5.11 THE WEIRD AND WONDERFUL CHEMISTRY OF AUDIOACTIVE DECAY

J. H. Conway

Department of Mathematics
Princeton University
Princeton, NJ 08544

1. Introduction

Suppose we start with a string of numbers (i.e., positive integers), say

5 5 5 5 5.

We might describe this in words in the usual way as "five fives," and write down the *derived* string

5 5.

This we describe as "two fives," so it yields the next derived string

2 5

which is "one two, one five," giving

1 2 1 5

namely, "one one, one two, one one, one five," or

1 1 1 2 1 1 1 5

and so on. What happens when an arbitrary string of positive integers is repeatedly derived like this?

I note that more usually one is given a sequence such as

55555 ; 55 ; 25 ; 1215 ; 11121115 ;

and asked to guess the generating rule or the next term.

The numbers in our strings are usually single-digit ones, so we'll call them *digits* and usually cram them together as we have just done. But occasionally we want to indicate the way the number in the string was obtained, and we can do this neatly by inserting commas recalling the commas and quotes in our verbal descriptions, thus:

$$
\begin{gathered}
5\ 5\ 5\ 5\ 5 \\
,5\ 5, \\
,2\ 5, \\
,1\ 2,1\ 5, \\
,1\ 1,1\ 2,1\ 1,1\ 5, \\
\cdots
\end{gathered}
$$

The insertions of these commas into a string or portion thereof is called *parsing*.

We'll often denote repetitions by indices in the usual way, so that the derivation rule is

$$a^{\alpha}\, b^{\beta}\, c^{\gamma}\, d^{\delta}\ \cdots\ \to \alpha a \beta b \gamma c \delta d\ \cdots$$

When we do this, it is always to be understood that the repetitions are collected maximally, so that we must have

$$a \neq b,\ b \neq c,\ c \neq d\ , \ldots\ .$$

Since what we write down is often only a *chunk* of the entire string (i.e., a consecutive subsequence of its terms), we often use the square brackets "[" or "]" to indicate that the apparent left or right end really is the end. We also introduce the formal digits

0, as an index, to give an alternative way of indicating the ends (see below)

X for an arbitrary nonzero digit, and

$\neq n$ for any digit (maybe 0) other than n .

Thus	$X^0\, a^{\alpha}\, b^{\beta}\, c^{\gamma}$	means the same as $[\ a^{\alpha}\, b^{\beta}\, c^{\gamma}$
	$a^{\alpha}\, b^{\beta}\, c^{\gamma}\, X^0$	means the same as $a^{\alpha}\, b^{\beta}\, c^{\gamma}]$
	$a^{\alpha}\, b^{\beta}\, c^{\gamma}\, X^{\neq 0}$	means $a^{\alpha}\, b^{\beta}\, c^{\gamma}$ followed by at least another digit,
and	$a^{\alpha}\, b^{\beta}\, c^{\gamma} (\neq 2)^{\neq 0}$	means that this digit is not a 2.

I'm afraid that this heap of conventions makes it quite hard to check the proofs, since they cover many more cases than one naively expects. To separate these cases would make this article very long and tedious, and the reader who really wants to check all the details is advised first to spend some time practicing the derivation process. Note that when we write $L \to L' \to L'' \to \cdots$ we mean just that every string of type L derives to one of type L', every string of type L' derives to one of type L'', and so on. So when in our proof of the Ending Theorem we have

$$n^n] \overset{(n\neq 2)}{\to} n^{\neq n}] \to n']$$

the fact that the left arrow is asserted only when $n \neq 2$ does not excuse us from checking the right arrow for $n = 2$. (But, since $n > 1$ is enforced at that stage in the proof, we needn't check either of them for $n = 1$.)

By applying the derivation process n times to a string L, we obtain what we call its nth *descendant,* L_n. The string itself is counted among its descendants, as the 0th.

Sometimes a string factors as the product LR of two strings L and R whose descendants never interfere with each other, in the sense that $(LR)_n = L_nR_n$ for all n. In this case, we say the LR *splits* as $L.R$ (dots in strings will always have this meaning). It is plain that this happens just when (L or R is empty or) the last digit of L_n always differs from the first one of R_n. Can you find a simple criterion for this to happen? (When you give up, you'll find the answer in our Splitting Theorem.)

Obviously, we call a string with no nontrival splittings an *atom,* or *element.* Then every string is the split product, or *compound,* of a certain number of elements, which we call the elements it *involves.* There are infinitely many distinct elements, but most of them only arise from specially chosen starting strings. However, there are some very interesting elements that are involved in the descendants of every string except the boring ones [ ] and [22]. Can you guess how many of these *common elements* there are? (Hint: we have given them the names Hydrogen, Helium, Lithium , . . . , Uranium.)

It's also true (but ASTONISHINGLY hard to prove) that *every* string eventually decays into a compound of these elements, together with perhaps a few others (namely, isotopes of Plutonium and Neptunium, as defined below). Moreover, all strings except the two boring ones increase in length exponentially at the same constant rate. (This rate is roughly 1.30357726903: it can be precisely defined as the largest root of a certain algebraic equation of degree 71.) Also, the relative abundances of the elements settle down to fixed numbers (zero for Neptunium and Plutonium). Thus, of every million atoms about 91790 on average will be of Hydrogen, the commonest element, while about 27 will be of Arsenic, the rarest one.

You should get to know the common elements, as enumerated in our Periodic Table. The abundance (in atoms per million) is given first, followed by the atomic number and symbol as in ordinary chemistry. The actual digit-string defining the element is the numerical part of the remainder of the entry, which, when read in full, gives the derivate of the element of next highest atomic number, split into atoms. Thus, for example, the last line of the Periodic Table tells us that Hydrogen (H) is our name for the digit-string 22, and that the next higher element, Helium (He), derives to the compound

Hf.Pa.H.Ca.Li

which we might call

"Hafnium-Protactinium-Hydrogen-Calcium-Lithide"!

Not everything is in the Periodic Table! For instance, the single digit string "1" isn't. But watch:

1
11
21
1211
111221
312211
13112221
11132.13211 = Hf.Sn

after a few moves it has become Hafnium Stannide! This is an instance of our Cosmological Theorem, which asserts that the exotic elements (such as "1") all disappear soon after the Big Bang.

The Periodic Table (Uranium to Silver)

abundance	n	E_n	E_n inside the derivate of E_{n+1}
102.56285249	92	U	3
9883.5986392	91	Pa	13
7581.9047125	90	Th	1113
6926.9352045	89	Ac	3113
5313.7894999	88	Ra	132113
4076.3134078	87	Fr	1113122113
3127.0209328	86	Rn	311311222113
2398.7998311	85	At	Ho.1322113
1840.1669683	84	Po	1113222113
1411.6286100	83	Bi	3113322113
1082.8883285	82	Pb	Pm.123222113
830.70513293	81	Tl	111213322113
637.25039755	80	Hg	31121123222113
488.84742982	79	Au	132112211213322113
375.00456738	78	Pt	111312212221121123222113
287.67344775	77	Ir	3113112211322112211213322113
220.68001229	76	Os	1321132122211322212221121123222113
169.28801808	75	Re	111312211312113221133211322112211213322113
315.56655252	74	W	Ge.Ca.312211322212221121123222113
242.07736666	73	Ta	13112221133211322112211213322113
2669.0970363	72	Hf	11132.Pa.H.Ca.W
2047.5173200	71	Lu	311312
1570.6911808	70	Yb	1321131112
1204.9083841	69	Tm	11131221133112
1098.5955997	68	Er	311311222.Ca.Co
47987.529438	67	Ho	1321132.Pm
36812.186418	66	Dy	111312211312
28239.358949	65	Tb	3113112221131112
21662.972821	64	Gd	Ho.13221133112
20085.668709	63	Eu	1113222.Ca.Co.
15408.115182	62	Sm	311332
29820.456167	61	Pm	132.Ca.Zn
22875.863883	60	Nd	111312
17548.529287	59	Pr	31131112
13461.825166	58	Ce	1321133112
10326.833312	57	La	11131.H.Ca.Co
7921.9188284	56	Ba	311311
6077.0611889	55	Cs	13211321
4661.8342720	54	Xe	11131221131211
3576.1856107	53	I	311311222113111221
2743.3629718	52	Te	Ho.1322113312211
2104.4881933	51	Sb	Eu.Ca.3112221
1614.3946687	50	Sn	Pm.13211
1238.4341972	49	In	11131221
950.02745646	48	Cd	3113112211
728.78492056	47	Ag	132113212221

The Periodic Table (Palladium to Hydrogen)

abundance	n	E_n	E_n inside the derivate of E_{n+1}
559.06537946	46	Pd	111312211312113211
428.87015041	45	Rh	311311222113111221131221
328.99480576	44	Ru	Ho.132211331222113112211
386.07704943	43	Tc	Eu.Ca.311322113212221
296.16736852	42	Mo	13211322211312113211
227.19586752	41	Nb	1113122113322113111221131221
174.28645997	40	Zr	Er.12322211331222113112211
133.69860315	39	Y	1112133.H.Ca.Tc
102.56285249	38	Sr	3112112.U
78.678000089	37	Rb	1321122112
60.355455682	36	Kr	11131221222112
46.299868152	35	Br	3113112211322112
35.517547944	34	Se	13211321222113222112
27.246216076	33	As	11131221131211322113322112
1887.4372276	32	Ge	31131122211311122113222.Na
1447.8905642	31	Ga	Ho.13221133122211332
23571.391336	30	Zn	Eu.Ca.Ac.H.Ca.312
18082.082203	29	Cu	131112
13871.124200	28	Ni	11133112
45645.877256	27	Co	Zn.32112
35015.858546	26	Fe	13122112
26861.360180	25	Mn	111311222112
20605.882611	24	Cr	31132.Si
15807.181592	23	V	13211312
12126.002783	22	Ti	11131221131112
9302.0974443	21	Sc	3113112221133112
56072.543129	20	Ca	Ho.Pa.H.12.Co
43014.360913	19	K	1112
32997.170122	18	Ar	3112
25312.784218	17	Cl	132112
19417.939250	16	S	1113122112
14895.886658	15	P	311311222112
32032.812960	14	Si	Ho.1322112
24573.006696	13	Al	1113222112
18850.441228	12	Mg	3113322112
14481.448773	11	Na	Pm.123222112
11109.006821	10	Ne	111213322112
8521.9396539	9	F	31121123222112
6537.3490750	8	O	132112211213322112
5014.9302464	7	N	111312212221121123222112
3847.0525419	6	C	3113112211322112211213322112
2951.1503716	5	B	1321132122211322212221121123222112
2263.8860325	4	Be	111312211312113221133211322112211213322112
4220.0665982	3	Li	Ge.Ca.312211322212221121123222112
3237.2968588	2	He	13112221133211322112211213322112
91790.383216	1	H	Hf.Pa.22.Ca.Li

2. The Theory

We start with some easy theorems that restrict the possible strings after the first few moves. Any chunk of a string that has lasted at least n moves will be called an *n-day-old* string.

The One-Day Theorem. Chunks of types

$$,a\,x,\;\; b\,x,\;\; x^{4 \text{ or more}} \text{ and } x^3\,y^3$$

don't happen in day-old strings. (Note that the first one has a given parsing.)

Proof. The first possibility comes from $x^a x^b$, which, however, should have been written x^{a+b}, in the previous day's string. The other two, whichever way they are parsed, imply cases of the first.

The Two-Day Theorem. No digit 4 or more can be born on or after the second day. Also, a chunk 3×3 (in particular 3^3) can't appear in any 2-day-old list.

Proof. The first possibility comes from a chunk x^4 or more, while the second, which we now know must parse $,3x,3y,$ can only come from a chunk x^3y^3 , of the previous day's string.

When tracking particular strings later, we'll use these facts without explicit mention.

The Starting Theorem. Let R be any chunk of a 2-day-old string, considered as a string in its own right. Then the starts of its descendants ultimately cycle in one of the ways

$$[\;] \quad \text{or} \quad [1^1X^1 \to [1^3 \to [3^1X^{\neq 3}$$

$$\text{or} \quad [2^2] \quad \text{or} \quad [2^21^1X^1 \to [2^21^3 \to [2^23^1X^{\neq 3}$$

If R is not already in such a cycle, at least three distinct digits appear as initial digits of its descendants.

Proof. If R is nonempty and doesn't start with 2^2, then it *either* starts with a 1 and is of one of the types

$$[1^1X^{0 \text{ or } 1} \quad \text{or} \quad [1^1(2^{2 \text{ or } 3} \text{ or } 3^2) \quad \text{or} \quad [1^2X^{1 \text{ or } \neq 1} \quad \text{or} \quad [1^3$$

or starts with a 2 and is of one of the types $[2^1X^{2 \text{ or } \neq 2}$ or $[2^3$
or starts with a 3 and is of one of the types $[3^1X^{3 \text{ or } \neq 3}$ or $[3^2X^{3 \text{ or } \neq 3}$
or starts with some $n > 3$ and has form $[n^1$.

It is therefore visible in

$$\begin{tikzcd}
 & {[1^12^3} \arrow[dr] & {[3^2X^3} \arrow[dr] & & {[1^1]} \arrow[dr] & {[3^2X^{\neq 3}} \arrow[dr] & {[n^1} \arrow[dr] & & {[2^3} \arrow[dr] & \\
{[3^1X^3} \arrow[r] & {[1^13^2} \arrow[r] & {[1^2X^1} \arrow[r] & {[2^1X^2} \arrow[r] & {[1^12^2} \arrow[r] & {[1^2X^{\neq 1}} \arrow[r] & {[2^1\,X^{\neq 2}} \arrow[r] & {[1^1X^1} \arrow[r] & {[1^3} \arrow[r] & {[3^1X^{\neq 3}} \arrow[ll, bend left]
\end{tikzcd}$$

which establishes the desired results for it.

This proves the theorem except for strings of type $[2^2R'$ all of whose descendants start with 2^2. This happens only if no descendant of R' starts with a 2, and so we can complete the proof by applying the results we've just found to R'.

The Splitting Theorem. A 2-day-old string LR splits as $L.R$ just if one of L and R is empty or L and R are of the types shown in one of

L	R
$n]$	$[m$
$2]$	$[1^1X^1$ *or* $[1^3$ *or* $[3^1X^{\neq 3}$ *or* $[n^1$
$\neq 2]$	$[2^21^1X^1$ *or* $[2^21^3$ *or* $[2^23^1X^{\neq 3}$ *or* $[2^2n^{(0 \text{ or } 1)}$
	$(n \geq 4\ ,\ m \leq 3)$

Proof. This follows immediately from the Starting Theorem applied to R and the obvious fact that the last digit of L is constant.

Now we investigate the evolution of the end of the string!

The Ending Theorem. The end of a string ultimately cycles in one of the ways:

$$\begin{tikzcd}
2.311322113212221] \arrow[r] & 2.13211322211312113211] \arrow[d] \\
2.12322211331222113112211] \arrow[u] & 2.1113122113322113111221131221] \arrow[l]
\end{tikzcd}$$

$$\begin{tikzcd}
2.31221132221222112112322211n] \arrow[d, shift left=4ex] \\
2.1311222113321132211221121332211n] \arrow[u, shift right=4ex]
\end{tikzcd} \qquad (n > 1)$$

$$\text{or} \qquad \begin{tikzcd} 2^2\,] \arrow[loop above] \end{tikzcd}$$

(Note: our splitting theorem shows that these strings actually do split at the dots, although we don't use this.)

Proof. A string with last digit 1 must end in one of the ways visible in

$$1^{\geq 3}\,] \rightarrow (\neq 2)^X\, 1^1\,] \rightarrow (\neq 2)^X\, 1^2\,] \rightarrow 2^{X \neq 2}\, 1^1\,] \rightarrow$$

$$2^{X \neq 2}\, 1^2\,] \rightarrow 2^2\, 1^1\,] \rightarrow 2^2\, 1^2\,] \rightarrow 2^3\, 1^1\,]$$

and its subsequent evolution is followed on the right-hand side of Figure 1.

A string with last digit $n > 1$ must end $n^n]$ or $n^{\neq n}]$ and so evolves via

$$\begin{tikzcd}
n^n] \arrow[loop above, "(n=2)"] \arrow[r, "(n \neq 2)"] & n^{\neq n}\,] \arrow[r] & n^1] \arrow[r] & 1n] \arrow[r] & 11n] \arrow[r] & (\neq 1)11n] \arrow[r] & 211n] \arrow[r] & 2211n]
\end{tikzcd}$$

and the last string here is the first or second on the left of Figure 1.

(≠2)2211*n*] (*n* > 1)	(≠2)2221]	
(≠2)22211*n*]	3211]	
32211*n*]	31221]	
322211*n*]	3112211]	
(≠3)332211*n*]	3212221]	
2322211*n*]	312113211]	
21332211*n*]	3111221131221]	
2112322211*n*]	(≠3)331222113112211]	
221121332211*n*]	2.311322113212221]	(*period* 4)
2211211232221l*n*]	2.1321132221131211321 1]	
2211221121332211*n*]	2.111312211332211311122113 1221]	
2212221121123222l1*n*]	2.311311222.1232221133122211 3112211]	
2113221122112133221 1*n*]	2.1112133.22.12.311322113212221]	
22113222122211211232221 1*n*]		
221133211322112211213322l1*n*]		
22.12.3122113222122211211232221l*n*]		
2.1311222113321132211221121332211*n*]		(*period* 2)
2.11132.13.22.12.31221132221222112112322211*n*]		

Figure 1. The evolution of endings other than 2^2].

This figure proves the theorem except for the trivial case 2^2]. (When any of these strings contains a dot, its subsequent development is only followed from the digit just prior to the rightmost dot.)

We are now ready for our first major result.

The Chemical Theorem.

(a) The descendants of any of the 92 elements in our Periodic Table are compounds of those elements.

(b) All sufficiently late descendants of any of these elements other than Hydrogen involve all 92 elements simultaneously.

(c) The descendants of any string other than [ ] or [22] also ultimately involve all 92 elements simultaneously.

(d) These 92 elements are precisely the common elements as defined in the introduction.

Proof.

(a) follows instantly from the form in which we have presented the Periodic Table.

(b) It also follows that if the element E_n of atomic number n appears at some time t, then for any $m < n$, all elements on the E_m line of the table will appear at the later time $t + n - m$. In particular,

$$E_n \text{ at } t \to \text{Hf \& Li at } t+n-1 \quad (\text{if } n \geq 2)\,,$$

$$\text{Hf \& Li at } t \to \text{Hf \& Li at } t+2 \text{ and } t+71\,,$$

$$\text{Hf at } t \to \text{Sr \& U at } t+72-38\,,$$

$$\text{U at } t \to E_n \text{ at } t+92-n\,.$$

From these we successively deduce that if any of these 92 elements other than Hydrogen is involved at some time t_0, Hafnium and Lithium will simultaneously be involved at some strictly later time $\leq t_0 + 100$, and then both will exist at all times $\geq t_0 + 200$, Uranium at all times $\geq t_0 + 300$, and every other one of these 92 elements at all times $\geq t_0 + 400$.

In other words, once you can fool some of the elements into appearing some of the time, then soon you'll fool some of them all of the time, and ultimately you'll be fooling all of the elements all of the time!

(c) If L is not of form $L'2^2$], this now follows from the observation that Calcium (digit-string 12) is a descendant of L, since it appears in both the bottom lines of Figure 1. Otherwise we can replace L by L', which does not end in a 2.

(d) follows from (a), (b), (c) and the definition of the common elements.

Now we'll call an arbitrary string *common* just if it is a compound of common atoms.

The Arithmetical Theorem.

(a) The lengths of all common strings other than boring old [ ] and [22] increase exponentially at the same rate $\lambda > 1$.

(b) The relative abundances of the elements in such strings tend to certain fixed values, all strictly positive.

Notes. Since each common element has at least 1 and at most 42 digits we can afford to measure the lengths by either digits or atoms: we prefer to use atoms. The numerical value of λ is 1.30357726903; the abundances are tabulated in the Periodic Table.

Proof. Let $\mathbf{v}$ be the 92-component vector whose (i)-entry is the number of atoms of atomic number i in some such string. Then at each derivation step, $\mathbf{v}$ is multiplied by the matrix $\mathbf{M}$ whose (i, j)-entry is the number of times E_j is involved in the derivate of E_i . Now our Chemical Theorem shows that some power of $\mathbf{M}$ has strictly positive (i, j)-entries for all $i \neq 1$ (the $(1, j)$-entry will be 0 for $j \neq 1$, 1 for $j = 1$, since every descendant of a single atom of Hydrogen is another such).

Let λ be an eigenvalue of $\mathbf{M}$ with the largest possible modulus, and $\mathbf{v}_0$ a corresponding eigenvector. Then the nonzero entries of $\mathbf{v}_0\,\mathbf{M}^n$ are proportional to λ^n, while the entries in the successive images of all other vectors grow at most this rate. Since the 92 coordinate vectors (which we'll call **H, He** , . . . , **U** in the obvious way) span the space, at least one of them must increase at rate λ

On the other hand, our Chemical Theorem shows that the descendants of each of **He, Li**, . . . , **U** increase as fast as any of them, and that this is at some rate > 1 , while **H** is a fixed vector (rate 1). These remarks establish our Theorem.

(We have essentially proved the Frobenius-Perron Theorem, that the dominant eigenvalue of a matrix with positive entries is positive and occurs just once, but I didn't want to frighten you with those long names.)

The Transuranic Elements.

For each number $n \geq 4$, we define two particular atoms:

an isotope of *Plutonium* (Pu) : $31221132221222112112322211n$

an isotope of *Neptunium* (Np): $1311222113321132211221121332211n$

For $n = 2$, these would be Lithium (Li) and Helium (He); for $n = 3$, they would be Tungsten (W) and Tantalum (Ta), while for $n \geq 4$ they are called the transuranic elements. We won't bother to specify the number n in our notation.

We can enlarge our 92-dimensional vector space by adding any number of new pairs of coordinate vectors **Pu**, **Np** corresponding to pairs of transuranic elements.

Our proof of the Ending Theorem shows that every digit 4 or more ultimately lands up as the last digit in one of the appropriate pair of transuranic elements, and (see the bottom left of Figure 1) that we have the decomposition

$$Pu \rightarrow Np \rightarrow Hf.Pa.H.Ca.Pu.$$

Now $\mathbf{Pu} \pm \mathbf{Np}$ is an eigenvector of eigenvalue ± 1 *modulo the subspace corresponding to the common elements,* since $\mathbf{Pu} \rightarrow\leftarrow \mathbf{Np}$ modulo that space. Because these eigenvalues are strictly less than λ in modulus, the relative abundances of the transuranic elements tend to 0.

So far, I can proudly say that this magnificent theory is essentially all my own work. However, the next theorem, the finest achievement so far in Audioactive Chemistry, is the result of the combined labors of three brilliant investigators.

The Cosmological Theorem.

Any string decays into a compound of common and transuranic elements after a bounded number of derivation steps. As a consequence, *every* string other than the two boring ones increases at the magic rate λ, and the relative abundances of the atoms in its descendants approach the values we have already described.

Proof of the Cosmological Theorem would fill the rest of this book! Richard Parker and I found a proof over a period of about a month of very intensive work (or, rather, play!). We first produced a very subtle and complicated argument, which (almost) reduced the problem to tracking a few hundred cases, and then handled these on dozens of sheets of paper (now lost). Mike Guy found a simpler proof that used tracking and back-tracking in roughly equal proportions. Guy's proof still filled lots of pages (almost all lost) but had the advantage that it found the longest-lived of the exotic elements, namely, the isotopes of *Methuselum* (2233322211n ; see Figure 2). Can you find a proof in only a few pages? Please!

2233322211n ($n > 1$)
223332211n
223322211n
222332211n
322322211n
13221332211n
111322112322211n
311322211213322211n
13211332211211232221 1n
La.H.1232221122112133221 1n
1112133221222112112322211n
Sr.32211322112211213322211n
132221132221222112112322211n
1113322113321132211221121332211n
3123222.Ca.(Li or W or Pu)
1311121332
11133112112.Zn
Zn.321122112
131221222112
1113112211322112
3113212221132221 12
132113121132211332 2112
1113122113111221 1322.Na
311311222113312221 1332
Ho.Pa.H.Ca.Ac.H.Ca.Zn

Figure 2. The descendants of Methuselum.

The Degree of λ .

Plainly, λ is an algebraic number of degree at most 92. We first reduce this bound to 71 by exhibiting a 21-dimensional invariant subspace on which the eigenvalues of $\mathbf{M}$ are 0 or ± 1.

$$\mathbf{v}_1 = \mathbf{H},\ \mathbf{v}_2 = \mathbf{He} - \mathbf{Ta},\ \mathbf{v}_3 = \mathbf{Li} - \mathbf{W},\ \ldots,\ \mathbf{v}_{20} = \mathbf{Ca} - \mathbf{Pa}\ ,$$

or, in atomic number notation,

$$\mathbf{v}_1 = \mathbf{E}_1,\ \mathbf{v}_2 = \mathbf{E}_2 - \mathbf{E}_{73},\ \ \mathbf{v}_3 = \mathbf{E}_3 - \mathbf{E}_{74},\ \ldots,\ \mathbf{v}_{20} = \mathbf{E}_{20} - \mathbf{E}_{91}\ ,$$

and also define

$$\mathbf{v}_{21} = \{\ \mathbf{Sc} + \mathbf{Sm} - \mathbf{H} - \mathbf{Ni} - \mathbf{Er} - 3\mathbf{U}\ \}/2\ ,$$

then observe that

$$\mathbf{v}_{21} \to \mathbf{v}_{20} \to \mathbf{v}_{19} \to\ \cdots\ \to \mathbf{v}_4 \to \mathbf{v}_3 \to \mathbf{v}_2,\ \mathbf{v}_1 \to \mathbf{v}_1\ .$$

An alternate base for this space consists of the eigenvectors

$$\mathbf{v}_1 \text{ and } \mathbf{v}_3 \pm \mathbf{v}_2$$

of $\mathbf{M}$ with the respective eigenvalues

$$1 \text{ and } \pm 1\ ,$$

together with the following Jordan block of size 18 for the eigenvalue 0

$$\mathbf{v}_{21} - \mathbf{v}_{19} \to \mathbf{v}_{20} - \mathbf{v}_{18} \to \mathbf{v}_5 - \mathbf{v}_3 \to \mathbf{v}_4 - \mathbf{v}_2 \to 0.$$

(This shows that $\mathbf{M}$ is one of those "infinitely rare" matrices that cannot be diagonalized. Don't expect to follow these remarks unless you've understood more of linear algebra than I fear most of your colleagues have!)

Richard Parker and I have recently proved that the residual 71st degree equation for λ is irreducible, even when it is read modulo 5. We use the fact that the numbers in a finite field of order q all satisfy $x^q = x$ (since the nonzero ones form a group of order $q - 1$, and so satisfy $x^{q-1} = 1$) .

Working always modulo 5, we used a computer to evaluate the sequence of matrices.

$$\mathbf{M}_0 = \mathbf{M},\ \mathbf{M}_1 = \mathbf{M}_0^5,\ \mathbf{M}_2 = \mathbf{M}_1^5,\ \mathbf{M}_3 = \mathbf{M}_2^5,\ \ldots,\ \mathbf{M}_{73} = \mathbf{M}_{72}^5,$$

and to verify that the nullity (modulo 5) of $\mathbf{M}_{n+2} - \mathbf{M}_2$ was 21 for $1 \leq n \leq 70$, but 92 for $n = 71$. Note that the 21 vectors of the above "alternate base" are *eigenvectors* of $\mathbf{M}_2$ whose eigenvalues (modulo 5) lie in the field of order 5.

If the 71st degree equation were reducible modulo 5, then $\mathbf{M}_2$ would have an eigenvector linearly independent of these with eigenvalue lying in some extension field of order $q = 5^n$ $(1 \leq n \leq 70)$. But then the eigenvalues ϕ of these 22 eigenvectors would all satify $\phi^q = \phi$, and the 22 eigenvectors would be nullvectors for

$$(\mathbf{M}_2)^q - \mathbf{M}_2 = \mathbf{M}_{n+2} - \mathbf{M}_2,$$

contradicting our computer calculations.

It is rather nice that we were able to do this without being able to write down the polynomial. However, Professor Oliver Atkin of Chicago has since kindly calculated the polynomial explicitly and has also evaluated its largest root λ as

$$1.303577269034296391257099112152549 8$$

approximately. The polynomial is

$$\begin{aligned}
&x^{71} - x^{69} - 2x^{68} - x^{67} + 2x^{66} + 2x^{65} + x^{64} - x^{63} - x^{62} - x^{61} \\
&- x^{60} - x^{59} + 2x^{58} + 5x^{57} + 3x^{56} - 2x^{55} - 10x^{54} - 3x^{53} - 2x^{52} + 6x^{51} \\
&+ 6x^{50} + x^{49} + 9x^{48} - 3x^{47} - 7x^{46} - 8x^{45} - 8x^{44} + 10x^{43} + 6x^{42} + 8x^{41} \\
&- 5x^{40} - 12x^{39} + 7x^{38} - 7x^{37} + 7x^{36} - x^{35} - 3x^{34} + 10x^{33} + x^{32} - 6x^{31} \\
&- 2x^{30} - 10x^{29} - 3x^{28} + 2x^{27} + 9x^{26} - 3x^{25} + 14x^{24} - 8x^{23} - 7x^{21} \\
&+ 9x^{20} + 3x^{19} - 4x^{18} - 10x^{17} - 7x^{16} + 12x^{15} + 7x^{14} + 2x^{13} - 12x^{12} - 4x^{11} \\
&- 2x^{10} + 5x^{9} + x^{7} - 7x^{6} + 7x^{5} - 4x^{4} + 12x^{3} - 6x^{2} + 3x - 6
\end{aligned}$$