

A NEW GEMATRIA

by Lee Sallows

"You never learn anything, you only get used to it". — Laurent Siklóssy

Septivigesimal Notation

The method of representing numbers by means of place-value notations that differ in their *radix* or *base*, such as binary (base 2), octal (base 8), decimal (base 10), and so on, is nowadays widely familiar. We remind ourselves that the numeral "123" interpreted in base b means $1 \times b^2 + 2 \times b^1 + 3 \times b^0$, the case $b = 10$ then yielding $1 \times 10^2 + 2 \times 10^1 + 3 \times 10^0 = 100 + 20 + 3 = 123$, or one hundred twenty-three in our conventional decimal notation. In each case the total number of distinct symbols or *digits* called for is simply the same as the base indicated: two symbols for binary, eight for octal, etc. Standard practice is to use the required number of decimal digits (0-9) for all bases up to ten and to supplement these with letters of the alphabet for higher bases. Thus, in addition to 0-9, hexadecimal (base 16) employs A for 10, B for 11, etc., up to F for 15, to complete its sixteen digits, a hexadecimal numeral such as AB9F then indicating $10 \times 16^3 + 11 \times 16^2 + 9 \times 16^1 + 15 \times 16^0$ (= 43935 in decimal). This recurrent use of the familiar decimal digits is a convenience for us decimal-oriented users, yet inessential. Our choice of what symbols to use as digits in hexadecimal or in any other system is entirely free. Sixteen *runes* (ranked in an agreed order) would serve equally. Or the first sixteen letters of the alphabet, say. From the standpoint of logology however, to neglect the remaining ten letters would seem a job half done.

An obvious idea then is a number system using base 26, with its digits comprising A to Z in their natural order. This implies a notation in which every natural number is represented by a unique string of letters, while every string of letters could be interpreted as a unique natural number; an intriguing prospect offering scope for developments beyond reach of traditional gematria. However, there is a problem. For in that case, as first digit the letter A will represent zero, which is the place-holder or empty position indicator, a key element in any positional number system. The trouble is that this not only conflicts with the more natural $A = 1$, $B = 2$, etc., the behavior of zero is peculiar in that leading zeroes may be appended to or deleted from any numeral without affecting its value. Thus in decimal notation, 12, 012, 0012,... all stand for twelve. It seems arbitrary that A should be singled out to exhibit an aberrant property not shared by the other letters. Still more serious, the numerical value of string ABC will then be the same as that of BC, while AARON will equal RON (!), and so on. Our hope of a system in which every distinct letter string stands for a unique number is thus *not* realized in this scheme. Even reordering the digits—which is undesirable—merely shifts the problem to a new letter.

Having pondered this predicament on and off over a period of years, a solution recently occurred to me. It is a simple step, but one impossible to make as long as base 26 is the starting point—a

central assumption difficult to shake off. The answer is to introduce an extra symbol for zero. I shall use "**␣**", the (bold) underscore, which is a near approach to that typographical cousin of zero, the blank. With the addition of this extra character we arrive thus at a base 27 or *septivigesimal* (SV) system whose digits comprise the underscore and the upper case letters, so that $\underline{ }=0, A=1, B=2, \dots Z=26$. Henceforth, if the underscore is a leading character it may be dropped or ignored; elsewhere it must be retained. Note that use of the blank itself would have left any *trailing* zeroes invisible and thus undetectable, a fatal flaw. On the other hand, there is nothing to prevent us from *interpreting* underscores as blanks, and vice versa, should we so choose. The advantage of doing so will emerge.

The SV notation for numbers beyond 26 is then best illustrated through visualizing a suitable *odometer*, which is a perfect model of a positional number system, the Number made Flesh, as it were. Its simple mechanism is familiar to us in the tape counter or car mileometer: a row of rotatable discs viewed edge-on, each bearing the ten digits evenly inscribed around their periphery. In an odometer designed for SV, however, 0-9 are replaced by $\underline{ }-Z$, the discs then rotating in 27 rather than 10 steps. Each successive input or event to be counted advances the rightmost disc by one digit. With the completion of a cycle its left hand neighbor is advanced one digit also. And so on with the remaining discs. Each new input thus gives rise to a new combination of digits in the viewing aperture. Starting from the reset or zero position, which is a row of underscores, the SV notation for a given number n is then found on the readout after entering n successive inputs:

$n:$ 0, 1, 2, .. 26, 27, 28, 29, ... 53, 54, 55, 56, 728, 729, 730, 731, .. SV: $\underline{ }$, A, B, .. Z, A $\underline{ }$, AA, AB, ... AZ, B $\underline{ }$, BA, BB, ZZ, A $\underline{ } \underline{ }$, A $\underline{ }$ A, A $\underline{ }$ B, ..

The sequence speaks for itself. As the odometer advances, the readout progresses systematically through every possible combination of digits, all the 1-digit strings followed by all the 2-digit strings, etc., the first yielding a numeral for one, the second a numeral for two, and so on. Our goal is achieved: every distinct string of *letters* is now a unique code or label for a distinct positive integer, while leading zeroes may be added or deleted without affecting any letter; the behavior of A is no longer aberrant. At the same time, the codes for many integers include underscores, a fact which need cause no concern and will even prove useful in a moment. Furthermore, negative integers, fractions, real numbers, and so on, can all be represented by bringing in minus signs, "decimal points" (*unit* points, or *separators* as they are better called), and other signs, in the usual way. Aside from its offbeat radix and digit symbols SV is an entirely conventional number system.

To convert a letter string to its decimal equivalent is easy. Take CAT. Noting C=3, A=1, T=20, we write: $CAT = 3 \times 27^2 + 1 \times 27^1 + 20 \times 27^0 = 2187 + 27 + 20 = 2234$. The reverse process of deriving the SV notation for a given number calls for successive divisions by 27. Consider 74417. Then $74417 \div 27 = 2756$ with remainder 5. E=5 is thus the final digit of the SV numeral. The process is then repeated with the previous quotient: $2756 \div 27 = 102$ with remainder 2. B=2, becomes the next to last digit. Then $102 \div 27 = 3$ with remainder 21. U=21 precedes B. The end is reached when the quotient falls below 27, its value then standing for the leading digit: C=3 in our example. So, $74417 = CUBE$. Test yourself: what is 1492 in SV? See Answers and Solutions.

The logological implications of SV now begin to unfold. For as the sequence above extends it will come to include every English word, along with all the words of those languages using the Roman alphabet. Every word is an integer written in SV notation, and as such, every word is greater or lesser than any other given word, is an odd or an even word, is a composite or a prime word, is a perfect square or cube or whatever, and so on. For instance, the smallest English word that is prime is AN, which is 41. Another is PRIME, which is 8864267. Henceforth it becomes natural to identify a word with the properties of the integer it represents. Likewise, a dictionary now reveals itself as a list of (essentially random) integers whose lexicographical ordering is different to what it would be were they ordered by magnitude. Moreover, the advantage of interpreting underscores as blanks now appears. For then the series of natural numbers must also include every finite *sequence* of words, such as WORD_WAYS, THE_NUMBER_OF_THE_BEAST, and INTERPRET_UNDERSCORES_AS_BLANKS. That is, neglecting punctuation and other non-alphabetic signs, every sentence, indeed every *text* corresponds to a unique integer, also. The ramifications of all this in relation to logology become nothing short of momentous in the realization that from here on *relations among words (texts) are expressible in mathematical equations.*

Adding and Multiplying Words

If every word is a number then two words can be added together and their sum expressed in SV notation. Chances are the result is a meaningless string of letters. Might cases exist that yield a *word*? There are thousands.

We need a name for integers whose representation in SV is a word. I propose *wint*, a contraction of *word-integer*. Consider the addition of two wints, 399749 and 85817, which are THIN and DISK, respectively. Starting from zero and advancing our SV odometer 399749 steps the readout will be THIN. Advancing its thin discs a further 85817 steps must show a result that matches the decimal number $399749 + 85817 = 485566$. It is another wint: XRAY. That is, $485566 = 399749 + 85817$ or $XRAY = THIN + DISK$, which are just equivalent statements in different notations. Starting with XRAY and *reversing* the odometer by 399749 backward steps thus produces $XRAY - THIN$, which is DISK, as we could have predicted from elementary arithmetic.

The above suggests two approaches to calculating with SV numerals. One is to convert from SV into decimal, perform computations as usual, and then translate the answer back into SV. The other is to work directly in SV itself, a method requiring addition and multiplication tables, as provided on the next page. Consider an example:

$$\begin{array}{r} \text{ONLY} \\ \text{REST} \\ \hline \text{AFTER} \end{array} +$$

We proceed exactly as with ordinary sums, beginning with the right-hand digits. The tables show $Y + T = AR$. Thus, write down the unit's R and carry the twenty-seven's A. Next, $L + S = AD$, to which must be added the carry: $AD + A = AE$. Write down the E and carry the A. Then $N + E + A$ (the second carry) = T. Lastly, $O + R = AF$. $ONLY + REST = AFTER$ you have mastered this example. The procedure for subtraction, multiplication, and (long) division is entirely analogous. The following sums can be checked in the same way:

ANTI + BULK = DIET	BEST + POLE = RUDY	DUET + BULL = GORE
HAVE + GIFT = OKAY	BEST + HOLE = JUDY	ACID + CASH = DEAL
CHAT + TALK = WIND	BEDS + JOYS = LUCK	AIDS + KEYS = LOCK
IDEA + KEGS = TILT	AVOW + JILT = LEAP	FULL + GOAT = NINE
FEAR + OMEN = URGE	AVOW + WILL = YEAH	FUCK + PUNT = WORD
AMEN + CUBE = EGGS	EDEN + FATE = KEYS	FLAW + MASK = SMUG
EGGS + GULP = MATH	CUBE + GOLF = KINK	INCH + MAIL = VOLT
CHEF + EGGS = HOLY	CUBE + POLO = TINT	SEND + MORE = AEUEI

Semantic cohesion lends charm to some of these examples. If you are *anti-bulk* then *diet*; to *have a gift* is *okay*; *chat*, *talk*, and *wind* are synonyms; *Rudy* and *Judy* doubtless know that the *joys* of sound sleep in comfortable *beds* is a matter of *luck*; SEND + MORE is *not* equal to MONEY, contrary to cryptarithmic superstition. But how were these specimens derived? Doing sums in SV would be a lot easier if only we knew our SV tables by heart. Yet even then calculating by hand would remain tedious and slow. Prohibitively slow. Using a computer for the job is not merely quicker, its ability to perform fast searches through a data base makes it a virtually indispensable tool for research in this field. The above examples, for instance, are among more than 7000 discovered by a simple program that, using a stored list of words, took each 4-letter pair in turn, calculated their sum and then checked to see if it was present in the list. This process need not be limited to 4-letter words of course, or just pairs, or addition only.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	+		
A	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A_	A	
B	B	D		D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A_	AA	B
C	C	F	I		F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A_	AA	AB	C
D	D	H	L	P		H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A_	AA	AB	AC	D
E	E	J	O	T	Y		J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A_	AA	AB	AC	AD	E
F	F	L	R	X	AC	AI		L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A_	AA	AB	AC	AD	AE	F
G	G	N	U	AA	AH	AO	AV		N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A_	AA	AB	AC	AD	AE	AF	G
H	H	P	X	AE	AM	AU	BB	BJ		P	Q	R	S	T	U	V	W	X	Y	Z	A_	AA	AB	AC	AD	AE	AF	AG	H
I	I	R	A_	AI	AR	B_	BI	BR		R	S	T	U	V	W	X	Y	Z	A_	AA	AB	AC	AD	AE	AF	AG	AH	AI	I
J	J	T	AC	AM	AW	BF	BP	BZ	CI	CS		T	U	V	W	X	Y	Z	A_	AA	AB	AC	AD	AE	AF	AG	AH	AI	J
K	K	V	AF	AQ	BA	BL	BW	CG	CR	DB	DM		V	W	X	Y	Z	A_	AA	AB	AC	AD	AE	AF	AG	AH	AI	AJ	K
L	L	X	AI	AU	BF	BR	CC	CO	D_	DL	DX	EI		X	Y	Z	A_	AA	AB	AC	AD	AE	AF	AG	AH	AI	AJ	AK	L
M	M	Z	AL	AY	BK	BX	CJ	CW	DI	DV	EH	EU	FG		Z	A_	AA	AB	AC	AD	AE	AF	AG	AH	AI	AJ	AK	AL	M
N	N	AA	AO	BB	BP	CC	CQ	DD	DR	EE	ES	FF	FT	GG		AA	AB	AC	AD	AE	AF	AG	AH	AI	AJ	AK	AL	AM	N
O	O	AC	AR	BF	BU	CI	CX	DL	E_	EO	FC	FR	GF	GU	HI		AC	AD	AE	AF	AG	AH	AI	AJ	AK	AL	AM	AN	O
P	P	AE	AU	BJ	BZ	CO	DD	DT	EI	EY	FN	GC	GS	HH	HX	IM		AE	AF	AG	AH	AI	AJ	AK	AL	AM	AN	AO	P
Q	Q	AG	AX	BN	CD	CU	DK	EA	ER	FH	FY	GO	HE	HV	IL	JB	JS		AG	AH	AI	AJ	AK	AL	AM	AN	AO	AP	Q
R	R	AI	B_	BR	CI	D_	DR	EI	F_	FR	GI	H_	HR	II	J_	JR	KI	L_		AI	AJ	AK	AL	AM	AN	AO	AP	AQ	R
S	S	AK	BC	BV	CN	DF	DY	EQ	FI	GA	GT	HL	ID	IW	JO	KG	KZ	LR	MJ		AK	AL	AM	AN	AO	AP	AQ	AR	S
T	T	AM	BF	BZ	CS	DL	EE	EY	FR	GK	HD	HX	IQ	JJ	KC	KW	LP	MI	NB	NV		AM	AN	AO	AP	AQ	AR	AS	T
U	U	AO	BI	CC	CX	DR	EL	FF	G_	GU	HO	II	JC	JX	KR	LL	MF	N_	NU	OO	PI		AO	AP	AQ	AR	AS	AT	U
V	V	AQ	BL	CG	DB	DX	ES	FN	GI	HD	HZ	IU	JP	KK	LF	MA	MW	NR	OM	PH	QC	QY		AQ	AR	AS	AT	AU	V
W	W	AS	BO	CK	DG	EC	EZ	FV	GR	HN	IJ	JF	KB	KY	LU	MQ	NM	OI	PE	QA	QX	RT	SP		AS	AT	AU	AV	W
X	X	AU	BR	CO	DL	EI	FF	GC	H_	HX	IU	JR	KO	LL	MI	NF	OC	P_	PX	QU	RR	SO	TL	UI		AU	AV	AW	X
Y	Y	AW	BU	CS	DQ	EO	FM	GK	HI	IG	JE	KC	LA	LZ	MX	NV	OT	PR	QP	RN	SL	TJ	UH	VF	WD		AW	AX	Y
Z	Z	AY	BX	CW	DV	EU	FT	GS	HR	IQ	JP	KO	LN	MM	NL	OK	PJ	QI	RH	SG	TF	UE	VD	WC	XB	YA		AY	Z
x	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z			

Above are addition and multiplication tables for SV numbers. Example: P+Q=AF, P×Q=JB.

Either of the two methods described above may be implemented in a program of this kind, which of them is best? Decimal operations are a built-in feature of every programming language, while routines for converting from SV to decimal and back again are simple enough to write, so this would seem easier and fast. So it is, provided your software can handle large enough integers. In Pascal, for instance, the largest integer allowed is 2147483647, which is ENQWLTI, so the highest win you could process would be ENQUIRY; a similar limit is met when using a pocket calculator. Routines for calculating directly in SV evade this problem but are more difficult to write and slower in execution. Incidentally, programmers will find "@" preferable to "_" in representing zero, as a glance at an ASCII code chart will make clear.

Question: Given two wins a and b , what is the likelihood that $(a+b)$ is a win? The answer depends on its length, and hence on that of a and b . The sum of two n -digit numbers contains n or $n+1$ digits. Suppose we add two smallish 3-letter wins to produce a 3-letter result and ignore cases that contain underscores. The total number of possible 3-letter strings is $26^3 = 17,576$. How many of these are words? A modest lexicon of 25,000 words stored on my hard disk reveals 747 items of 3 letters. The chance of a random 3-letter string forming a word is thus something like 747 in 17,576 or about 4.25 per cent. Equivalent figures for some other string lengths are as follows:

No. of digits	1	2	3	4	5	6	7
Percent words	7.7%	13.4%	4.25%	0.46%	0.026%	0.001%	0.00005%
No. of words	2	91	747	2140	3091	3791	4043

So whereas the chance of two smallish 3-letter words adding to form another word is around one in twenty, for two smallish 7-letter words it is only about one in two million, despite there being over five times more 7-letter words than 3-letter words, as seen in the bottom line above. On the other hand, more 7-letter words mean more candidate pairs to try. Nonetheless, solution counts diminish rapidly beyond 4-letter words. Instances that combine words whose meanings make them in some sense apposite then become very much rarer. A random selection of 5 and 6-letter word sums follow. Sometimes the sheer incongruity of the word combinations produces quite a comical effect as the brain struggles to invent scenarios to "explain" the equations.

AMEND + BEING = CROAK	ADAPT + DIODE = EMPTY	ANGER + MANIA = NOUNS
ATONE + EXTRA = GRIEF	AUDIO + JOKED = LIONS	BEAST + MADLY = OFFER
BYRON + FEMME = IDEAS	CACTI + EXILE = HYMEN	COMIC + LUMPY = PIZZA
CRIES + DREAM = HINGE	DWARF + MINUS = REPLY	EXIST + QUOTE = WRYLY
FIFTH + NINTH = TRUMP	GREAT + MEDAL = TWICE	HAPPY + ISLAM = QUARK
HOVEL + VENOM = ACUITY	WALTZ + WASTE = ASCEND	ODDLY + POKED = ADSORB
ANCHOR + PROVEN = RESCUE	ANTICS + REVEAL = STONED	
FITFUL + JESUIT = POMADE	ACROSS + RANDOM = SEETHE	
FORMAT + LYRICS = SNIVEL	ACROSS + RANCID = SEESAW	
CLOSER + RAPING = UNEASY	ELICIT + LOVERS = RADIAL	
HILTON + MEEKLY = UNREAL	GAUCHE + HAIRDO = OCCULT	
REGAIN + WINKLE = ANNULUS	SADIST + THIEVE = ALIMONY	

The focus here on word pairs of equal length is explicable. Consider AFFAIR + CUB = AFFECT. Leading digits in the longer word survive in the total, a result less pleasing than when a contrasting word is produced. The smaller the difference in lengths the less glaring is this effect. Single letter differences hardly show up, and sometimes not at all:

BOWS + TOTEM = TRIBE	COIN + HALVE = HEADS	BUMS + SQUAT = STOOL
CATS + DOGMA = DRIFT	LUST + COYLY = DATER	GRIT + SAUCE = SILLY
MIRE + SLICK = SYRUP	RIDE + ORGAN = PIPES	JEST + DEATH = DOGMA
HOLDS + SEPTUM = SNEEZE	DIANA + SPIDER = STRESS	
DOWRY + DECEIT = DISBAR	NOBLE + KNIGHT = LAXITY	
QUART + ARMPIT = BIGRAM	MOANS + BANANA = BOBCAT	

After $a+b$, what about $a \times b$? The product of two n -digit numbers contains $2n$ or $2n + 1$ digits. This is n digits greater than in their sum. The likelihood of $(a \times b)$ forming a wint is thus drastically reduced, a fact reflected in the extreme rarity of instances. The problem of leading digits surviving in the result disappears with products so that word lengths are not important. Even so, a program that ran through my lexicon and tested every single pair a, b formable from words using 6 or less letters discovered fewer than 100 cases for which $a \times b$ is a wint. A few of the better examples, by which I mean those using no single letter words, acronyms, proper names, etc., follow. The funeral tone in the right hand group (*dig trench..bury body..rugged coffin..cram corpses*) is pure chance!

IS × GEM = BOXED	JOB × JIG = DAEMON	DIG × DEFT = REVOKE
BY × SAD = BASIS	TAB × HOP = FIXATE	HAW × BODY = TRENCH
BE × FRET = NOUNS	BAY × BREW = ENOUGH	FRY × BURY = RUGGED
AT × KING = STOVE	ARE × BUMS = DRIVEN	FRY × NIT = COFFIN
BE × DYKE = JUNKY	TO × ZINC = TARTAR	ZIP × CRAM = CORPSES

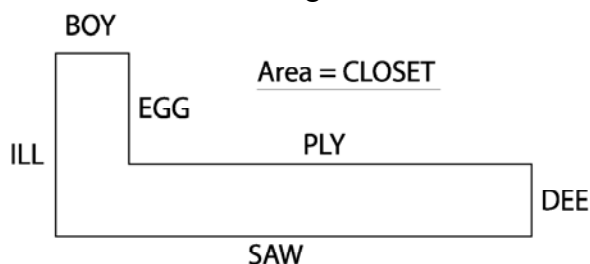
The equations looked at so far are of simplest type: $y = a+b$ and $y = ab$. Substitutions and transposals lead to more intricate forms; e.g, by combining (BE × FRET = NOUNS) with (FRET = SLID – LUCK) and (NOUNS = ANGER + MANIA), we get BE = (ANGER + MANIA) / (SLID – LUCK). And so on to any degree of complexity. *Any* degree of complexity? Peering beyond these elementary expressions, at this point in our progress all kinds of possibilities begin to suggest themselves. To employ a metaphor, until now we have been examining the contours of an unfamiliar object recovered from the ground. There are suggestions it may contain explosive material. Now we have realized the thing is *ticking*. This brings us to the Big Bang.

The New Gematria

On my bookshelf is *The Penguin Dictionary of Mathematics*. I open it at random and select an entry at will. The item reads, "**GCD**. Abbreviation for *greatest common divisor*." Of course: the GCD of 12 and 18 is 6, for instance, because 6 is the highest number that divides both 12 and 18. GCDs for larger numbers can be found using a simple process called Euclid's algorithm. But if every word is a number then every pair (or more) of words must have a GCD as well. Could some of these GCDs themselves be words? A program similar to the one described above but now incorporating Euclid's algorithm pours out examples by the score: The GCD of JET and PIE is BAD, of BELCH and DRYLY is LAW, of FEEDS and EARLY is ADD, etc. The GCD of STEWS and PENNY is GCD. Verification of these instances is tedious but straightforward, easier if you convert to decimal first. Easier still performed on your PC.

Re-opening the *Dictionary* elsewhere I land by chance on **Pythagoras Theorem**. Let's see, could

a right-angled triangle have sides whose lengths are wints? That is, can three words be found to satisfy $a^2 = b^2 + c^2$? A program that took pairs of words, calculated the square root of the sum of their squares and then checked the result against a word list has identified a cheerful fact about the square on one hypotenuse: $AN^2 = I^2 + AM^2$. This is the sole instance discovered; can any reader track down a second? But if lengths can be wints, so can areas. Take the square on the side of length AM in the triangle just mentioned; its area is $AM^2 = BEG$. In fact this is the *only* such square I have found, although the products we looked at earlier yield rectangles with the required property; e.g., If length \times breadth = JIG \times JOB, then area = DAEMON square units. Turning instead to triangles in general, armed with Heron's formula, which expresses area in terms of 3 sides, I went fishing for wint-edged specimens whose area is also a wint, but met with no success. Not to be outdone I tried a different figure:



The six sides of the L-shape above are all wints, while its area is given by $(BOY \times EGG) + (DEE \times SAW) = (BOY \times ILL) + (DEE \times PLY) = CLOSET$.

Can any reader improve on this, perhaps by finding a different shape that uses a more apt set of words? How about a wint-sided figure whose area is equal to AREA, or some other appropriate word? The possible variations are endless.

It is time for another lucky dip in the *Dictionary*. I find: "**Congruence modulo n** . A relation expressing the fact that two integers differ by a multiple of a chosen number n ." That's right: we write $a \bmod n \equiv b$, (spoken, " a modulo n is congruent to b ") when the result of dividing a by n leaves a remainder of b . Hence: $10 \bmod 4 \equiv 2$, and $9 \bmod 3 \equiv 0$, and well, you can see it coming:

CIRCLE mod LOVE \equiv FLAG	MUSES mod BLUE \equiv ARMY
LEEWAY mod HATE \equiv ANEW	HANGS mod GREEN \equiv JIBE
OUTCRY mod THIS \equiv FOXY	MODULO mod ARTHUR \equiv DEAF
LEGATO mod THAT \equiv SNOW	PERFORM mod MODULO \equiv CROCUS

Ten samples taken at random from among xillions the computer finds. This abundance encourages specialization. Consider: $(FOUR \times SIX) \bmod TEN \equiv ART$, and $(ZERO + FOUR + EIGHT + TWELVE) \bmod SIX \equiv GOD$. Almost every wint is a *number-word*. Can it be done using number-words only? It can: $(ELEVEN \times NINETY SIX) \bmod TEN \equiv ONE$. But alas, $(11 \times 96) \bmod 10 \equiv 6 (\neq 1)$! What we really seek is an expression that works on *both* levels of interpretation. Three (non-trivial) instances are as follows:

(FORTYTWO + FORTYNINE) mod SIX \equiv ONE	and (42+49) mod 6 \equiv 1
(EIGHTYFIVE - EIGHTYTWO) mod TWO \equiv ONE	and (85-82) mod 2 \equiv 1
(NINETY NINE - TWELVE) mod FIVE \equiv TWO	and (99-12) mod 5 \equiv 2

It would be nice to find two-level *equations* rather than congruences. However, the difficulties involved are horrendous. To date I have not succeeded in finding even a single-level equation. Immortal fame awaits the first logologist to succeed.

What has the *Dictionary* in store for us next? I pick a new page at random and my glance alights on "**Magic square**. A square array of numbers whose sum in any row, column, or main diagonal is the same." The contingency this suggests is exciting: a magic array using *wints*! Can it really be done? A program for seeking squares formed from any set of special numbers is not difficult to write. Contrary to intuition tighter internal constraints make it harder to find wints that will satisfy 3×3 squares than it is for larger types. The rhyming word triples that appear in solutions reflect these closer restrictions:

HOG	BIG	HUG	=	6244	1708	6406	LASH	ARK	LATE	
FUG	FOG	FIG		4948	4786	4624		HATE	HASH	HARK
DIG	JUG	DOG		3166	7864	3328		DARK	PATE	DASH

The magic constants here are SRU = 14358 (left) and XECX = 476142 (right). The challenge this suggests is obvious: can a square can be found whose constant total is itself a wint? Moving on to 4×4 squares the problem has yielded to attack:

EWE	HOE	TOG	LOW	magic constant = ATOM	DOT	RAJ	TIT	PAX	magic constant = BEAT
NUN	RIP	FIN	HAW		PIT	TAX	FOR	PAL	
PIN	RAW	DUN	HIP		TAR	RIP	NAP	FOX	
JOG	BOW	OWE	ROE		RAP	BOX	RAT	TIN	

FAR	FOB	OAT	BUS	magic constant = ACME	DIM	OWE	TUG	RAP	magic constant = BEAN
GAY	JUG	FAT	FOG		RIG	TAP	DOT	PAY	
HUN	MAR	ON	HAM		THE	TIP	NAP	DID	
HOB	BE	HUE	MAT		PAP	BUD	SPY	TOW	

Here the top squares have a (hidden) "graeco-latin" structure, which entails that the four words in each quadrant and the four corner words of each 3×3 subsquare also total to ATOM and BEAT. Evidently the potential of SV magic squares to combine both mathematical and logological properties provides enormous scope for future investigations. For the time being, however, I propose a specific challenge: can any reader discover a square with MAGIC as its constant total?

Meanwhile, turning again to the *Dictionary* my eye falls on: "**Irrational number**. A real number that cannot be written as an integer or as a quotient of two integers." The classic example is the diagonal of a unit square: $\sqrt{2}$, or 1.4142135..., the dots indicating an endless string of digits. But can irrationals be expressed in SV? Of course they can, SV is just another number system. Using a recursive formula for finding the square root of any number k , such as $x_{n+1} = \frac{1}{2}(x_n + k/x_n)$, in which x_1 is set to an initial guess, while x_2, x_3, \dots are successively closer approximations, the SV expansion of $\sqrt{2}$, or as we should now say, \sqrt{B} , turns out to be A.KDYZAWWXDUAFBSV-

OTSTLYCBVUNUTPBFGSSEW.. (as I trust A.K. Dewdney will note with interest). The decimal interpretation of this is then $1 \times 27^0 + 11 \times 27^{-1} + 4 \times 27^{-2} + 25 \times 27^{-3} + \dots$, (since A=1, K=11, D=4, etc.) = 1.4142135... . Note that although we call the latter a "decimal", the term becomes a confusing misnomer when applied to SV expansions like A.KDYZA..., i.e., SV numerals that include separators or unit (not "decimal"!) points.

Looking over this result we find that the first English word to occur in the string is NUT, closely followed by SEW. As it happens this initial fragment contains no underscores. However, if $\sqrt{2}$ is a so-called "normal" number, as mathematicians believe, then its infinite (strictly, *chaotic*) sequence will eventually include every possible pattern of n digits, for every n . An odd-seeming yet perfectly serious question is thus: What is the first English *sentence* to occur in the septavigesimal expansion of $\sqrt{2}$? And the same can be asked of all similar numbers, including say, π , e , and Euler's constant, γ . Who knows? Perhaps the first sentence to appear in π is .._GOD_EXISTS_.. , while that in e might be .._PROVE_IT_.., and that in γ might be .._THIS_NUMBER_IS_IRRATIONAL_.., which would be fun since nobody yet knows whether this is true or not.

Another intriguing number is phi (ϕ), the golden ratio, which equals $\frac{1}{2}(1 + \sqrt{5})$. Its SV expansion begins thus: A.PRNTPFCUCRKDYGRYLLCQNBIGOVQTRTLYIKM..., in which BIG is the first word to occur. This is an interesting case to consider since my *Dictionary of Mathematics* has now fallen open at: "**Reciprocal**. The number produced by dividing 1 by a given number." The reciprocal of C is thus $1/C$ or A/C . Simple long division reveals the answer: C into A won't go, so write ".", append a zero to A and repeat: C=3 into A_ = 27 goes I=9 times exactly. Thus the reciprocal of C is _I, which is $9 \times 27^{-1} = 9/27 = 1/3 = 0.3333\dots$, in decimal. In fact the result of *any* long division is always a *recurring* quotient, although we tend to overlook this when the repetend is made up of zeroes only. The reciprocals of the alphabet illustrate this point:

A/A = A...	A/N = _AYAYAYAYA..
A/B = _MMMMMM..	A/O = _AUPEJUPEJ..
A/C = _I..	A/P = _AROEAROE..
A/D = _FTFTFT..	A/Q = _AOWVFINGYKCDTQLSAOWV..
A/E = _EJUPEJUPEJ..	A/R = _AMMMMMMMM..
A/F = _DMMMMM..	A/S = _AKIYOQAKIYOQ..
A/G = _CWCWCW..	A/T = _AILDAILDA..
A/H = _CJCJCJ..	A/U = _AGSGSGSGS..
A/I = _C	A/V = _AFCRKAFCRK..
A/J = _BRXHBRXHBR..	A/W = _ADRUCNBIJOGADRU..
A/K = _BLGIVBLGIV..	A/X = _ACJCJCJCJ..
A/L = _BFTFTFT..	A/Y = _ABDHQGOCFLYXVRISKWTNABD..
A/M = _BBBBBBB..	A/Z = _AAAAAAAAA..

What has all this to do with ϕ ? Simply this: the reciprocal of ϕ has a fascinating property, namely: $1/\phi = 1 - \phi$. That is, $1/\phi =$ _PRNTPFCUCRKD.., which is the same string as above but minus its leading A! Some readers may like to test this by taking a small portion of ϕ , say A.PRNT, and dividing it into A by long division (using the multiplication table as an aid). The exercise requires patience but is instructive.

Before leaving these reciprocals take another look at $1/E =$ _EJUPEJUP.. . Now E, or 5, is a prime number, while the length of the repetend, EJUP, is 4. Students of recreational math may

recall that when the reciprocal of a prime p has a period of $p-1$ the repeating sequence forms a so-called *cyclic number*. A cyclic number of n digits has the intriguing property that when multiplied by any number from 1 to n , the resulting product reveals the self-same n digits arranged in the same cyclic order. Thus:

$$\begin{aligned} A \times \text{EJUP} &= \text{EJUP} \\ B \times \text{EJUP} &= \text{JUPE} \\ C \times \text{EJUP} &= \text{PEJU} \\ D \times \text{EJUP} &= \text{UPEJ} \end{aligned}$$

In fact *cyclic numeral* would be a better term for these curiosities since their periods obviously depend on the radix of the number system in use. Cyclicity is thus not a property of numbers but only of numbers-written-in-a-certain-base. EJUP is the smallest cyclic numeral in base 27, i.e., SV. After $1/E$, the next prime reciprocal to produce a new cyclic is $1/17 = 1/Q = \text{..AOWVFINGYKCDTQLSAOW..}$, with period 16. Better yet, though, $1/29 = 1/AB$, which produces the next case, results in a *pangram* of period 28: $\text{..YCSNXEOVIHJFMZAWGLBUKDQRPTMYC..}$. M alone occurs twice. So if c is the integer represented by these 28 digits (27 discounting the leading blank), we have:

$A \times c = \text{..YCSNXEOVIHJFMZAWGLBUKDQRPTM}$	$O \times c = \text{MZAWGLBUKDQRPTM..YCSNXEOVIHJF}$
$B \times c = \text{AWGLBUKDQRPTM..YCSNXEOVIHJFMZ}$	$P \times c = \text{NXEOVIHJFMZAWGLBUKDQRPTM..YCS}$
$C \times c = \text{BUKDQRPTM..YCSNXEOVIHJFMZAWGL}$	$Q \times c = \text{OVIHJFMZAWGLBUKDQRPTM..YCSNXE}$
$D \times c = \text{CSNXEOVIHJFMZAWGLBUKDQRPTM..Y}$	$R \times c = \text{PTM..YCSNXEOVIHJFMZAWGLBUKDQR}$
$E \times c = \text{DQRPTM..YCSNXEOVIHJFMZAWGLBUK}$	$S \times c = \text{QRPTM..YCSNXEOVIHJFMZAWGLBUKD}$
$F \times c = \text{EOVIHJFMZAWGLBUKDQRPTM..YCSNX}$	$T \times c = \text{RPTM..YCSNXEOVIHJFMZAWGLBUKDQ}$
$G \times c = \text{FMZAWGLBUKDQRPTM..YCSNXEOVIHJ}$	$U \times c = \text{SNXEOVIHJFMZAWGLBUKDQRPTM..YC}$
$H \times c = \text{GLBUKDQRPTM..YCSNXEOVIHJFMZAW}$	$V \times c = \text{TM..YCSNXEOVIHJFMZAWGLBUKDQRP}$
$I \times c = \text{HJFMZAWGLBUKDQRPTM..YCSNXEOVI}$	$W \times c = \text{UKDQRPTM..YCSNXEOVIHJFMZAWGLB}$
$J \times c = \text{IHJFMZAWGLBUKDQRPTM..YCSNXEOV}$	$X \times c = \text{VIHJFMZAWGLBUKDQRPTM..YCSNXEO}$
$K \times c = \text{JFMZAWGLBUKDQRPTM..YCSNXEOVIH}$	$Y \times c = \text{WGLBUKDQRPTM..YCSNXEOVIHJFMZA}$
$L \times c = \text{KDQRPTM..YCSNXEOVIHJFMZAWGLBU}$	$Z \times c = \text{XEOVIHJFMZAWGLBUKDQRPTM..YCSN}$
$M \times c = \text{LBUKDQRPTM..YCSNXEOVIHJFMZAWG}$	$A \times c = \text{YCSNXEOVIHJFMZAWGLBUKDQRPTM..}$
$N \times c = \text{CSNXEOVIHJFMZAWGLBUKDQRPTM..Y}$	$AA \times c = \text{ZAWGLBUKDQRPTM..YCSNXEOVIHJFM}$

Note how the initial digits of successive products run through the alphabet! For an insight into this, as well as for much more on the lore of cyclic numerals, see Martin Gardner's uniformly marvellous *Mathematical Circus* (Pelican).

* * * * *

Life is short and so is space. The *Dictionary* has played its part; its message should be clear. We began with a novel system for writing numbers with letters; every word was revealed as a unique number; elementary arithmetic became applicable to words. BOOM! A Big Bang marks the explosion into being of an entire universe of mathematical word-play. Almost anywhere we look in a dictionary of mathematics, a new variation, a new topic, a new area for research, opens up. Thus far we have not even scratched the surface. Just think for a moment of the subjects not even touched on here: logarithms, trigonometry, progressions (complete: AM, BOB, EAR, ..), group theory, probability, calculus, polynomials (solve: $Ix^2 + NOWx - SAYING = _$; x is a wint), complex numbers, self-descriptive numbers (remember Gödel coding?), and on and on.. Our new gematria is no mere rich vein but a vast mine as endless and ramified as mathematics itself; it can never be exhausted.

Sematria

"Septivigesimal gematria" is a jawbreaker; *sematria* is more musical, and even fits fairly well etymologically: *sema* ($\sigma\eta\mu\alpha$)=sign + *matria* ($\mu\alpha\tau\rho\acute{\iota}\alpha$)=measuring, i.e., measure expressed in signs. *Sematria* grew out of my dissatisfaction with standard gematria, a system whose central feature is really its greatest weakness: different words *equate*. But every word is a unique ordering of letters. What then can we expect of a numerology that is insensitive to letter order? It seemed to me that if we are going to have a system that identifies words with numbers then distinct words must correlate with distinct values. A new approach was evidently needed. To accept A=1, B=2,.. as a starting point was of course natural, but thereafter a *place-value* system would be demanded. The trouble was that my back-of-the-envelope try-outs would always run into the problem created by zero. The place-holder was vital, I knew that. A to Z were already spoken for. On intuitive grounds the blank ought to stand for zero; it just seemed right. But the blank was invisible and that would prove ruinous! The riddle looked insoluble. And then the penny dropped. Once an extra symbol for zero was brought in everything fell into place. The result, I submit, is less an invention than a discovery. It is not just a system for interpreting words as numbers, it is *the* system: simple, natural, inevitable; the only wonder is that it has not appeared earlier*. But that is explicable, perhaps. For simple as it is, there is no denying that calculating on paper in SV is plain arduous. Pretty excruciating in fact, if carried out at length: try seeking two words whose sum is a word *by hand*. No, it has taken until now, the age of the personal computer, before *sematria* could come into its own.

One alternative to SV (suggested to me, aptly enough, in the parking lot of the Max-Planck Institute for Psycho-Linguistics in Nijmegen, by Doug Hofstadter) deserves mention. It is base 26 but *without zero*: A=1, B=2, ... Z=26, AA=27, AB=28, etc. That's all. Et voilà: every letter string is a unique integer! However, loss of zero exerts a price, for although we can still add and multiply, we cannot, in general, divide. This is because expressions like A/B have no solution: the smallest number that can be written in this notation is A, or .A if we bring in the unit point. The resulting system is a so-called *integral domain*, whereas SV has the structure known as a *field*, a property that allows us to work with any real number. Of course the present system could be extended to include more typographical signs: .=0, A=1,.., Z=26, !=27, ?=28, \$=29,.., but then where do we stop? Base 36? Base 50? Base 100? Without a uniform standard, word values will differ from system to system. The alphabet furnishes our only natural standard.

Here and there SV can illuminate an existing topic. Writing in *Word Ways* for February 1992, Christopher McManus introduced **halfway words**, which are word trios in which letters in the centre word lie midway in the alphabet between corresponding letters in the outer words, e.g., AGE - JIG - SKI. It is a neat invention. However, now look at it thus: JIG is the *average* or *arithmetic mean* of AGE and SKI, because $\frac{1}{2}(\text{AGE}+\text{SKI}) = \text{JIG}$, a particular instance of $\frac{1}{2}(a+b) = c$, with *a*, *b* and *c* wints. Halfway trios thus form arithmetic progressions. The converse however need not be true: CAUSES is the mean of BINARY and CUBISM, CIGARS is the mean of LIES and FRAUDS, GHOST is the mean of FAITH and HOUSE, and HINDOO is the mean of LOOP and PROTON (to mention but a few), none of which conform to McManus' definition. So halfway trios are actually special instances of arithmetic means, a result of their definition in terms of form (digit properties) rather than content (numerical value).

Sematria offers enormous opportunities for research in new realms of mathematical wordplay,

but the difficulties entailed in tracking down *worthwhile* finds should not be underrated. Programming computers to perform arithmetic with words is fairly easy, getting them to recognize and hunt down *fruity* correlations is something else again. Limitations of space have made it impossible to discuss algorithms here, but I shall be glad to exchange ideas, great or small, on programs or other aspects of sematria with anyone interested. Meanwhile I can only encourage others in their explorations and wonder into what strange worlds this new development may in time lead us.

Nijmegen, February 1993

* Ross Eckler informs me that credit is due to Philip Cohen for first introducing a base 26 system (in which zero was A), in *Word Ways* for February 1977. I regret having been unaware of Cohen's pioneering work which might otherwise have been discussed here.

The above article was published under the title *Base 27: The Key to a New Gematria* in *Word Ways* May 1993, Vol. 26, No. 2, pp. 67-77.