Calendars and the uniform passage of time

Halley's comet passed perihelion on February 9, 1986 at 11:00 a.m. Where is it now (March 27, 1997 at 10:00 a.m.)?

This is a question about both space and time, but in this article we shall be concerned only with time. That's because the first step in answering this question is to determine exactly how much time has elapsed between Halley's perihelion passage and now. Of course we can proceed in an entirely elementary manner by counting explicitly:

Period	Elapsed time	Basis of calculation
the rest of February 9	13 hours	24 hours in a day
the rest of February	19 days	28 days in February in 1986 (not a leap year)
the rest of 1986	306 days	365 days in all of 1986, $31 + 28$ gone already
1987	365 days	
1988	366 days	1988 is a leap year
1989	365 days	
1990	365 days	
1991	365 days	
1992	366 days	1992 is a leap year
1993	365 days	
1994	365 days	
1995	365 days	
1996	366 days	1996 is a leap year
January & February, 1997	59 days	28 days in February again
March 1 – March 27	26 days	
today	10 hours	

This makes a total of 4063 days and 23 hours.

Julian days

Astronomers have devised a more efficient procedure to do this sort of calculation. They measure the passage of time in days from a somewhat arbitrary date in the past. The elapsed time since this date is called the **Julian date**. It is calculated in the following fashion. Let Y, M, D stand for year, month, and day, which we shall assume to be in terms of Greenwich mean time.

- *If $M > 2$ let*

  $$y = Y, \quad m = M$$

  otherwise

  $$y = Y - 1, \quad m = M + 12 \ .$$

- *If the date is earlier than October 15, 1582 let*

  $$B = 0$$

  otherwise

  $$A = \mathrm{Int}(y/100), \quad B = 2 - A + \mathrm{Int}(A/4) \ .$$

- *Then the Julian date is*

$$\mathrm{Int}(365.25y) + \mathrm{Int}(30.6001\,(m+1)) + B + D + 1720994.5$$

Recall that $\mathrm{Int}(x)$ is the integer just below (or equal to) x, so that if $n = \mathrm{Int}(x)$ then

$$n \le x < n + 1 \ .$$

For the Julian date of the perihelion passage of Halley's comet, we have

$$\begin{aligned}
y &= 1985 \\
m &= 14 \\
D &= 9.4583 \\
A &= \mathrm{Int}(1985/100) \\
&= 19 \\
B &= 2 - A + \mathrm{Int}(A/4) \\
&= -13 \\
JD &= \mathrm{Int}(725021.25) + \mathrm{Int}(459.0015) - 13 + 9.4583 + 1720994.5 \\
&= 2446470.9583
\end{aligned}$$

and for the current date we have

$$\begin{aligned}
y &= 1997 \\
m &= 3 \\
D &= 27.4167 \\
A &= 19 \\
B &= -13 \\
JD &= \mathrm{Int}(729404.25) + \mathrm{Int}(122.4004) - 13 + 27.4167 + 1720994.5 \\
&= 2450534.9167
\end{aligned}$$

so that the elapsed time between the two is 4063.9583 days.

Our aim now is to understand this recipe for calculating the Julian date. It involves the sum of the number D, which is straightforwardly related to the passage of days, and several other parts:

Part I. *If* $M > 2$ *let*

$$y = Y, \quad m = M$$

otherwise

$$y = Y - 1, \quad m = M + 12 \ .$$

This is natural enough. In effect we are changing the beginning of the year from January 1 to March 1. From the standpoint of calculation this is a good idea, because it means that the elapsed time between any two dates in the same year is independent of the exact year (i.e. does not depend on the number of days in the month of February). This is such a great idea that you might wonder why it hasn't been done already, once and for all! Well, our calendar is largely derived from the ancient Roman calendar, and in fact the original Roman calendar began on March 1. The month of January, for example, only acquired its present name when the change was made—Janus was the Roman god with two faces, looking both backwards and forwards, and a good choice for the month at the end of one year and the beginning of another. Furthermore, the names of the months September, October, November, and December reflect in the Latin roots of their names that they were once the seventh, eighth, ninth, and tenth months. The year's beginning on January 1 was apparently not adopted uniformly throughout the Roman world, and even in recent times there were vestiges of this tradition still in evidence in Western culture.

Part II. *If the date is earlier than October 15, 1582 let*

$$B = 0$$

otherwise

$$A = \mathrm{Int}(y/100), \quad B = 2 - A + \mathrm{Int}(A/4) \ .$$

The ancient Roman calendar was quite chaotic. It was Julius Caesar, acting with the advice of the Alexandrian mathematician Sosigenes, who reformed it, in about 46 B.C. The main effect was to bring the length of the calendar year into better synchronization with the true length of a solar year. Since the number of days in a solar year is not an integer, in order to maintain this synchronization the length of a year must vary from time to time. In the Julian calendar, three out of every four years had 365 days and the fourth year 366.

The Julian calendar, in effect, assumes the true length of a solar year (say from vernal equinox to vernal equinox) to be 365.25 days. This is not correct, and overshoots the mark a bit, since the true solar year is very close to 365.2422 days. Even in Roman times the better estimate of 365.2467 days was known to experts, and I have not seen any explanation of why the Julian scheme was accepted when it was known not to be as accurate as it might have been.

By the year 1500 or so, it was obvious to anybody who knew his history that the date of the vernal equinox, which is a natural candidate for marking the beginning of the solar year by all inhabitants of Earth, was about 10 days earlier than it had been in Roman times. After a great deal of discussion, Pope Gregory XIII, acting with the advice of several Jesuit mathematicians and astronomers, decreed that a change would be made official. In Catholic Europe, the next day after October 4, 1582 was October 15. In the rest of Europe this reform was apparently considered an attempt by the Roman Catholic Church to subvert the true order of things, and in effect to steal days away from people. There seems even to have been current a common opinion that they were days removed permanently from one's life span. The change was, however, adopted gradually over a period of several centuries—by England in the middle of the 18th century, by Russia only after the revolution of 1917. Of course by the time these countries adopted the Gregorian calendar, even more than ten days had to be given up.

Pope Gregory's decree arranged that the calendar of his time would agree with that of ancient times, but it also modified the Julian scheme of leap years in order to maintain synchronization. It was decreed that every 100th year would *not* be a leap year, except that every fourth century break would be. Thus 1600 was a leap year in the Gregorian scheme, while 1700, 1800, and 1900 were not, and 2000 will be. In our shifted calendar, these years must be shifted back 1, so that year y normally has 366 days if $y + 1$ is divisible by 4.

It is still conventional to use the Julian calendar up to this change-over date. The calculation of the constant B takes this into account.

Exercise. *How long does the Gregorian calendar assume the solar year to be?*

Exercise. *How much time would elapse from October 15, 1582 until the true solar calendar and the Gregorian calendar differ by a whole day?*

Part III. *Add* $\mathrm{Int}(365.25y)$.

This term and the exact formula for B measure the number of days between the start of different (adjusted) years. We can understand this more directly if we make a table of the number of elapsed days from March 1, 1599 to March 1 of adjusted year y.

Year y	Days in that year $d(y)$	Days elapsed to March 1 of that year $e(y)$
1599	366	0
1600	365	366
1601	365	731
1602	365	1096
1603	366	1461

1604	365	1827
. . .		
1699	365	36525
1700	365	36890
. . .		

Thus $e(y)$ is the sum of $e(y-1)$ and $d(y)$. The pattern of calculations can be understood better if we write the term $d(y)$ in the second column as the sum of four terms

$$d_0(y) + d_1(y) - d_2(y) + d_3(y)$$

where

$$\begin{aligned} d_0(y) &= 365 \\ d_1(y) &= 1 \quad (y+1 \equiv 0 \bmod 4) \\ &= 0 \quad \text{otherwise} \\ d_2(y) &= 1 \quad (y+1 \equiv 0 \bmod 100) \\ &= 0 \quad \text{otherwise} \\ d_3(y) &= 1 \quad (y+1 \equiv 0 \bmod 400) \\ &= 0 \quad \text{otherwise} \end{aligned}$$

That is to say, the number of days in year $y-1$ is the sum of (a) 365 plus (b) 1 if y is divisible by 4 (c) minus 1 if y is divisible by 100 plus (d) 1 if y is divisible by 400. The sum $e(y)$ has a corresponding decomposition

$$e(y) = e_0(y) + e_1(y) - e_2(y) + e_3(y)$$

where (letting $z = y - 1599$)

$$\begin{aligned} e_0(y) &= 365z \\ e_1(y) &= \mathrm{Int}(z/4) \\ e_2(y) &= \mathrm{Int}(z/100) \\ e_3(y) &= \mathrm{Int}(z/400) \end{aligned}$$

Exercise. *Why is this effectively the same as the rule given earlier?*

Part IV. *The only term with the month in it is now*

$$\mathrm{Int}(30.6001\,(m+1)) \ .$$

This by far the most interesting term. I will look at it in detail in the next section.

Part V. *Add in the constant* 1720994.5.

This somewhat arbitrary choice was made by the sixteenth century scholar Joseph Scaliger, based on some arithmetic combining three cycles each of several years in length—the *Julian* cycle, in which days of the week repeat every 28 Julian years, the *Metonic* cycle, which measures an approximate cycle of the relative positions of the Sun and the Moon (which determine the occurrence of Easter in the Roman Catholic Church), and the *indictio*, a period important in Roman tax administration. For simple arithmetical reasons, the combination of these cycles forced the starting date of the new Julian reckoning to be January 1, 4713 B.C. The details can be found in the book by Otto Neugebauer listed among the references, This scheme is surely one of the first examples of ‘legacy code’, but it is probably still used in modern astronomy because essentially all astronomical data originated after Julian date 0. Incidentally, the father of Joseph Scaliger was Julius Caesar Scaliger. It is often said that Julian days are named after the elder Scaliger, but I have not seen any convincing evidence for this assertion. (My edition of the *Encyclopaedia Britannica* asserts that Joseph Scaliger was the most famous scholar of all time! So fickle is fame.)

What about the 0.5? This is tacked on because this system is used primarily by astronomers, who want all dates in one observing session to be the same. Hence the Julian day begins at noon.

Months and days

Where does the expression

$$f(m) = \mathrm{Int}(30.6001(m+1))$$

come from?

Let's make a table. Recall that in applying this formula, m will be in the range 3 to 14. We therefore have:

Month m	$30.6001(m+1)$	$f(m)$
3	122.4004	122
4	153.0005	153
5	183.6006	183
6	214.2007	214
7	244.8008	244
8	275.4009	275
9	306.0010	306
10	336.6011	336
11	367.8013	367
12	397.8013	397
13	428.4014	428
14	459.0015	459

The first thing to strike us is that the 0.0001 has no practical effect whatsoever! The factor 30.6 would have done as well, and we shall use it from now on. The reason for the extra bit is probably that in base 2 arithmetic the fraction $0.6 = 3/5$ is representable only as an infinite repeating series, and therefore arithmetic involving it is not exact on most computers. The extra 0.0001 is a small fudge factor put in to ensure that rounding errors will not affect the calculation seriously. We'll see later why it this fudging is safe, and get a rough idea of how much we can fudge without disturbing the validity of this formula. As for the rest, we now look at the differences in the last column, and shift the month numbering back by 3:

m	$f(m)$	Difference
0	122	0
1	153	31
2	183	30
3	214	31
4	244	30
5	275	31
6	306	31
7	336	30
8	367	31
9	397	30
10	428	31
11	459	31

The last column is now seen to be the sequence of days in the months March, April, May, June, July, etc. Thus the third column is the sum of the initial value 122 and the number of days elapsed from midnight of March 1 to midnight on the first of month m. We can simplify things somewhat by subtracting off the constant 122 from the second column and the constant 30 from the last column, making an adjustment in the other columns as well:

n	$0.4 + 0.6\,n$	a_n	b_n
0	0.4	0	0
1	1.0	1	1
2	1.6	1	0
3	2.2	2	1
4	2.8	2	0
5	3.4	3	1

6	4.0	4	1
7	4.6	4	0
8	5.2	5	1
9	5.8	5	0
10	6.4	6	1
11	7.0	7	1

where

$$a_n = b_0 + b_1 + \cdots + b_n$$

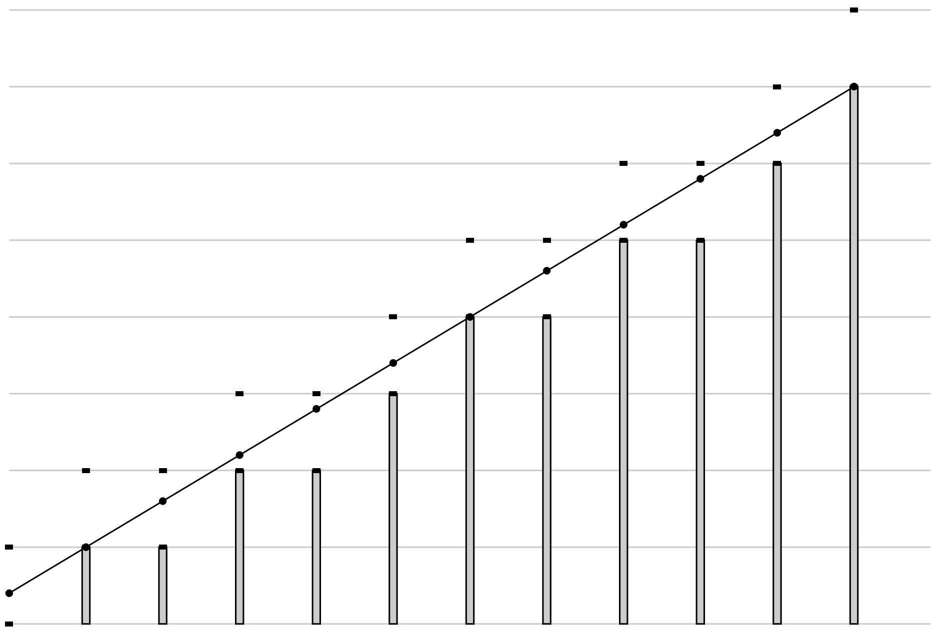
and also

$$a_n = \mathrm{Int}(0.6\,n + 0.4)\ .$$

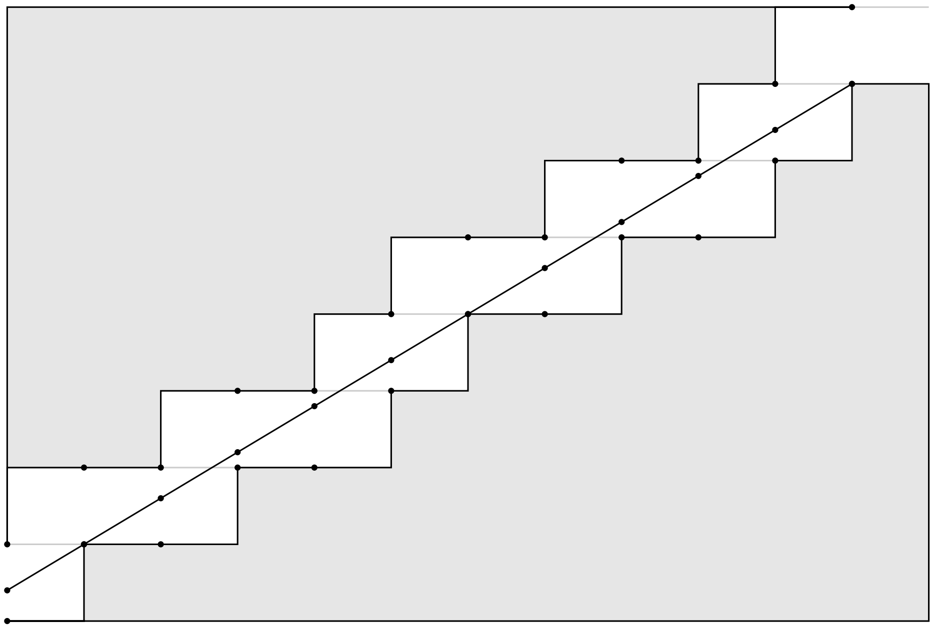
These modifications correspond to the fact that we can rewrite (with $n = m - 3$)

$$\mathrm{Int}(30.6(m+1)) = \mathrm{Int}(30.6(n+4)) = \mathrm{Int}(30n + 120 + 0.6n + 2.4) = 122 + 30n + \mathrm{Int}(0.6n + 0.4)\ .$$

In other words, *we have an empirical sequence of terms* $a_0 = 0$, $a_1 = 1$, $a_2 = 1$, $a_3 = 2$, $\ldots$, $a_{11} = 7$*, and it happens that we can find a simple mathematical formula for it, namely the integer part of* $0.6\,n + 0$. We can therefore pose a similar question for any range of integer data. To understand better when and how a scheme like this works, here is a graph of the empirical data and the ‘approximating’ straight line:



What we can see here is that we can calculate the sequence a_n from the expression $\mathrm{Int}(0.6\,n + 0.4)$ because the line $y = ax + b$ with $a = 0.6$, $b = 0.4$ passes between (n, a_n) and $(n, a_n + 1)$ for $n = 0$ to $n = 11$. We can illustrate this even better by drawing a pair of ‘stairs’, the one below marking the original data and the one above it marking one unit higher:



Any straight line $y = ax + b$ will do the approximation we want as long as it lies below the stairs above and on or above the stairs below. Now you can see from the picture that if a line lies above the points $(1, 1)$ and $(11, 7)$ it will also lie above all the other lower stairs. Something similar happens for the upper stairs—there are extremal points on each of the stairs which are all that have to be considered. All in all, we can picture the possible approximating lines as those lying between the two grey regions in the next picture:

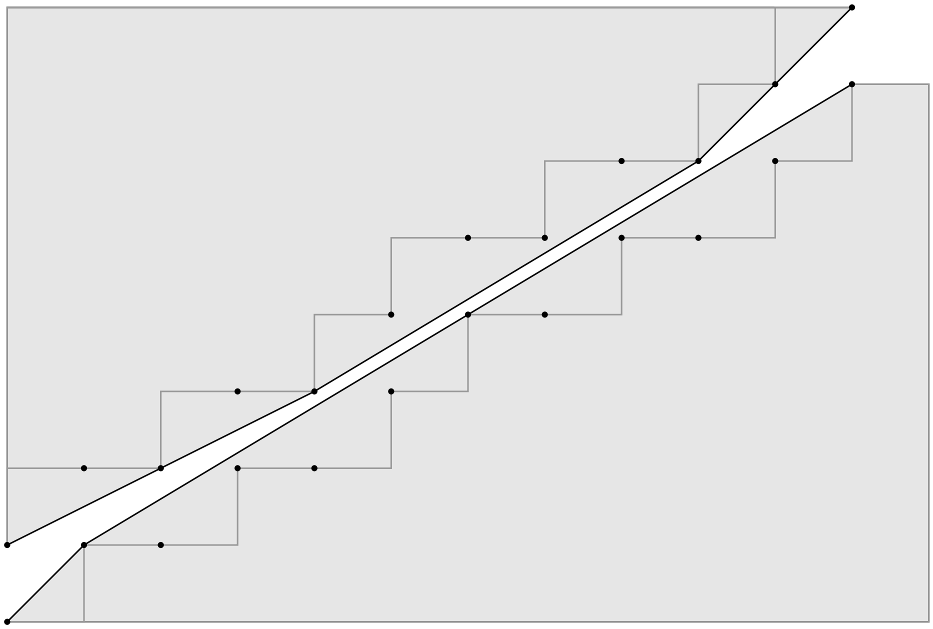
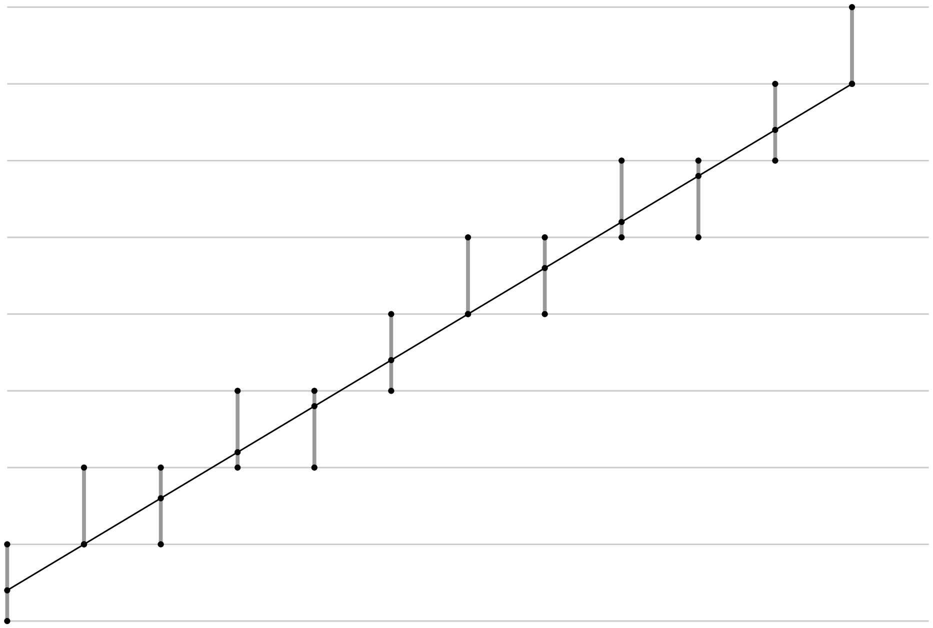


Figure 3. *The lines $y = ax + b$ such that $a_n = \mathrm{Int}(an + b)$ are those lying between the two grey regions.*

There is another, dual, way to understand the possible formulas $y = ax + b$ approximating the data we are given. The line must pass, as we have seen, between the two sets of stairs. This means that it must pass inside the segments marked below:



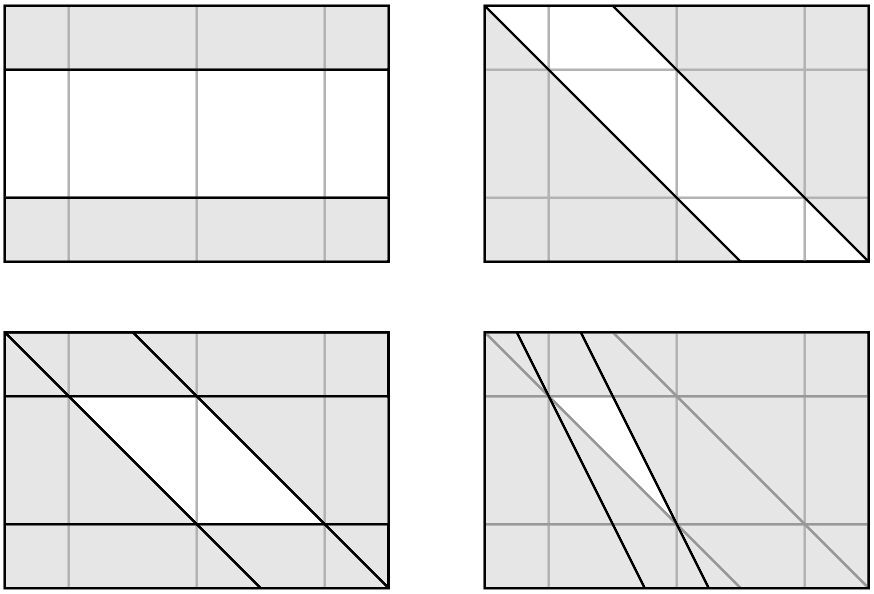
This condition also gives us 11 separate simple conditions on the coefficients of the line: $a_n \le an + b < a_n + 1$ for $n = 0$ to $n = 11$:

$$\begin{aligned} &0 \le b < 1 \\ &1 \le a + b < 2 \\ &1 \le 2a + b < 2 \\ &\quad \dots \\ &7 \le 11a + b < 8 \end{aligned}$$

Each of these conditions corresponds to a region in the (a, b) plane, and each of them imposes more and more severe restrictions on that region. The region $0 \le b < 1$ is a horizontal band, and the region

$$0 \le b < 1, \quad 1 \le a + b < 2$$

is a parallelogram.



If we add in the conditions

$$1 \leq 2a + b < 2$$

we get a triangle, shown at lower right in the figure above. And if we add in all the constraints $a_n \leq an + b < a_n + 1$ together we get the small region below, which could *a priori* be rather complicated, but is actually just a quadrilateral:

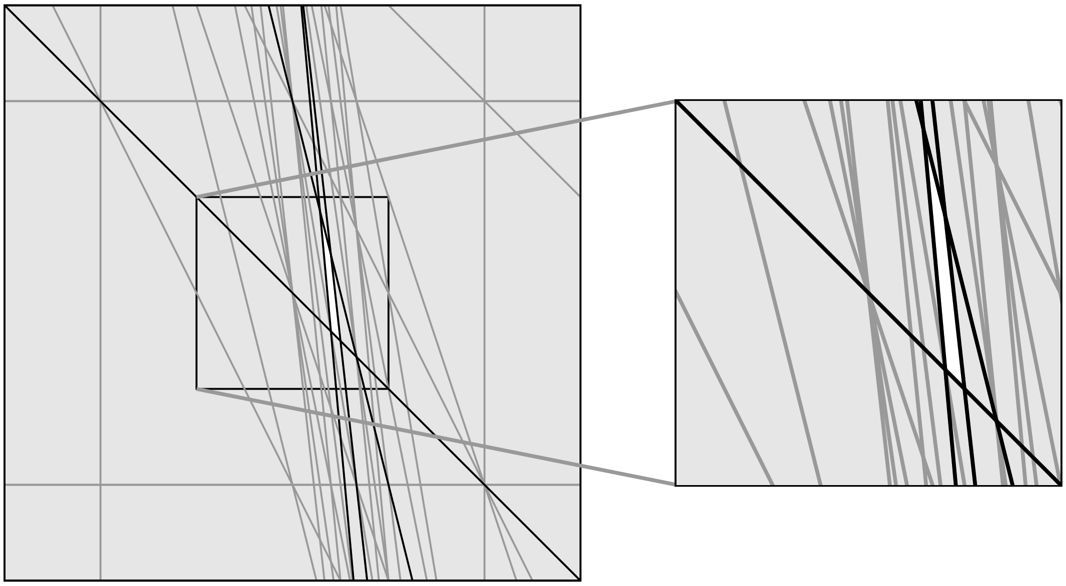


Figure 6. *The lines* $y = ax + b$ *such that* $a_n = \mathrm{Int}(an + b)$ *are those with* (a, b) *lying inside the white quadrilateral.*

The corner at the lower left is $(0.6, 0.4)$, for example. This explains the exact formula used for calculating the Julian date.

This region is a kind of 'window of calculation'. You can ask very generally for which sequences a similar formula holds, and it may very well happen that the corresponding region is empty.

Another historical aside: the original calendar proposed by Julius Caesar was considerably more rational than the one we now use. The month of August (which was called *Sextilis* in those days before Augustus became emperor, as July had been called *Quintilis* before Julius came along) had 30 days, September 31, and so on in alternation. But the month of July was dedicated to Julius Caesar, and when in turn the month of August was dedicated to his successor Augustus, it was insisted that August be as long as July. A day was stolen from February, and then the alternation of sizes was changed to make September shorter than August. It's not very amazing that these changes took place, but rather more amazing that we still live with the minor annoyances the vanity of a Roman emperor created.

Convexity

We have seen two ways of understanding what formulas of the form

$$a_m = \mathrm{Int}(am + b)$$

are acceptable, where a_m is the number of days elapsed between midnight on March 1 and midnight of the first of month m in any one adjusted year. Here an *adjusted* year is one that begins on March 1 and ends with the last day of February. The best answer to this question can be read off from the last figure: any pair (a, b) in the white region of this figure may be used. (In a world without rounding errors some points on the boundary of the region might also be used, but this is not a serious modification of the claim.) How can we describe this region in a simple manner? It has four corners, and *this region is the unique quadrilateral region with these corners*. This quadrilateral region is not a random quadrilateral. It has one property which it does not possess accidentally–namely, it is **convex**, which means that a line segment joining any two of its points lies completely in the region.

The convexity of this figure could have been predicted, since it is the intersection of the bands $a_i \le ai+b < a_i+1$, which are clearly convex, and it is a simple and general fact that the intersection of any number of convex regions is again convex (if not empty). Thus, our question is a special case of a more general question: *Given a set of inequalities $f_i(x) \le 0$, find the set of extremal vertices of the region where all of these inequalities are satisfied simultaneously.* Similarly, the first description involves questions of two general types: (1) *Given a set of points in the plane, find the set of extremal vertices of their convex hull*, and (2) *describe the intersection of two convex regions in the plane*. There is a large modern literature concerned with questions like this (the subject is called **combinatorial geometry**), but here I shall only lay out a simple procedure which will at least enable you to find the four corners of the slim quadrilateral in the last figure.

It is simple to see that the parallelogram

$$0 \le b < 1, \quad 1 \le a + b < 2$$

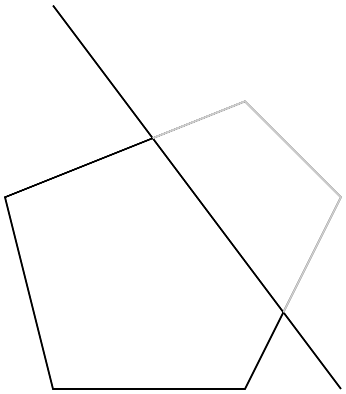
has corners $(1, 0)$, $(0, 1)$, $(1, 1)$, and $(2, 0)$. We want to chop away from this successively the regions

$$a_i > ai + b$$

and

$$ai + b \ge a_i + 1 \ .$$

By induction, we must answer the following question: *Given a convex polygon with vertices $P_1, P_1, \ldots, P_m$ in the (a, b) plane, find the vertices of the convex set of points inside or on this polygon where in addition $pa + qb \le r$.* I'll leave as an exercise to work out an algorithm to do this.



Exercise. *What are all the corners of the quadrilateral in Figure 6, exactly?*

In a way, this whole investigation is concerned with convexity—even the first few diagrams showing which linear functions would be acceptable is really concerned with those large upper and lower convex regions on one of the figures. Convexity plays an important role in many mathematical applications, and lies at the mathematical core of **linear programming** and the **simplex method**.

Exercise. *How are the two types of diagrams we have seen related? In Figure 3, plot the lines corresponding to the corners of the region in Figure 6.*

Exercise. *Write down the necessary and sufficient conditions on a sequence of three integers* f_0*,* f_1*,* f_2 *that no equation*

$$f_m = \mathrm{Int}(am + b)$$

be valid.

Exercise. *Write down a sequence of numbers* f_i *for which the region in the* (a, b) *plane has more than* 4 *vertices.*

Finally, I should point out that, except indirectly, the motion of the Moon has played no role in the discussion above. Historically it has played a major role in many important calendars, and a minor role in the one we use in so far as the date on which Easter is celebrated depends at least formally on it. But that is another story.

References

1. Herbert Edelsbrunner, *Algorithms in Combinatorial Geometry*. Springer-Verlag, 1987.

2. *Encyclopaedia Brittanica*, 11th Edition, 1910. Articles on Joseph *Scaliger*, the *Calendar*, even *Sosigenes*. The article on the calendar explains the reasons for starting the Julian date era in 4713 B.C. In general, the 11th edition is well known for its detailed mathematical articles, and worth looking into.

3. Explanatory supplement to the *Astronomical Ephemeris*. HMSO, 1961. The section on the calendar is a good addition to the corresponding article in the Encyclopaedia Britannica. The sections on the Julian calendar and the Gregorian calendar are short and clear.

4. Jean Meeus, *Astronomical Formulae for Calculators*. Willmann-Bell, Richmond, Virginia, 1988. Chapter 3 covers Julian dates. There are several books of this type. Those written for hand-held calculators are generally the simplest.

5. Otto Neugebauer, *History of Ancient Mathematical Astronomy*. Springer-Verlag, 1975. The beginning of Volume III discusses Julian dates in detail.

6. T. H. O’Beirne, *Puzzles and Paradoxes.* Oxford University Press, 1965. Chapter 10 has an interesting account of the problems of Easter date determination.