we saw above.

In[381]:=
```
solveHyperbolicBySearch[6 / 17, 2]
```

Out[381]//TableForm=

x	y
280	945
945	3186
3186	10 738
1 932 490	6 512 346
6 512 346	21 946 113
21 946 113	73 956 736

Although there are 9 classes and it went through 2 Pell recurrence iterations, the table has only 6 values. Here’s why:

In[382]:=
```
solveHyperbolicBySearch[6 / 17, 1]
```
Out[382]//TableForm=
```
{}
```

The fundamental solutions of all 9 classes do not yield admissible solutions. But as we saw for the solutions obtained via Pell, the next iteration yields admissible solutions.

The ratio 4/11 has only the trivial solution classes. The first Pell iteration yields fractional (x, y) so 3 iterations are needed to get to the first admissible solution.

In[383]:=
```
solveHyperbolicBySearch[4 / 11, 3]
```
Out[383]//TableForm=

x	y
105	336
336	1072
1072	3417

11.11 Feasibility of search

The method of direct search turns out to be quite practical for many cases. Even the example 6/17 which has a rather large base solution of the Pell equation is quickly solved on a modest computer. However, much larger base solutions can occur. Weisstein provides a list of solutions to the Pell equation for D up to 102. The solution for $D == 61$ is by far the largest. Since 61 is prime, this requires $q - 2p == 1$, so the ratio $p/q == 30/61$. Let’s look at it.

In[384]:=
```
solvePell[61]
```
Out[384]=
```
{1 766 319 049, 226 153 980}
```

The bound on v for this case is

In[385]:=
```
vboundNagell[30 / 61]
```
Out[385]=
```
114 149
```

This is not out of the question for a computer to solve, but much more computationally expensive than other methods we will look at next. This example has quite a few classes, and yields admissible solutions directly from the fundamental solutions, so we run it for just 1 iteration. The formula has been disabled so that it won’t be re-executed when the notebook is reopened and evaluated.

In[∘]:=
```
Timing[solveHyperbolicBySearch[30 / 61, 1]]
```

Out[∘]= {6.85486,

"x"	"y"
25	36
78	105
105	140
530	690
690	897
1674	2170
7684	9945
9945	12 870
12 870	16 654
58 720	75 969
140 712	182 040
182 040	235 505

}

It took only a few seconds to solve. Since this is the worst case among Pell solutions in Weisstein's list, all D values up to 102 are within reach of search. And since $q \leq D$ and $p < q/2$, this means all ratios with p, q of 1 or 2 digits can be solved quickly by search. However, ultimately there will be cases that are not practical. The next difficult case is $D == 109$, which results for $p/q == 54/109$.

In[386]:=
```
solvePell[109]
```
Out[386]=
```
{158 070 671 986 249, 15 140 424 455 100}
```

In[387]:=
```
vboundNagell[54 / 109]
```
Out[387]=
```
45 982 349
```

Assuming the time to search is proportional to the bound, using the timing for 30/61, we predict a search time of this many hours:

In[388]:=
```
(vboundNagell[54 / 109] / vboundNagell[30 / 61]) 6.85 / 3600
```
Out[388]=
```
0.766491
```

That is feasible, but a rather long time to dedicate to solving one instance. It is actually an underestimate, since the computations take more time as the integers grow larger. I went ahead and ran this on my laptop, with the following result.

```
Timing[uvsolnsbysearch54o109 = solveuvBySearch[54 / 109]]
```
Out[∘]=
```
{10 288.4, {{54, 0}, {55, 1}, {1035, 99}, {1254, 120},
  {39 621, 3795}, {48 015, 4599}, {1 517 770, 145 376}, {1 839 321, 176 175},
  {2 228 995, 213 499}, {70 459 290, 6 748 776}, {85 386 621, 8 178 555}}}
```

Convert time to hours.

In[389]:=
```
10 288.4 / 3600
```
Out[389]=
```
2.85789
```

Almost 3 hours. So this instance is at the border of what is feasible by search.

Here is one that is definitely out of reach, $D == 421$ (it has the largest Pell solution for $D \leq 500$).

In[390]:=
```
solvePell[421]
```
Out[390]= {3 879 474 045 914 926 879 468 217 167 061 449, 189 073 995 951 839 020 880 499 780 706 260}

In[391]:=
```
vboundNagell[210 / 421]
```
Out[391]= 450 764 467 341 464 849

Extrapolate the running time in seconds, again assuming proportionality, using the $D == 109$ time.

In[392]:=
```
(vboundNagell[210 / 421] / vboundNagell[54 / 109]) 10 288.4
```
Out[392]= 1.00857×10^{14}

Convert to years.

In[393]:=
```
% / (3600 × 24 × 365.25)
```
Out[393]= 3.19597×10^{6}

And this is probably an underestimate, because of the extra time it takes to do arithmetic on such large integers. One could imagine writing an optimized program and running it on a supercomputer to perhaps reduce this to a human timescale, but clearly this is not the way to go. This ratio actually has a modest-sized solution found by the reverse search, $(x, y) == (196, 225)$.

In[394]:=
```
probdifferent[{196, 225}]
```
Out[394]= $\frac{210}{421}$

Conclusion: while the method of direct search is simple and reasonably fast for many cases, it will be impractical for some cases.

Now we present a method of solution that is efficient and (I believe) complete.

11.12 Method of solution by reduction of RHS to 1

Alpern and Hua both provide methods of solving Equation (11) using continued fractions. (Alpern acknowledges Iain Davidson as his source for the method. It is probably owing to some earlier mathematician, but I have not located the original source.)

If $f < \sqrt{D}$, then if Equation (11) has a solution, the solution must be found among the convergents of $\sqrt{D}$, but if $f \neq 1$, the convergent(s) giving the solution(s) will not be located at the end of a repeat cycle. If $f \geq \sqrt{D}$, solutions are not guaranteed to be found among these convergents. So an approach is to transform the equation to bring the RHS to a magnitude less than $\sqrt{D}$. Hua's method recursively reduces f until it is less than $\sqrt{D}$, while keeping the form of the equation the same. Alpern's method reduces f to 1 in one step, at the cost of introducing a cross term involving $u\,v$. However, the solution can still be found using continued fractions.

Hua's method is complete. I believe Alpern's method is also complete, though I have not seen a proof.

Here I present Alpern's method. Since I don't know the correct person to attribute it to, I will call it the method of "reduction of RHS to 1," or just "reduction" for short. Hua's method, which I call the method of "recursive reduction," is given in Section 11.13.

The method of reduction requires the coefficients of u^2 and v^2 to be relatively prime, which they are. Let

$$u == s\,v - f\,w \tag{29}$$

where w is a new variable, and s is a parameter to be determined as described below.

In[395]:=
```
Collect[Expand[u^2 - D v^2 == f /. u → s v - f w], {w^2, v w, v^2}]
```

Out[395]= $\left(-D + s^2\right) v^2 - 2\,f\,s\,v\,w + f^2\,w^2 == f$

Dividing by f, and rearranging,

$$\frac{\left(s^2 - D\right)}{f}\,v^2 - 2\,s\,v\,w + f\,w^2 == 1 \tag{30}$$

To make the coefficient of v^2 integer, we need to choose s so that $s^2 - D$ is divisible by f:

$$s^2 - D \equiv 0 \pmod f \tag{31}$$

Values of s satisfying this congruence can be found by searching between 0 and $f - 1$. The range of search can be cut in half by noting that if s is a solution, then $f - s$ is also a solution, since s^2 is congruent to $(f - s)^2$ mod f. Hence it suffices to search up to $f/2$. $f/2$ itself cannot be a solution, since it would imply $f^2/4 - D \equiv 0 \pmod f$. Multiply by 4, to obtain $f^2 - 4\,D \equiv 0 \pmod f$. For our problem, f is square. So unless $f == 4$, this implies f and D have a common divisor. If $f == 4$, the congruence is $4 - D \equiv 0 \pmod 4$, which requires D even, again not relatively prime to f. So we can search $0 \le s < f/2$ and obtain the rest of the solutions by symmetry. It is interesting that if f is even, then if s is a solution, $f/2 - s$ is also a solution. This would allow the search range to be cut in half again. However, it is useful only for even f, so in order not to complicate the logic, the solution routine below does not try to take advantage of this.

- **Comment:** solving $s^2 - D \equiv 0 \pmod f$ is equivalent to finding n such that $D + n\,f == s^2$, i.e. a multiple of f added to D yields a square. When searching for values of n giving squares, only values $n < f/4$ need to be tested. If s is a solution, then $f - s$ is also a solution, since $(f - s)^2 \equiv s^2 \pmod f$. Since only $s < f$ are sought, either s or $f - s$ is less than or equal to $f/2$, hence $s^2 \le f^2/4$, so $n\,f < s^2 \le f^2/4$, or $n < f/4$. I tried this method of solving the congruence, but found that despite having a search range half the size of the search on s, the need to take a square root for the test, rather than simply a modulus, causes it to run slower in Mathematica than a search on s. There is a clever approach called the method of excludents that allows one to rule out large portions of the search space, but it does not seem readily turned into an automated method. It is not worth putting a lot of effort into speeding up the solving of the congruence anyway. It dominates the compute time only for large values of p, larger than 500 or so (as implemented here).

Observe that $s == q - p$ is always a solution to the congruence $s^2 - D == p^2$. $D == q(q - 2\,p) == (q - p)^2 - p^2$. In the congruence, the term p^2 can be dropped, leaving $s^2 - (q - p)^2 \equiv 0 \left(\text{mod } p^2\right)$. Clearly $s == q - p$ satisfies this. Then also $p^2 - s == p^2 - (q - p)$ is another solution. So 2 solutions always exist. (If q is large, $q - p$ may exceed the bound $f - 1$. In that case, its equivalent mod f will be found instead.) We show below that these lead to the trivial solutions.

For each value of s (positive or negative) satisfying Equation(31), Equation (30) can be solved using the method of continued fractions. (See Hua, Section 11.5, Exercise 2.) Solutions, if they exist, will be given by v/w equal to a convergent of the roots r of the companion quadratic equation obtained by setting $w == 1$ and $v == r$ in the LHS of Equation (30) and equating to 0:

$$\frac{s^2 - D}{f} r^2 - 2 s r + f == 0 \tag{32}$$

In[396]:= `Simplify[Solve[(s^2 - D)/f r^2 - 2 s r + f == 0, r]]`

Out[396]= $\left\{\left\{r \to -\frac{f}{\sqrt{D} - s}\right\}, \left\{r \to \frac{f}{\sqrt{D} + s}\right\}\right\}$

These roots can be expressed in a single formula

$$r \to \frac{f}{s \pm \sqrt{D}} \tag{33}$$

Changing the sign of s amounts to changing the overall sign of r, which changes only the signs of the convergents. Since changing sign of s and one of w or v leaves Equation (30) unchanged, we do not need to test negative values of s. Let $\delta = \pm 1$ hold the sign on the radical. Then we have

$$r \to \frac{f}{s + \delta \sqrt{D}}$$

where s is a positive solution of the congruence (31) and $\delta = \pm 1$.

In order for the method to be complete, one more step is needed. The method outlined above only finds solutions (u, v) that are relatively prime. Other solutions may exist that are not relatively prime. Suppose $\gcd(u, v) == d > 1$, then d^2 must divide f. To find those solutions, then, one can divide Equation (11) by d^2, solve, and then multiply the solution by d. (This is, of course, the same method as used earlier to find solutions by the Pell method using $d == p$.) Thus, for each square divisor d^2 of f, solve

$$\left(\frac{u}{d}\right)^2 - D \left(\frac{v}{d}\right)^2 == \frac{f}{d^2}$$

then multiply the solution by d. Since $f == p^2$, it has a square divisor for each divisor of p.

11.12.1 Trivial solutions are found by this method

This section can be skipped without loss of continuity. We show that the trivial solutions will be found by this method.

Observe that when $d == p$, $f/d^2 == 1$ is the new f, and the congruence (31) will always be satisfied, but the search range is limited to $s == 0$. Then Equation (30) is satisfied by $\{v \to 0, w \to 1\}$.

In[397]:= $\texttt{Simplify}\left[\frac{(s^2 - D)}{f} v^2 - 2\, s\, v\, w + f\, w^2 == 1 \;/.\; \{f \to 1,\, s \to 0,\, v \to 0,\, w \to 1\}\right]$

Out[397]= True

Then $\frac{u}{d} == s\, v - f\, w == w$ so this solution maps to the trivial solution $\{u \to \pm p,\, v \to 0\}$. We can show it will arise in the course of solution. When $s == 0$, the root $r == \pm 1 \big/ \sqrt{D}$ is less than 1 in magnitude so it has 0 as its first convergent, giving $v == 0$, $w == 1$. This trivial solution will be found twice, once for $\delta == 1$ and once for $\delta == -1$.

In[398]:= `Simplify[d {s v - f w, v} /. {d → p, s → 0, f → 1, v → 0, w → 1}]`

Out[398]= {-p, 0}

When $d == 1$, $f == p^2$ and with Mathematica's help we can see the other trivial solutions $\{u \to \pm(q - p),\, v \to \pm 1\}$ arising.

In[399]:= $\texttt{Reduce}\left[\left\{\frac{(s^2 - D)}{f} v^2 - 2\, s\, v\, w + f\, w^2 == 1 \;/.\; \{f \to p^2,\, D \to q\,(q - 2\,p),\, v \to 1,\, w \to 1\},\right.\right.$
$\left.\left. p > 0,\, q > 2\,p,\, 0 \le s < p^2\right\},\, s\right]$

Out[399]= $p > 1 \,\&\&\, 2\,p < q \le p + p^2 \,\&\&\, s == p + p^2 - q$

This solution is within the range $0 \le s < f == p^2$ since $q > 2\,p$. (It is the solution $s == p^2 - (q - p)$ obtained in the previous section using other arguments.)

Verify that this solution satisfies the congruence:

In[400]:= $\texttt{Simplify}\left[s^2 - D \;/.\; \{s \to p + p^2 - q,\, D \to q\,(q - 2\,p)\}\right]$

Out[400]= $p^2\,(1 + 2\,p + p^2 - 2\,q)$

This is divisible by p^2 as required. Now verify that the trivial solution $\{v \to 1,\, w \to 1\}$ satisfies Equation (30):

In[401]:= $\texttt{Simplify}\left[\frac{(s^2 - D)}{f} v^2 - 2\, s\, v\, w + f\, w^2 == 1 \;/.\; \{s \to p^2 + p - q,\, f \to p^2,\, D \to q\,(q - 2\,p),\, v \to 1,\, w \to 1\}\right]$

Out[401]= True

This solution maps to a trivial $\{u,\, v\}$ solution.

In[402]:= $\texttt{Simplify}\left[d\,\{s\, v - f\, w,\, v\} \;/.\; \{d \to 1,\, s \to p^2 + p - q,\, f \to p^2,\, v \to 1,\, w \to 1\}\right]$

Out[402]= {p - q, 1}

This solution will be found because if $\delta == 1$ the root r will be in the range $1 < r < 2$ so its first convergent will be 1, giving $v / w == 1/1$.

In[403]:=
```
Reduce[1 < (f/(s + √D) /. {f → p^2, s → p^2 + p - q, D → q (q - 2 p)}) < 2 && p ≥ 1 && q > 2 p,
  {p, q}, Integers]
```

Out[403]= $(p \mid q) \in \mathbb{Z}$ && ((p == 1 && q ≥ 3) || (p == 2 && q ≥ 5) || (p ≥ 3 && q > 2 p))

The conditions on p and q agree with p/q being any hyperbolic ratio. This trivial solution will also be found the second time, when $\delta == -1$, only for certain ratios.

In[404]:=
```
Reduce[1 < (f/(s - √D) /. {f → p^2, s → p^2 + p - q, D → q (q - 2 p)}) < 2 && p ≥ 1 && q > 2 p,
  {p, q}, Integers]
```

Out[404]= $(p \mid q) \in \mathbb{Z}$ && $p \geq 3$ && $2\,p < q < \frac{1}{4}\left(4 + 4\,p + p^2\right)$

Changing signs gives the other trivial solutions, $\{\pm(q-p), \pm 1\}$. So we see that the method developed here will generate the trivial solutions.

11.12.2 Example: $p/q == 4/11$ solved manually by method of reduction

This section can be skipped, unless you are interested in the gory details.

To see how the method of reduction works in practice, we solve a small example manually (with arithmetical assistance from Mathematica, of course). The process is rather lengthy, so we just do enough of it to obtain a few solutions.

Calculate D for this example.

In[405]:=
```
D4o11 = q (q - 2 p) /. {p → 4, q → 11}
```

Out[405]= 33

The divisors of 4 are 1, 2, and 4. Begin with divisor 1, so $f == 4^2 == 16$. We are solving

$u^2 - 33\,v^2 == 16$

Find values of s satisfying the congruence $s^2 - D \equiv 0 \pmod{f}$, $s^2 - 33 \equiv 0 \pmod{16}$. Search s between 0 and 7.

In[406]:=
```
Table[{s, Mod[s^2 - 33, 16] == 0}, {s, 0, 7}]
```

Out[406]=
```
{{0, False}, {1, True}, {2, False},
 {3, False}, {4, False}, {5, False}, {6, False}, {7, True}}
```

The values of s satisfying the congruence are 1, 7. By symmetry, $16 - s$ are also solutions, giving additionally 9 and 15. Take each in turn.

- $s == 1$

$$\frac{s^2 - D}{f} == \frac{1 - 33}{16} == -2$$

The roots *r* whose convergents will provide the solution are

$$r \to \frac{f}{s \pm \sqrt{D}} = \left\{\frac{16}{1+\sqrt{33}}, \frac{16}{1-\sqrt{33}}\right\}$$

In[407]:= **Convergents**$\left[\frac{16}{1+\sqrt{33}}\right]$

Out[407]= $\left\{2, \frac{5}{2}, \frac{7}{3}, \frac{19}{8}, \frac{1}{2}\left(-1+\sqrt{33}\right)\right\}$

In[408]:= **Convergents**$\left[\frac{16}{1-\sqrt{33}}\right]$

Out[408]= $\left\{-3, -\frac{7}{2}, -\frac{10}{3}, -\frac{27}{8}, \frac{1}{2}\left(-1-\sqrt{33}\right)\right\}$

Both have even repeat length, so we use just the first cycle.

In[409]:= **cv4o11s1 = Join**$\left[\text{Convergents}\left[\frac{16}{1+\sqrt{33}}, 4\right], \text{Convergents}\left[\frac{16}{1-\sqrt{33}}, 4\right]\right]$

Out[409]= $\left\{2, \frac{5}{2}, \frac{7}{3}, \frac{19}{8}, -3, -\frac{7}{2}, -\frac{10}{3}, -\frac{27}{8}\right\}$

See which of these provide a solution to Equation (30).

$$\frac{\left(s^2 - D\right)}{f} v^2 - 2\, s\, v\, w + f\, w^2 == 1, \text{ or}$$

$$-2\, v^2 - 2\, v\, w + 16\, w^2 == 1$$

In[410]:= **Table**$\left[\left\{cv, -2\, v^2 - 2\, v\, w + 16\, w^2 == 1\ /.\ \{v \to \text{Numerator}[cv], w \to \text{Denominator}[cv]\}\right\}, \{cv, cv4o11s1\}\right]$

Out[410]= $\left\{\{2, \text{False}\}, \left\{\frac{5}{2}, \text{False}\right\}, \left\{\frac{7}{3}, \text{False}\right\}, \left\{\frac{19}{8}, \text{False}\right\}, \{-3, \text{False}\}, \left\{-\frac{7}{2}, \text{False}\right\}, \left\{-\frac{10}{3}, \text{False}\right\}, \left\{-\frac{27}{8}, \text{False}\right\}\right\}$

None of these work. Try the next *s*.

- $s == 7$

$$\frac{s^2 - D}{f} == \frac{49 - 33}{16} == 1$$

The roots *r* whose convergents will provide the solution are

$$r \to \frac{f}{s \pm \sqrt{D}} = \left\{\frac{16}{7+\sqrt{33}}, \frac{16}{7-\sqrt{33}}\right\}$$

```
In[411]:= Convergents[16/(7 + √33)]
```

$$\left\{1, \frac{4}{3}, \frac{5}{4}, \frac{54}{43}, \frac{59}{47}, 7-\sqrt{33}\right\}$$

```
In[412]:= Convergents[16/(7 - √33)]
```

$$\left\{12, 13, \frac{38}{3}, \frac{51}{4}, 7+\sqrt{33}\right\}$$

The first has odd repeat length, the second is even. Use 2 cycles and 1 cycle respectively.

```
In[413]:= cv4o11s7 = Join[Convergents[16/(7 + √33), 10], Convergents[16/(7 - √33), 4]]
```

$$\left\{1, \frac{4}{3}, \frac{5}{4}, \frac{54}{43}, \frac{59}{47}, \frac{172}{137}, \frac{231}{184}, \frac{2482}{1977}, \frac{2713}{2161}, \frac{7908}{6299}, 12, 13, \frac{38}{3}, \frac{51}{4}\right\}$$

See which of these provide a solution to Equation (30).

$$\frac{(s^2 - D)}{f} v^2 - 2\, s\, v\, w + f\, w^2 == 1, \text{ or}$$

$$v^2 - 14\, v\, w + 16\, w^2 == 1$$

```
In[414]:= Table[{cv, v^2 - 14 v w + 16 w^2 == 1 /. {v → Numerator[cv], w → Denominator[cv]}},
  {cv, cv4o11s7}]
```

$$\left\{\{1, \text{False}\}, \left\{\frac{4}{3}, \text{False}\right\}, \left\{\frac{5}{4}, \text{True}\right\}, \left\{\frac{54}{43}, \text{False}\right\}, \left\{\frac{59}{47}, \text{False}\right\}, \left\{\frac{172}{137}, \text{False}\right\}, \left\{\frac{231}{184}, \text{True}\right\}, \left\{\frac{2482}{1977}, \text{False}\right\}, \left\{\frac{2713}{2161}, \text{False}\right\}, \left\{\frac{7908}{6299}, \text{False}\right\}, \{12, \text{False}\}, \{13, \text{False}\}, \left\{\frac{38}{3}, \text{False}\right\}, \left\{\frac{51}{4}, \text{True}\right\}\right\}$$

We have 3 solutions: $(v, w) == (5, 4), (231, 184)$, and $(51, 4)$. Convert to (u, v) using $u == s\, v - f\, w == 7\, v - 16\, w$.

```
In[415]:= Table[{7 v - 16 w, v} /. {v → vw[[1]], w → vw[[2]]},
  {vw, {{5, 4}, {51, 4}, {231, 184}}}]

Out[415]= {{-29, 5}, {293, 51}, {-1327, 231}}
```

Verify that these satisfy Equation (11).

```
In[416]:= DeleteDuplicates[Table[u^2 - 33 v^2 == 16 /. {u → uv[[1]], v → uv[[2]]}, {uv, %}]]

Out[416]= {True}
```

Let’s stop here.

11.12.3 Function to solve for (u, v) by method of reduction

We first write a function to find solutions (u, v) of Equation (11) by the method of reduction. The Pell recurrence can be used later to find additional solutions as desired. These solutions can be transformed into solutions (x, y) of Equation (2). This function does not make use of any other locally defined functions.

Outline of function:

- Make sure D is positive and nonsquare.
- Initialize list of solutions to empty set.
- For each divisor d of p, whose square is a square divisor of $f == p^2$:
  - Divide p^2 by d^2 to use as RHS f of $u^2 - D v^2 == f$, to seek relatively prime solutions u, v.
  - Make a list of s values satisfying the congruence $s^2 - D \equiv 0 \pmod f$, $0 \le s < f$. Note that if $s == 0$, (as it is if and only if $f == 1$ since D and f are relatively prime), the symmetry rule gives the other solution $f - s == f$, which is the same as 0 mod f. So handle $f == 1$ as a special case.
  - Make a list of the form $\left\{\left\{s, \frac{h}{k}, e\right\}, \ldots\right\}$ where s are the values in the list from the previous step, h/k is one of the convergents of the roots r of the companion quadratic (32) for that value of s, given by (33), and e is True or False according to whether $\{v \to h, w \to k\}$ satisfies Equation (30). Search the first cycle of convergents if the repeat length is even, two cycles if it is odd. (First calculate the continued fraction, then if repeat length is odd, concatenate repeat portion onto end of list to get second cycle. Make use of `Convergents` option to compute convergents from a list of continued fraction coefficients.) Note that in Equation (30), the factor $\left(s^2 - D\right)/f$ is assured to be integer by the choice of s. The list initially has depth 4 because there is a triple for each value of s and sign δ and each convergent. Use Flatten to remove two levels.
  - Extract the triples in the list where e is true. (A search is unnecessary if $f == 1$, since in that case, the solution is always obtained from the r-th convergent, where r is the repeat length if that is even, or twice that if odd. Also, if $f == -1$, there is a solution only if the repeat length is odd, and it is found at the end of the first repeat cycle. However, the routine does not take advantage of these economizations.)
  - Append the solutions $\{u, v\} == d\{h, k\}$ from this list to the list of u, v solutions.
- Sort and remove duplicates from the list and return it as the result. Since either sign of u or v satisfies Equation (11), we take the absolute value.

In[417]:=
```
solveuvByReduction[z_] := Module[
  {p, q, D, f, d, svalues, vwsolns, svwlist, uvlist},
  p = Numerator[z]; q = Denominator[z];
  D = q (q - 2 p);
  If[D > 0 && √D ∉ Rationals,
   uvlist = {};
   Do[ (* loop on square divisors of RHS *)
    f = (p / d)^2;
    If[f == 1,
     svalues = {0},
     svalues =
```

```
    Flatten[Position[Table[Mod[s^2 - D, f] == 0, {s, Ceiling[f / 2 - 1]}], True]];
   svalues = Join[svalues, f - svalues];
  ];
  vwsolns = Flatten[Table[
     Module[{r, convergents, contfrac, repeat, replen},
      r = f/(s + δ Sqrt[D]);
      contfrac = ContinuedFraction[r];
      (* List is of form {a0, a1, ... an,{b0, b1, ... bm}} *)
      repeat = contfrac[[-1]]; (* extract the repeat cycle at end *)
      replen = Length[repeat];
      If[Mod[replen, 2] == 0,
       convergents =
        Convergents[Flatten[contfrac][[ ;; replen]]], (* even replen *)
       convergents = Convergents[Join[Flatten[contfrac], repeat][[ ;; 2 replen]]]
       (* odd *)
      ];
      Table[{s, cv, ((s^2 - D)/f) v^2 - 2 s v w + f w^2 == 1 /.
          {v → Numerator[cv], w → Denominator[cv]}},
       {cv, convergents}]
      (* end Module *) ],
     {s, svalues}, {δ, {-1, 1}}], (* end of outer Table *)
    (* Flatten to remove two Table layers *) 2];
  svwlist = Cases[vwsolns, {_, _, True}];
  uvlist = Join[uvlist,
    d Table[{s v - f w, v} /.
       {s → svwlist[[i]][[1]],
        v → Numerator[svwlist[[i]][[2]]],
        w → Denominator[svwlist[[i]][[2]]]},
      {i, Length[svwlist]}]];
  , (* end do *)
  {d, Divisors[p]}
 ];
 DeleteDuplicates[Sort[Abs[uvlist]]]
 , (* else *) Print["D=", D, " not OK"]
 ]
]
```

Exercise the check on validity of D.

In[418]:=
```
solveuvByReduction[4 / 9] (* square *)
```
```
D=9 not OK
```

In[419]:=
```
solveuvByReduction[5 / 7] (* elliptical *)
```
```
D=-21 not OK
```

11.12.4 Examples

Example: $p / q == 7 / 18$

Try it out on the example 7/18 done earlier.

In[420]:=
```
uvsolnsbyreduction7o18 = solveuvByReduction[7 / 18]
```
Out[420]=
```
{{7, 0}, {11, 1}, {331, 39}}
```

Verify that these satisfy the equation.

In[421]:=
```
DeleteDuplicates[(uveqn /@ uvsolnsbyreduction7o18) /. {p → 7, q → 18}]
```
Out[421]=
```
{True}
```

This example has only the trivial solutions as fundamental. There are 3 solution classes.

In[422]:=
```
uvfundsolns7o18
```
Out[422]=
```
{{7, 0}, {11, -1}, {11, 1}}
```

The other solutions found by the method of reduction must be members of these classes.

In[423]:=
```
Position[Table[uvSameClass[uvbyreduction, uvfund],
  {uvbyreduction, uvsolnsbyreduction7o18}, {uvfund, uvfundsolns7o18}], True]
```
Out[423]=
```
{{1, 1}, {2, 3}, {3, 3}}
```

The first number of each pair is the index in the list of solutions found by the method of reduction. The second number is the index of the matching solution in the list of fundamental solutions. All 10 are members of one of the three classes.

Example: $p / q == 6 / 17$

Now let’s do 6/17, which has 9 classes of solutions. These were found by search earlier:

In[424]:=
```
uvsolnsbysearch6o17
```
Out[424]=
```
{{6, 0}, {11, 1}, {74, 8}, {249, 27}, {839, 91}}
```

Include the conjugate solutions to have all 9 classes represented.

In[425]:=
```
uvfundsolns6o17 = uvIncludeConjugates[uvsolnsbysearch6o17]
```
Out[425]=
```
{{6, 0}, {11, -1}, {11, 1}, {74, -8},
 {74, 8}, {249, -27}, {249, 27}, {839, -91}, {839, 91}}
```

In[426]:=
```
Timing[uvsolnsbyreduction6o17 = solveuvByReduction[6 / 17]]
```

Out[426]=
```
{0.086155, {{6, 0}, {11, 1}, {74, 8}, {249, 27}, {839, 91}, {6131, 665},
  {20 661, 2241}, {69 626, 7552}, {508 799, 55 187}, {5 778 119, 626 725}}}
```

Above, the timing for direct search was 0.0055 seconds, so this method is actually slower for this case.

Check class membership for the solutions found by reduction. In this list, the first index is the position of the solution in the reduction results, and the second index is the position in the list of fundamental solutions, which we can identify with the class number.

In[427]:=
```
Position[Table[uvSameClass[uvbyreduction, uvfund],
  {uvbyreduction, uvsolnsbyreduction6o17}, {uvfund, uvfundsolns6o17}], True]
```

Out[427]=
```
{{1, 1}, {2, 3}, {3, 5}, {4, 7}, {5, 9}, {6, 8}, {7, 6}, {8, 4}, {9, 2}, {10, 3}}
```

The 2nd and 9th solutions are in the same class, leaving one class missing, class 8. But not to worry, we find it among the conjugate solutions. This is not the method's fault: it is implemented to return only positive values, so members of conjugate classes can be missed. Class 8 is (839, −91} and the method of reduction yielded its conjugate {839, 91} as its 5th solution. Augment the list of solutions by search with their conjugates.

In[428]:=
```
uvallsolnsbyreduction6o17 = uvIncludeConjugates[uvsolnsbyreduction6o17]
```

Out[428]=
```
{{6, 0}, {11, -1}, {11, 1}, {74, -8}, {74, 8}, {249, -27}, {249, 27},
 {839, -91}, {839, 91}, {6131, -665}, {6131, 665}, {20 661, -2241},
 {20 661, 2241}, {69 626, -7552}, {69 626, 7552}, {508 799, -55 187},
 {508 799, 55 187}, {5 778 119, -626 725}, {5 778 119, 626 725}}
```

See which classes these correspond to.

In[429]:=
```
Position[Table[uvSameClass[uvbyreduction, uvfund],
  {uvbyreduction, uvallsolnsbyreduction6o17}, {uvfund, uvfundsolns6o17}], True]
```

Out[429]=
```
{{1, 1}, {2, 2}, {3, 3}, {4, 4}, {5, 5}, {6, 6}, {7, 7}, {8, 8}, {9, 9}, {10, 9},
 {11, 8}, {12, 7}, {13, 6}, {14, 5}, {15, 4}, {16, 3}, {17, 2}, {18, 2}, {19, 3}}
```

All of the solutions by reduction, now that conjugates are included, match a fundamental solution class, and all 9 classes are represented.

Example: $p / q == 210 / 421$

Finally, to show that this method can solve a problem that is out of reach of direct search, we do 210/421. The estimated time to search for all fundamental solutions calculated above was over a million years.

In[430]:=
```
Timing[uvsolnsbyreduction210o421 = solveuvByReduction[210 / 421]]
```

Out[430]=
```
{0.972827, {{210, 0}, {211, 1}, {631, 29}, {2737, 133}, {8210, 400},
  {9051, 441}, {52 835, 2575}, {213 657, 10 413}, {235 550, 11 480},
  {707 491, 34 481}, {1 375 197, 67 023}, {4 553 746, 221 936}, {18 414 750, 897 480},
  {20 301 673, 989 443}, {107 509 717, 5 239 703}, {118 526 025, 5 776 605},
```

```
{479 303 659, 23 359 831}, {1 587 137 793, 77 352 363}, {3 085 024 639, 150 354 901},
{9 266 090 225, 451 601 605}, {10 215 568 158, 497 876 328}, {41 310 414 290, 2 013 346 400},
{241 180 083 879, 11 754 398 061}, {265 893 329 915, 12 958 847 975},
{798 629 467 678, 38 922 818 648}, {3 560 478 408 139, 173 527 099 849},
{18 854 900 581 519, 918 931 626 821}, {20 786 924 940 915, 1 013 092 732 575},
{22 916 920 014 034, 1 116 902 340 544}, {121 359 041 951 074, 5 914 677 797 816},
{541 047 213 992 173, 26 369 027 743 807}, {1 625 073 666 335 915, 79 201 244 337 175},
{1 791 591 749 822 994, 87 316 839 149 904}, {10 459 741 349 140 415, 509 776 601 183 125},
{42 297 818 566 346 853, 2 061 469 539 873 927},
{46 631 992 351 362 425, 2 272 704 245 141 645}, {140 062 495 137 574 354,
 6 826 228 330 237 664}, {272 248 730 012 590 638, 13 268 591 223 208 992},
{901 508 345 253 952 699, 43 936 828 344 187 559}, {3 645 581 200 412 088 225,
 177 674 756 157 894 645}, {4 019 136 693 473 448 862, 195 880 736 903 465 992},
{21 283 775 406 627 444 718, 1 037 307 742 607 533 712}, {23 464 681 763 981 726 325,
 1 143 598 614 746 747 895}, {94 888 087 462 270 858 711, 4 624 562 415 518 882 749},
{314 206 801 656 955 658 142, 15 313 502 511 270 644 472},
{610 744 529 635 869 713 506, 29 765 867 062 721 040 904},
{1 834 414 495 264 963 422 125, 89 403 892 060 302 336 895},
{2 022 383 319 325 212 218 577, 98 564 931 999 946 459 557},
{8 178 252 201 184 118 248 835, 398 583 623 779 632 649 775},
{47 746 593 341 298 636 362 466, 2 327 026 573 520 417 291 544},
{52 639 092 297 457 656 915 335, 2 565 472 382 638 603 261 325},
{158 105 245 716 433 219 542 457, 7 705 578 187 855 453 906 637},
{704 870 451 654 608 722 190 866, 34 353 283 807 388 815 527 856},
{3 732 718 125 299 524 173 485 026, 181 921 550 024 390 859 564 584},
{4 115 202 366 659 372 154 676 335, 200 562 691 335 457 121 626 725},
{4 536 879 011 511 214 156 388 371, 221 113 953 516 384 880 085 311},
{24 025 536 151 793 618 226 387 871, 1 170 932 985 075 215 410 628 939},
{107 111 503 111 871 200 945 687 342, 5 220 295 242 623 060 183 163 328},
{321 716 993 576 973 450 818 253 335, 15 679 526 869 180 246 811 552 125},
{354 682 697 412 696 581 096 737 611, 17 286 176 966 542 883 466 932 151},
{2 070 722 460 247 302 890 094 362 210, 100 920 837 575 508 382 473 521 800},
{8 373 729 330 505 647 117 184 750 062, 408 110 596 127 065 735 628 789 808},
{9 231 768 808 124 696 268 693 822 725, 449 928 881 491 983 415 880 282 855},
{27 728 272 128 209 811 936 359 952 451, 1 351 393 294 573 312 884 296 228 591},
{53 897 275 390 772 899 381 086 550 077, 2 626 792 474 557 417 748 631 330 193},
{178 472 250 537 165 715 507 907 634 631, 8 698 205 266 023 747 353 806 950 221},
{721 718 312 180 732 207 620 075 796 925, 35 174 398 287 138 392 151 688 815 855},
{795 671 359 757 236 947 160 461 831 553, 38 778 649 289 366 347 388 697 218 773},
{4 213 564 207 970 069 775 182 078 955 397, 205 356 554 154 407 426 041 418 290 073},
{4 645 319 796 098 524 994 889 789 331 950, 226 399 033 974 047 793 666 292 086 480},
{18 785 062 398 716 209 733 360 282 487 394, 915 527 922 056 834 109 203 058 904 696},
{62 203 744 780 648 610 649 733 348 944 873, 3 031 625 021 755 202 337 739 947 136 893},
{120 909 530 434 431 052 484 619 267 618 199, 5 892 769 946 990 822 074 214 206 770 691},
```

```
{363 160 346 891 421 612 673 565 513 231 150, 17 699 352 320 792 106 590 267 494 108 480},
{400 372 669 148 302 081 660 118 753 864 013, 19 512 969 936 088 362 396 285 746 524 383},
{1 619 054 425 275 115 675 454 252 552 848 415, 78 907 884 477 953 590 546 708 432 015 925},
{9 452 427 161 618 730 372 363 793 368 237 039, 460 683 111 612 224 143 919 744 498 260 251},
{10 420 998 671 858 564 157 899 870 640 312 290, 507 888 398 627 609 641 616 332 773 815 000},
{31 300 208 337 832 572 942 686 165 161 569 733, 1 525 478 813 498 125 180 655 016 573 860 903},
{139 543 706 396 316 255 014 423 209 773 045 979,
 6 800 944 114 077 542 671 017 063 780 868 039},
{738 968 871 392 325 343 777 103 414 572 630 279,
 36 015 139 099 923 106 526 317 236 561 959 411},
{898 169 175 983 773 799 358 484 229 927 301 199,
 43 774 087 191 752 960 285 253 670 896 082 309}}}
```

This took less than a second on my laptop. We observe that many of these are larger than the Nagell bound. We can presume that many of the large solutions are later-generation solutions that would be given by the Pell recurrence from the fundamental solutions. This is verified below.

11.12.5 Function to reduce set of solutions to representatives of distinct classes

The examples of the previous section show that the method of reduction yields many solutions that are not fundamental. In fact, by construction, it only yields positive solutions, and by our convention fundamental solutions come in pairs with opposite signs of v, so the conjugates of the fundamental solutions won't occur. Besides this there are solutions that are in the second or later Pell generations. Let's write a function to accept a list of solutions such as given by the method of reduction and return a list including just one member of each class drawn from the given solutions and their conjugates. If the original list contains at least all the positive fundamental solutions, the result will be all and only the fundamental solutions, including conjugates.

Definition of fundamental solution used here is the member of a class for which $u \geq 0$ is least, and conjugates are obtained by changing the sign of v.

The function is complicated somewhat by the desire to avoid needing to provide D as an argument. D can be found from a pair of non-conjugate solutions. For the set of solutions returned by `solveuv-ByReduction`, the first two solutions are always non-conjugate, but in order to allow this function to work on any valid input list, we search for a non-conjugate pair of solutions to guarantee we can calculate D.

Outline of the function:

- Ensure $u \geq 0$ for all solutions and augment list with conjugates by negating v.
- Sort into ascending order so that the fundamental solutions are first. Delete duplicates such as any with $v == 0$ or if the list already contained conjugates.
- Find a pair of solutions that do not have the same $|v|$ and use them to calculate D and f. This function does not check that the remaining solutions have the same f for that D.
- Initialize list `uvclass` of distinct class members to empty.

- For each {*u*, *v*} in the list,
  - If it is not in the same class as any element of `uvclass`, append it to `uvclass`.
- Return the list of distinct class members. If the given list included only a pair of conjugate solutions, return the given list. If the given list included only one solution, return that with its conjugate.

In[431]:=
```
uvGetClasses[uvlist_] := Module[{Df, D, f, uvall, uvclass},
  uvall = Sort[DeleteDuplicates[Join[
      Table[{Abs[uv[[1]]], uv[[2]]}, {uv, uvlist}],
      Table[{Abs[uv[[1]]], -uv[[2]]}, {uv, uvlist}]]]];
  Module[{i, u, v, U, V}, (* finds a pair of givens that yield D,f *)
   Df = {};
   {u, v} = uvall[[1]];
   i = 1;
   While[Df == {} && i < Length[uvall],
    i = i + 1; {U, V} = uvall[[i]];
    If[v^2 ≠ V^2,
     Df = {(u^2 - U^2)/(v^2 - V^2), (u^2 - D v^2)};
    ]
   ]
  ];
  If[Df ≠ {},
   {D, f} = Df;
   uvclass = {};
   Do[
    If[Position[Table[(* test each given in
          uvall against growing set of class representatives *)
         ((u U - D v V) / f ∈ Integers && (v U - u V) / f ∈ Integers ) /.
          {u → uv[[1]], v → uv[[2]], U → UV[[1]], V → UV[[2]]}, {UV, uvclass}]
       , True] == {},
     uvclass = Append[uvclass, uv]
    ]
    , {uv, uvall}]; (* end Do *)
   (* return *) uvclass
   , (* else only given one solution
    or a pair of conjugate solutions, can't find D,f *)
   (* return *) uvall
  ]
 ]
```

Test the section that deals with a given set that cannot determine *D*. The function still returns the list of class members.

```
In[432]:= uvGetClasses[{{11, 1}}]
Out[432]= {{11, -1}, {11, 1}}
```

```
In[433]:= uvGetClasses[{{11, -1}, {11, 1}}]
Out[433]= {{11, -1}, {11, 1}}
```

```
In[434]:= uvGetClasses[{{7, 0}}]
Out[434]= {{7, 0}}
```

```
In[435]:= uvGetClasses[{{7, 0}, {11, 1}}]
Out[435]= {{7, 0}, {11, -1}, {11, 1}}
```

11.12.6 Examples

Example: $p / q == 7 / 18$

This case has only 3 classes, the trivial solutions, but the reduction method generates several more. Here is the solution set obtained by the reduction method earlier:

```
In[436]:= uvsolnsbyreduction7o18
Out[436]= {{7, 0}, {11, 1}, {331, 39}}
```

Reduce this to representatives of each class.

```
In[437]:= uvclasssolns7o18 = uvGetClasses[uvsolnsbyreduction7o18]
Out[437]= {{7, 0}, {11, -1}, {11, 1}}
```

This matches the set of fundamental solutions found earlier.

```
In[438]:= uvfundsolns7o18
Out[438]= {{7, 0}, {11, -1}, {11, 1}}
```

```
In[439]:= {h7o18, k7o18} = solvePell[D7o18]
Out[439]= {17, 2}
```

If we run the Pell recurrence on these solutions we will see the other solutions appear. The largest one is in the 5th generation.

In[440]:=
```
Sort[Flatten[Table[RecurrenceTable[{
      u[i + 1] == h u[i] + D k v[i],
      v[i + 1] == k u[i] + h v[i],
      u[1] == uv[[1]],
      v[1] == uv[[2]]} /. {D → D7o18, h → h7o18, k → k7o18},
    {u, v}, {i, 5}],
   {uv, uvfundsolns7o18}], 1]]
```

Out[440]=
```
{{7, 0}, {11, -1}, {11, 1}, {43, 5}, {119, 14}, {331, 39}, {1451, 171},
 {4039, 476}, {11 243, 1325}, {49 291, 5809}, {137 207, 16 170}, {381 931, 45 011},
 {1 674 443, 197 335}, {4 660 999, 549 304}, {12 974 411, 1 529 049}}
```

Example: $p / q == 6 / 17$

This example, worked earlier, has 9 classes.

In[441]:=
```
uvsolnsbyreduction6o17
```

Out[441]=
```
{{6, 0}, {11, 1}, {74, 8}, {249, 27}, {839, 91}, {6131, 665},
 {20 661, 2241}, {69 626, 7552}, {508 799, 55 187}, {5 778 119, 626 725}}
```

In[442]:=
```
uvclasssolns6o17 = uvGetClasses[uvsolnsbyreduction6o17]
```

Out[442]=
```
{{6, 0}, {11, -1}, {11, 1}, {74, -8},
 {74, 8}, {249, -27}, {249, 27}, {839, -91}, {839, 91}}
```

This agrees with the list of fundamental solutions found earlier, supporting the claim that the method of reduction is complete.

In[443]:=
```
uvfundsolns6o17
```

Out[443]=
```
{{6, 0}, {11, -1}, {11, 1}, {74, -8},
 {74, 8}, {249, -27}, {249, 27}, {839, -91}, {839, 91}}
```

In[444]:=
```
uvclasssolns6o17 == uvfundsolns6o17
```

Out[444]=
```
True
```

This one only requires 2 Pell generations to obtain all the solutions produced by the reduction method.

In[445]:=
```
Sort[Flatten[Table[RecurrenceTable[{
      u[i + 1] == h u[i] + D k v[i],
      v[i + 1] == k u[i] + h v[i],
      u[1] == uv[[1]],
      v[1] == uv[[2]]} /. {D → D6o17, h → h6o17, k → k6o17},
    {u, v}, {i, 2}],
   {uv, uvclasssolns6o17}], 1]]
```

Out[445]=
```
{{6, 0}, {11, -1}, {11, 1}, {74, -8}, {74, 8}, {249, -27},
 {249, 27}, {839, -91}, {839, 91}, {6131, 665}, {20 661, 2241},
 {69 626, 7552}, {508 799, 55 187}, {1 714 614, 185 976}, {5 778 119, 626 725},
 {42 224 186, 4 579 856}, {142 292 301, 15 433 767}, {479 514 251, 52 010 623}}
```

It is interesting to see how the solutions map from one generation to the next. Get the *u*, *v* solutions grouped by generation rather than sorted by size.

In[446]:=
```
uv2gens6o17 = Table[RecurrenceTable[{
      u[i + 1] == h u[i] + D k v[i],
      v[i + 1] == k u[i] + h v[i],
      u[1] == uv[[1]],
      v[1] == uv[[2]]} /. {D → D6o17, h → h6o17, k → k6o17},
    {u, v}, {i, 2}],
   {uv, uvclasssolns6o17}];
```

In[447]:=
```
TableForm[uv2gens6o17, TableDepth → 1]
```

Out[447]//TableForm=
```
{{6, 0}, {1 714 614, 185 976}}
{{11, -1}, {508 799, 55 187}}
{{11, 1}, {5 778 119, 626 725}}
{{74, -8}, {69 626, 7552}}
{{74, 8}, {42 224 186, 4 579 856}}
{{249, -27}, {20 661, 2241}}
{{249, 27}, {142 292 301, 15 433 767}}
{{839, -91}, {6131, 665}}
{{839, 91}, {479 514 251, 52 010 623}}
```

Each row has 2 generations of one class. Now convert these to *x*, *y*.

In[448]:=
```
TableForm[Table[xyfromuv /@ uv /. {p → 6, q → 17}, {uv, uv2gens6o17}], TableDepth → 1]
```

Out[448]//TableForm=

$\left\{\{0, 0\}, \left\{\frac{392364}{5}, \frac{1322244}{5}\right\}\right\}$

$\left\{\{1, 0\}, \left\{\frac{116429}{5}, \frac{392364}{5}\right\}\right\}$

$\left\{\{0, 1\}, \left\{\frac{1322244}{5}, \frac{4455869}{5}\right\}\right\}$

$\left\{\left\{\frac{54}{5}, \frac{14}{5}\right\}, \{3186, 10738\}\right\}$

$\left\{\left\{\frac{14}{5}, \frac{54}{5}\right\}, \{1932490, 6512346\}\right\}$

$\left\{\left\{\frac{189}{5}, \frac{54}{5}\right\}, \{945, 3186\}\right\}$

$\left\{\left\{\frac{54}{5}, \frac{189}{5}\right\}, \{6512346, 21946113\}\right\}$

$\left\{\left\{\frac{644}{5}, \frac{189}{5}\right\}, \{280, 945\}\right\}$

$\left\{\left\{\frac{189}{5}, \frac{644}{5}\right\}, \{21946113, 73956736\}\right\}$

Each row has 2 generations of one class. Classes 1-3 (the trivial solution classes) are integer in the 1st generation but fractional in the second. Classes 4-9 are fractional in the 1st generation but integer in the second. Classes 1-3 form recycling triplets as always. Classes 4, 6, and 8 form another triplet, and classes 5, 7, and 9 a third triplet.

Example: $p / q == 8 / 19$

This example has two classes of solutions that yield no admissible solutions.

In[449]:=
```
D8o19 = q (q - 2 p) /. {p → 8, q → 19}
```

Out[449]= 57

In[450]:=
```
{h8o19, k8o19} = solvePell[D8o19]
```

Out[450]= {151, 20}

In[451]:=
```
uvclasssolns8o19 = uvGetClasses[solveuvByReduction[8 / 19]]
```

Out[451]= {{8, 0}, {11, -1}, {11, 1}, {46, -6}, {46, 6}}

There are 5 classes. Obtain the first 4 Pell generations.

In[452]:=
```
uv4gens8o19 = Table[RecurrenceTable[{
      u[i + 1] == h u[i] + D k v[i],
      v[i + 1] == k u[i] + h v[i],
      u[1] == uv[[1]],
      v[1] == uv[[2]]} /. {D → D8o19, h → h8o19, k → k8o19},
    {u, v}, {i, 4}],
  {uv, uvclasssolns8o19}]; TableForm[uv4gens8o19, TableDepth → 1]
```

Out[452]//TableForm=
```
{{8, 0}, {1208, 160}, {364 808, 48 320}, {110 170 808, 14 592 480}}
{{11, -1}, {521, 69}, {157 331, 20 839}, {47 513 441, 6 293 309}}
{{11, 1}, {2801, 371}, {845 891, 112 041}, {255 456 281, 33 836 011}}
{{46, -6}, {106, 14}, {31 966, 4234}, {9 653 626, 1 278 654}}
{{46, 6}, {13 786, 1826}, {4 163 326, 551 446}, {1 257 310 666, 166 534 866}}
```

Each row has 4 generations of one class. Convert to (x, y).

In[453]:=
```
TableForm[Table[xyfromuv /@ uv /. {p → 8, q → 19}, {uv, uv4gens8o19}], TableDepth → 1]
```

Out[453]//TableForm=
```
{{0, 0}, {120, 280}, {36 640, 84 960}, {11 065 560, 25 658 040}}
{{1, 0}, {51, 120}, {15 801, 36 640}, {4 772 251, 11 065 560}}
{{0, 1}, {280, 651}, {84 960, 197 001}, {25 658 040, 59 494 051}}
```

$$\left\{\left\{\frac{28}{3}, \frac{10}{3}\right\}, \left\{\frac{28}{3}, \frac{70}{3}\right\}, \left\{\frac{9628}{3}, \frac{22\,330}{3}\right\}, \left\{\frac{2\,908\,828}{3}, \frac{6\,744\,790}{3}\right\}\right\}$$

$$\left\{\left\{\frac{10}{3}, \frac{28}{3}\right\}, \left\{\frac{4150}{3}, \frac{9628}{3}\right\}, \left\{\frac{1\,254\,490}{3}, \frac{2\,908\,828}{3}\right\}, \left\{\frac{378\,853\,030}{3}, \frac{878\,457\,628}{3}\right\}\right\}$$

No integer solutions appear for the 4th and 5th classes. So although this example has 5 classes, only 3 of them give rise to admissible solutions of Equation (2). The proof in Section 12.2 only shows the existence of integer solutions if the starting solutions in the Pell recurrence are integer, whereas here the fundamental solutions are fractional. The example 6/17 above shows that fractional solutions can be followed by integer solutions in the next generation, but that does not happen here. The proof in Section 12.2 implies that if two successive generations are fractional, all are fractional, so the 4th and 5th classes yield no admissible solutions.

Example: p / q == 9 / 22

The reverse search found a recycling doublet for the smallest admissible (x, y) solutions, namely (3, 9) and (9, 24). It is not part of a triplet.

In[454]:=
```
recycleSolutions[{{3, 9}, {9, 24}}]
```

Out[454]//TableForm=

x	y
3	9
9	24

Let’s work it out using classes and Pell generations.

In[455]:=
```
D9o22 = q (q - 2 p) /. {p → 9, q → 22}
```

Out[455]= 88

In[456]:=
```
{h9o22, k9o22} = solvePell[D9o22]
```
Out[456]=
```
{197, 21}
```

In[457]:=
```
uvclasssolns9o22 = uvGetClasses[solveuvByReduction[9 / 22]]
```
Out[457]=
```
{{9, 0}, {13, -1}, {13, 1}, {57, -6}, {57, 6}}
```

There are 5 classes. Obtain the first 3 Pell generations.

In[458]:=
```
uv3gens9o22 = Table[RecurrenceTable[{
      u[i + 1] == h u[i] + D k v[i],
      v[i + 1] == k u[i] + h v[i],
      u[1] == uv[[1]],
      v[1] == uv[[2]]} /. {D → D9o22, h → h9o22, k → k9o22},
    {u, v}, {i, 3}],
   {uv, uvclasssolns9o22}]; TableForm[uv3gens9o22, TableDepth → 1]
```
Out[458]//TableForm=
```
{{9, 0}, {1773, 189}, {698553, 74466}}
{{13, -1}, {713, 76}, {280909, 29945}}
{{13, 1}, {4409, 470}, {1737133, 185179}}
{{57, -6}, {141, 15}, {55497, 5916}}
{{57, 6}, {22317, 2379}, {8792841, 937320}}
```

Each row has 3 generations of one class. Convert to (x, y).

In[459]:=
```
TableForm[Table[xyfromuv /@ uv /. {p → 9, q → 22}, {uv, uv3gens9o22}], TableDepth → 1]
```
Out[459]//TableForm=
```
{{0, 0}, {126, 315}, {50085, 124551}}
{{1, 0}, {50, 126}, {20140, 50085}}
{{0, 1}, {315, 785}, {124551, 309730}}
{{9, 3}, {9, 24}, {3978, 9894}}
{{3, 9}, {1599, 3978}, {630444, 1567764}}
```

All solutions are integer. As always, the trivial solutions give rise to a triplet in each generation. But interestingly, the doublets cross generations: the first doublet (with $x < y$) is generation 1 of class 5 with generation 2 of class 4.

Example: $p / q == 210 / 421$

Now apply this to an example that has a large number of classes.

In[460]:=
```
uvclasssolns210o421 = uvGetClasses[uvsolnsbyreduction210o421]
```

Out[460]=
```
{{210, 0}, {211, -1}, {211, 1}, {631, -29}, {631, 29}, {2737, -133},
 {2737, 133}, {8210, -400}, {8210, 400}, {9051, -441}, {9051, 441},
 {52 835, -2575}, {52 835, 2575}, {213 657, -10 413}, {213 657, 10 413},
 {235 550, -11 480}, {235 550, 11 480}, {707 491, -34 481}, {707 491, 34 481},
 {1 375 197, -67 023}, {1 375 197, 67 023}, {4 553 746, -221 936}, {4 553 746, 221 936},
 {18 414 750, -897 480}, {18 414 750, 897 480}, {20 301 673, -989 443},
 {20 301 673, 989 443}, {107 509 717, -5 239 703}, {107 509 717, 5 239 703},
 {118 526 025, -5 776 605}, {118 526 025, 5 776 605}, {479 303 659, -23 359 831},
 {479 303 659, 23 359 831}, {1 587 137 793, -77 352 363}, {1 587 137 793, 77 352 363},
 {3 085 024 639, -150 354 901}, {3 085 024 639, 150 354 901}, {9 266 090 225, -451 601 605},
 {9 266 090 225, 451 601 605}, {10 215 568 158, -497 876 328}, {10 215 568 158, 497 876 328},
 {41 310 414 290, -2 013 346 400}, {41 310 414 290, 2 013 346 400},
 {241 180 083 879, -11 754 398 061}, {241 180 083 879, 11 754 398 061},
 {265 893 329 915, -12 958 847 975}, {265 893 329 915, 12 958 847 975},
 {798 629 467 678, -38 922 818 648}, {798 629 467 678, 38 922 818 648},
 {3 560 478 408 139, -173 527 099 849}, {3 560 478 408 139, 173 527 099 849},
 {18 854 900 581 519, -918 931 626 821}, {18 854 900 581 519, 918 931 626 821},
 {20 786 924 940 915, -1 013 092 732 575}, {20 786 924 940 915, 1 013 092 732 575},
 {22 916 920 014 034, -1 116 902 340 544}, {22 916 920 014 034, 1 116 902 340 544},
 {121 359 041 951 074, -5 914 677 797 816}, {121 359 041 951 074, 5 914 677 797 816},
 {541 047 213 992 173, -26 369 027 743 807}, {541 047 213 992 173, 26 369 027 743 807},
 {1 625 073 666 335 915, -79 201 244 337 175}, {1 625 073 666 335 915, 79 201 244 337 175},
 {1 791 591 749 822 994, -87 316 839 149 904}, {1 791 591 749 822 994, 87 316 839 149 904},
 {10 459 741 349 140 415, -509 776 601 183 125}, {10 459 741 349 140 415, 509 776 601 183 125},
 {42 297 818 566 346 853, -2 061 469 539 873 927},
 {42 297 818 566 346 853, 2 061 469 539 873 927},
 {46 631 992 351 362 425, -2 272 704 245 141 645},
 {46 631 992 351 362 425, 2 272 704 245 141 645},
 {140 062 495 137 574 354, -6 826 228 330 237 664},
 {140 062 495 137 574 354, 6 826 228 330 237 664},
 {272 248 730 012 590 638, -13 268 591 223 208 992},
 {272 248 730 012 590 638, 13 268 591 223 208 992},
 {901 508 345 253 952 699, -43 936 828 344 187 559},
 {901 508 345 253 952 699, 43 936 828 344 187 559},
 {3 645 581 200 412 088 225, -177 674 756 157 894 645},
 {3 645 581 200 412 088 225, 177 674 756 157 894 645},
 {4 019 136 693 473 448 862, -195 880 736 903 465 992},
 {4 019 136 693 473 448 862, 195 880 736 903 465 992}}
```

In[461]:=
```
Length[uvclasssolns210o421]
```

Out[461]=
```
81
```

Assuming completeness, this is the number of classes for this example. Verify that all are within the

Nagell bound.

In[462]:=
```
vboundNagell210o421 = vboundNagell[210 / 421]
```

Out[462]= 450 764 467 341 464 849

We can see that this bound is larger than the v in the last solution, but let’s verify.

In[463]:=
```
DeleteDuplicates[
 Table[Abs[uv[[2]]] ≤ vboundNagell210o421, {uv, uvclasssolns210o421}]]
```

Out[463]= {True}

In fact the last few solutions are a significant fraction of the Nagell bound. They are out of reach of a direct search.

In[464]:=
```
DeleteDuplicates[
 Table[N[Abs[uv[[2]]] / vboundNagell210o421], {uv, uvclasssolns210o421}]]
```

Out[464]= $\{0., 2.21845\times10^{-18}, 6.43352\times10^{-17}, 2.95054\times10^{-16}, 8.87381\times10^{-16}, 9.78338\times10^{-16}, 5.71252\times10^{-15}, 2.31008\times10^{-14}, 2.54678\times10^{-14}, 7.64945\times10^{-14}, 1.48687\times10^{-13}, 4.92355\times10^{-13}, 1.99102\times10^{-12}, 2.19503\times10^{-12}, 1.1624\times10^{-11}, 1.28151\times10^{-11}, 5.18227\times10^{-11}, 1.71603\times10^{-10}, 3.33555\times10^{-10}, 1.00186\times10^{-9}, 1.10452\times10^{-9}, 4.46652\times10^{-9}, 2.60766\times10^{-8}, 2.87486\times10^{-8}, 8.63485\times10^{-8}, 3.84962\times10^{-7}, 2.03861\times10^{-6}, 2.2475\times10^{-6}, 2.4778\times10^{-6}, 0.0000131214, 0.0000584985, 0.000175704, 0.000193708, 0.00113092, 0.00457327, 0.00504189, 0.0151437, 0.0294358, 0.0974718, 0.394163, 0.434552\}$

The fact that these range up to close to the Nagell bound suggests that all classes have been found. Running the Pell recurrence on the set of class representatives would generate the other solutions found by the reduction method, but we will skip that for reasons of space.

11.12.7 Function to solve hyperbolic case for (x, y) by reduction method

The final task is to convert the solutions (u, v) found by the reduction method to (x, y). To avoid the arbitrariness of the solution set found by the method of reduction, which includes various Pell generations, we will convert that set into the set of class representatives, and then apply any desired number of Pell recurrence steps to generate additional solutions. These are then converted to (x, y) and sifted to obtain admissible solutions. This function uses functions `solvePell`, `solveuvByReduction`, and `uvGetClasses`.

The function has a mandatory argument $z == p/q$ ($0 < z < 1/2$ and giving nonsquare D), an optional `iters` which is the number of Pell generations to use (default 3), and optional `tableform` to specify whether to print the results in tabular form vs. a list.

Outline of the function:

- Obtain $\{h, k\}$ the base solution of the Pell equation for this D.
- Obtain a set of solutions $\{u, v\}$ of Equation (11) using `solveuvByReduction`.

- Convert the set of solutions into a minimal set of representatives of each class using `uvGetClasses`. (Assuming the method of reduction is complete, these are all the classes for the instance. The solution method always finds the fundamental solutions as well as others so the result is the set of fundamental solutions.)
- Run the Pell recurrence to produce additional generations as specified by `iters`.
- Convert the set of $\{u, v\}$ solutions to $\{x, y\}$.
- Sift the $\{x, y\}$ solution set to retain only admissible solutions, i.e. integer and positive. Sort so that $x \le y$ and then sort solutions into ascending order, deleting duplicates. Duplicates always appear in the first generation because the conjugate pairs yield $\{x, y\}$ pairs that differ only in order.

In[465]:=
```
solveHyperbolicByReduction[z_, iters_: 3, tableform_: True] :=
 Module[{p, q, D, hkpell, h, k, uvsolns, xysolns},
  p = Numerator[z];
  q = Denominator[z];
  D = q (q - 2 p);
  hkpell = solvePell[D]; (* get the base solution *)
  If[Length[hkpell] == 2,
   {h, k} = hkpell;
   uvsolns = uvGetClasses[Evaluate[solveuvByReduction[z]]];
   uvsolns = Flatten[Table[RecurrenceTable[{
         u[i + 1] == h u[i] + D k v[i],
         v[i + 1] == k u[i] + h v[i],
         u[1] == uv[[1]],
         v[1] == uv[[2]]},
        {u, v}, {i, iters}],
      {uv, uvsolns}]
     , 1]; (* Flatten to produce list of {u,v} pairs *)
   xysolns = Table[1/2 {((p - u)/(2 p - q) - v), ((p - u)/(2 p - q) + v)} /. {u → uv[[1]], v → uv[[2]]},
     {uv, uvsolns}];
   xysolns = DeleteDuplicates[
     Sort[Cases[Sort /@ xysolns, {_Integer?Positive, _Integer?Positive}]]];
   If[tableform,
    TableForm[xysolns, TableHeadings → {None, {"x", "y"}}],
    xysolns]
   ]
  ]
```

Exercise the test for valid z.

Elliptical ratio.

In[466]:=
```
solveHyperbolicByReduction[4 / 7]
```
```
D=-7 not OK
```

Square *D*.

In[467]:=
```
solveHyperbolicByReduction[4 / 9]
```
```
D=9 not OK
```

11.12.8 Examples

Example: $p / q == 7 / 18$

In[468]:=
```
solveHyperbolicByReduction[7 / 18]
```
Out[468]//TableForm=

x	y
2	7
7	21
21	60
95	266
266	742
742	2067

There are 3 classes, and the solver ran 3 generations of the Pell recurrence. Solutions are integer in each generation. However, the first generation is just the 3 trivial solutions, removed by the sift at the end, resulting in 6 solutions.

Example: $p / q == 4 / 11$

This example has 3 classes, but only every other generation yields admissible solutions. So using 3 generations we get just one triplet, the third generation.

In[469]:=
```
solveHyperbolicByReduction[4 / 11]
```
Out[469]//TableForm=

x	y
105	336
336	1072
1072	3417

Example: $p / q == 3 / 43$

An extreme case of a ratio involving small p and q but whose smallest solution is very large. I found this example by looking for the ratio with $q \leq 50$ having the largest bound on v but only the trivial solution classes, so that the Pell recurrence would generate a large triplet as the smallest admissible solutions. It turns out the second generation are also not admissible, so the solutions finally appear in the third generation and are even bigger.

In[470]:=
```
solveHyperbolicByReduction[3 / 43]
```

Out[470]//TableForm=

x	y
29 356 663 716 856 580 847 741 762 556	781 741 938 418 732 926 066 926 219 805
781 741 938 418 732 926 066 926 219 805	20 817 095 027 449 354 780 936 957 432 245
20 817 095 027 449 354 780 936 957 432 245	554 340 792 126 897 394 565 585 271 973 396

Use scientific notation to see how big these are.

In[471]:=
```
% // N
```

Out[471]= $\{\{2.93567\times 10^{28}, 7.81742\times 10^{29}\}, \{7.81742\times 10^{29}, 2.08171\times 10^{31}\}, \{2.08171\times 10^{31}, 5.54341\times 10^{32}\}\}$

Example: $p/q == 6/17$

In[472]:=
```
solveHyperbolicByReduction[6 / 17]
```

Out[472]//TableForm=

x	y
280	945
945	3186
3186	10 738
1 932 490	6 512 346
6 512 346	21 946 113
21 946 113	73 956 736
13 309 062 481	44 850 530 088
44 850 530 088	151 142 881 176
151 142 881 176	509 340 034 225

This example has 9 classes. If one views the x, y values before sifting one finds that the first generation consists of the 3 trivial solutions, and fractional solutions for the other classes, so no admissible solutions. The second generation yields fractional values from the trivial solutions, and integer solutions for the rest, giving 6 admissible solutions, all distinct. The third generation gives admissible values from the trivial solutions, but fractional values from the others, giving the last 3 admissible solutions. The pattern continues to further generations, adding 6 solutions in the even-numbered generations and 3 solutions in the odd-numbered ones. Observe that there are three triplets of solutions related by the recycling recurrence. In the table, the first two triplets come from the six non-trivial classes in the second generation, and the last triplet from the trivial solutions in the third generation.

Example: $p/q == 210/421$

Now let’s see the results for the challenging 210/421 case. It generates a lot of solutions, so let’s proceed cautiously. First generate the list for 3 Pell cycles, but don’t print.

In[473]:=
```
xysolnsbyreduction210o421 = solveHyperbolicByReduction[210 / 421, 3, False];
```

In[474]:=
```
Length[xysolnsbyreduction210o421]
```

Out[474]= 201

Recall that there are 81 classes for this example. The number of solutions we obtained is

In[475]:= $\frac{(81-3)}{2} + 2 \times 81$

Out[475]= 201

This number results as follows. The first 3 classes are the trivial solutions, which are suppressed. The rest of the first generation yield integer solutions, but since negative *v* values yield the same solutions as positive except with *x* and *y* swapped, these non-distinct solutions are suppressed. So the total of distinct admissible solutions for the first generation is (81 – 3)/2 == 39. Then each of the next generations yields a complete set of 81 distinct admissible solutions. (In the second and later generations, the initial solutions with negative *v* values give rise to solutions with positive *v*.)

Make a table of just the first 50 solutions, representing 39 from the the first generation and 11 from the second.

In[476]:=
```
TableForm[xysolnsbyreduction210o421[[1 ;; 50]], TableHeadings → {None, {"x", "y"}}]
```

Out[476]//TableForm=

x	y
196	225
1197	1330
3800	4200
4200	4641
25 025	27 600
101 517	111 930
111 930	123 410
336 400	370 881
653 982	721 005
2 165 800	2 387 736
8 758 530	9 656 010
9 656 010	10 645 453
51 134 902	56 374 605
56 374 605	62 151 210
227 971 809	251 331 640
754 892 610	832 244 973
1 467 334 764	1 617 689 665
4 407 244 205	4 858 845 810
4 858 845 810	5 356 722 138
19 648 533 840	21 661 880 240
114 712 842 804	126 467 240 865
126 467 240 865	139 426 088 840
379 853 324 410	418 776 143 058
1 693 475 654 040	1 867 002 753 889
8 967 984 477 244	9 886 916 104 065
9 886 916 104 065	10 900 008 836 640
10 900 008 836 640	12 016 911 177 184
57 722 182 076 524	63 636 859 874 340
257 339 093 124 078	283 708 120 867 885
772 936 210 999 265	852 137 455 336 440
852 137 455 336 440	939 454 294 486 344

4 974 982 373 978 540	5 484 758 975 161 665
20 118 174 513 236 358	22 179 644 053 110 285
22 179 644 053 110 285	24 452 348 298 251 930
66 618 133 403 668 240	73 444 361 733 905 904
129 490 069 394 690 718	142 758 660 617 899 710
428 785 758 454 882 465	472 722 586 799 070 024
1 733 953 222 127 096 685	1 911 627 978 284 991 330
1 911 627 978 284 991 330	2 107 508 715 188 457 322
10 123 233 832 009 955 398	11 160 541 574 617 489 110
11 160 541 574 617 489 110	12 304 140 189 364 237 005
45 131 762 523 375 987 876	49 756 324 938 894 870 625
149 446 649 572 842 506 730	164 760 152 084 113 151 202
290 489 331 286 574 336 196	320 255 198 349 295 377 100
872 505 301 602 330 542 510	961 909 193 662 632 879 405
961 909 193 662 632 879 405	1 060 474 125 662 579 338 962
3 889 834 288 702 242 799 425	4 288 417 912 481 875 449 200
22 709 783 383 889 109 535 356	25 036 809 957 409 526 826 900
25 036 809 957 409 526 826 900	27 602 282 340 048 130 088 225
75 199 833 764 288 882 817 805	82 905 411 952 144 336 724 442

Verify that these all yield the desired probability.

In[477]:=
```
DeleteDuplicates[probdifferent /@ xysolnsbyreduction210o421]
```

Out[477]= $\left\{\frac{210}{421}\right\}$

11.13 Method of solution by recursive reduction of RHS

Hua, Section 11.5, presents a method of solving Equation (11) by recursively reducing the RHS until it comes into the range for which a solution (if any exists) is guaranteed to be found among the convergents of $\sqrt{D}$. I provide it here as an alternative to the method of reduction, because for p less than about 100, it is significantly faster. Although both methods take less than one second on most cases with small p, if one wants to solve a large set of cases, say to test a hypothesis (as I did numerous times), the savings in compute time can be significant. For large p, this method becomes the slower of the two, at least as implemented here.

We start with Equation (11):

$$u^2 - D\,v^2 == f$$

If $|f| < \sqrt{D}$, then the equation can be solved by using the method of continued fractions. If this inequality is not satisfied, then we proceed as follows.

The method requires seeking integers l, h satisfying

$$h == \frac{l^2 - D}{f}$$

This requires that $l^2 - D \equiv 0 \pmod{f}$, which can be solved by searching. Observe that this is the same congruence as Equation (31) with l taking the place of s. However, in this method, it is required that $|l| < |f|/2$.

Hua recasts the congruence in the form

$$l^2 == D + f\,h$$

It suffices to search for perfect squares using the range $-h_{max} < h < h_{max}$ where

$$h_{max} == \mathrm{Max}\left(\left|\frac{f}{4}\right|, \left|\frac{D}{f}\right|\right)$$

Since $f \geq \sqrt{D}$ it is guaranteed that $h_{max} < |f|$. Then solve

$$u^2 - D\,v^2 == h$$

If $|h| < \sqrt{D}$ solve directly using continued fractions; otherwise repeat recursively. Since h is reduced on each step, the recursion is guaranteed to terminate. Once one has a solution

$$x^2 - D\,y^2 == h$$

then solutions to $u^2 - D\,v^2 == f$ are given by

$$u == \frac{D\,y \pm l\,x}{h}, \quad v == \frac{x \pm l\,y}{h}$$

The same sign must be used for each. Proof that this works:

In[478]:= $\mathtt{Simplify}\left[u^2 - D\,v^2 \;/.\; \left\{u \to \frac{D\,y + l\,x}{h},\; v \to \frac{x + l\,y}{h}\right\}\right]$

Out[478]= $\frac{\left(-D + l^2\right)\left(x^2 - D\,y^2\right)}{h^2}$

In[479]:= $\mathtt{Simplify}\left[u^2 - D\,v^2 \;/.\; \left\{u \to \frac{D\,y - l\,x}{h},\; v \to \frac{x - l\,y}{h}\right\}\right]$

Out[479]= $\frac{\left(-D + l^2\right)\left(x^2 - D\,y^2\right)}{h^2}$

Use the fact that x, y are solutions with RHS h, and that l was picked to satisfy the congruence:

In[480]:= $\mathtt{Simplify}\left[\%\;/.\; l^2 \to D + f\,h\right]$

Out[480]= $\frac{f\left(x^2 - D\,y^2\right)}{h}$

In[481]:= $\mathtt{Simplify}\left[\%\;/.\;\left\{h \to x^2 - D\,y^2\right\}\right]$

Out[481]= f

It is not guaranteed that u, v will be integer. However, non-integer results obtained deeper within the recurrence may yield integer solutions when the recurrence is unwound.

For each stage of the recursive process, more than one l, h pair may be found satisfying the congruence, and all of them may lead to solutions.

It is worth noting that for our problem, in the first stage of the recursion there is always a solution to the congruence $l == q - p$, $h == 1$.

$$l^2 == D + f\,h == q\,(q - 2\,p) + p^2\,h == q^2 - 2\,q\,p + p^2 + (h - 1)\,p^2 == (q - p)^2 + (h - 1)\,p$$

This yields the Pell equation to solve. However, if q is large, it may not be in the range of l less than $f/2$, so it may not be used. There may be other solutions to the congruence as well.

11.13.1 Example: $p/q = 5/11$ solved manually by method of recursive reduction

This section can be skipped, unless you are interested in the gory details.

This example was chosen more or less at random as a ratio with a fairly small numerator, that yields small solutions. Here is what the Pell equation method gives.

```
In[482]:= solveHyperbolicByPell[5 / 11]
```

Out[482]//TableForm=

x	y
15	30

In this example, p is prime, so we have no sub-equations that can be obtained by dividing Equation (11) by the squares of nontrivial divisors of p. In general those need to be solved as well to obtain all solutions. The only sub-equation here is the Pell equation, which gives the solution shown above. Here we look for additional solutions that are not obtained from the Pell equation.

```
In[483]:= D5o11 = q (q - 2 p) /. {p → 5, q → 11}
Out[483]= 11

In[484]:= f5o11 = p^2 /. p → 5
Out[484]= 25
```

Determine the maximum value of the RHS for which any solution must be given by a convergent.

```
In[485]:= fmax5o11 = Floor[Sqrt[D5o11]]
Out[485]= 3
```

The equation being solved is

```
In[486]:= u^2 - D5o11 v^2 == f5o11
Out[486]= u^2 - 11 v^2 == 25
```

We will need a table of values of $u^2 - D\,v^2$ for u/v the convergents of $\sqrt{D}$.

```
In[487]:= cv5o11 = Convergents[Sqrt[D5o11]]
```

Out[487]= $\left\{3, \frac{10}{3}, \sqrt{11}\right\}$

```
In[488]:= Length[cv5o11] - 1
Out[488]= 2
```

This is even. Get just the first cycle of convergents.

In[489]:=
```
cv5o11 = Convergents[√D5o11 , 2]
```

Out[489]= $\left\{3, \frac{10}{3}\right\}$

In[490]:=
```
TableForm[
 Table[
  {u, v, u^2 - D5o11 v^2} /. {u → Numerator[cv5o11[[i]]], v → Denominator[cv5o11[[i]]]},
  {i, Length[cv5o11]}],
 TableHeadings → {None, {"u", "v", "f"}}]
```

Out[490]//TableForm=

u	v	f
3	1	-2
10	3	1

Now seek to reduce the RHS of the equation. Determine the range of *h* values that need to be scanned for satisfying the congruence:

In[491]:=
```
hmax5o11 = Floor[Max[Abs[f5o11] / 4, D5o11 / Abs[f5o11]]]
```

Out[491]= 6

Find the values that work. Note need to offset position values that start at 1 so that *h* values start at $-h_{\max}$.

In[492]:=
```
hvalues5o11 = Flatten[Position[
    Table[√(D5o11 + f5o11 h) , {h, -hmax5o11, hmax5o11}], _Integer, {1}]] - hmax5o11 - 1
```

Out[492]= {1}

Find the corresponding l value:

In[493]:=
```
lvalues5o11 = √(D5o11 + f5o11 hvalues5o11)
```

Out[493]= {6}

The solution $h \to 1$, $l \to 6$ can be solved right away since $h < 6$. Since this is the Pell equation, the solution is at the end of the convergents table above, $u == 10$, $v == 3$.

In[494]:=
```
u^2 - D5o11 v^2 == 1 /. {u → 10, v → 3}
```

Out[494]= True

Unwind to the top-level equation

In[495]:=
```
{u → (D5o11 v0 + l u0) / h, v → (u0 + l v0) / h} /. {u0 → 10, v0 → 3, h → 1, l → 6}
```

Out[495]= {u → 93, v → 28}

Verify:

In[496]:=
```
u^2 - D5o11 v^2 == f5o11 /. {u → 93, v → 28}
```
Out[496]=
```
True
```

There is also a solution for the other sign on *l*

In[497]:=
```
{u → (D5o11 v0 - l u0) / h, v → (u0 - l v0) / h} /. {u0 → 10, v0 → 3, h → 1, l → 6}
```
Out[497]=
```
{u → -27, v → -8}
```

These are negative but the absolute values satisfy the equation too:

In[498]:=
```
u^2 - D5o11 v^2 == f5o11 /. {u → 27, v → 8}
```
Out[498]=
```
True
```

Convert (u, v) to (x, y)

In[499]:=
```
xyfromuv[{93, 28}] /. {p → 5, q → 11}
```
Out[499]=
```
{30, 58}
```

In[500]:=
```
xyfromuv[{27, 8}] /. {p → 5, q → 11}
```
Out[500]=
```
{7, 15}
```

Verify:

In[501]:=
```
probdifferent[{30, 58}]
```
Out[501]= $\frac{5}{11}$

In[502]:=
```
probdifferent[{7, 15}]
```
Out[502]= $\frac{5}{11}$

In sum, then, the method yielded the following solutions

x	y
7	15
30	58

The Pell equation solution (15, 30) was found at the start of this section. It is the middle member of a recycling triplet that these two solutions belong to.

11.13.2 Putting the method of recursive reduction into a function

First we define a function to find the solutions (u, v) using recursive reduction. That function is called by a function that uses those solutions to get (x, y) values satisfying the original problem, and includes an optional parameter n for steps of recurrence to run. If $n == 1$ the solution of the Pell equation (used only for the recurrence) is skipped.

As was done above for solving the reduced equation (method given by Alpern, Section 11.12), it is necessary to also solve Equation (11) divided by the square divisors of f, in order to find all solutions.

The recursive reduction routine includes a stanza that checks whether the RHS of the equation being solved is a perfect square, and if so includes the trivial solution $\left(\sqrt{f}, 0\right)$. This is not necessary, but if it is not included, then minimal members of some classes are not found. In order to produce the same result from `uvGetClasses` as `solveuvByReduction` without running the Pell recurrence in reverse, this step is included. It does not have a large impact on running time. (I tried simply joining the three trivial solutions to the output of the function before feeding to `uvGetClasses`. This yields the minimal class members in many cases but not all.)

Either sign of u, v satisfies Equation (11). The method seeks only positive values. For the hyperbolic case, negative u values do not yield admissible x, y. Only positive values of v are used, so that only distinct solutions with $x < y$ are generated. The admissibility tests on (x, y) use `Positive` as the criterion to exclude the trivial solutions; replacing that by `NonNegative` would include them.

The following function implements the recursion to find (u, v) solutions. It is meant to be called by the next function, a driver that collects all the solutions for Equation (11) and the sub-equations obtained dividing by square divisors of f. Finally, there is a function to take these solutions, run the Pell recurrence to find as many solutions as desired, and convert to (x, y).

Note: above, the congruence to be solved, $l^2 - D == 0 \pmod f$, was recast as $l^2 == D + h\,f$, and solutions were searched for by h. This turns out to be less efficient in Mathematica than searching by l, because search by h requires taking square roots, whereas search by l involves only the modulus operation. If $f == \pm 1$ the only solution is $l == 0$. However, this case is never encountered, since the condition $|f| < D$ always holds, so the recursion is not done. Thus the search bounds are $0 < l < f/2$.

This function does not call any functions that are not built-in to Mathematica. The arguments are D and f. The function does not check them for validity, since it is assumed it will only be called by the driver routine, which ensures that $D > 0$ and non-square. Note that although in Equation (11) $f == p^2$ is always positive, during the recursion, f may be positive or negative. This function does not sift the solutions for integer values, since fractional solutions produced at lower levels of the recursion may become integer at higher levels. The driver routine tests the final solutions for admissibility.

Outline of function:

- If $f < \sqrt{D}$, terminal case:
  - Calculate convergents of $\sqrt{D}$. If repeat length is odd, use two cycles.
  - Search the list convergents for u/v such that $u^2 - D\,v^2 == f$.
  - Also, if f is a square, include the solution $\left(\sqrt{f}, 0\right)$. Return list of (u, v) pairs as result.
- Otherwise, recursive reduction of RHS:
  - Find l values satisfying the congruence $l^2 - D \equiv 0 \pmod f$.
  - For each l, calculate $h == \left(l^2 - D\right)/f$ and call self, replacing f by h. It is guaranteed that $|h| < |f|$ so the recursion will terminate.

- Convert the (x, y) pairs satisfying $x^2 - D\,y^2 == h$ into (u, v) pairs satisfying $u^2 - D\,v^2 == f$ using $u == (D\,y + \delta\,l\,x)/h$, $v == (x + \delta\,l\,y)/h$ where $\delta == \pm 1$. Return the list of (u, v) as result.

A few economizations would be possible to speed up the terminal case. Since the convergents used are always those of $\sqrt{D}$, they could be calculated once at the outset and re-used each time the terminal case is reached. Also, not infrequently the same value of f appears at different stages of the calculation. For instance, $f == 1$ will be encountered when the routine is called by the driver for the divisor $d == p$, but not infrequently the recursive reduction step also reduces the RHS to 1. The solutions of $u^2 - D\,v^2 == f$ for a given f could be cached and re-used when that value is encountered again. Finally, when $f = \pm 1$, the solution is always found in a known location among the convergents, so searching them all is not necessary. These economizations are not done, since they would mainly improve the already speedy solution for small p, and don't help much for large p, when solving the congruence (time proportional to p^2) is dominant.

Caching of solutions of the congruence for a given $|f|$ could be used as well to speed up the recursive step, which would be helpful when p is large. Unfortunately this would not help as much as one would hope, since the recurring values of f are the small ones, and the initial value $f == p^2$, which takes the longest, only occurs once, at the top level.

In[503]:=
```
solveRecursiveReduction[D_, f_] := Module[
  {l, h, lvalues, hvalues, contfrac, repeat, replen,
   convergents, rruvvalues, uvvalues, uvlist, r, s, unextgen, wnextgen},
  If[f^2 < D, (* solution by convergents is guaranteed *)
   contfrac = ContinuedFraction[√D];
   (* List is of form {a0, a1, ... an,{b0, b1, ... bm}} *)
   repeat = contfrac[[-1]]; (* extract the repeat cycle at end *)
   replen = Length[repeat];
   If[Mod[replen, 2] == 0,
    convergents = Convergents[Flatten[contfrac][[ ;; replen]]], (* even replen *)
    convergents =
     Convergents[Join[Flatten[contfrac], repeat][[ ;; 2 replen]]] (* odd *)
   ];
   (* turn convergent fractions into {u,v} values *)
   uvvalues = Transpose[{Numerator /@ convergents, Denominator /@ convergents}];
   (* find {u,v} values satisfying the equation *)
   uvvalues = Select[uvvalues, u^2 - D v^2 == f /. {u → #[[1]], v → #[[2]]} &];
   (* If f is square, include the trivial solution u==√f,v=0 *)
   If[√f ∈ Integers,
    uvvalues = Append[uvvalues, {√f, 0}];
   ];
   uvvalues (* return result *)
```

```
    , (* else |f|≥√D *)
    (* recursive reduction of RHS.  Seek l, h satisfying l²==D+f h *)
    lvalues =
     Flatten[Position[Table[Mod[l² - D, f] == 0, {l, Ceiling[f / 2 - 1]}], True]];
    hvalues = Table[(l² - D) / f, {l, lvalues}];
    uvlist = {};
    Do[(* solve with reduced RHS for each l,h pair *)
     l = lvalues[[j]]; h = hvalues[[j]];
     rruvvalues = solveRecursiveReduction[D, h];
     uvvalues = Flatten[Table[{(D v + δ l u) / h, (u + δ l v) / h} /.
          {u → rruvvalues[[i]][[1]], v → rruvvalues[[i]][[2]]},
        {δ, {-1, 1}}, {i, Length[rruvvalues]}], 1];
     (* Sign of u, v can be either, so take absolute
      value.  Remove duplicates to fend off combinatorial explosion. *)
     uvlist = Sort[DeleteDuplicates[Abs[Join[uvvalues, uvlist]]]]
     , {j, Length[hvalues]} (* Do control *)
    ];
    uvlist (* return result *)
   ]
  ]
```

Next is the top-level driver of the recursive routine.

This function does not call any functions that are not built-in to Mathematica, except `solveRecursiveReduction`.

Outline of function:

- Calculate D and ensure that it is positive and nonsquare. Calculate $f == p^2$.
- Obtain the list of square divisors of p. For each divisor d:
  - Call `solveRecursiveReduction` to obtain solutions (x, y) of $x^2 - D\,y^2 == f / d^2$.
  - Convert each solution to $(u, v) == (p\,x, p\,y)$ which satisfies $u^2 - D\,v^2 == f$.
  - Sift for integer solutions, sort and delete duplicates.

In[504]:=
```
solveuvByRecursiveReduction[z_] := Module[
  {p, q, D, f, uvvalues, divisorlist},
  p = Numerator[z]; q = Denominator[z];
  D = q (q - 2 p);  f = p^2;
  If[ D > 0 && Sqrt[D] ∉ Rationals ,
   divisorlist = Divisors[p];
   uvvalues = {};
   Do[
    uvvalues = Join[uvvalues,
       Cases[d (solveRecursiveReduction[D, f / d^2]), {_Integer, _Integer}]];
    , {d, divisorlist}];
   DeleteDuplicates[Sort[uvvalues]] (* return result *)
   , (* else *) Print["D=", D, " not OK"]
  ]
 ]
```

Now the function to generate (x, y) solutions. It is the same as `solveHyperbolicByReduction` except for calling the recursive reduction method routine instead. Optional argument `iters` is the number of Pell iterations to run (default 3). Optional argument `tableform` is `True` (default) to format solutions in a table; otherwise they are output in a list suitable for input to other functions. This function uses `solvePell`, `solveRecursiveReduction`, `solveuvByRecursiveReduction`, and `uvGetClasses`.

In[505]:=
```
solveHyperbolicByRecursiveReduction[z_, iters_: 3, tableform_: True] :=
 Module[{p, q, D, hkpell, h, k, uvsolns, xysolns},
  p = Numerator[z];
  q = Denominator[z];
  D = q (q - 2 p);
  hkpell = solvePell[D]; (* get the base solution *)
  If[Length[hkpell] == 2,
   {h, k} = hkpell;
   uvsolns = uvGetClasses[Evaluate[solveuvByRecursiveReduction[z]]];
   uvsolns = Flatten[Table[RecurrenceTable[{
        u[i + 1] == h u[i] + D k v[i],
        v[i + 1] == k u[i] + h v[i],
        u[1] == uv[[1]],
        v[1] == uv[[2]]},
       {u, v}, {i, iters}],
      {uv, uvsolns}]
     , 1]; (* Flatten to produce list of {u,v} pairs *)
   xysolns = Table[1/2 {((p - u)/(2 p - q) - v), ((p - u)/(2 p - q) + v)} /. {u → uv[[1]], v → uv[[2]]},
     {uv, uvsolns}];
   xysolns = DeleteDuplicates[
     Sort[Cases[Sort /@ xysolns, {_Integer?Positive, _Integer?Positive}]]];
   If[tableform,
    TableForm[xysolns, TableHeadings → {None, {"x", "y"}}],
    xysolns]
  ]
 ]
```

Try it out on an example done earlier.

In[506]:=
```
solveHyperbolicByRecursiveReduction[4 / 11]
```

Out[506]//TableForm=

x	y
105	336
336	1072
1072	3417

Going to 5th Pell generation yields two sets of recycling triplets.

In[507]:=
```
solveHyperbolicByRecursiveReduction[4 / 11, 5]
```
Out[507]//TableForm=

x	y
105	336
336	1072
1072	3417
223 377	711 712
711 712	2 267 616
2 267 616	7 224 945

Here is an example that has solutions missed by the Pell equation method. I verified by direct search (not shown here) that the method of recursive reduction does not miss any solutions in this range.

In[508]:=
```
solveHyperbolicByRecursiveReduction[20 / 41, 2]
```
Out[508]//TableForm=

x	y
16	25
85	120
120	168
552	760
760	1045
2905	3984
12 609	17 280
17 280	23 680
23 680	32 449
102 544	140 505
388 245	531 960
531 960	728 872

Note on runtime efficiency

For small *p* the method of recursive reduction has the advantage over the method of reduction that it is much faster (10 or more times faster in tests). This is mainly because the method of reduction computes convergents more times. Of course, for small *p* both methods take much less than 1 second on a typical laptop to solve the equation, but if one wants to run a loop to, say, do a search to test a hypothesis, the advantage of this other method can be significant. However, for *p* larger than about 1000, recursive reduction is slower than reduction.

In[509]:=
```
Timing[solveuvByRecursiveReduction[6 / 17];]
```
Out[509]= {0.005659, Null}

In[510]:=
```
Timing[solveuvByReduction[6 / 17];]
```
Out[510]= {0.098075, Null}

In[511]:=
```
Timing[solveuvByRecursiveReduction[97 / (2 × 97 + 1)];]
```
Out[511]= {0.008322, Null}

In[512]:=
```
Timing[solveuvByReduction[97 / (2 × 97 + 1)];]
```
Out[512]=
```
{0.041556, Null}
```

In[513]:=
```
Timing[solveuvByRecursiveReduction[541 / (2 × 541 + 1)];]
```
Out[513]=
```
{0.180772, Null}
```

In[514]:=
```
Timing[solveuvByReduction[541 / (2 × 541 + 1)];]
```
Out[514]=
```
{0.237254, Null}
```

11.14 Prime values of *p* are special

In explorations of solutions for a large number of hyperbolic ratios, I noticed there were no exceptions to the following statement:

- If p is prime, $q > 2\,p$, and D nonsquare, there are no solutions to Equation (2) except those arising from applying the Pell recurrence to the trivial solutions. In other words, these cases have only the 3 classes of solutions that always exist.

A proof of this statement is given in Section 12.5.

For composite p, there are often additional classes, yielding singlets, doublets, or triplets. (Except for $p == 4$; see next section.)

This result greatly simplifies the solution when p is prime: applying the Pell recurrence to the three trivial solutions generates all solutions. There is no need to use one of the reduction methods or direct search, which generally take much more computation than solving the Pell equation. It puts ratios with very large p and q values in range to be completely solved.

One can understand that prime p should yield fewer solutions than composite p, since there will be no subsidiary equations obtained by dividing Equation (11) by the squares of non-trivial divisors of p, which can yield additional solutions that have the divisor as gcd. However, there is nothing that says in general Equation (11) itself cannot have additional primary solutions. For instance,

In[515]:=
```
uvGetClasses[solveuvByRecursiveReduction[6 / 17]]
```
Out[515]=
```
{{6, 0}, {11, -1}, {11, 1}, {74, -8},
 {74, 8}, {249, -27}, {249, 27}, {839, -91}, {839, 91}}
```

In[516]:=
```
uvGetClasses[solveuvByReduction[6 / 17]]
```
Out[516]=
```
{{6, 0}, {11, -1}, {11, 1}, {74, -8},
 {74, 8}, {249, -27}, {249, 27}, {839, -91}, {839, 91}}
```

The first three are in the trivial-solution class. The next two pairs are solutions from subsidiary equations and have common factors of 2 and 3 respectively. The last pair consists of relatively prime values and hence comes from Equation (11) directly.

In[517]:=
```
{GCD[74, 8],
 GCD[249, 27],
 GCD[839, 91]}
```
Out[517]=
```
{2, 3, 1}
```

So something must be happening to prevent this when p is prime. What that is is shown in Section 12.5.

11.14.1 The case $p == 4$ is also special

I also observed that when $p == 4$, $p/q < 1/2$, and D is nonsquare, the only solutions are those in the trivial classes, the same as for prime p. I consider this behavior unexpected. It requires that the only relatively prime solutions of Equation (11) be those in the trivial classes, and also that Equation (11) divided by 4 has no solutions. (Those would be solutions of Equation (11) with u, v even.)

Stated formally:

- If $p == 4$, $q > 8$, and D nonsquare, there are no solutions to Equation (2) except those arising from applying the Pell recurrence to the trivial solutions. In other words, these cases have only the 3 classes of solutions that always exist.

A proof of this statement is given in Section 12.6.

11.15 Near-triangular solutions for hyperbolic cases

This section can be skipped without loss of continuity, but I thought it showed some interesting relationships among solutions and between the elliptical and hyperbolic cases.

In Section 8.5.3 we explored cases yielding solutions in which x, y are 1 plus successive triangular numbers. These turn out to all be for elliptical ratios of the form $p/(2p-1)$, and also have vertex solutions where $x == y == p$. There are similarities to the cases where x, y are 1 less than successive triangular numbers, which lie in the hyperbolic region.

We can derive an expression for $z == p/q$ for which solutions are 1 less than successive triangular numbers.

In[518]:=
$$\texttt{zfortriangminus1 = Simplify}\left[\texttt{probdifferent}\left[\left\{v\,\frac{(v-1)}{2}-1,\ v\,\frac{(v+1)}{2}-1\right\}\right]\right]$$

Out[518]=
$$\frac{4-5\,v^2+v^4}{2\left(6-5\,v^2+v^4\right)}$$

The numerator and denominator factor nicely. I don’t know how to make Mathematica give what I think is the ideally simplest form. Start with the numerator.

In[519]:=
```
Factor[Numerator[zfortriangminus1]]
```
Out[519]=
```
(-2 + v) (-1 + v) (1 + v) (2 + v)
```

This is $(v^2-1)(v^2-4)$. Whether v is even or odd, it is a multiple of 4.

In[520]:= `Simplify[(v^2 - 1) (v^2 - 4) /. v → 2 n]`

Out[520]= $4\left(1-5n^2+4n^4\right)$

In[521]:= `Simplify[(v^2 - 1) (v^2 - 4) /. v → 2 n + 1]`

Out[521]= $4n(1+n)\left(-3+4n+4n^2\right)$

In[522]:= `Factor[Denominator[zfortriangminus1] / 2]`

Out[522]= $\left(-3+v^2\right)\left(-2+v^2\right)$

Whether *v* is even or odd, this is a multiple of 2, but not 4.

In[523]:= `Simplify[(-3 + v^2) (-2 + v^2) /. v → 2 n]`

Out[523]= $6-20n^2+16n^4$

In[524]:= `Simplify[Expand[(-3 + v^2) (-2 + v^2) /. v → 2 n + 1]]`

Out[524]= $2\left(1-6n+2n^2+16n^3+8n^4\right)$

So we have expressions for *p* and *q* separately.

In[525]:= `pfortriangminus1 = 1/4 Numerator[zfortriangminus1]`

Out[525]= $\frac{1}{4}\left(4-5v^2+v^4\right)$

In[526]:= `qfortriangminus1 = 1/4 Denominator[zfortriangminus1]`

Out[526]= $\frac{1}{2}\left(6-5v^2+v^4\right)$

These obey $q == 2p + 1$, which implies they are relatively prime.

In[527]:= `Simplify[qfortriangminus1 - 2 pfortriangminus1]`

Out[527]= 1

List the first few ratios in this category. Omit $v \le 2$ which give 0 for *p* or *q*.

In[528]:= `Table[zfortriangminus1, {v, 3, 10}]`

Out[528]= $\left\{\frac{10}{21}, \frac{45}{91}, \frac{126}{253}, \frac{280}{561}, \frac{540}{1081}, \frac{945}{1891}, \frac{1540}{3081}, \frac{2376}{4753}\right\}$

In[529]:= `solveHyperbolicByReduction[10 / 21, 1]`

Out[529]//TableForm=

x	y
2	5
5	10
10	18

The smallest solution is (3 – 1, 6 – 1) where 3 and 6 are successive triangular numbers. The larger solutions are not of that form.

In[530]:=
```
solveHyperbolicByReduction[45 / 91, 1]
```

Out[530]//TableForm=

x	y
5	9
33	45
45	60
120	153
324	405
405	505

The smallest solution is (6 – 1, 10 – 1).

In[531]:=
```
solveHyperbolicByReduction[126 / 253, 1]
```

Out[531]//TableForm=

x	y
9	14
105	126
126	150
234	273
826	945
945	1080
1548	1764
1764	2009
5460	6201
8769	9954
9954	11 298
30 226	34 290
48 374	54 873
54 873	62 244
62 244	70 604
99 576	112 945
301 224	341 649
341 649	387 498
546 390	619 710
1 652 294	1 873 998
1 873 998	2 125 449

The smallest solution is (10 – 1, 15 – 1).

Notice that in all three examples the second solution has $y == p$, and the third solution has $x == p$. We can show that these solutions always occur for this case (though not necessarily as second and third solutions).

We need to show that assuming $y == p$ when p/q is of the required form yields integer x.

In[532]:= `Solve[xyeqn[{x, y}], x] /. y → p`

Out[532]= $\left\{\left\{x\to\frac{p-2p^2+2pq-\sqrt{p^2+4p^2q-8p^3q+4p^2q^2}}{2p}\right\},\left\{x\to\frac{p-2p^2+2pq+\sqrt{p^2+4p^2q-8p^3q+4p^2q^2}}{2p}\right\}\right\}$

In[533]:= `Simplify[% /. {q → 2 p + 1}, Assumptions → p > 2]`

Out[533]= $\left\{\left\{x\to\frac{3}{2}+p-\frac{1}{2}\sqrt{9+16p}\right\},\left\{x\to\frac{1}{2}\left(3+2p+\sqrt{9+16p}\right)\right\}\right\}$

In[534]:= `Simplify[% /. {p → pfortriangminus1}, Assumptions → v > 2]`

Out[534]= $\left\{\left\{x\to\frac{1}{4}\left(20-9v^2+v^4\right)\right\},\left\{x\to\frac{1}{4}v^2\left(-1+v^2\right)\right\}\right\}$

The first of these factors:

In[535]:= $\mathbf{Factor}\left[\frac{1}{4}\left(20-9v^2+v^4\right)\right]$

Out[535]= $\frac{1}{4}(-2+v)(2+v)\left(-5+v^2\right)$

$$== \frac{1}{4}\left(v^2-4\right)\left(v^2-5\right)$$

Try this out for the first case.

In[536]:= `% /. v → 3`

Out[536]= 5

The first of these gives the second solution (5, 10), and the second gives (18, 10) which is the symmetric partner of the third solution.

So now we just need to show that these are integer for all *v*.

For even *v*

In[537]:= $\mathbf{Simplify}\left[\left\{\frac{1}{4}\left(20-9v^2+v^4\right),\frac{1}{4}v^2\left(-1+v^2\right)\right\}\ /.\ v\to 2n\right]$

Out[537]= $\left\{5-9n^2+4n^4,\ n^2\left(-1+4n^2\right)\right\}$

This is integer.

For odd *v*

In[538]:= $\mathbf{Simplify}\left[\left\{\frac{1}{4}\left(20-9v^2+v^4\right),\frac{1}{4}v^2\left(-1+v^2\right)\right\}\ /.\ v\to 2n+1\right]$

Out[538]= $\left\{3-7n-3n^2+8n^3+4n^4,\ n(1+n)(1+2n)^2\right\}$

Also integer. QED.

Remarks

In this section we showed that a subset of ratios of the form $p/(2p+1)$ have a solution consisting of two consecutive triangular numbers minus 1, and two more solutions in which one member is p.

There are some similarities to the elliptical special cases in Section 8.5.3, which are a subset of ratios of the form $p/(2p-1)$. These have a solution consisting of two triangular numbers plus 1, and two more solutions in which one member is p. There are some differences: for the elliptical case, all ratios of the form $p/(2p-1)$ have solutions in which one member is p, but for the hyperbolic case, in general ratios of the form $p/(2p+1)$ do not have such a solution. (As shown in Section 5.2.1, they do have inadmissible solutions around the negative vertex in which one member is $-p$.)

I have the feeling that there is something much bigger going on here, and we are only seeing the ripples on the surface from some enormous creature lurking in the mysterious depths.

12 Various proofs

This section contains proofs of various claims made in the previous sections. They are placed here to avoid cluttering the exposition.

12.1 No solutions exist between the three ellipse circum-vertex points

When examining the exhaustive enumeration of solutions for the elliptical case, in Section 8.3, we calculated the range of $z == p/q$ that would be covered by enumeration for a given maximum value of x and y. The minimum z was calculated conservatively there as the value that gives the maximum x at the vertex of the ellipse. This is conservative since somewhat smaller values of z (longer ellipses) would still fall short of $x+1$. That is, if

$$\frac{x_e}{2x_e-1} \le z \le 1, \quad x_e == x_{max}$$

then this range of z values is definitely covered by a search over $x \le y \le x_{max}$. But in fact some slightly smaller z values are also covered, for which the ellipse still does not reach $x_{max}+1$. We'd like a more precise bound. One could use instead

$$\frac{x_e}{2x_e-1} < z \le 1, \quad x_e == x_{max}+1$$

(note the change to a nonstrict inequality at the smaller bound) but now there seems to be the possibility that some solutions involving $x_{max}+1$ could exist and would be missed. The ellipse extends slightly further in x or y near the vertex, in the arc to the points (x_e, x_e-1) and (x_e-1, x_e). Here we prove the claim stated in Section 8.3.1 that no solutions can exist in the arc between the points at the vertex and

1 unit away from it in x or y, so this bound is safe to use.

The question is whether an integer solution can exist in the interval between the endpoint where $x == y$ and the points that are 1 unit away in x or y. Let x_e, not necessarily integer, be the endpoint, so that (x_e, x_e) lies on the ellipse. Then the points $(x_e - 1, x_e)$ and $(x_e, x_e - 1)$ also lie on the ellipse. This can be proved as follows:

In[539]:=
```
Reduce[probdifferent[{x, x}] == probdifferent[{x - 1, x}], x]
```

Out[539]=
```
True
```

Suppose an integer point (x, y) lies on the ellipse strictly between $(x_e - 1, x_e)$ and (x_e, x_e).

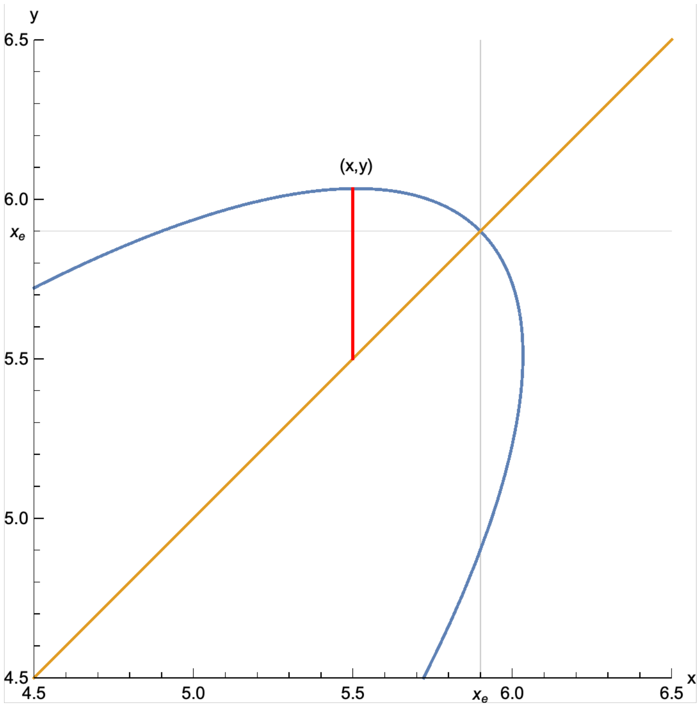
Here is a plot showing the situation. (Ignore the actual numbers, which are just for the illustration.) The ellipse (blue) extends above the horizontal line passing through the vertex. Drop a line segment (red) from (x, y) to (x, x) on the line (yellow) $y == x$. The length $y - x$ of this line segment needs to be integer if (x, y) is to be integer, since (y, x) must also be integer.

In[540]:=
```
xyforlineseg = Join[{{5.5, y}} /.
   Solve[{xyeqn[{5.5, y}] /. {p → 59, q → 108}, y > 5.5}][[1]], {{5.5, 5.5}}]
```

Out[540]=
```
{{5.5, 6.03349}, {5.5, 5.5}}
```

```
In[541]:= yvaluesforplot = Solve[xyeqn[{x, y}] /. {p → 59, q → 108}, y];
Show[Plot[{Labeled[y /. yvaluesforplot, "(x,y)", Above], x}, {x, 4.5, 6.5},
   PlotRange → {{4.5, 6.5}, {4.5, 6.5}}, AspectRatio → 1,
   GridLines → {{5.9}, {5.9}},
   AxesLabel → {"x", "y"}],
  ListLinePlot[xyforlineseg, PlotStyle → {Red}],
  Graphics[Text[Style["xe"], {5.9, 4.45}]],
  Graphics[Text[Style["xe"], {4.45, 5.9}]]
]
```

Out[542]=



Assume $x_e - 1 < x < x_e$ and then $y == x_e + \delta$ with $0 < \delta \leq \delta_{max}$ where δ_{max} is the maximum vertical excursion of the ellipse above the horizontal line $y == x_e$ through the endpoint. Let $x == x_e - 1 + \xi$ with $0 < \xi < 1$. Then $y - x$ is

```
In[543]:= Simplify[y - x /. {x → xe - 1 + ξ, y → xe + δ}]
Out[543]= 1 + δ - ξ
```

Now, using the limits on ξ and δ, the minimum is

```
In[544]:= 1 + δ - ξ /. {δ → 0, ξ → 1}
Out[544]= 0
```

The maximum is

```
In[545]:= 1 + δ - ξ /. {δ → δmax, ξ → 0}
Out[545]= 1 + δmax
```

Thus

$0 < y - x < 1 + \delta_{max}$

But $0 < \delta_{max} < 1$ as we will show. (In fact $\delta_{max} \leq \frac{1}{2}\left(\sqrt{2} - 1\right) \simeq 0.21$.) Therefore since $y - x = 0$ is excluded by the strict inequality, integer solutions must obey $y - x == 1$. This is $(y - 1, y)$ which implies (y, y) is also a solution, contrary to the strict inequality. We conclude that there can be no solutions in the interval strictly between the endpoint and its neighbors 1 unit away.

We need to put a bound on δ_{max} to show that $y - x \geq 2$ is not a possibility. The maximum value of y is found by solving the ellipse equation for x and finding where the two solutions are equal.

In[546]:=
```
xvsy = Solve[probdifferent[{x, y}] == z, x]
```

Out[546]= $\left\{\left\{x \to \frac{2 y + z - 2 y z - \sqrt{4 y^2 + 4 y z - 8 y^2 z + z^2}}{2 z}\right\}, \left\{x \to \frac{2 y + z - 2 y z + \sqrt{4 y^2 + 4 y z - 8 y^2 z + z^2}}{2 z}\right\}\right\}$

Making these equal requires the term inside the radical to be zero. Here is the right sequence of parts to pluck it out:

In[547]:=
```
terminradical = Part[xvsy, 1, 1, 2, 3, 4, 2, 1]
```

Out[547]= $4 y^2 + 4 y z - 8 y^2 z + z^2$

In[548]:=
```
ymaxvsz = Solve[terminradical == 0, y]
```

Out[548]= $\left\{\left\{y \to \frac{z - \sqrt{2}\, z^{3/2}}{2 (-1 + 2 z)}\right\}, \left\{y \to \frac{z + \sqrt{2}\, z^{3/2}}{2 (-1 + 2 z)}\right\}\right\}$

The negative branch gives the minimum y, so we need the positive branch to get maximum y.

In[549]:=
```
ymax = Part[ymaxvsz, 2, 1, 2]
```

Out[549]= $\frac{z + \sqrt{2}\, z^{3/2}}{2 (-1 + 2 z)}$

Verify that this is positive for elliptical ratios.

In[550]:=
```
Reduce[ymax > 0, z]
```

Out[550]= $z > \frac{1}{2}$

Now take the difference between y_{max} and the endpoint value $y_e == x_e$.

In[551]:= $\text{deltamax} = \text{Simplify}\left[\text{ymax} - \frac{z}{2 z - 1}\right]$

Out[551]= $\frac{z - \sqrt{2}\, z^{3/2}}{2 - 4 z}$

Determine for what range of z this is less than 1.

In[552]:=
```
Reduce[deltamax < 1, z]
```

Out[552]= $0 \leq z < \frac{1}{2} \;||\; \frac{1}{2} < z < 2 \left(3 + 2 \sqrt{2}\right)$

There is an indefinite value at $z == 1/2$, but for all values of $1/2 < z \leq 1$ the inequality holds. Look for a tighter bound. Here is δ_{max} at the maximum $z == 1$.

In[553]:=
```
deltamax /. z → 1
```

Out[553]= $\frac{1}{2}\left(-1+\sqrt{2}\right)$

In[554]:=
```
% // N
```

Out[554]= 0.207107

So the maximum difference between the endpoint and max y is about 0.2. Now look at it as $z \to 1/2$.

In[555]:=
```
deltamaxvseps = Simplify[deltamax /. z → 1/2 + ϵ]
```

Out[555]= $\frac{(1+2\,\epsilon)\left(-1+\sqrt{1+2\,\epsilon}\right)}{8\,\epsilon}$

For small ϵ the term $(1+2\,\epsilon) \to 1$. The term $\left(-1+\sqrt{1+2\,\epsilon}\right) \to (-1+(1+\epsilon)) == \epsilon$. So this difference goes to $1/8$. Here is the Taylor series expansion up to the linear term:

In[1116]:=
```
Series[deltamaxvseps, {ϵ, 0, 1}]
```

Out[1116]= $\frac{1}{8}+\frac{3\,\epsilon}{16}+O[\epsilon]^2$

Hence the excursion is limited to between 0.125 and about 0.2.

The strict upper bound occurs at $z == 1$ and was found above. Verify that it holds over the whole elliptical range.

In[557]:=
```
Reduce[{1/2 < z ≤ 1, deltamax ≤ 1/2 (√2 - 1)}, z]
```

Out[557]= $\frac{1}{2} < z \leq 1$

We have shown that $\delta_{max} < 1$, which in turn implies that there can be no solutions in between (x_e, x_e) and $(x_e - 1, x_e)$ or $(x_e, x_e - 1)$. Hence solutions cannot exceed the value of x_e, and it is safe to use the bound $z > x_e/(2\,x_e - 1)$ where the vertex $x_e == x_{max} + 1$ where x_{max} is the largest value of x or y used in the exhaustive search.

12.2 Number of solutions for hyperbolic nonsquare *D* is infinite

We have seen that solutions (u, v) to Equation (11) sometimes yield fractional (x, y) values. For instance, the trivial solutions are integer, but often the next Pell generation is fractional. Here we prove that if one generation has integer (x, y), then two generations later (x, y) will also be integer. Since the trivial solutions are always integer, this guarantees that admissible solutions will appear in the third generation and every odd generation thereafter, proving that the number of admissible solutions is infinite.

Assume that (x, y) is an integer solution to Equation (2). We do not require this solution to be admissible ($x \geq 0$, $y \geq 0$, $x + y \geq 2$), just integer.

Apply two generations of the Pell recurrence:

In[558]:=
```
{unplus2, vnplus2} =
  Simplify[nextPell[nextPell[uvfromxy[{x, y}]]], Assumptions → pelleqn[{h, k}, D]]
```

Out[558]=
$$\{2\,D\,h\,k\,(-x+y)+h^2\,(p+(-2\,p+q)\,(x+y))+D\,k^2\,(p+(-2\,p+q)\,(x+y)),\ h^2\,(-x+y)+D\,k^2\,(-x+y)+2\,h\,k\,(p+(-2\,p+q)\,(x+y))\}$$

This simplifies a bit further. I don't know how to get Mathematica to put it in this form, but it can verify that it is the same as the above:

In[559]:=
```
Simplify[{unplus2, vnplus2} == {(h^2 + D k^2) (p + (q - 2 p) (x + y)) - 2 D h k (x - y),
    -(h^2 + D k^2) (x - y) + 2 h k (p + (q - 2 p) (x + y))}]
```

Out[559]= True

Convert back to (x, y).

In[560]:=
```
{xnplus2, ynplus2} = Simplify[xyfromuv[{unplus2, vnplus2}] /. D → q (q - 2 p),
   Assumptions → {pelleqn[{h, k}, q (q - 2 p)], xyeqn[{x, y}]}]
```

Out[560]=
$$\left\{\frac{q\,(1-2\,h^2+2\,h\,k\,q)\,x+p\,(1-2\,x+h^2\,(-1+4\,x)+h\,k\,q\,(1-6\,x-2\,y))+2\,h\,k\,p^2\,(-1+2\,x+2\,y)}{2\,p-q},\ \frac{q\,(1-2\,h^2-2\,h\,k\,q)\,y-2\,h\,k\,p^2\,(-1+2\,x+2\,y)+p\,(1-2\,y+h^2\,(-1+4\,y)+h\,k\,q\,(-1+2\,x+6\,y))}{2\,p-q}\right\}$$

This has not used the Pell equation to simplify the terms involving h^2, which can introduce factors of $2\,p - q$ to cancel the denominator. Try again where we explicitly substitute h^2.

In[561]:=
```
{xnplus2, ynplus2} =
  Simplify[xyfromuv[{unplus2, vnplus2}] /. {h^2 → 1 + q (q - 2 p) k^2, D → q (q - 2 p)},
   Assumptions → {pelleqn[{h, k}, q (q - 2 p)], xyeqn[{x, y}]}]
```

Out[561]=
$$\{x+k^2\,q\,(p-4\,p\,x+2\,q\,x)+h\,k\,(-2\,q\,x+p\,(-1+2\,x+2\,y)),\ y+k^2\,q\,(p-4\,p\,y+2\,q\,y)+h\,k\,(p-2\,p\,x-2\,p\,y+2\,q\,y)\}$$

Again, this simplifies a bit further.

In[562]:=
```
Simplify[{xnplus2, ynplus2} == {x + k^2 q (p + 2 (q - 2 p) x) + h k (p (2 (x + y) - 1) - 2 q x),
    y + k^2 q (p + 2 (q - 2 p) y) - h k (p (2 (x + y) - 1) - 2 q y)}]
```

Out[562]= True

This expression is clearly integer if x, y and all other variables are integer. QED.

- Note

Since the Pell recurrence can be run forward or backward, this result implies that if a given class of solutions yields fractional (x, y) for two successive generations, then all generations of that class will be fractional. An example was seen in Section 11.12.6, for $p/q == 8/19$. It has two classes that yield no admissible solutions.

12.3 Pell recurrence from trivial solutions gives recycling triplets

In Section 11.8 we noted that applying the Pell recurrence to the solution $(u, v) == (p, 0)$, which corresponds to the trivial solution $(x, y) == (0, 0)$ yields the solution $(u, v) == (p\,h, p\,k)$ where (h, k) is the base solution of the Pell Equation.

We now show that this holds in general, and that the companion solutions $(u, v) == (q - p, \pm 1)$, which correspond to the other trivial solutions $(x, y) == (0, 1)$ and $(1, 0)$ yield the neighbors given by applying the recycling recurrence to that solution.

We will show algebraically, using a proof by induction, that the x, y recurrence starting from the three trivial solutions yields a series of solution triplets related by the recycling recurrence. We will work with a single step of the iteration, since it is simpler and it does not matter that the solutions are not guaranteed to be integer.

The proof is in the form of an induction proof, and has two parts:

- Base case: showing that the solutions found by applying the x, y recurrence to the three trivial solutions form a recycling triplet.
- Inductive step: showing that applying the Pell recurrence to members of a recycling pair yields another recycling pair. This result is stronger than needed for the present claim. It not only implies that a recycling triplet goes to a recycling triplet, but that any recycling pair, regardless of its origin, goes to another recycling pair.

The proof involves a lot of algebra, but fortunately we have Mathematica to do the heavy lifting.

We will often use the formula for obtaining (x, y) from (u, v) defined in Section 4.2.5:

```
In[563]:= xyfromuv[{u, v}]
```

Out[563]= $\left\{\frac{1}{2}\left(\frac{p-u}{2\,p-q}-v\right), \frac{1}{2}\left(\frac{p-u}{2\,p-q}+v\right)\right\}$

and the recycling recurrence defined in Section 5.4:

```
In[564]:= recycle[{x, y}]
```

Out[564]= $\left\{y, \frac{(-1+y)\,y}{x}\right\}$

12.3.1 Base case

First, recall the correspondence between trivial solutions in (u, v) and in (x, y):

In[565]:=
```
Simplify[xyfromuv /@
  {{q - p, -1}, {p, 0}, {q - p, 1}}]
```

Out[565]= {{1, 0}, {0, 0}, {0, 1}}

We want to show that the same sequence is followed when the recycling recurrence is applied to the next generation solutions.

Here is the Pell recurrence in (u, v), repeated from Equation (24):

$$u_{n+1} == h\, u_n + D\, k\, v_n$$
$$v_{n+1} == k\, u_n + h\, v_n$$

Apply the Pell recurrence to each of the trivial solutions (p, 0), ($q - p$, ± 1). The variable names here reflect the values of (x, y) corresponding to the (u, v) solutions.

In[566]:=
```
uvfrom10 = Simplify[{h u + D k v, k u + h v} /. {u → q - p, v → -1, D → q (q - 2 p)}]
```

Out[566]= $\{k\,(2\,p - q)\,q + h\,(-p + q),\ -h + k\,(-p + q)\}$

In[567]:=
```
uvfrom00 = Simplify[{h u + D k v, k u + h v} /. {u → p, v → 0, D → q (q - 2 p)}]
```

Out[567]= $\{h\,p,\ k\,p\}$

In[568]:=
```
uvfrom01 = Simplify[{h u + D k v, k u + h v} /. {u → q - p, v → 1, D → q (q - 2 p)}]
```

Out[568]= $\{k\,q\,(-2\,p + q) + h\,(-p + q),\ h + k\,(-p + q)\}$

We observe that the solution coming from (p, 0) is the solution obtained by the Pell equation method developed in Section 11.3.

Convert these to (x, y).

In[569]:=
```
xyfrom10 = Simplify[xyfromuv[uvfrom10]]
```

Out[569]= $$\left\{\frac{p + 3\,h\,p + 2\,k\,p^2 - 2\,h\,q - 5\,k\,p\,q + 2\,k\,q^2}{4\,p - 2\,q},\ \frac{p\,(1 - h + k\,(-2\,p + q))}{4\,p - 2\,q}\right\}$$

In[570]:=
```
xyfrom00 = Simplify[xyfromuv[uvfrom00]]
```

Out[570]= $$\left\{\frac{p\,(1 - h + k\,(-2\,p + q))}{4\,p - 2\,q},\ \frac{1}{2}\left(k\,p + \frac{p - h\,p}{2\,p - q}\right)\right\}$$

In[571]:=
```
xyfrom01 = Simplify[xyfromuv[uvfrom01]]
```

Out[571]= $$\left\{\frac{p\,(1 - h + 2\,k\,p - k\,q)}{4\,p - 2\,q},\ \frac{-2\,k\,p^2 - 2\,q\,(h + k\,q) + p\,(1 + 3\,h + 5\,k\,q)}{4\,p - 2\,q}\right\}$$

Verify that these yield a recycling triplet for the example 7/18.

In[572]:=
```
{h7o18, k7o18} = solvePell[q (q - 2 p) /. {p → 7, q → 18}]
```
Out[572]=
```
{17, 2}
```

In[573]:=
```
{xyfrom10, xyfrom00, xyfrom01} /. {p → 7, q → 18, h → h7o18, k → k7o18}
```
Out[573]=
```
{{2, 7}, {7, 21}, {21, 60}}
```

Indeed, they are a recycling triplet.

Now verify algebraically. First, the pair of solutions corresponding to (1, 0) → (0, 0). Show that recycling the solution obtained by applying the Pell recurrence to (1, 0) yields the one from (0, 0).

In[574]:=
```
Simplify[recycle[xyfrom10] == xyfrom00]
```
Out[574]=
$$\left\{0, \frac{p\left(-1 + h^2 + k^2 (2 p - q) q\right)}{p + 3 h p + 2 k p^2 - 2 h q - 5 k p q + 2 k q^2}\right\} == \{0, 0\}$$

Observe the numerator of the *y* expression contains the Pell Equation. Use the fact that *h*, *k* satisfy the Pell Equation.

In[575]:=
```
Simplify[%, Assumptions → {h^2 - q (q - 2 p) k^2 == 1}]
```
Out[575]=
```
True
```

Now the same for (0, 0) → (0, 1).

In[576]:=
```
Simplify[recycle[xyfrom00] == xyfrom01]
```
Out[576]=
$$\left\{0, \frac{1 - h^2 + k^2 q (-2 p + q)}{-1 + h + 2 k p - k q}\right\} == \{0, 0\}$$

In[577]:=
```
Simplify[%, Assumptions → {h^2 - q (q - 2 p) k^2 == 1}]
```
Out[577]=
```
True
```

Q.E.D.

So we have proved that the Pell recurrence applied to the trivial solutions gives a recycling triplet.

12.3.2 Inductive step

We now prove that if (X, Y) and (Y, Z) are a recycling pair of solutions of (2), then applying the Pell recurrence to them yields another recycling pair. First, convert these to (u, v) for applying the Pell recurrence.

In[578]:=
```
uvforXY = Simplify[uvfromxy[{X, Y}]]
```
Out[578]=
```
{p + (-2 p + q) (X + Y), -X + Y}
```

Put (Y, Z) in terms of (X, Y) using the recycling recurrence.

In[579]:=
```
recycle[{X, Y}]
```
Out[579]=
$$\left\{Y, \frac{(-1 + Y) Y}{X}\right\}$$

Convert to (u, v).

In[580]:=
```
uvforYZ = Simplify[uvfromxy[%]]
```

Out[580]=
$$\left\{p+\frac{(-2\,p+q)\,Y\,(-1+X+Y)}{X},\ \frac{Y\,(-1-X+Y)}{X}\right\}$$

Apply the Pell recurrence to obtain the next generation solutions. First, the (u, v) corresponding to (X, Y).

In[581]:=
```
uvfromXY =
  Simplify[{h u + D k v, k u + h v} /. {u → uvforXY[[1]], v → uvforXY[[2]], D → q (q - 2 p)}]
```

Out[581]=
$$\{k\,q\,(-2\,p+q)\,(-X+Y)+h\,(p+(-2\,p+q)\,(X+Y)),\ h\,(-X+Y)+k\,(p+(-2\,p+q)\,(X+Y))\}$$

And now the next generation of (u, v) corresponding to (Y, Z).

In[582]:=
```
uvfromYZ =
  Simplify[{h u + D k v, k u + h v} /. {u → uvforYZ[[1]], v → uvforYZ[[2]], D → q (q - 2 p)}]
```

Out[582]=
$$\left\{\frac{k\,q\,(-2\,p+q)\,Y\,(-1-X+Y)}{X}+h\left(p+\frac{(-2\,p+q)\,Y\,(-1+X+Y)}{X}\right),\ \frac{h\,Y\,(-1-X+Y)}{X}+k\left(p+\frac{(-2\,p+q)\,Y\,(-1+X+Y)}{X}\right)\right\}$$

Convert these back to (x, y).

In[583]:=
```
xyfromXY = Simplify[xyfromuv[uvfromXY]]
```

Out[583]=
$$\left\{\frac{p+2\,q\,(-h+k\,q)\,X+h\,p\,(-1+4\,X)+k\,p\,q\,(1-6\,X-2\,Y)+2\,k\,p^2\,(-1+2\,X+2\,Y)}{4\,p-2\,q},\ \frac{p-2\,q\,(h+k\,q)\,Y-2\,k\,p^2\,(-1+2\,X+2\,Y)+h\,p\,(-1+4\,Y)+k\,p\,q\,(-1+2\,X+6\,Y)}{4\,p-2\,q}\right\}$$

In[584]:=
```
xyfromYZ = Simplify[xyfromuv[uvfromYZ]]
```

Out[584]=
$$\left\{\frac{1}{4\,p\,X-2\,q\,X}\big(2\,q\,(-h+k\,q)\,X\,Y+2\,k\,p^2\,(2\,(-1+Y)\,Y+X\,(-1+2\,Y))+p\,(-2\,k\,q\,(-1+Y)\,Y+X\,(1-h+k\,q+4\,h\,Y-6\,k\,q\,Y))\big),\ \frac{1}{4\,p\,X-2\,q\,X}\big(-2\,q\,(h+k\,q)\,(-1+Y)\,Y+2\,k\,p^2\,(X-2\,X\,Y-2\,(-1+Y)\,Y)+p\,(2\,(2\,h+3\,k\,q)\,(-1+Y)\,Y-X\,(-1+h+k\,q-2\,k\,q\,Y))\big)\right\}$$

Apply the recycling recurrence to the first solution. The result should equal the second solution.

```
In[585]:= xyfromXYrecycle = Simplify[recycle[xyfromXY]]
```

Out[585]=
$$\left\{\frac{p-2q(h+kq)Y-2kp^2(-1+2X+2Y)+hp(-1+4Y)+kpq(-1+2X+6Y)}{4p-2q},\right.$$
$$\left(\left(p-2q(h+kq)Y-2kp^2(-1+2X+2Y)+hp(-1+4Y)+kpq(-1+2X+6Y)\right)\left(-1+\frac{p-2q(h+kq)Y-2kp^2(-1+2X+2Y)+hp(-1+4Y)+kpq(-1+2X+6Y)}{4p-2q}\right)\right)\Big/$$
$$\left.\left(p+2q(-h+kq)X+hp(-1+4X)+kpq(1-6X-2Y)+2kp^2(-1+2X+2Y)\right)\right\}$$

Show that these two expressions for the second generation (*Y*, *Z*) are the same.

```
In[586]:= Simplify[xyfromXYrecycle == xyfromYZ]
```

Out[586]=
$$\left\{2kqY+\frac{kp\left(X-X^2+Y-2XY-Y^2\right)}{X},\right.$$
$$\left(\left(p-2q(h+kq)Y-2kp^2(-1+2X+2Y)+hp(-1+4Y)+kpq(-1+2X+6Y)\right)\left(-1+\frac{p-2q(h+kq)Y-2kp^2(-1+2X+2Y)+hp(-1+4Y)+kpq(-1+2X+6Y)}{4p-2q}\right)\right)\Big/$$
$$\left(p+2q(-h+kq)X+hp(-1+4X)+kpq(1-6X-2Y)+2kp^2(-1+2X+2Y)\right)+$$
$$\frac{1}{4pX-2qX}\left(2q(h+kq)(-1+Y)Y+2kp^2(2(-1+Y)Y+X(-1+2Y))+\right.$$
$$\left.\left.p(-2(2h+3kq)(-1+Y)Y+X(-1+h+kq-2kqY))\right)\right\}=\{0,0\}$$

We can see terms from Equation (2) appearing in these expressions. Use that equation to cancel terms.

```
In[587]:= Simplify[%, Assumptions → p X^2 - 2 (q - p) X Y + p Y^2 - p X - p Y == 0]
```

Out[587]=
$$\left\{0,\frac{p\left(-1+h^2+k^2(2p-q)q\right)}{p+2q(-h+kq)X+hp(-1+4X)+kpq(1-6X-2Y)+2kp^2(-1+2X+2Y)}\right\}=\{0,0\}$$

Now we see the Pell Equation (20) in the numerator.

```
In[588]:= Simplify[%, Assumptions → {h^2 - q (q - 2 p) k^2 == 1}]
```

```
Out[588]= True
```

QED.

Note that the only assumption made here was that (*X*, *Y*) and (*Y*, *Z*) are a recycling pair, so this implies that any such pair is taken to another recycling pair by the Pell recurrence.

There is no guarantee that the new solutions will be admissible, but it was shown in Section 12.2 that if the starting solutions are admissible, at least every other generation thereafter is admissible.

12.4 Recycling recurrence is complete for $p == 1$ and $p == 2$

In Section 11.9.3 we showed that for $p == 1$ all solutions belong to one class, which implies that the recycling recurrence gives the same set of solutions as the Pell recurrence, and is complete. We now look at $p == 1$ from another viewpoint, and also show that for $p == 2$ the recycling recurrence is complete.

12.4.1 The case $p == 1$

Here we need to require $q > 2$ since $p/q == 1/1$ is an elliptical case and $p/q == 1/2$ is parabolic. For these, $D == -1$ and 0 respectively, for which the assumption $D > 0$ required for the Pell recurrence does not hold. In Section 10.1 we showed that for $p == 1$ and $q > 2$, D is never square, so the Pell recurrence can be used.

Nagell p. 199 notes that if $D == a^2 - 1$, the Pell equation $r^2 - D\,s^2 == 1$ has as its base solution

$r == a, \quad s == 1$

```
In[589]:= Simplify[r^2 - (a^2 - 1) s^2 == 1 /. {r → a, s → 1}]
Out[589]= True
```

This condition, $D + 1$ square, always holds when $p == 1$:

```
In[590]:= Simplify[q (q - 2 p) + 1 /. p → 1]
Out[590]= (-1 + q)^2
```

Hence $D == (q - 1)^2 - 1$ and the base solution of the Pell equation is

$r == q - 1, \quad s == 1$

```
In[591]:= Simplify[r^2 - q (q - 2 p) s^2 == 1 /. {p → 1, r → q - 1, s → 1}]
Out[591]= True
```

When $p == 1$, $(u, v) == (r, s)$ and Equation (8) is the same as the Pell equation (20). So we have precisely Nagell's special case. Note that there is a trivial solution $(u, v) == (1, 0)$. This converts to the trivial solution $(x, y) == (0, 0)$:

```
In[592]:= xyfromuv[{1, 0}] /. p → 1
Out[592]= {0, 0}
```

The base solution converts to the trivial solution (0, 1):

```
In[593]:= xyfromuv[{q - 1, 1}] /. p → 1
Out[593]= {0, 1}
```

Reversing the sign of u gives negative (x, y), not admissible. Reversing the sign of v swaps x and y. Thus we obtain all three trivial solutions from the Pell equation.

Applying the Pell recurrence gives the first nontrivial solution:

In[594]:=
```
Simplify[xyfromuv[nextPell[{q - 1, 1}] /. {D → q (q - 2), h → q - 1, k → 1}] /. p → 1]
```
Out[594]=
```
{1, -1 + 2 q}
```

Since the Pell recurrence generates all solutions, continuing will generate the complete set of the solutions.

In recurrence form, $\{u_{n+1}, v_{n+1}\}$ is given by

In[595]:=
```
nextPell[{u_n, v_n}] /. {D → q (q - 2), h → q - 1, k → 1}
```
Out[595]=
```
{(-1 + q) u_n + (-2 + q) q v_n, u_n + (-1 + q) v_n}
```

These are integer in form. Get this in terms of (t, v).

In[596]:=
```
Simplify[Solve[{u_{n+1} == u_n (q - 1) + (q - 2) q v_n,
      v_{n+1} == u_n + (q - 1) v_n} /. {u_{n+1} → p + (q - 2 p) t_{n+1}, u_n → p + (q - 2 p) t_n}, {t_{n+1}, v_{n+1}}] /. p →
```
Out[596]=
```
{{t_{1+n} → 1 + (-1 + q) t_n + q v_n, v_{1+n} → 1 + (-2 + q) t_n + (-1 + q) v_n}}
```

Now convert to (x, y).

In[597]:=
```
Simplify[Solve[{t_{1+n} == 1 + (-1 + q) t_n + q v_n, v_{1+n} == 1 + (-2 + q) t_n + (-1 + q) v_n} /.
    {t_{n+1} → x_{n+1} + y_{n+1} , v_{n+1} → y_{n+1} - x_{n+1}, t_n → x_n + y_n, v_n → y_n - x_n}, {x_{n+1}, y_{n+1} }]]
```
Out[597]=
```
{{x_{1+n} → y_n, y_{1+n} → 1 - x_n + 2 (-1 + q) y_n}}
```

This is the recycling recurrence, since it has $x_{n+1} == y_n$ and $y_{n+1} \neq x_n$. The recycling recurrence cannot be started with the trivial solution, but beyond that this must be the same.

Hence we have shown that the recycling recurrence gives the same sequence of solutions (after the trivial solutions) as the Pell recurrence, and therefore is complete.

Verify with an example $p/q == 1/5$ that this gives the same sequence as solution of the Pell equation. First, the recycling recurrence derived in Section 5.4:

In[598]:=
```
TableForm[RecurrenceTable[
  {x[n + 1] == y[n], y[n + 1] == y[n] (y[n] - 1) / x[n], x[1] == 1, y[1] == 9}, {x, y}, {n, 6}]]
```
Out[598]//TableForm=

1	9
9	72
72	568
568	4473
4473	35 217
35 217	277 264

Now the Pell recurrence:

In[599]:=
```
TableForm[RecurrenceTable[
  {x[n + 1] == y[n], y[n + 1] == 1 - x[n] + 2 (q - 1) y[n], x[1] == 0, y[1] == 1} /. q → 5,
  {x, y}, {n, 7}]]
```

Out[599]//TableForm=

0	1
1	9
9	72
72	568
568	4473
4473	35 217
35 217	277 264

This is the same as given by the method of solution via Pell equation.

In[600]:=
```
solveHyperbolicByPell[1 / 5, 7]
```

Out[600]//TableForm=

x	y
0	1
1	9
9	72
72	568
568	4473
4473	35 217
35 217	277 264

12.4.2 The case $p == 2$

Here we need to require q odd so that p and q are coprime, and $q \geq 5$ since $p/q == 2/3$ ($D == -3$) is an elliptical case (and of course 2/1 is not a probability). For $q == 3$ the results actually hold, but the recurrence terminates. In Section 10.1 we showed that for $p == 2$ and $q > 4$, D is never square, so the Pell recurrence can be used.

A generalization of $D == a^2 - 1$ for the Pell equation is $D == a^2 - p^2$ for the u, v equation. In fact, this always holds, with $a == p - q$.

In[601]:=
```
Simplify[q (q - 2 p) + p^2]
```

Out[601]= $(p - q)^2$

So the smallest positive solution is

$u == q - p, \quad v == 1$

In[602]:=
```
Simplify[u^2 - q (q - 2 p) v^2 == p^2 /. {u → q - p, v → 1}]
```

Out[602]= True

Map to x, y

In[603]:=
```
Simplify[xyfromuv[{q - p, 1}]]
```

Out[603]= {0, 1}

So the fact that for $D == (q-p)^2 - p^2$ Equation (8) has this solution is simply a manifestation of the trivial solution. It is not helpful for finding nontrivial solutions since one still needs to solve the Pell equation to obtain additional solutions via the Pell recurrence. However, for $p == 2$ there is another recurrence that does not require solving the Pell equation, as we now develop.

Hua Theorem 11.4.4 states that the complete set of solutions for

$$u^2 - D\,v^2 == 4$$

are given by

$$\frac{u + v\sqrt{D}}{2} == \pm\left(\frac{u_0 + v_0\sqrt{D}}{2}\right)^n,\quad n == 0,\ \pm 1,\ \pm 2,\ \ldots$$

where u_0, v_0 is the smallest positive solution. This is our equation when $p == 2$. In this case the smallest positive solution is the trivial solution $(u_0, v_0) == (q-2, 1)$.

Using Hua's recurrence, $u_{n+1} + v_{n+1}\sqrt{D}$ is given by

In[604]:=
```
Collect[
 Simplify[Expand[2 ((u_n + v_n Sqrt[q (q - 4)])/2) ((u_0 + v_0 Sqrt[q (q - 4)])/2)] /. {u_0 -> q - 2, v_0 -> 1}],
 {Sqrt[(-4 + q) q]}]
```

Out[604]=
$$\frac{1}{2}\sqrt{(-4+q)\,q}\;(u_n - 2\,v_n + q\,v_n) + \frac{1}{2}\left(-2\,u_n + q\,u_n - 4\,q\,v_n + q^2\,v_n\right)$$

Equating rational and irrational parts,

$$u_{n+1} == \frac{1}{2}\left(u_n\,(q-2) - 4\,q\,v_n + q^2\,v_n\right),$$

$$v_{n+1} == \frac{1}{2}\left(u_n + (q-2)\,v_n\right)$$

These are not manifestly integer, but they are integer per Hua's theorem 11.4.4. Put it in terms of (t, v).

In[605]:=
```
Simplify[Solve[{u_(n+1) == 1/2 (u_n (q - 2) - 4 q v_n + q^2 v_n),
    v_(n+1) == 1/2 (u_n + (q - 2) v_n)} /. {u_(n+1) -> p + (q - 2 p) t_(n+1), u_n -> p + (q - 2 p) t_n}, {t_(n+1), v_(n+1)}] /.
  p -> 2]
```

Out[605]=
$$\left\{\left\{t_{1+n} \to \frac{1}{2}\,(2 + (-2+q)\,t_n + q\,v_n),\ v_{1+n} \to \frac{1}{2}\,(2 + (-4+q)\,t_n + (-2+q)\,v_n)\right\}\right\}$$

And now in terms of (x, y).

In[606]:=
```
Simplify[Solve[{t_{n+1} == 1/2 (2 + (-2 + q) t_n + q v_n), v_{1+n} == 1/2 (2 + (-4 + q) t_n + (-2 + q) v_n)} /.
   {t_{n+1} → x_{n+1} + y_{n+1} , v_{n+1} → y_{n+1} - x_{n+1}, t_n → x_n + y_n, v_n → y_n - x_n}, {x_{n+1}, y_{n+1} }]]
```

Out[606]= $\{\{x_{1+n} \to y_n,\ y_{1+n} \to 1 - x_n + (-2 + q)\ y_n\}\}$

In[607]:=
```
Simplify[% /. q → 3]
```

Out[607]= $\{\{x_{1+n} \to y_n,\ y_{1+n} \to 1 - x_n + y_n\}\}$

This is the same as the recycling recurrence, since it has $x_{n+1} == y_n$ and $y_{n+1} \neq x_n$. Verify with an example $p/q = 2/5$. First, the recycling recurrence:

In[608]:=
```
TableForm[RecurrenceTable[
  {x[n + 1] == y[n], y[n + 1] == y[n] (y[n] - 1) / x[n], x[1] == 1, y[1] == 4}, {x, y}, {n, 9}]]
```

Out[608]//TableForm=

1	4
4	12
12	33
33	88
88	232
232	609
609	1596
1596	4180
4180	10 945

Now, the recurrence derived here:

In[609]:=
```
TableForm[RecurrenceTable[
  {x[n + 1] == y[n], y[n + 1] == 1 - x[n] + (q - 2) y[n], x[1] == 0, y[1] == 1} /. q → 5,
  {x, y}, {n, 10}]]
```

Out[609]//TableForm=

0	1
1	4
4	12
12	33
33	88
88	232
232	609
609	1596
1596	4180
4180	10 945

Unlike the $p == 1$ case, the Pell solutions should match every third solution, since the three trivial solutions belong to different classes.

In[610]:=
```
solveHyperbolicByPell[2 / 5, 3]
```
Out[610]//TableForm=

x	y
4	12
88	232
1596	4180

12.4.3 Conclusion

Since we have shown that the proven complete Pell recurrences for the solutions in the cases $p == 1$ and $p == 2$ are equivalent to the recycling recurrence (except for the trivial solutions), the recycling recurrence starting from the smallest nontrivial solution gives the complete set of solutions for these cases.

Note that the recurrences for $p == 1$ and $p == 2$ are the same in terms of Q where $Q == q$ for p even and $Q == 2\,q$ for p odd.

For $p == 1$:

In[611]:= $\text{Simplify}[\{\{x_{1+n} \to y_n,\ y_{1+n} \to 1 - x_n + 2\,(-1 + q)\ y_n\}\}\ /.\ q \to Q / 2]$

Out[611]= $\{\{x_{1+n} \to y_n,\ y_{1+n} \to 1 - x_n + (-2 + Q)\ y_n\}\}$

For $p == 2$:

In[612]:= $\text{Simplify}[\{\{x_{1+n} \to y_n,\ y_{1+n} \to 1 - x_n + (-2 + q)\ y_n\}\}\ /.\ q \to Q]$

Out[612]= $\{\{x_{1+n} \to y_n,\ y_{1+n} \to 1 - x_n + (-2 + Q)\ y_n\}\}$

However, the two cases are fundamentally different since for $p == 1$ there is only one class of solutions, while for $p == 2$ there are three.

12.5 If $p > 2$ is prime, the only solution classes are the trivial-solution classes

(To avoid an overly lengthy section title, the conditions that this be a hyperbolic case with nonsquare D are implicit.

The claim being proved here (stated earlier in Section 11.14) is that if $p > 2$ is prime, with $p/q < 1/2$ and D nonsquare, then the only solutions that exist are those obtained by applying the Pell recurrence to the trivial solutions. We work entirely in (u, v)-space in this section.

12.5.1 Introduction

To distinguish among the classes arising from the three trivial solutions, let us call the class to which the solution $(u, v) == (q, 0)$ belongs $\mathbb{C}_0$, the class to which $(q - p, -1)$ belongs $\mathbb{C}_{-1}$, and the one to which $(q - p, 1)$ belongs $\mathbb{C}_1$.

Outline of the proof:

- Rewrite Equation (11) in the form of a product of two terms on the LHS equaling a multiple of p^2 on the RHS.
- Show that for solutions in class $\mathbb{C}_{-1}$ or $\mathbb{C}_1$, one of the two terms on the LHS is itself a multiple of p^2.
- Show that if one of the two terms on the LHS is a multiple of p^2, then the solution is in class $\mathbb{C}_{-1}$ or $\mathbb{C}_1$. With the previous result, this is an if and only if.
- Show that there cannot be solutions of Equation (11) for which each of the two terms on the LHS is a multiple of p. Hence for prime p the only possible partitioning between the two terms is to have p^2 divide one of them and not the other.

From these steps we conclude that for prime p, since all solutions must have one of the two terms a multiple of p^2, all relatively prime solutions are in class $\mathbb{C}_{-1}$ or $\mathbb{C}_1$. The only other solution class is that obtained by dividing Equation (11) by p^2, which is $\mathbb{C}_0$. Including this solution class, there are only three solution classes, the trivial-solution classes.

In all that follows in this section, it is assumed that $p > 2$, $p/q < 1/2$, and D is nonsquare. We also assume that the solutions (u, v) we are working with have $\gcd(u, v) == 1$, i.e. excluding solutions obtained by solving Equation (11) divided by divisors of p^2.

12.5.2 Rewriting Equation (11) in factored form with p on RHS only

The initial goal is to rewrite Equation (11) with p absent from the left side.

$$u^2 - D v^2 == f$$
$$u^2 - q (q - 2 p) v^2 == p^2$$
$$u^2 - \left((q - p)^2 - p^2\right) v^2 == p^2$$

Let $a == q - p$. This change replaces the parameters p, q by a new pair p, a, so we can treat a as independent of p. The only constraints on a are $a > p$ and $\gcd(a, p) == 1$. We also require that $a^2 - p^2 == D$ not be a square, so that the Pell recurrence is applicable. In terms of p and a, Equation (11) takes the form

$$u^2 - \left(a^2 - p^2\right) v^2 == p^2$$
$$u^2 - a^2 v^2 == \left(1 - v^2\right) p^2$$
$$(u - a v) (u + a v) == \left(1 - v^2\right) p^2$$

Rewrite so all terms are non-negative.

$$(a v - u) (a v + u) == \left(v^2 - 1\right) p^2 \tag{34}$$

Observe that changing the sign of u or v yields the same equation. Changing one changes the solution to the conjugate class, while changing both keeps the solution in the same class. For brevity, in the following section headings and bodies, the word “term” means $a v - u$ or $a v + u$.

12.5.3 The trivial classes give a term that is a multiple of p^2

This step is not essential to the proof, since later we show the converse, which suffices.

For the fundamental solutions of classes $\mathbb{C}_{-1}$ and $\mathbb{C}_1$, $v^2 - 1 == 0$, so Equation (34) simply gives $u == \pm a$, which is not helpful. To get a constraint involving *p* we need to use the second Pell generation.

If (*h*, *k*) is the primary solution to the Pell equation, the Pell recurrence is

$u_{n+1} == h\, u_n + D\, k\, v_n == h\, u_n + \left(a^2 - p^2\right) k\, v_n$

$v_{n+1} == k\, u_n + h\, v_n$

This gives the second generation solution in $\mathbb{C}_{-1}$ as

In[613]:= `Simplify[{h u + (a^2 - p^2) k v, k u + h v} /. {u → a, v → -1}]`

Out[613]= $\left\{a\, h - a^2\, k + k\, p^2,\ -h + a\, k\right\}$

The two terms on the LHS of Equation (34) for this solution are

In[614]:= `Simplify[{a v - u, a v + u} /. {u → h a + k (p^2 - a^2), v → -h + k a}]`

Out[614]= $\left\{-2\, a\, h + 2\, a^2\, k - k\, p^2,\ k\, p^2\right\}$

This shows that $a\, v + u$ is a multiple of p^2.

The second generation solution in $\mathbb{C}_1$ is

In[615]:= `Simplify[{h u + (a^2 - p^2) k v, k u + h v} /. {u → a, v → 1}]`

Out[615]= $\left\{a\, h + a^2\, k - k\, p^2,\ h + a\, k\right\}$

The two terms on the LHS of Equation (34) for this solution are

In[616]:= `Simplify[{a v - u, a v + u} /. {u → h a + k (-p^2 + a^2), v → h + k a}]`

Out[616]= $\left\{k\, p^2,\ 2\, a\, h + 2\, a^2\, k - k\, p^2\right\}$

This shows that $a\, v - u$ is a multiple of p^2.

So both classes give one term that is a multiple of p^2. The other term is not, as we prove later. We have not shown that this holds for succeeding generations, though it must.

12.5.4 If one term is a multiple of p^2, the solution must be in a trivial class

First, assume that $a\, v + u == n\, p^2$, where *n* is an integer. We can show that this solution belongs to class $\mathbb{C}_{-1}$. Use Equation (26), replacing the denominator by the equivalent *f*:

$$\left\{r \to \frac{u_1\, u_2 - D\, v_1\, v_2}{f},\ \frac{s \to u_1\, v_2 - u_2\, v_1}{f}\right\}$$

$$\left\{r \to \frac{u_1\, u_2 - \left(a^2 - p^2\right) v_1\, v_2}{p^2},\ s \to \frac{u_1\, v_2 - u_2\, v_1}{p^2}\right\}$$

Set $(u_1, v_1) == (a, -1)$, the fundamental solution of class $\mathbb{C}_{-1}$. Set $(u_2, v_2) == (u, v)$ with $u == n\, p^2 - a\, v$.

In[617]:= Simplify[

$$\left\{r \to \frac{u_1\, u_2 - \left(a^2 - p^2\right) v_1\, v_2}{p^2},\ s \to \frac{u_1\, v_2 - u_2\, v_1}{p^2}\right\} \,/.\, \left\{u_1 \to a,\ v_1 \to -1,\ u_2 \to n\, p^2 - a\, v,\ v_2 \to v\right\}\Big]$$

Out[617]= $\{r \to a\, n - v,\ s \to n\}$

Both r and s are integer, so this shows (u, v) belongs to the same class as $(a, -1)$, namely $\mathbb{C}_{-1}$.

Now show it for the conjugate class. Assume $a\, v - u == n\, p^2$.

In[618]:= Simplify[

$$\left\{r \to \frac{u_1\, u_2 - \left(a^2 - p^2\right) v_1\, v_2}{p^2},\ s \to \frac{u_1\, v_2 - u_2\, v_1}{p^2}\right\} \,/.\, \left\{u_1 \to a,\ v_1 \to 1,\ u_2 \to -n\, p^2 + a\, v,\ v_2 \to v\right\}\Big]$$

Out[618]= $\{r \to -a\, n + v,\ s \to n\}$

Again, integer, so if $a\, v - u == n\, p^2$, the solution is in $\mathbb{C}_1$.

12.5.5 Both terms cannot be divisible by p

Note we are not assuming p prime for this section of the proof.

Suppose that each of the terms on the LHS of Equation (34) is divisible by p. We can write

$a\, v + u == d_1\, p,\quad a\, v - u == d_2\, p$

where $d_1\, d_2 == v^2 - 1$. Adding,

$2\, a\, v == (d_1 + d_2)\, p$

Subtracting,

$2\, u == (d_1 - d_2)\, p$

If p is odd, and using $\gcd(a, p) == 1$, these imply that both u and v are multiples of p, contradicting the assumption that they are relatively prime. If $p > 2$ is even, then both u and v are multiples of $p/2$, again contradicting the assumption that they are relatively prime.

If p is composite, it is possible for one term to be divisible by p. If d is a divisor of p, $1 < d < p$, then one term can be divisible by $p\, d$, and the other term divisible by p/d. See the Examples section below.

If p is prime, the only partitioning allowed is for one term to be a multiple of p^2 and the other not divisible by p. We showed that this implies that any solution is a member of class $\mathbb{C}_{-1}$ or $\mathbb{C}_1$.

This concludes the proof that for prime p, the only solutions are those in the three classes to which the trivial solutions belong.

Examples

- An example where p is prime: $5/11$. $a == q - p == 6$.

In[619]:= `uvGetClasses[solveuvByRecursiveReduction[5 / 11]]`

Out[619]= {{5, 0}, {6, -1}, {6, 1}}

Find the next generation. Get (h, k).

In[620]:= `D5o11 = q (q - 2 p) /. {p → 5, q → 11}`

Out[620]= 11

In[621]:= `solvePell[11]`

Out[621]= {10, 3}

The 2nd generation from (a, −1) is

In[622]:= $\{h\, a + k\, (p^2 - a^2),\ -h + k\, a\}$ /. {p → 5, a → 6, h → 10, k → 3}

Out[622]= {27, 8}

Calculate $a\, v \pm u$ directly:

In[623]:= `{a v + u, a v - u} /. {a → 6, u → 27, v → 8}`

Out[623]= {75, 21}

Calculate using the formulas in terms of h, k:

In[624]:= $\{k\, p^2,\ -k\, p^2 - 2\, h\, a + 2\, k\, a^2\}$ /. {p → 5, a → 6, h → 10, k → 3}

Out[624]= {75, 21}

In[625]:= `FactorInteger /@ %`

Out[625]= {{{3, 1}, {5, 2}}, {{3, 1}, {7, 1}}}

Thus

$\{a\, v + u == 75 == 3\, p^2,\ a\, v - u == 21 == 3 \times 7\}$

In[626]:= $v^2 - 1$ /. v → 8

Out[626]= 63

So Equation (34) is

$(a\, v + u)\ (a\, v - u) == (v^2 - 1)\ p^2$

$75 \times 21 == 63 \times 25$

$3 \times 25 \times 21 == 63 \times 25$

The 2nd generation from (a, 1) is

In[627]:= $\{h\, a + k\, (-p^2 + a^2),\ h + k\, a\}$ /. {p → 5, a → 6, h → 10, k → 3}

Out[627]= {93, 28}

Calculate $a\, v \pm u$ directly:

In[628]:= `{a v + u, a v - u} /. {a → 6, u → 93, v → 28}`

Out[628]= {261, 75}

Calculate using the formulas in terms of *h*, *k*:

In[629]:= $\{-k\,p^2 + 2\,h\,a + 2\,k\,a^2,\ k\,p^2\}$ /. {p → 5, a → 6, h → 10, k → 3}

Out[629]= {261, 75}

In[630]:= `FactorInteger /@ %`

Out[630]= {{{3, 2}, {29, 1}}, {{3, 1}, {5, 2}}}

$\{a\,v + u == 3^2 \times 29,\ a\,v - u == 75 == 3\,p^2\}$

In[631]:= $v^2 - 1$ /. v → 28

Out[631]= 783

In[632]:= `FactorInteger[%]`

Out[632]= {{3, 3}, {29, 1}}

So Equation (34) is

$(a\,v + u)\,(a\,v - u) == (v^2 - 1)\,p^2$

$261 \times 75 == 783 \times 25$

$261 \times 3 \times 25 == 783 \times 25$

- Now a case in which p is not prime, 6/17, $a == 11$.

In[633]:= `uvGetClasses[solveuvByReduction[6 / 17]]`

Out[633]= {{6, 0}, {11, -1}, {11, 1}, {74, -8}, {74, 8}, {249, -27}, {249, 27}, {839, -91}, {839, 91}}

The relatively prime pair of solutions other than the trivial one is {839, ±91}.

In[634]:= `D6o17 = q (q - 2 p) /. {p → 6, q → 17}`

Out[634]= 85

In[635]:= `solvePell[85]`

Out[635]= {285 769, 30 996}

First, look at the solutions in the trivial classes $\mathbb{C}_{-1}$ and $\mathbb{C}_1$.

The 2nd generation from $(a, -1)$ is

In[636]:= $\{h\,a + k\,(p^2 - a^2),\ -h + k\,a\}$ /. {p → 6, a → 11, h → 285 769, k → 30 996}

Out[636]= {508 799, 55 187}

Calculate $a\,v \pm u$ directly:

In[637]:= `{a v + u, a v - u} /. {a → 11, u → 508 799, v → 55 187}`

Out[637]= {1 115 856, 98 258}

Calculate using the formulas in terms of *h*, *k*:

In[638]:= $\{k\,p^2, -k\,p^2 - 2\,h\,a + 2\,k\,a^2\}$ /. {p → 6, a → 11, h → 285 769, k → 30 996}

Out[638]= {1 115 856, 98 258}

In[639]:= FactorInteger /@ %

Out[639]= {{{2, 4}, {3, 5}, {7, 1}, {41, 1}}, {{2, 1}, {73, 1}, {673, 1}}}

Thus

$\{a\,v + u == 2^2 \times 3^5 \times 7 \times 41 == 6^2 \times 3^3 \times 7 \times 41,\ a\,v - u == 2 \times 73 \times 673\}$

So $a\,v + u$ is a multiple of 6^2 but $a\,v - u$ has only one divisor in common with 6.

In[640]:= {1 115 856, 98 258} / 6^2

Out[640]= $\left\{30996, \frac{49129}{18}\right\}$

The 2nd generation from $(a, 1)$ is

In[641]:= $\{h\,a + k\,(-p^2 + a^2), h + k\,a\}$ /. {p → 6, a → 11, h → 285 769, k → 30 996}

Out[641]= {5 778 119, 626 725}

Calculate $a\,v \pm u$ directly:

In[642]:= {a v + u, a v - u} /. {a → 11, u → 5 778 119, v → 626 725}

Out[642]= {12 672 094, 1 115 856}

Calculate using the formulas in terms of h, k:

In[643]:= $\{-k\,p^2 + 2\,h\,a + 2\,k\,a^2, k\,p^2\}$ /. {p → 6, a → 11, h → 285 769, k → 30 996}

Out[643]= {12 672 094, 1 115 856}

In[644]:= FactorInteger /@ %

Out[644]= {{{2, 1}, {829, 1}, {7643, 1}}, {{2, 4}, {3, 5}, {7, 1}, {41, 1}}}

Thus

$\{a\,v + u == 2 \times 829 \times 7643,\ a\,v - u == 2^4 \times 3^5 \times 7 \times 41 == 6 \times 2^2 \times 3^3 \times 7 \times 41\}$

So $a\,v - u$ is a multiple of 6^2 while $a\,v + u$ has only a divisor of 2 in common with 6.

In[645]:= {12 672 094, 1 115 856} / 6^2

Out[645]= $\left\{\frac{6336047}{18}, 30996\right\}$

Turning now to the solutions (839, ±91), which are not in classes $\mathbb{C}_{-1}$ or $\mathbb{C}_1$.

In[646]:= {a v + u, a v - u} /. {a → 11, u → 839, v → 91}

Out[646]= {1840, 162}

In[647]:= FactorInteger /@ %

Out[647]= {{{2, 4}, {5, 1}, {23, 1}}, {{2, 1}, {3, 4}}}

$\{a\,v + u == 2^4 \times 5 \times 23,\ a\,v - u == 2 \times 3^4 == 6 \times 3^3\}$

Thus $a\,v - u$ is a multiple of 6 but not of 6^2, while $a\,v + u$ has only a divisor of 2 in common with 6. Neither term is divisible by p^2.

In[648]:=
```
{1840, 162} / 6^2
```

Out[648]= $\left\{\frac{460}{9}, \frac{9}{2}\right\}$

The other member of the class is (839, –91). It gives negative terms in Equation (34) so use their absolute values.

In[649]:=
```
Abs[{a v + u, a v - u} /. {a → 11, u → 839, v → -91}]
```

Out[649]= {162, 1840}

In[650]:=
```
FactorInteger /@ %
```

Out[650]= {{{2, 1}, {3, 4}}, {{2, 4}, {5, 1}, {23, 1}}}

In[651]:=
```
{162, 1840} / 6^2
```

Out[651]= $\left\{\frac{9}{2}, \frac{460}{9}\right\}$

Neither term is a multiple of 6^2. In this case it is $a\,v + u$ that is a multiple of 6 and $a\,v - u$ is only a multiple of 2.

In[652]:=
```
v^2 - 1 /. v → 91
```

Out[652]= 8280

In[653]:=
```
FactorInteger[%]
```

Out[653]= {{2, 3}, {3, 2}, {5, 1}, {23, 1}}

Then Equation (34) is

$162 \times 1840 == 8280 \times 36$

$(2 \times 3^4)\ (2^4 \times 5 \times 23) == (2^3 \times 3^2 \times 5 \times 23)\ (2^2 \times 3^2)$

$2^5 \times 3^4 \times 5 \times 23 == 2^5 \times 3^4 \times 5 \times 23$

Let's look at how $v^2 - 1 == (v - 1)(v + 1)$ matches up with the terms in the numerator:

In[654]:=
```
FactorInteger[90]
```

Out[654]= {{2, 1}, {3, 2}, {5, 1}}

In[655]:=
```
FactorInteger[92]
```

Out[655]= {{2, 2}, {23, 1}}

So neither $v - 1$ nor $v + 1$ divides either $a\,v - u$ or $a\,v + u$ separately.

These examples illustrate how for the trivial class, one of the two terms $a\,v \pm u$ is a multiple of p^2. For other classes, the divisors of p^2 are spread between the two terms such that one term is not a multiple of p.

12.5.6 Generalizing the partitioning constraint

The result of Section 12.5.5 can be generalized to composite p. Suppose $p == p_1\, p_2$ where p_1 and p_2 are divisors of p with $1 < p_{1,2} < p$. Suppose the terms on the LHS of Equation (34) partition as

$$a\,v + u == d_1\,p_1,\quad a\,v - u == d_2\,p_2\,p == d_2\,p_1\,p_2^2$$

where $d_1\,d_2 == v^2 - 1$. Adding,

$$2\,a\,v == \left(d_1 + d_2\,p_2^2\right)\,p_1$$

Subtracting,

$$2\,u == \left(d_1 - d_2\,p_2^2\right)\,p_1$$

If $p_1 == 2$, these reduce to

$$a\,v == d_1 + d_2\,p_2^2$$
$$u == d_1 - d_2\,p_2^2$$

which do not constrain gcd(u, v).

If $p_1 > 2$, then u, v have a common divisor of p_1 (if odd) or $p_1/2$ (if even). Therefore this partitioning is disallowed if $p_1 > 2$.

The following partitioning is allowed if gcd(p_1, p_2) == 1:

$$a\,v + u == d_1\,p_1^2,\quad a\,v - u == d_2\,p_2^2$$

These results are not very useful for ruling out solutions, but they do put some constraints. For instance, for $p == 9$, the only partitionings of p^2 are

$$\{1,\,3^4\},\ \{3,\,3^2\},\ \{3^2,\,3^2\}$$

Only the first is allowable. So ratios with $p == 9$ have no other relatively prime solutions besides those in $\mathbb{C}_{-1}$ and $\mathbb{C}_1$. They may, and often do, have solutions with gcd of 3. For example, the last two here:

```
In[656]:= uvGetClasses[solveuvByRecursiveReduction[9 / 19]]

Out[656]= {{9, 0}, {10, -1}, {10, 1}, {66, -15}, {66, 15}}
```

12.6 If p == 4, the only solution classes are the trivial-solution classes

(To avoid an overly lengthy section title, the assumption that the case is hyperbolic with nonsquare D is implicit.)

The claim formally stated is:

- If $p == 4$, $q > 8$, and D nonsquare, there are no solutions to Equation (2) except those arising from applying the Pell recurrence to the trivial solutions. In other words, these cases have only the 3 classes of solutions that always exist.

The proof has two parts:

1. If $p == 4$, $q > 8$, and D nonsquare, then the only relatively prime solutions are those belonging to classes $\mathbb{C}_{-1}$ and $\mathbb{C}_1$ (as defined in Section 12.5). These are the solutions of Equation (11) directly.
2. If $p == 4$, $q > 8$, and D nonsquare, then there are no solutions with $\gcd(u, v) == 2$. These would be obtained from relatively prime solutions of Equation (11) divided by 4.

The other solutions arise from Equation (11) divided by 16. These solutions necessarily belong to class $\mathbb{C}_0$.

12.6.1 Proof that the relatively prime solutions for $p == 4$ belong to the trivial-solution classes

Suppose $u^2 - D v^2 == p^2 == 16$, with $\gcd(u, v) == 1$. As in Section 12.5.2, define $a == q - p$, then $D == a^2 - p^2 == a^2 - 16$. Since $\gcd(a, p) == 1$, here a must be odd, hence D odd. Then $u^2 - D v^2 == 16$ requires that u, v be of the same parity. They cannot both be even, since they are relatively prime. Hence both u and v are odd.

In Section 12.5.2 we showed that Equation (11) can be rewritten in factored form (Equation (34)):

$$(a\, v - u)\, (a\, v + u) == \left(v^2 - 1\right) p^2 == 16 \left(v^2 - 1\right)$$

In Section 12.5.4 we showed that if one of the terms $a\, v - u$ or $a\, v + u$ is a multiple of p^2, then $(u, v) \in \mathbb{C}_{\pm 1}$, and in Section 12.5.5 we showed that both terms cannot simultaneously be multiples of p. These results hold for composite p. Therefore, the only partitioning of p^2 between the two terms that can give a solution not in a trivial class is if one term is divisible by a non-trivial divisor d of p but not by p itself, and the other term is divisible by p^2 / d but not by p^2. The only non-trivial divisor of 4 is 2. Let

$$a\, v - u == 2\, d_1, \quad a\, v + u == 8\, d_2$$

where $d_1\, d_2 == v^2 - 1$. Both d_1 and d_2 must be odd to prevent $a\, v - u$ being a multiple of 4 and $a\, v + u$ a multiple of 16. But $d_1\, d_2 == v^2 - 1$ is even since v is odd, requiring d_1 or d_2 to be even. Hence this partitioning is disallowed, and the relatively prime solutions must be in $\mathbb{C}_{-1}$ or $\mathbb{C}_1$.

12.6.2 Proof that for $p == 4$ there are no solutions with $\gcd(u, v) == 2$

The other potential source of solutions that are not in a trivial class is from Equation (11) divided by 4. We can rule out the existence of solutions for $p == 4$.

Assume $(2\, u, 2\, v)$ satisfies Equation (11), where $\gcd(u, v) == 1$. Then (u, v) satisfies Equation (11) divided by 4.

$$u^2 - D\, v^2 == 4$$

From the fact that D is odd and $\gcd(u, v) == 1$ we again have that u and v must both be odd.

Putting $D == a^2 - p^2 == a^2 - 16$, we can factor Equation (11) as follows:

$$u^2 - \left(a^2 - 16\right) v^2 == 4$$
$$u^2 - a^2 v^2 == 4 \left(1 - 4 v^2\right)$$
$$(a v - u) (a v + u) == 4 \left(4 v^2 - 1\right)$$

Here $4 v^2 - 1$ is odd, so the product of the two terms on the left is a multiple of 4 but not of 8. They can partition with each being a multiple of 2, or with one being odd and the other a multiple of 4. We deal with each possibility in turn.

Suppose the terms partition as

$$a v + u == 2 d_1, \quad a v - u == 2 d_2$$

where $d_1 d_2 == 4 v^2 - 1$. Since $4 v^2 - 1$ is odd, both d_1 and d_2 must be odd. Adding,

$$2 a v == 2 (d_1 + d_2)$$
$$a v == (d_1 + d_2)$$

The sum $d_1 + d_2$ is even, and since a is odd, v must be even, contradicting what was shown above.

Subtracting,

$$2 u == 2 (d_1 - d_2)$$
$$u == (d_1 - d_2)$$

Again, this is even, contradicting u odd.

The other partitioning is

$$a v + u == d_1, \quad a v - u == 4 d_2$$

Since $a v$ and u are odd, their sum is even, requiring d_1 to be even, a contradiction.

Hence neither partitioning is allowed, and there can therefore be no solutions. This concludes the proof that for $p == 4$, $p/q < 1/2$, and D nonsquare, the only solutions are those in the three trivial classes.

13 Solving using Mathematica

Mathematica is able to find solutions when they exist, working directly with x, y numbers of red and blue balls. Individual solutions, e.g. for $p/q == 4/11$:

```
In[657]:= FindInstance[{probdifferent[{x, y}] == 4 / 11, x > 0, y > 0}, {x, y}, Integers]
Out[657]= {{x → 1072, y → 3417}}
```

Constraining the solutions to be positive excluded inadmissible trivial and negative solutions.

A case that has only one distinct admissible solution:

```
In[658]:= FindInstance[{probdifferent[{x, y}] == 126 / 247, x > 0, y ≥ x}, {x, y}, Integers, 2]
Out[658]= {{x → 18, y → 21}}
```

Admissible solutions do not always exist, e.g. for $p/q == 4/9$.

In[659]:=
```
FindInstance[{probdifferent[{x, y}] == 4 / 9, x > 0, y > 0}, {x, y}, Integers]
```

Out[659]= { }

Mathematica only gives the empty set as a result when it can prove there are no solutions.

These results are not systematic. Mathematica can also give a formula for systematically finding all solutions.

In[660]:=
```
mmareduce4o11 = Reduce[probdifferent[{x, y}] == 4 / 11 && x > 0 && y > 0, {x, y}, Integers]
```

Out[660]=

$$\Big(c_1 \in \mathbb{Z} \,\&\&\, c_1 \geq 1 \,\&\&\, x == \frac{1}{132}\Big(-7\Big(22+\frac{1}{2}\Big(-55\left(23-4\sqrt{33}\right)^{2c_1}+9\sqrt{33}\left(23-4\sqrt{33}\right)^{2c_1}-55\left(23+4\sqrt{33}\right)^{2c_1}-9\sqrt{33}\left(23+4\sqrt{33}\right)^{2c_1}\Big)\Big)+33\Big(2+\frac{1}{6}\Big(-27\left(23-4\sqrt{33}\right)^{2c_1}+5\sqrt{33}\left(23-4\sqrt{33}\right)^{2c_1}-27\left(23+4\sqrt{33}\right)^{2c_1}-5\sqrt{33}\left(23+4\sqrt{33}\right)^{2c_1}\Big)\Big)\Big) \,\&\&\, y == \frac{1}{33}\Big(-22+\frac{1}{2}\Big(55\left(23-4\sqrt{33}\right)^{2c_1}-9\sqrt{33}\left(23-4\sqrt{33}\right)^{2c_1}+55\left(23+4\sqrt{33}\right)^{2c_1}+9\sqrt{33}\left(23+4\sqrt{33}\right)^{2c_1}\Big)\Big)\Big) \,||$$

$$\Big(c_1 \in \mathbb{Z} \,\&\&\, c_1 \geq 1 \,\&\&\, x == \frac{1}{132}\Big(-7\Big(22-11\left(23-4\sqrt{33}\right)^{2c_1}+\sqrt{33}\left(23-4\sqrt{33}\right)^{2c_1}-11\left(23+4\sqrt{33}\right)^{2c_1}-\sqrt{33}\left(23+4\sqrt{33}\right)^{2c_1}\Big)+33\Big(2+\frac{1}{3}\Big(-3\left(23-4\sqrt{33}\right)^{2c_1}+\sqrt{33}\left(23-4\sqrt{33}\right)^{2c_1}-3\left(23+4\sqrt{33}\right)^{2c_1}-\sqrt{33}\left(23+4\sqrt{33}\right)^{2c_1}\Big)\Big)\Big) \,\&\&\, y == \frac{1}{33}\Big(-22+11\left(23-4\sqrt{33}\right)^{2c_1}-\sqrt{33}\left(23-4\sqrt{33}\right)^{2c_1}+11\left(23+4\sqrt{33}\right)^{2c_1}+\sqrt{33}\left(23+4\sqrt{33}\right)^{2c_1}\Big)\Big) \,||$$

$$\Big(c_1 \in \mathbb{Z} \,\&\&\, c_1 \geq 1 \,\&\&\, x == \frac{1}{132}\Big(-7\Big(22-11\left(23-4\sqrt{33}\right)^{2c_1}-\sqrt{33}\left(23-4\sqrt{33}\right)^{2c_1}-11\left(23+4\sqrt{33}\right)^{2c_1}+\sqrt{33}\left(23+4\sqrt{33}\right)^{2c_1}\Big)+33\Big(2+\frac{1}{3}\Big(3\left(23-4\sqrt{33}\right)^{2c_1}+\sqrt{33}\left(23-4\sqrt{33}\right)^{2c_1}+3\left(23+4\sqrt{33}\right)^{2c_1}-\sqrt{33}\left(23+4\sqrt{33}\right)^{2c_1}\Big)\Big)\Big) \,\&\&\, y == \frac{1}{33}\Big(-22+11\left(23-4\sqrt{33}\right)^{2c_1}+\sqrt{33}\left(23-4\sqrt{33}\right)^{2c_1}+11\left(23+4\sqrt{33}\right)^{2c_1}-\sqrt{33}\left(23+4\sqrt{33}\right)^{2c_1}\Big)\Big) \,||$$

$$\Big(c_1 \in \mathbb{Z} \,\&\&\, c_1 \geq 1 \,\&\&\, x == \frac{1}{132}\Big(-7\Big(22-11\left(23-4\sqrt{33}\right)^{2c_1}-\sqrt{33}\left(23-4\sqrt{33}\right)^{2c_1}-11\left(23+4\sqrt{33}\right)^{2c_1}+\sqrt{33}\left(23+4\sqrt{33}\right)^{2c_1}\Big)+33\Big(2+\frac{1}{3}\Big(-3\left(23-4\sqrt{33}\right)^{2c_1}-\sqrt{33}\left(23-4\sqrt{33}\right)^{2c_1}-3\left(23+4\sqrt{33}\right)^{2c_1}+\sqrt{33}\left(23+4\sqrt{33}\right)^{2c_1}\Big)\Big)\Big) \,\&\&$$

$$y == \frac{1}{33}\left(-22+11\left(23-4\sqrt{33}\right)^{2\,c_1}+\sqrt{33}\left(23-4\sqrt{33}\right)^{2\,c_1}+11\left(23+4\sqrt{33}\right)^{2\,c_1}-\sqrt{33}\left(23+4\sqrt{33}\right)^{2\,c_1}\right)\Bigg)\;||$$

$$\Bigg(c_1\in\mathbb{Z}\;\&\&\;c_1\geq 1\;\&\&\;x == \frac{1}{132}\Bigg(-7\left(22-11\left(23-4\sqrt{33}\right)^{2\,c_1}+\sqrt{33}\left(23-4\sqrt{33}\right)^{2\,c_1}-11\left(23+4\sqrt{33}\right)^{2\,c_1}-\sqrt{33}\left(23+4\sqrt{33}\right)^{2\,c_1}\right)+33\left(2+\frac{1}{3}\left(3\left(23-4\sqrt{33}\right)^{2\,c_1}-\sqrt{33}\left(23-4\sqrt{33}\right)^{2\,c_1}+3\left(23+4\sqrt{33}\right)^{2\,c_1}+\sqrt{33}\left(23+4\sqrt{33}\right)^{2\,c_1}\right)\right)\Bigg)\;\&\&$$

$$y == \frac{1}{33}\left(-22+11\left(23-4\sqrt{33}\right)^{2\,c_1}-\sqrt{33}\left(23-4\sqrt{33}\right)^{2\,c_1}+11\left(23+4\sqrt{33}\right)^{2\,c_1}+\sqrt{33}\left(23+4\sqrt{33}\right)^{2\,c_1}\right)\Bigg)\;||\;\Bigg(c_1\in\mathbb{Z}\;\&\&\;c_1\geq 1\;\&\&$$

$$x == \frac{1}{132}\Bigg(33\left(2+\frac{1}{6}\left(-27\left(23-4\sqrt{33}\right)^{2\,c_1}-5\sqrt{33}\left(23-4\sqrt{33}\right)^{2\,c_1}-27\left(23+4\sqrt{33}\right)^{2\,c_1}+5\sqrt{33}\left(23+4\sqrt{33}\right)^{2\,c_1}\right)\right)-7\left(22+\frac{1}{2}\left(-55\left(23-4\sqrt{33}\right)^{2\,c_1}-9\sqrt{33}\left(23-4\sqrt{33}\right)^{2\,c_1}-55\left(23+4\sqrt{33}\right)^{2\,c_1}+9\sqrt{33}\left(23+4\sqrt{33}\right)^{2\,c_1}\right)\right)\Bigg)\;\&\&$$

$$y == \frac{1}{33}\left(-22+\frac{1}{2}\left(55\left(23-4\sqrt{33}\right)^{2\,c_1}+9\sqrt{33}\left(23-4\sqrt{33}\right)^{2\,c_1}+55\left(23+4\sqrt{33}\right)^{2\,c_1}-9\sqrt{33}\left(23+4\sqrt{33}\right)^{2\,c_1}\right)\right)\Bigg)$$

In[661]:= **Length[mmareduce4o11]**

Out[661]= 6

Not pretty. There are 6 families of solutions. I did not use `Simplify` because it groups two *y* solutions for each *x* solution, making the next steps more complicated. But `Simplify` does make the expressions less fearsome, so I use it in the next step.

Now pick out the *x* and *y* solutions from this mess. The first index selects one solution, the next index selects *x* or *y*, and the third index selects the RHS of the equals sign.

In[662]:=
```
mmaxsolns4o11 =
  Table[Simplify[mmareduce4o11[[i]][[3]][[2]]], {i, Length[mmareduce4o11]}]
```

Out[662]=
$$\left\{\frac{1}{33}\left(-22-\left(23-4\sqrt{33}\right)^{2c_1}\left(-11+\sqrt{33}\right)+\left(11+\sqrt{33}\right)\left(23+4\sqrt{33}\right)^{2c_1}\right),\right.$$
$$\frac{1}{33}\left(-22+\left(23-4\sqrt{33}\right)^{2c_1}\left(11+\sqrt{33}\right)-\left(-11+\sqrt{33}\right)\left(23+4\sqrt{33}\right)^{2c_1}\right),$$
$$\frac{1}{66}\left(-44+\left(55-9\sqrt{33}\right)\left(23+4\sqrt{33}\right)^{2c_1}+\left(23-4\sqrt{33}\right)^{2c_1}\left(55+9\sqrt{33}\right)\right),$$
$$\frac{1}{33}\left(-22-\left(23-4\sqrt{33}\right)^{2c_1}\left(-11+\sqrt{33}\right)+\left(11+\sqrt{33}\right)\left(23+4\sqrt{33}\right)^{2c_1}\right),$$
$$\frac{1}{66}\left(-44+\left(55-9\sqrt{33}\right)\left(23-4\sqrt{33}\right)^{2c_1}+\left(23+4\sqrt{33}\right)^{2c_1}\left(55+9\sqrt{33}\right)\right),$$
$$\left.\frac{1}{33}\left(-22+\left(23-4\sqrt{33}\right)^{2c_1}\left(11+\sqrt{33}\right)-\left(-11+\sqrt{33}\right)\left(23+4\sqrt{33}\right)^{2c_1}\right)\right\}$$

In[663]:=
```
mmaysolns4o11 =
  Table[Simplify[mmareduce4o11[[i]][[4]][[2]]], {i, Length[mmareduce4o11]}]
```

Out[663]=
$$\left\{\frac{1}{66}\left(-44+\left(55-9\sqrt{33}\right)\left(23-4\sqrt{33}\right)^{2c_1}+\left(23+4\sqrt{33}\right)^{2c_1}\left(55+9\sqrt{33}\right)\right),\right.$$
$$\frac{1}{33}\left(-22-\left(23-4\sqrt{33}\right)^{2c_1}\left(-11+\sqrt{33}\right)+\left(11+\sqrt{33}\right)\left(23+4\sqrt{33}\right)^{2c_1}\right),$$
$$\frac{1}{33}\left(-22+\left(23-4\sqrt{33}\right)^{2c_1}\left(11+\sqrt{33}\right)-\left(-11+\sqrt{33}\right)\left(23+4\sqrt{33}\right)^{2c_1}\right),$$
$$\frac{1}{33}\left(-22+\left(23-4\sqrt{33}\right)^{2c_1}\left(11+\sqrt{33}\right)-\left(-11+\sqrt{33}\right)\left(23+4\sqrt{33}\right)^{2c_1}\right),$$
$$\frac{1}{33}\left(-22-\left(23-4\sqrt{33}\right)^{2c_1}\left(-11+\sqrt{33}\right)+\left(11+\sqrt{33}\right)\left(23+4\sqrt{33}\right)^{2c_1}\right),$$
$$\left.\frac{1}{66}\left(-44+\left(55-9\sqrt{33}\right)\left(23+4\sqrt{33}\right)^{2c_1}+\left(23-4\sqrt{33}\right)^{2c_1}\left(55+9\sqrt{33}\right)\right)\right\}$$

Now calculate the first 5 families of solutions. Use Flatten to remove the grouping into 6 solutions per family.

In[664]:=
```
mmaxvalues4o11 = Flatten[Table[Simplify[mmaxsolns4o11 /. C[1] → n], {n, 1, 5}]]
```

Out[664]=
```
{1072, 336, 105, 1072, 3417, 336, 2 267 616, 711 712, 223 377, 2 267 616, 7 224 945,
 711 712, 4 793 740 560, 1 504 560 240, 472 220 281, 4 793 740 560, 15 273 531 721,
 1 504 560 240, 10 133 965 277 632, 3 180 639 637 056, 998 273 452 065, 10 133 965 277 632,
 32 288 238 834 657, 3 180 639 637 056, 21 423 197 803 174 896, 6 723 870 688 177 552,
 2 110 349 605 446 537, 21 423 197 803 174 896, 68 257 321 622 934 585, 6 723 870 688 177 552}
```

```
In[665]:= mmayvalues4o11 = Flatten[Table[Simplify[mmaysolns4o11 /. C[1] → n], {n, 1, 5}]]

Out[665]= {3417, 1072, 336, 336, 1072, 105, 7 224 945, 2 267 616, 711 712, 711 712, 2 267 616,
  223 377, 15 273 531 721, 4 793 740 560, 1 504 560 240, 1 504 560 240, 4 793 740 560,
  472 220 281, 32 288 238 834 657, 10 133 965 277 632, 3 180 639 637 056, 3 180 639 637 056,
  10 133 965 277 632, 998 273 452 065, 68 257 321 622 934 585, 21 423 197 803 174 896,
  6 723 870 688 177 552, 6 723 870 688 177 552, 21 423 197 803 174 896, 2 110 349 605 446 537}
```

Now display the same results in a nice tabular form, sorted from smallest to largest. Sort the individual solutions and remove duplicates to obtain only distinct solutions with $x \le y$.

```
In[666]:= TableForm[Sort[DeleteDuplicates[
    Table[Sort[{mmaxvalues4o11[[i]], mmayvalues4o11[[i]]}],
     {i, 1, Length[mmaxvalues4o11]}]]], TableHeadings → {None, {x, y, "p/q"}}]

Out[666]//TableForm=
```

x	y
105	336
336	1072
1072	3417
223 377	711 712
711 712	2 267 616
2 267 616	7 224 945
472 220 281	1 504 560 240
1 504 560 240	4 793 740 560
4 793 740 560	15 273 531 721
998 273 452 065	3 180 639 637 056
3 180 639 637 056	10 133 965 277 632
10 133 965 277 632	32 288 238 834 657
2 110 349 605 446 537	6 723 870 688 177 552
6 723 870 688 177 552	21 423 197 803 174 896
21 423 197 803 174 896	68 257 321 622 934 585

14 Open questions

The most important questions, about the existence of solutions and methods that can find all solutions, have been answered. I am sure there are many more interesting features of this problem that remain to be discovered. Here are a few possible directions to explore. I would welcome hearing of any results you find.

14.1 Special cases

We examined some special cases for which solutions could be found without sophisticated methods. The *Varsity Math* case $p/q == 1/2$ is one. In Sections 6.3 and 12.4 we examined the special cases $p == 1$ and $p == 2$. Like all p, these have a trivial solution $(u, v) == (q - p, 1)$, but unlike $p > 2$, for each of these there is a recurrence that generates all solutions without solving the Pell equation.

In the elliptical regime, the family of probability ratios of the form $p/(2\,p - 1)$ have three solutions at and adjacent to the far vertex, as discussed in Section 5.2.1 and Section 8.5. This does not say anything about the possible existence of other solutions. In Section 8.5.2 we saw some sub-cases that have solutions around the midsection of the ellipse.

In the hyperbolic regime, Section 11.15 analyzed a subset of ratios of the form $p/(2\,p + 1)$ for which three solutions can be found from simple formulas. There does not seem to be a simple formula giving admissible solutions for all ratios of this form. (They do have integer solutions at the vertex $(x, y) == (-p, -p)$ and its neigbors 1 unit away.)

Are there other families of probability ratios that similarly guarantee certain solutions?

14.2 Properties of the recycling recurrence

Recycling triplets are common for cases in the hyperbolic regime, since, as shown in Section 12.3 they always arise by applying the Pell recurrence to the three trivial solutions. For $p == 1$ and $p == 2$, the recycling recurrence yields an infinite series of solutions. But for $p > 2$, I have not encountered any cases for which the solutions include a series of more than 3 solutions related by the recycling recurrence. Applying the recycling recurrence to the end members of these series in either direction yields fractional values. For some cases there are also singlets, solutions that are not related to any other by the recycling recurrence. There are also doublets, pairs of solutions that are recycling neighbors without a third. Triplets can arise among solutions that are not the in the trivial-solution classes.

Can it be proved that there are no quartets or larger series of recycling neighbors for $p > 2$?