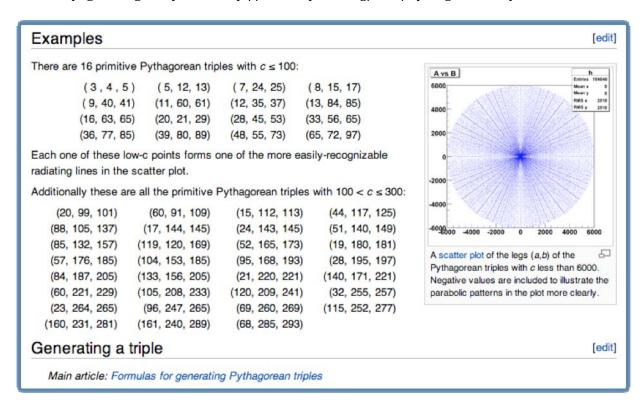
PYTHAGOREAN TRIPLES

APR. 28, 2012 at the CMC³ Recreational Mathematics Conference in S. Tahoe. By Mark Harbison. harbism@scc.losrios.edu c: 916-475-9461

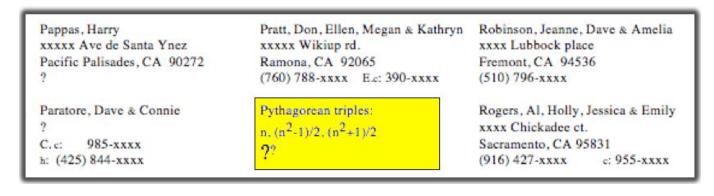
This web page changed my life. http://en.wikipedia.org/wiki/Pythagorean_triple



I spent every "spare moment" for about 3 months in early 2011 trying to generate that image myself. It plots points $(-6000 \le a \le 6000)$ by $(-6000 \le b \le 6000)$ such that $a^2 + b^2$ is a perfect square.

I first tried to just find a large data set that I could copy to use for my own scatterplot. But every web page that mentioned "Pythagorean Triples" discussed the properties of them or showed just a few examples. It's easier to get just a sample than a complete population of the #s.

There are many formulas available for generating Primitive Pythagorean Triples (PPTs). PPT's are like (3, 4, 5) where all pairs of #s are relatively-prime. A non-primitive example is (6, 8, 10) since it's twice (3, 4, 5). I tried my old favorite (from my personal address book):



It would be nice if PPTs depended on only 1 variable n. But my formula should have said "for odd n: { n, $(n^2-1)/2$, $(n^2+1)/2$ }". Ex. $(6^2-1)/2 = 35/2$ is not a whole #. And start at $n \ge 3$ (not 1), since $(1^2-1)/2 = 0$ is not a triangle length.

3	4	5	$3, (3^2-1)/2, (3^2+1)/2$	also	$2^2 \pm 1^2$ and $2*2*1$
5	12	13	$5, (5^2-1)/2, (5^2+1)/2$	also	$3^2 \pm 2^2$ and $2*3*2$
7	24	25	$7, (7^2-1)/2, (7^2+1)/2$	also	$4^2 \pm 3^2$ and $2*4*3$
9	40	41	$9, (9^2-1)/2, (9^2+1)/2$	also	$5^2 \pm 4^2$ and $2*5*4$
11	60	61	$11, (11^2-1)/2, (11^2+1)/2$	also	$6^2 \pm 5^2$ and $2*6*5$
13	84	85	$13, (13^2-1)/2, (13^2+1)/2$	also	$7^2 \pm 6^2$ and $2*7*6$
15	112	113	$15, (15^2-1)/2, (15^2+1)/2$	also	$8^2 \pm 7^2$ and $2*8*7$
17	144	145	$17, (17^2-1)/2, (17^2+1)/2$	also	$9^2 \pm 8^2$ and $2*9*8$
19	180	181	$19, (19^2-1)/2, (19^2+1)/2$	also	$10^2 \pm 9^2$ and $2*10*9$
	etc.				

This is a good start. But personally, I was not yet satisfied because I knew that other PPTs existed that were not found with this method. For example, why did this method skip (8, 15, 17)? I wanted a complete list, not just a partial list. Especially for variety: c - b may be 2 (e.g. 17-15), not necessarily a difference of 1. Or like how (36, 77, 85) have no differences of 1 at all.

Then I noticed that $15 = 4^2 - 1^2$ and $17 = 4^2 + 1^2$ and $8 = 2 \cdot 4 \cdot 1$. It's better to have 2 variables (m & n) than just n. Originally, I did not allow for m & n to be more than 1 apart (ex. m = 4 & n = 1). My address book formula was just a *special case* of the more-general **Euclid's formula** for a PPT: (a = $m^2 - n^2$, b = 2*m*n, c = $m^2 + n^2$) for any relatively-prime m > n with opposite parity.

The case "not both even" is already covered by the relatively-prime condition, but for reason's that I am not going into today, it's also necessary that m & n are "not both odd". So the conditions could have said (with more words) "relatively-prime m > n that are not both odd".

Note:
$$a^2 + b^2 = (m^2 - n^2)^2 + (2mn)^2 = (m^2 + n^2)^2 = c^2$$
.

Euclid's formula is nice, but I kept reading http://en.wikipedia.org/wiki/Pythagorean_triple for other ideas, anyway. I guess that I had too much spare time on my hands. Here are just a few examples (of the 36 "elementary properties" listed),

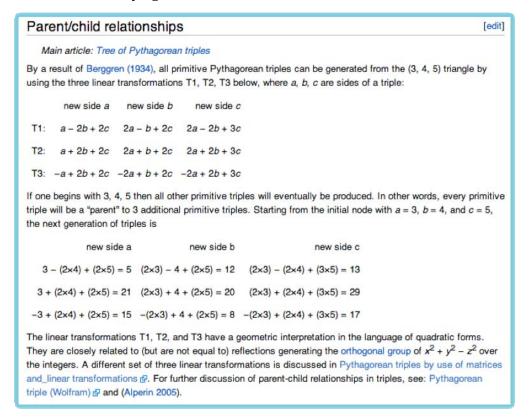
- At most one of *a*, *b*, *c* is a perfect square.
- All prime factors of c is of the form 4n+1.
- Exactly one of a, b, (a + b), (b a) is divisible by 7.
- (I'll skip the similar properties for divisibility by 3, 4, 5, 8, 9, 11 and 13).
- Every integer greater than 2 that is not congruent to 2 mod 4 is part of a PPT. (But where?)

I'll skip the relationships to areas, perimeters, inscribed circles and the Platonic sequence (though it is interesting to see that both Plato and Euclid spend so much time on them).

Neither am I personally interested in how PPT's relate to stereographic projections of unit circles to the x-axis, or spinors for the Lorenz group SO(1, 2) and group theory. I also skipped the sections on Gaussian Integers and generalizations to Pythagorean n-tuples or nth powers.

Someday, I do want to read more about Heronian triangles (not necessarily 90° triangles), but I've got a rather large to-do list, already. And there was no scatterplot image of these that caught my attention.

Then I made the mistake of trying to use this information—which was new to me:

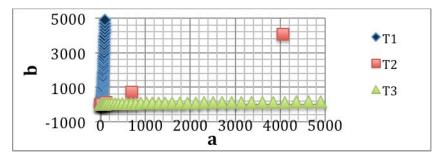


That sounds easy. And wouldn't it be cool to see a complete list of every PPT generated from the same root (3, 4, 5)? I'm excited. I typed 3 in cell R2, 4 in S2, 5 in T2, and = R2-2*S2+2*T2 in cell R5, =2*R2-S2+2*T2 in R6, and =2*R2-2*S2+3*T2 in T5. Then after a few days of cut-and-paste, I can see all of the generations.

	0th. Step	_			2r	nd. Ste	p				4	th. Ste	₽p	_		
gen. = 0	3 4 5	3	4	5	3	4	5	3	4	5	3	4	5	3	4	5
gen. = 1	T1	5	12	13	5	12	13	5	12	13	5	12	13	5	12	13
	T2	21	20	29	21	20	29	21	20	29	21	20	29	21	20	29
	Т3	15	8	17	15	8	17	15	8	17	15	8	17	15	8	17
					Ī											
gen. = 2			Т	1	7	24	25	7	24	25	7	24	25	7	24	25
			Т	2	55	48	73	55	48	73	55	48	73	55	48	73
			Т	3	45	28	53	45	28	53	45	28	53	45	28	53
			Т	1	39	80	89	39	80	89	39	80	89	39	80	89
			Т	2	119	120	169	119	120	169	119	120	169	119	120	169
			Т	3	77	36	85	77	36	85	77	36	85	77	36	85
			Т	1	33	56	65	33	56	65	33	56	65	33	56	65
			Т	2	65	72	97	65	72	97	65	72	97	65	72	97
			Т	3	35	12	37	35	12	37	35	12	37	35	12	37

OK, I changed my mind. This **TI T2 T3 method** is producing some reduntant PPTs after a lot of work. And after 6 generations, it missed the PPT (65, 72, 97). I don't have the energy to go past gen 6 since the next level would take $3^7 = 2187$ rows and quite a few columns (3 columns per triple).

One more try at this T1 T2 T3 method used just the same transormation repeatedly. It was disappointing.

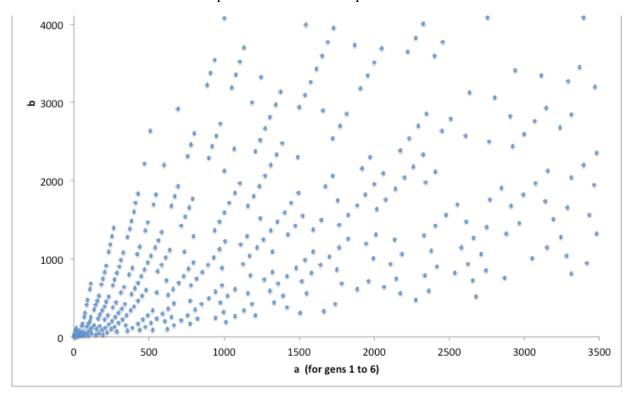


Well, at least that got me started. I can use those results in a new "sheet" on my Excel file that I decided to call "40 multiples". Maybe I'll understand Pythagorean Triples better if I allow for non-primitive ones and I can find patterns in the scatterplot. I decided to rank the (a, b) by the highest # (a < b only).

gens.	1 to 4	max(a,b)		2000	
3	4	4	one		
5	12	12	two		
15	8	15	three	1800	
21	20	21	four	1000	
7	24	24	five		
35	12	35	six		
9	40	40	seven	1600	
45	28	45	eight		
55	48	55	nine		
33	56	56	ten	4.400	
63	16	63	eleven	1400	
65	72	72	twelve		00 000
77	36	77	thirteen		
39	80	80	fourteen	1200	
91	60	91	fifteen	1200	0 000 M a 8 000 M
105	88	105	sixteen		
117	44	117	seventeen		
119	120	120	eighteen	1000	010 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
85	132	132	nineteen		
51	140	140	twenty		000 00 00 00 00 00 00 00 00 00 00 00 00
133	156	156	twenty-one	000	
165	52	165	twenty-two	800	000 000 000
95	168	168	twenty-three		00000
57	176	176	twenty-four		
187	84	187	twenty-five	600	
105	208	208	twenty-six	000	
209	120	209	twenty-seven		
207	224	224	twenty-eight		
115	252	252	twenty-nine	400	
273	136	273	thirty		100 pt 10
275	252	275	thirty-one		
175	288	288	thirty-two	000	
299	180	299	thirty-three	200	
297	304	304	thirty-four		
319	360	360	thirty-five		37
377	336	377	thirty-six	0	
403	396	403	thirty-seven	U	THE PROPERTY AND THE PR
217	456	456	thirty-eight		0 500 1000 1500 2000
459	220	459	thirty-nine		
697	696	697	forty		

No, that didn't help me get very close to the original scatterplot on page 1 of this handout. It was fun (of course). But if there really are predictable patterns here, I'll have to look somewhere else to find them.

At this point, I thought that I should leave out non-PPTs. I started using the GCD() function in Excel to return the Greatest Common Divisor of a pair of numbers. A PPT has GCD = 1. This helped me make a scatterplot that had a good mix of patterns and randomness, but left out so many blanks. Would this match the Wiki scatterplot if I included multiples of them or not?



I now had a way to get the exact values of points in a scatterplot (by moving the mouse over a point). My most-exciting breakthrough was when I guessed that these points were on the same parabola:

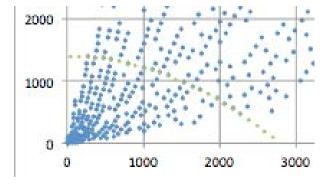
It wouldn't be good enough to be just "close" to a parabola. They need to be exact.

I used the QuadReg feature on a TI-83 to get $y = (-0.0001652893)x^2 + 0x + 1512.5$ and $R^2 = 1$. I'm not sure how I did it (since the \rightarrow Frac button fails), but I guessed that 0.0001652893 = 1/(2*55*55) and $\sqrt{2*1512.5} = 55$.

Now in the function $y = (-1/(2*55^2)) x^2 + 1512.5$ the Domain must use multiples of 55: $\{55, 165, 275, ..., 2915, 3025\}$ to get whole # Range $\{1512, 1508, 1500, ..., 108, 0\}$ but we'll ignore the 0 distance.

These points are plotted in green on top of the previous scatterplot in blue.

Hooray! Both the table and the graph matched!



To test my hypothesis, I tried another parabola. And by the 3rd parabola, I was convinced that I really did have a pattern—not just a coincidence.

				n-obvious = 16 p				ere obv		4.5 and 5618 =	- 7*52*52
	exactly on $y = (-1/15842)x^2+3960.5$ and $15842 = 2*89*89$ but 29 others were not yet on it $(89-1)/2 = 44$							(-1/3016)/X Z+140	4.5 and 5016	- 2 33 33
89	3960	3961	g > 6	1*89 = 89	2*44*(89-44) = 3960		(53-1)	/2 -	26		
267	3956	3965	g > 6	3*89 = 267	2*43*(89-43) = 3956	53	1404	1405	g > 6	1*53	
445	3948	3973	g > 6	5*89 = 445	etc.	159	1400	1409	g > 6	3*53	
623	3936	3985	g > 6	etc.	{44 nonzero's of these}	265	1392	1417	gen. = 6	5*53	
801	3920	4001	g > 6		,	371	1380	1429	gen. = 5	etc.	
979	3900	4021	g > 6			477	1364	1445	gen. = 6	ctc.	
1157	3876	4045	g > 6	sqrt(2*3960.5)	- 80	583	1344	1465			
		4073	ŭ	3411(2 3300.3)	- 65				gen. = 6		
1335	3848		g > 6			689	1320	1489	g > 6		
1513	3816	4105	g > 6			795	1292	1517	gen. = 5		
1691	3780		gen. = 6	also (89^2-1)/2	2 = 3960	901	1260	1549	g > 6		
1869	3740	4181	gen. = 5			1007	1224	1585	gen. = 5		
2047	3696	4225	gen. = 6			1113	1184	1625	gen. = 6		
2225	3648	4273	gen. = 5			1219	1140	1669	gen. = 5		
2403	3596	4325	gen. = 6			1325	1092	1717	gen. = 5		
2581	3540	4381	g > 6			1431	1040	1769	g > 6		
2759	3480	4441	g > 6			1537	984	1825	gen. = 5		
2937	3416		gen. = 5			1643	924	1885	gen. = 5		
3115	3348		gen. = 5			1749	860	1949	_		
									gen. = 5		
3293	3276		gen. = 5			1855	792	2017	g > 6		
3471	3200		gen. = 5			1961	720	2089	gen. = 5		
3649	3120	4801	gen. = 6			2067	644	2165	gen. = 5		
3827	3036	4885	g > 6			2173	564	2245	gen. = 6		
4005	2948	4973	g > 6			2279	480	2329	gen. = 6		
4183	2856	5065	g > 6			2385	392	2417	g > 6		
4361	2760		gen. = 6			2491	300	2509	g > 6		
4539	2660	5261	g > 6			2597	204	2605	g > 6		
4717 4895	2556 2448	5365 5473	g > 6 gen. = 5			2703 2809	104 0	2705 2809	g > 6		
						2003	U	2005	g > 6		
			gen. = 5			2*25*	(F2 26)	4.40.4			
5251	2220	5701	_				(53-26)				
5429	2100		gen. = 6				(53-25)				
5607	1976		gen. = 6			2*24*	(53-24)	= 1392			
5785	1848	6073	gen. = 5				etc.				
5963	1716	6205	g > 6			{26 no	nzero's	of these	2)		
6141	1580	6341	g > 6								
6319	1440	6481	g > 6			- 4 - 44					
6497	1296	6625	g > 6			•	53-4) = 3				
6675	1148	6773	g > 6				53-3) = 3 53-2) = 2				
6853	996	6925	g > 6		2*6*(89-6) = 996	-	53-2) = 2 53-1) = 2				
7031	840	7081	_		2*5*(89-5) = 840	0*(53-		104			
			g > 6			0 (33-	0j – U				
7209	680	7241	g > 6		2*4*(89-4) = 680						
7387	516	7405	g > 6		2*3*(89-3) = 516						
7565	348	7573	g > 6	85*89 = 7565	2*2*(89-2) = 348						

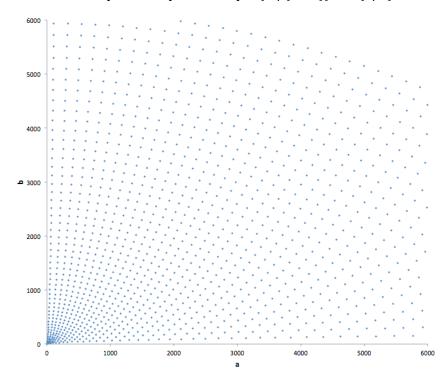
```
7743 176 7745 g > 6 87*89 = 7743 2*1*(89-1) = 176
7921 0 7921 g > 6 89*89 = 7921 0*(89-0) = 0
```

Now that I have a way to predict Pythagorean Triples, I made a "complete list" of all such parabola equations (up to my arbitray stopping point). Oops! I generated both PPTs and non-PPTs at once!

parabola equation	a	b	С	note
y = (-1/(2*3*3))x^2+4.5	3	4	5	_
y = (-1/(2*5*5))x^2+12.5	5	12	13	
	15	8	17	
y = (-1/(2*7*7))x^2+24.5	7	24	25	
	21	20	29	
	35	12	37	
y = (-1/(2*9*9))x^2+40.5	9	40	41	
	27	36	45	not PPT. (GCD=9)
	45	28	53	, ,
	63	16	65	
y = (-1/(2*11*11))x^2+60.5	11	60	61	
, , , , , , , , , , , , , , , , , , , ,	33	56	65	
	55	48	73	
	77	36	85	
	99	20	101	
y = (-1/(2*13*13))x^2+84.5	13	84	85	
, , , , , , , , , , , , , , , , , , , ,	39	80	89	
	65	72	97	
	91	60	109	
	117	44	125	
	143	24	145	
y = (-1/(2*15*15))x^2+112.5	15	112	113	
, (-, (45	108	117	not PPT. (GCD=9)
	75	100	125	not PPT. (GCD=25)
	105	88	137	
	135	72	153	not PPT. (GCD=9)
	165	52	173	
	195	28	197	
y = (-1/(2*17*17))x^2+144.5	17	144	145	
, (=, (= =, =, ,,,, = = = = , ,,, = = = , , , , = , = , , , , = , = , , , , = , = , , , , = , = , , , , = , = , , , , = , = , = , , , = , = , = , , , =	51	140	149	
	85	132	157	
	119	120	169	
	153	104	185	
	187	84	205	
	221	60	229	
	255	32	257	
y = (-1/(2*19*19))x^2+180.5	19	180	181	
, (1,(2 13 13)), 2:100:3	57	176	185	
	95	168	193	
	133	156	205	
	171	140	221	
	209	120	241	
	247	96	265	
	285	68	293	
	203	JO	233	

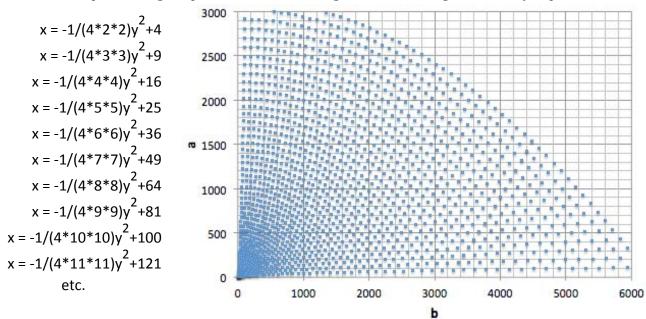
			8	
	323	36	325	
$y = (-1/(2*21*21))x^2+220.5$	21	220	221	
	63	216	225	not PPT. (GCD=9)
	105	208	233	
	147	196	245	not PPT. (GCD=49)
	189	180	261	not PPT. (GCD=9)
	231	160	281	
		etc.		

This is a scatterplot of all parabolas $y = (-1/(2*n^2))x^2 + (1/2)n^2$ for $3 \le n \le 111$



I like how it covers more space than the "gens 1 to 4" scatterplot. But now it's **too predictable** (it doesn't look random enough).

Another example of being too predictable is switching the axes to also get this family of parabolas



I set up rows 1, 2, 3, ..., 56 (which felt like a nice place to stop—not too big or small). And I set up a pair of columns for the same 1, 2, 3, ..., 56. Then I typed = $$J2^2-K1^2 in cell K2 and =2*J2*K\$1 in L2 and copied-and-pasted them for the entire $\frac{1}{2}$ of a matrix.

I needed to copy-and-pasteSpecial (just values) to make a single list of (a, b) values from the matrix. Since $(\frac{1}{2})56^2 = 1568$ then I used 1555 Pythagorean Triples.

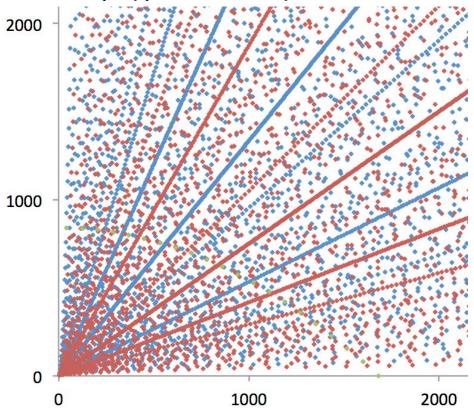
But I'm frustated that Popular (small) PPTs like (5, 12, 13) are at the beginning of many different clusters. I wish there was a single variable to list them all nicely in order.

		n =	1555	n = 1555		n = :	1555		
		M^:	2-N^2	M^2-N^2		M^2	-N^2		
m^2-n^2		PPTs	sorted	PPTs so	orted	PPTs	sorted		diff.
PP	Ts	b	y a	by k)	by	/ C		in c
3	4	3	4	3	4	3	4	5	-skip-
		5	12	15	8	5	12	13	8
15	8	7	24	5	12	15	8	17	4
		9	40	35	12	7	24	25	8
35	12	11	60	63	16	21	20	29	4
		13	84	21	20	35	12	37	8
63	16	15	8	99	20	9	40	41	4
		15	112	7	24	45	28	53	12
99	20	17	144	143	24	11	60	61	8
		19	180	45	28	63	16	65	4
143	24	21	20	195	28	33	56	65	0
		21	220	255	32	55	48	73	8
195	28	23	264	77	36	77	36	85	12
		25	312	323	36	13	84	85	0
255	32	27	364	9	40	39	80	89	4
222	26	29	420	399	40	65	72	97	8
323	36	31	480	117	44	99	20	101	4
200	40	33	56	483	44	91	60	109	8
399	40	33	544	55 575	48	15	112	113	4
400	4.4	35 35	12	575 165	48	117	44	125	12
483	44	35 37	612 684	165	52 52	105 143	88 24	137 145	12 8
575	48	73	2664	675 621	100	273	136	305	0 12
3/3	40	/3	2004	021	100	2/3	130	303	12
	•								
	•								
2499	100	75	308	2499	100	207	224	305	0
55		75	2812	153	104	25	312	313	8
2703	104	77	36	2703	104	75	308	317	4
		77	2964	725	108	323	36	325	8
2915	108	79	3120	2915	108	253	204	325	0
		81	3280	15	112	175	288	337	12
3135	112	83	3444	3135	112	299	180	349	12
<mark>5</mark>	<mark>12</mark>	85	132	837	116	225	272	353	4
		85	3612	119	120	357	76	365	12
21	20	87	416	209	120	27	364	365	0
		87	3784	391	120	275	252	373	8
45	28	89	3960	957	124	345	152	377	4
		91	60	85	132	135	352	377	0
77	36	91	4140	475	132	189	340	389	12
		93	476	1085	132	325	228	397	8

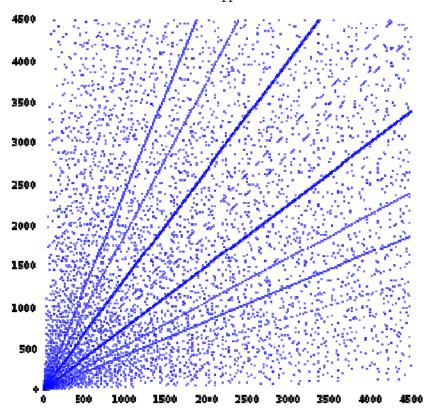
So I finally tried **brute force**. After setting up row 2 = { 1, 2, 3, ...} and column B = { 3, 4, 5, ...}, cell C3 = IF(\$B3>C\$2,IF(ROUND(SQRT(\$B3^2+C\$2^2),0)=SQRT(\$B3^2+C\$2^2), CONCATENATE(C\$2,", ",\$B3),""),"").

Paste this formula as much as computer memory can handle. Optionally re-size the results for visibility.

So I finally combined every Pythagorean Triple that I could find $(n = 633 \ 462 \ of them)$ and then I got the first 500 multiples (up to a reasonable max.) of them and made this scatterplot:



I hope that you feel that it looks "close enough" to the one on Wikipdia:



because I'm out-of-time now. Thanks... MH

Files in the folder "Pythagorean Triples" (in alphabetical order):

2000px-Pythagorean_triple_scatterplot.svg.png	Mar. 1, 2011	543 000 bytes
25.docx	Mar. 3, 2011	242 000 bytes
633462.xlsx	Apr. 13, 2011	46 262 000 bytes
a few Pythagorean Triples.doc	May 5, 2011	20 000 bytes
dtc.57.tif.gif	Mar. 1, 2011	22 000 bytes
$Formulas \ \ {\it for generating Pythagorean Triples - Wikipedia, the Free Encyclopedia.} html$	Apr. 6, 2011	60 000 byes
fun fun.xlsx	Feb. 27 to Apr. 13, 2012	419 000 bytes
IsPrime macro.xlsb	Mar. 31, 2011	565 000 bytes
pythagorean a b.tiff	Mar. 1, 2011	543 000 bytes
pythagorean work (version 1).xlsb	Mar. 31 (8:30 am) to Apr. 13, 2011	17 324 000 bytes
pythagorean work.xlsb	Mar. 31, 2011 at 8:27 am	554 000 bytes
pythagorean work.xlsx	Mar. 13 to Mar. 18, 2011	687 000 bytes
spare.xlsx	Apr. 11 to Apr. 12, 2011	9 577 000 bytes

Files in the folder "Pythagorean Triples" (in chronological order):

2000px-Pythagorean_triple_scatterplot.svg.png	Mar. 1, 2011 (9:12 am)	500 KB
pythagorean a b.tiff	Mar. 1, 2011 (9:17 am)	500 KB
dtc.57.tif.gif	Mar. 1, 2011 (9:29 am)	20 KB
25.docx	Mar. 3, 2011	200 KB
pythagorean work.xlsx	Mar. 13 to Mar. 18, 2011	600 KB
IsPrime macro.xlsb	Mar. 31, 2011	500 KB
pythagorean work.xlsb	Mar. 31, 2011 at 8:27 am	500 KB
pythagorean work (version 1).xlsb	Mar. 31 (8:30 am) to Apr. 13, 2011	17 MB
$Formulas \ \ {\it for generating Pythagorean Triples - Wikipedia, the Free Encyclopedia.} html$	Apr. 6, 2011	60 KB
spare.xlsx	Apr. 11 to Apr. 12, 2011	9 MB
633462.xlsx	Apr. 13, 2011	44 MB
a few Pythagorean Triples.doc	May 5, 2011	24 KB
fun fun.xlsx	Feb. 27 to Apr. 13, 2012	400 KB

Sheets on pythagorean work (version 1).xlsb:

- * PPT's is a summary w/ 3 scatterplots (each 3500×6000) and numbers from other Sheets. For ex., If a^2 / (4n) is an integer, then (a, | n a^2 / (4n) | , n + a^2 / (4n)) is a P.T.
- * 40 mult 1 scatterplot (5000 x 5000) and numbers based on a formula from Ernest J. Eckert's "The College Mathematics Journal" article (v. 23, no. 5, Nov., 1992).
- * gens 1 disappointing scatterplot and numbers based on TI, T2,T3 from generation 0 to 6.
- * parabolas 55 equations like $y = x^2$, their numbers and 1 non-chaotic scatterplot (6000 x 6000) which created the base for the Sheet "500 mult."
- * Sheet3 historical examples of my thinking process: the first equation I found and the first scatterplot (8000×8000) showing a pattern arising out of chaos.
- * m^2-n^2 55 equations like $x = -y^2$ and their numbers.
- * 500 mult. 501 pairs of columns by 1270 rows of just numbers.