Introduction
to
Linear Categories and Applications

F. William Lawvere

Foreword

The techniques revolving around the concept of linear category are very widely used in nearly all branches of pure and applied mathematics. In this text we exploit the clarification which results from making this concept explicit, by applying it to the study of linear control systems. After a review of basic matrix theory, based on the explicit concept of rig (although using only the bare rudiments of the developed theory of rigs and rings), the linear category of R-matrices (which has sets "of indices" as objects) and the category of R-linear spaces are explicitly introduced as are the extensive and intensive functors relating them

$$\mathrm{Mat}(R) \longrightarrow \mathrm{Lin}(R) \longleftarrow \mathrm{Mat}(R)^{op}$$

Of course, the simplified results in the case where R is a field, are reviewed as a basis for the study of specific linear systems. However, as much as possible is done for general rigs, because the latter often arise in applications: Not only constant, but also variable quantities, not only continuous, but also discrete quantities, not only positives with negatives, but also strictly nonnegative quantities, not only quantities with additive cancellation, but also additively idempotent quantities, all arise daily in physics, economics etc. as "scalars" where matrix techniques guided by linear concepts must be used.

An important example of additively idempotent quantities is the following rig R:

the scalar quantities are all <u>non</u>negative real numbers, including ∞ ; but the addition and multiplication are defined by

$$a + b \underset{\text{def}}{=} \text{ the minimum of } a,b$$

$$a \cdot b \underset{\text{def}}{=} \text{ the usual } \underline{\text{sum}} \text{ of } a,b$$

The fact that the usual sum distributes over min shows that this definition of R indeed satisfies the rig axioms, and hence (by general principles) that Mat(R) and Lin(R) satisfy the axioms for linear categories. Note that the real number 0 is the "1" of this rig and ∞ is its "0". If A and B are sets, then a matrix $A \xrightarrow{f} B$ in Mat(R) might have the following interpretation: the elements of A are indices for certain states or products or locations, likewise B, and we have in mind a specific process for transforming any $a \in A$ to any $b \in B$ at <u>cost</u> f(b/a) in R ; if f(b/a) = "1", it means the cost of getting b from a by our process is 0, whereas if f(b/a) = "0", it means that the cost is <u>infinite</u>, i.e. in practice impossible for our specific process. Then if we wish to consider a two-step process

$$A \xrightarrow{f} B \xrightarrow{g} C$$

the cost of gf (g following f) is given by the usual rule of matrix multiplication

$$(gf)(c/a) = \sum_{b \in B} g(c/b) \cdot f(b/a)$$

because (due to our definition of R) this means

$$(gf)(c/a) = \underset{b \in B}{\mathrm{MIN}} \left[g(c/b) \underset{\text{usual sum}}{+} f(b/a) \right]$$

and we would naturally choose the cheapest intermediate $b \in B$, for given $c \in C$, $a \in A$. For a one-point set 1, a matrix $1 \xrightarrow{\ \shortmid\ } A$ might specify the costs of acquiring the states or things indexed by the elements of A and another matrix $A \xrightarrow{\ \shortmid\ } 1$ might specify the costs of disposing of them; then the matrix product

$$\begin{tikzcd} 1 \arrow[r, "\shortmid" marking] & A \arrow[r, "\shortmid" marking] & 1 \\ \arrow[from=1-1, to=1-3, bend left, "\shortmid" marking] \end{tikzcd}$$

would be the single quantity in R obtained from choosing the index in A for which the usual sum of these two costs is cheapest, and associativity of

$$\begin{tikzcd} 1 \arrow[r, "\shortmid" marking] \arrow[rr, bend right, "\shortmid" marking] & A \arrow[r, "f", "\shortmid" marking] \arrow[rr, bend left, "\shortmid" marking] & B \arrow[r, "\shortmid" marking] & 1 \end{tikzcd}$$

means in particular that for any process f we can choose first the output which is cheapest to dispose of <u>or</u> the input which is cheapest to acquire and arrive at the same resultant cost.

While the general concept of category involves associativity of "multiplication", the additional special feature of <u>linear</u> categories is that the "maps" from any object A to any object B can also be <u>added</u>, in a way that satisfies the distributive laws below (also known as <u>bilinearity</u>)

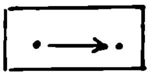
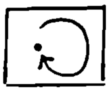
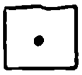
$$\begin{tikzcd} A' \arrow[r, "\alpha"] & A \arrow[r, shift left, "f_1"] \arrow[r, shift right, "f_2"'] & B \arrow[r, "\beta"'] & B' \end{tikzcd}$$

$$(f_1+f_2)\alpha = f_1\alpha + f_2\alpha \text{ as maps } A' \longrightarrow B$$

$$\beta(f_1+f_2) = \beta f_1 + \beta f_2 \text{ as maps } A \longrightarrow B'$$

The objects of interest in the theory of control systems are not R-linear spaces as such, but concrete realizations in Lin(R) of the abstract diagram

$$\mathbb{D} = \boxed{\begin{tikzcd} \bullet \arrow[r, bend left, "a"] & \bullet \arrow[l, bend left, "b"] \end{tikzcd}}$$

These form a new linear category $\mathrm{Lin}(R)^{\mathbb{D}}$; if we need to take into account the more detailed structure of the objective or subjective states of the system, $\mathbb{D}$ might be replaced by a more complicated abstract diagram or "directed graph", with a resulting richer category of concrete realizations with scalars R ; simpler directed graphs such as  and  have as their linear realizations the linear transformations (analyzed in terms of "rank") and the linear operators (analyzed in terms of "eigenvalues") which are the main objects of study of elementary linear algebra. The concrete realizations over R of the graph  are just the R-linear spaces themselves, (which, if R is a field, can be analyzed merely in terms of "dimensions").

The analysis just referred to consists, in $\mathrm{Lin}(R)^{\mathbb{D}}$ as in even more general linear categories of interest, in isolating certain "simple" concrete objects and determining how the arbitrary concrete objects can be expressed in terms of the simple ones via the "direct sum" operation or refinements of the latter; depending on the precise nature of R and $\mathbb{D}$, the distinct simple objects may form a "continuous" family (involving parameters from R such as eigenvalues to specify them) or just a discrete family; in the case $\mathbb{D}=$  , with R a field, there is just one simple object, concretized as R itself,(for every linear space is a direct sum of copies of R).

A different family of concrete objects in $\mathrm{Lin}(R)^{\mathbb{D}}$ is that of those which serve as domains in representing various types of elements in arbitrary objects. In the case of the one-dot graph, there is one main type of element of an object V and the single object R serves to represent these as linear maps $R \longrightarrow V$. In the more general case there is needed a small linear category $R[\mathbb{D}] \subset \mathrm{Lin}(R)^{\mathbb{D}}$ to represent even the main types of elements; its maps consist roughly of R-linear combinations of meaningful "words" of arrows from the graph $\mathbb{D}$. But for example in the case of linear control systems

$$\begin{tikzcd} X \arrow[r, shift left, "a"] & Y \arrow[l, shift left, "b"] \end{tikzcd}$$

there is still another type of "element" which is of primary concern: pairs $\langle x,y \rangle$ in which the subjective process y "controls" the objective process x in the sense that the equation

$$x = b(y - ax)$$

is satisfied; here the controlling intervention b is acting on the discrepancy in Y between the subjective process y and the observation a of the objective process x; that the elements of the R-linear spaces X,Y have the interpretation of processes in time is implicit in the set-up itself, as calculation of the representing object for such controlled elements reveals. The spectral analysis based on the idealized "simple" control systems plays a role in the engineering design of more complex control systems.

Buffalo
Sept 1992

Table of Contents

Foreword i - v

Preliminaries on Rigs 1

Rings 7

Fields, Nilpotents, Idempotents (Logical Operators in Algebra) 10

Examples of Commutative Rings 18

15 Exercises on 2 x 2 Matrices 22

Linear Spaces 27

Linear Combinations and Matrices 33

Linear Categories, Control Systems 41

Table of Some Categories 61

Rank of a Linear Transformation 62

Spectral Analysis of Operators and Systems 64

Determinants Computed by Linear Categories and Nilpotents 75

Test on Trace-Equivalence of Operators 77

Appendices:

The Geometric Series and Differential Equations 81

Multi-Dimensional Calculus 84

Preliminaries on Rigs

0. A central example of "rig" is the field $\mathbb{R}$ of real numbers, but for linear algebra it is helpful to explicitly recognize the possible role of other rigs as these occur in applications. Two roles of rigs in linear algebra are

(1) the "base" ring of "scalars" which is fundamental to the very definition of linear transformation, namely the extent of the <u>homogeneity</u> "$T(\lambda x) = \lambda T(x)$ for all inputs x of T" is the rig of scalars λ for which that equation is true, and

(2) in the spectral analysis of diagrams as simple as the "$\begin{tikzcd}\bullet \arrow[loop above]\end{tikzcd}$" denoting a single linear operator, each "color" is associated with specific rigs which are usually larger than the base rig.

Various rigs arise in applications for many different reasons:

1) The linear space of all smooth vector fields on a region U in 3-space actually admits as scalars the rig $\mathbb{R}^U$ of all smooth scalar fields, i.e. variable quantities as scalars, not just the constant scalars from $\mathbb{R}$.

2) Often it is important to concentrate on linear transformations representable by matrices with <u>whole-number</u> entries, so that the appropriate rig of scalars might be the rig $\mathbb{N}$ of natural numbers or the rig $\mathbb{Z}$ of integers.

3) In statistics, economics, etc. it is often important to concentrate on "positivity", in which case the rig $\mathbb{R}_+$ of non-negative reals (we usually consider that $\infty \in \mathbb{R}_+$) or a rig $\mathbb{R}_+^U$ of nonnegative functions may be the relevant supply of

scalars. The simplest results in linear algebra (such as that dimension is the only invariant) hold only when the base rig is actually a field, but even so the usual applications deal with <u>diagrams</u> of linear maps, not merely with the linear spaces themselves, and even in the simplest cases $\bullet\!\circlearrowleft$ and $\begin{smallmatrix}\bullet\\ \downarrow\\ \bullet\end{smallmatrix}$ the categories of such diagrams require a much richer collection of invariants than mere dimension.

The word "rig" was obtained by omitting the letter "n" from "ring", for a <u>ring</u> is nothing but a rig in which every quantity has a <u>negative</u> in the same rig.

1. A <u>rig</u> R is a specified set of quantities together with two specified quantities 0,1 in it and two specified binary operations $+,\cdot$ on it satisfying the conditions

$$0 + x = x \qquad\qquad 1\cdot x = x$$

$$(a+b)+(x+y) = (a+x)+(b+y) \qquad (a\cdot b)\cdot(x\cdot y) = (a\cdot x)\cdot(b\cdot y)$$

$$a\cdot 0 = 0$$

$$a\cdot(x+y) = a\cdot x+b\cdot y$$

for all a,b,x,y in R . . The four-variable combination of the commutative and associative laws is the form which is most likely to come up in practice. The distributive laws as stated imply more general forms

$$\left(\sum_{i=1}^{n} a_i\right)\cdot\left(\sum_{j=1}^{m} x_j\right) = \sum_{i,j} a_i x_j$$

for any two lists a,x of quantities in R , where we apply the usual sigma notation for repeated + and where the sum on the right has n×m terms. The obvious naming procedure 2 = 1+1, 3 = 2+1, etc. uniquely maps the rig $\mathbb{N}$ into any rig R being considered, and the distributive law also implies the binomial theorem

$$(x+y)^n = \sum_{i+j=n} c_{ij} x^i y^j$$

for any x,y in R , any n in $\mathbb{N}$, where $c_{ij} = \binom{i+j}{i}$ are in $\mathbb{N}$ and the sum has n + 1 terms.

DEFINITION: x has negative y iff $x + y = 0$

x has reciprocal y iff $x \cdot y = 1$

1) EXERCISE: If x_1, x_2 have negatives in R and if a is any quantity in R then x_1+x_2 and ax_1 (and ax_2) also have negatives in R . Hence if 1 has a negative in a rig R then R is actually a ring.

Even in rings many non zero quantities may fail to have reciprocals. For example, the ring of all rational functions (in one variable t), for which 2,5 are not poles, contains $\frac{1}{t}$, $\frac{t^5}{(t-3)^2}$ etc. but not $\frac{t^3}{(t-2}$, $\frac{1}{(t-5)^2}$ etc.

DEFINITION: An idempotent pair in a rig is a pair of quantities p,q for which $p+q = 1$ and $p \cdot q = 0$.

2) EXERCISE: In any idempotent pair, each of the two quantities satisfies $p^2 = p$. In a ring, an idempotent pair is determined by a single quantity p satisfying $p^2 = p$. Relative to any given idempotent pair p,q, the whole rig splits in two in the following sense: given any x there is a unique pair of quantities u,v satisfying the three conditions $x = u+v$ $pu = u$, $qv = v$

DEFINITION: A quantity x in a rig is <u>nilpotent of order $\leq n$</u> iff $x^{n+1} = 0$. Thus $x^2 = 0$ means x is nilpotent of order ≤ 1

$x^3 = 0$ but $x^2 \neq 0$ means x is nilpotent of order 2 etc.

3) EXERCISE: If x is nilpotent of order n and a is any quantity in the rig in question, then ax is nilpotent of order $\leq n$. If some nilpotent has a reciprocal, then $0 = 1$ and indeed there is only one quantity in the whole rig.

4) EXERCISE: If $x^2 = 0$, $y^2 = 0$ then $(x+y)^3 = 0$.
We could only say $(x+y)^2 = 0$ if we moreover know that $xy = 0$, but that is usually <u>not</u> true. More generally (use the binomial theorem)

if x is nilpotent of order $\leq n$ and y of order $\leq m$
then $x + y$ is nilpotent of order $\leq n+m$

Caution: this result is NOT true in general without the condition $xy = yx$, which will itself <u>not</u> be true when x,y denote maps in a linear category rather than merely quantities in a rig.

<u>DEFINITION:</u> A field is a ring in which every nonzero quantity has a reciprocal, and $0 \neq 1$.

5) EXERCISE: The only nilpotent quantity in a field is 0.

2. Given a rig R it is often important to consider new rigs which contain R but which also contain a further quantity x satisfying an equation such as

$$x^2 = a+bx$$

where a,b are some-previously-fixed quantities in R . This can always be done by considering the new rig to consist of <u>ordered pairs</u> of quantities from R with $0 \underset{\text{def}}{=} \langle 0,0\rangle$, $1 \underset{\text{def}}{=} \langle 1,0\rangle$, $x \underset{\text{def}}{=} \langle 0,1\rangle$ and the obvious pair-wise addition, so that every new quantity is uniquely expressed as u + v x for u,v in R , but with a special multiplication rule depending on the given a,b.

6) EXERCISE: Define the special multiplication of pairs u + v ϵ which extends a given rig R to the rig $R[\epsilon]$ of "dual numbers" over R , wherein the equation $\epsilon^2 = 0$ is satisfied.

7) EXERCISE: For a <u>ring</u> R ,define the multiplication for the ring $R[i]$ of "complex numbers" over R wherein the basic new element i satisfies $i^2 = -1$. Show that if $u^2 + v^2 = 1$, then u + v i has a <u>reciprocal</u> in $R[i]$. For what kind of u,v in is u + v i nilpotent of order = 1 in $R[i]$? If R is a <u>field</u>, what stronger property must it have to insure that $R[i]$ is also a field?

Note that in the rig $R[i]$ for R a ring and $i^2 = -1$, we have in particular that $i^4 = 1$; we can crudely picture that fact by pretending that R is a line (even if it isn't) and hence that $R[i]$ is a coordinatized plane, then noting that multiplication by i is rotation through a right angle. However, to obtain an extension rig $R[j]$ containing a quantity j with $j^4 = 1$ but $j^2+1 \neq 0$, we need more than two "dimensions"; in such, every quantity could be expressed as $u_0+u_1j + u_2j^2+u_3j^3$ where the four u's come from R .

8) EXERCISE: For the equation $\theta^3 = 1$, consider $\mathbb{N}[\theta]$ to consist of triples of natural numbers with $\theta = \langle 0,1,0\rangle$ and determine the path followed by $\langle k,n,m\rangle$ upon being multiplied by $\theta, \theta^2, \theta^3$.

9) EXERCISE: In any rig, if $x + y = 1$ and x is nilpotent, then y has a reciprocal.

Rings

A ring is any set R of "quantities" which is furnished with the structure of 0,1, addition + and multiplication $\cdot$ in such a way that 0,1 are distinguished elements of R and

$$R \times R \xrightarrow{+} R$$

$$R \times R \xrightarrow{\cdot} R$$

are given mappings subject to the following axioms: + assigns to any pair a of elements of R a unique <u>sum</u> a + b in R and $\cdot$ assigns to any pair a <u>product</u> ab in R (note that by contrast a "linear space" is closed under a given addition but has <u>no</u> given multiplication under which it is closed; we will often use several different multiplications on the same underlying additive system, which means we would be considering <u>different</u> rings; such a definitive of multiplication must be verified to satisfy the following equational axioms in order to be called a ring) so that

(associative) $a(bc) = (ab)c$ all a,b, c in R

(identity) $1\,a = a = a\,1$ all a in R

(distributive) $(a_1+a_2)(b_1+b_2) = a_1b_1+a\,b_2+a_2b_1+a_2b_2$ all a_i, b_i in R

$a\,0 = 0 = 0\,a$ for all a in R

and so that the addition by itself satisfies

$$0 + a = a = a + 0 \qquad \text{all a in R}$$

$$(a_1+a_2)+(b_1+b_2) = (a_1+b_1)+(a_2+b_2) \qquad \text{all } a_i, b_i \text{ in } R$$

EXERCISE The two axioms above for addition alone imply the usual commutative and associative laws for addition.

We will almost always assume that our ground rings are moreover <u>commutative</u> which means that the multiplication satisfies

$$\boxed{ab = ba} \qquad \text{all a,b in R}$$

and also usually that they have <u>negatives</u>, i.e. that there is a mapping $R \xrightarrow{-(\,)} R$ so that

$$a+(-a) = 0 \qquad \text{all a in R}$$

When R has negatives, we can define subtraction by

$$a - b \underset{\text{def}}{=} a + (-b) \qquad \text{all } a,b \text{ in } R$$

However, very important in probability and many other fields is the linear algebra over the system $\mathbb{R}^r$ of <u>non-negative</u> real numbers, so all our results which do not depend on negatives will apply to these fields directly.

$\boxed{\text{EXERCISE}}$ Many of the important formulas of algebra are valid in any ring sin they depend on the <u>distributive law</u> which is one of the defining axioms of rings. For example

$$\left(\sum_{i=1}^{n} a_i\right)\left(\sum_{j=1}^{m} b_j\right) = \sum_{i,j} a_i b_j$$

$$1-a^3 = (1-x)(1+x+x^2)$$

$$(a-b)(a+b) = a^2-b^2$$

$$(a+b)^2 = a^2+2ab+b^2$$

.

If we do <u>not</u> have commutativity at our disposal the last two equations are false and the correct version is a little more complicated, for example

$$(a+b)^2 = a^2+ab+ba+b^2$$

Of course we use the usual abbreviations

$$a^2=aa$$

$$2=1+1$$

$$2x=x+x=(1+1)x$$

etc

More information on rings is contained in the sections of these notes called "Examples of commutative rings" and"Use of logical operations in algebra".

A mapping $R \xrightarrow{T} S$ between rings is called a <u>ring homomorphism</u> if it satisfies all of

$$T(0)=0$$

$$T(x_1+x_2)=T(x_1)+T(x_2)$$

$$T(1)=1$$

$$T(x_1x_2)=T(x_1)T(x_2)$$

or all r_1, r_2 in R, where the ring operations on R are used on the left, and those from S on the right, these operations being denoted by the same symbol in all rings even though they have differing meaning depending on which ring involved.

EXERCISE If R,S have negatives and T is a ring homomorphism then it follows that
$$T(-r) = -T(r)$$
In a similar way if r happens to be an invertible element of R then T(r) is an invertible element of S and in fact their reciprocals correspond:
$$T\left(\frac{1}{r}\right) = \frac{1}{T(r)}$$

Only fragments of the category of (commutative) rings and ring homomorphisms are used in introductory linear algebra. The detailed study of this category is called Commutative Algebra or Algebraic Geometry.

Linear Algebra is an important tool in Commutative Algebra and hence in the study of algebraic spaces as well as differential geometry, etc., and on the other hand Commutative Algebra and Algebraic Geometry yield important tools for more advanced Linear Algebra and its applications to Circuit Theory, Systems Theory, Linear differential Equations, etc.

Fields, Nilpotents, Idempotents

The most basic properties of algebraic structures such as rings, linear spaces, categories are expressed by equations, for example the distributive property, nilpotence, associativity, the property of being a solution. However in working with these equations we must frequently use stronger logical operators both in stating stronger axioms on the ground ring in linear algebra as well a in summarizing the meaning of our complicated calculations. (It should be remarked however that most of this logic again becomes equational when we pass to a higher realm). For example the additional axioms which state that a given ring R is a $\boxed{\text{field}}$ is that R is nondegenerate

$$0 = 1 \vdash \text{false},$$

usually expressed by introducing "not" and saying

$$\text{true} \vdash 0 \neq 1)$$

and that every non-zero quantity in R has a reciprocal

$$x \neq 0 \vdash \exists y[xy=1]$$

When the law of excluded middle is valid the latter is equivalent to the (in general stronger) condition

$$x=0 \vee \exists y[xy=1]$$

being true_R (which has the virtue of being invariant under more geometrical transformations but the drawback, in these cases like continuous functions whe the law of excluded middle is false, of being less likely to be true). Usually one expresses this field axiom using $\forall, \Rightarrow$ as:(true $\vdash$)

$$\forall x[x \neq 0 \Rightarrow \exists y[xy=1]]$$

with the understanding that the universe over which both x, y vary is R. Thus $\mathbb{Z} = \{\ldots -3,-2,-1,0,1,2,3\ldots\}$ is a ring R which is not a field since for example $5 \neq 0$ but there is no y in $\mathbb{Z}$ for which 5y = 1. Since in any ring we can deduce purely equationally from hypotheses

$$xy_1 = 1$$
$$xy_2 = 1$$

that $y_1 = y_2$ [Here is the deduction, using only (commutative) ring axioms and the hypothesis:

$$y_1 = y_1 1 = y_1(xy_2) = (y_1x)y_2 = (xy_1)y_2 = 1\cdot y_2 = y_2 \;]$$

we can conclude that in any ring

$$\vdash \mathcal{U}n_y\,(xy=1) \quad \text{for all } x$$

and hence in any field that

$$\forall x[x \neq 0 \Rightarrow \exists! y[xy=1]]$$

Further (since R is nondegenerate if it is a field) the (just justified) reciprocal of x cannot be zero either.

EXERCISE If y is a reciprocal of x then x is a reciprocal of y; if, in any given ring R, 0 has a reciprocal then R is degenerate . Thus if we restri the universe to the set G of all non-zero elements of R (G is no longer a rin the slightly simpler statement

$$\forall x \exists! y[xy=1]$$

is true over G. Since this is the criterion for the existence of a mapping, there is a mapping

$$G \xrightarrow{(\,)^{-1}} G$$

called the reciprocal mapping whose graph is the statement

$$x\cdot y = 1,$$

that is $y = x^{-1}$ iff $x\cdot y = 1$.

EXERCISE A much better way of understanding the last construction is as follows. Let R be any commutative ring (not necessarily a field, maybe even degenerate). Define G to be the subset of R consisting of all elements x of R satisfying

$$\exists y[xy=1]$$

in R. Then there is a reciprocal mapping $G \longrightarrow G$, 1 is in G, and G is clos under multiplication, i.e. if x_1, x_2 are in G then x_1x_2 is in G, since $1^{-1}=1$ $(x_1x_2)^{-1}=x_2^{-1}x_1^{-1}$ This means that G is a (commutative) group called "the multiplicative group of R". If 0 is in G then R is degenerate. For any x in G, -x is also in G. But x_1,x_2 in G do not imply $x_1 + x_2$ in G. If $R = \mathbb{R}$ the real numbers, then $1 + x^2$ is always in G for any x, and the same is true if R = C(X) = the ring of all continuous real-valued functions on any continuous domain ("topological space") X. Now the condition that a ring R be a field is just that R be the disjoint union of $\{0\}$ and G, i.e. that (readir the $\vee$ form of the definition backwards)

$$\exists! x[\neg G(x)]$$

Since any ring R has a special element 1 and since R has an addition operation, there are elements in R which may as well be denoted

$$2 = 1 + 1$$

$$3 = 1 + 1 + 1$$

$$\vdots$$

(not all of these need be distinct). Even if R is a field, not all of these need have reciprocals, for example there is an important field with only thre elements in all in which 3 = 0. However most of the rings dealt with in detai in this course, even those which are not fields, will involve $\mathbb{R}$ in such a wa that all of the above do have reciprocals, which will be denoted as usual by $\frac{1}{2}$, 1/3,.... Thus ½ in G, 1/3 in G..... where G denotes the multiplicative group of any such ring R.

EXERCISE In any ring having $\frac{1}{2}$

$$\vdash \forall x \exists y [y+y=x]$$

etc.

By a subring of a given ring R is meant a subset S of the elements of which contains 0, 1 and is closed under the addition, subtraction, and the multiplication of R. Thus if p is any polynomial with coefficients in $\mathbb{Z}$ in several, say three, variables, and if x,y,z are in S then p(x,y,z) is also in

EXERCISE If R is a ring having ½ and if S is a subset containing 0,1, and closed under addition and the unary operation of multiplication by ½, then S is a subring if and only if S is closed under the unary operation of squaring (The answer is a frequently-used formula).

Now a subring is not necessarily closed under division, even to the extent to which the latter is defined in R. Thus $\mathbb{Z} \subset \mathbb{R}$ is a subring, but $\mathbb{Z}$ is not a field even though $\mathbb{R}$ is a field. But any subring of any field does have a special property not shared by all rings of interest, namely

$$xy=0 \Rightarrow [x=0 \vee y=0]$$

EXERCISE Prove the statement just made, in any subring of a field .

A nondegenerate ring S having this property for all x,y in S is called an **integral domain**. This is intimately related to the cancellation property

$$\forall x_1 x_2 [a x_1 = a x_2 \Rightarrow x_1 = x_2]$$

for an element a, which (using subtraction) is easily proved equivalent to the "non zero division" property of a

$$\forall x [ax=0 \Rightarrow x=0]$$

where all universal quantifiers range over all x, x_1, x_2 in the ring in which we are considering a (We might call a "monomorphic" in that ring). Note that the property uses the logical operators in an essential way, since when we want to prove

a is monomorphic $\vdash$ something else about a

we can't always eliminate the $\forall, \Rightarrow$ implicit on the left hand side. Of course if the "something else" is just an instance of the cancellation property, such proof will present no problem. Now clearly a = 0 can <u>not</u> be monomorphic in a nondegenerate ring, since

$$a\,x = 0 \cdot x = 0$$

for any x, yet we could take x = 1, hence it wouldn't follow that x = 0. Nov the idea of an integral domain is that (assuming the law of excluded middle) the <u>only</u> a which is not monomorphic in R is a = 0. That is, the validity <u>for all a in R</u> of any one of

$$a \neq 0 \Rightarrow \forall x[ax = 0 \Rightarrow x = 0]$$
$$\exists x[ax = 0 \wedge x \neq 0] \Rightarrow a = 0$$
$$\forall x[ax = 0 \Rightarrow [a = 0 \vee x = 0]]$$

is equivalent (using LEM) to the condition that R is an integral domain. The last form with $\vee$ is the one familiar from high school as a crucial step in the method of solving polynomial equations by factoring. This method is used proving

THEOREM In any integral domain the equation

$$x^2 = x$$

has precisely two solutions.

Proof: if $x^2 = x$ then $x^2 - x = 0$ and hence $x(x-1) = 0$, (since $x(x-1) = x^2-x$ in any ring). Now use the integral domain property to get $x = 0 \vee x-1 = 0$, i.e $x = 0 \vee x = 1$. We say "precisely two" because the ring is nondegenerate.

EXERCISE In any commutative ring, an element satisfying $x^2 = x$ is called **idempotent**. If x is an idempotent, then so is its "complement" 1 - x. The product of any two idempotents is an idempotent. If x and y are idempotents and if xy = 0, (o says x,y are "disjoint" or "orthogonal") then x + y is also an idempotent. One says a ring "has connected spectrum" if it has precisely two idempotents. In general the idempotents describe chunks of the "spectrum", for ex. of a linear transformation(which gives rise to a ring in a way we will study).

Very important in analyzing linear transformations will be the **nilpotent** elements in commutative rings, where x is nilpotent iff

$$\exists n\left[x^{n+1}=0\right]$$

Here the $\exists n$ does <u>not</u> range over the ring we are talking about but rather ove the set 0, 1,2,3,..... of natural numbers, which act as <u>exponents</u> on elements of any ring (or indeed of any system wherein at least multiplicatio is defined). In more detail we could say that x is nilpotent of order 1 if

$$x^2 = 0$$

while $x \neq 0$, that x is nilpotent of order 2 if

$$x^3 = 0$$

while $x^2 \neq 0$, etc. Of course 0 is nilpotent of order zero. In a nondegenerate ring x = 1 is <u>not</u> nilpotent of any order. Using commutativity, the product of a nilpotent with <u>any</u> element is again nilpotent. Again using commutativity, the sum of any two nilpotent elements is nilpotent, however we have to care for the order. For example if $x^2 = 0$ and $y^2 = 0$, then we can calculate that $\boxed{(x+y)^3 = 0.}$ As for $(x + y)^2$, it might be 0, but only in case xy = 0, which is not always true. Analysis of the calculation leads to the idea that to be sure we have to <u>add</u> the orders of nilpotency:

THEOREM If $x^{n+1} = 0$, $y^{m+1}=0$ in a commutative ring, then always

$$(x + y)^{n+m+1} = 0$$

PROOF: In any commutative ring the distributive law implies the binomial expansion

$$(x+y)^p = \sum_{i+j=p} C_{ij}\, x^i y^j$$

for any x,y in the ring and in p in $\mathbb{N}$ (note that $\mathbb{N} \subset \mathbb{Z}$ and that $\mathbb{Z}$ can be used as coefficients in any ring, indeed in any system where addition and subtraction are defined.

In fact

$$C_{ij} = \frac{(i+j)!}{i!\,j!}$$

is in $\mathbb{N}$, despite the denominators, where ! denotes "factorial", by Pascal) Thus the proof of the theorem reduces to the following fact about the elementa arithmetic of $\mathbb{N}$:

LEMMA $i+j = n+m+1 \vdash i \geq n+1 \vee j \geq m+1$

Proof of Lemma as **EXERCISE**

In any case, since

$$x^{n+1} = x^n \cdot x$$

is a product, it is immediately clear that

THEOREM In an integral domain the only nilpotent element is 0

An extremely important property (for analysis, linear algebra, compute science etc) is the following, showing that while the existence of nilpotent elements has the "negative" consequences that some elements (the ones "near" zero) are definitely not invertible, it also has the "positive" consequence t some other elements (those "near" 1) definitely <u>are</u> invertible, and there is even a specific formula for the reciprocals.

THEOREM If h is any nilpotent element in a ring, then 1-h has a reciprocal the same ring. In fact if $h^{n+1} = 0$, then

$$\frac{1}{1-h} = \sum_{k=0}^{n} h^k$$

PROOF Calculate that the right-hand side, multiplied by 1-h, gives 1.

<u>REMARK</u> In a ring furnished with a notion of convergence, the above can be generalized to h for which $h^n \to 0$ as $n \to \infty$, i.e to small h's not necessarily so small as "nilpotent". But the formula of the Theorem is surprisingly often useful even just for the nilpotent case.

EXERCISE If u has a reciprocal and h is nilpotent then $u \pm h$ has a reciproca (A formula, only slightly more complicated than that of the theorem, can either deduced from the theorem as a corollary or calculated and proved directly).

EXERCISE If $u_1 = 1-h_1$, $u_2 = 1-h_2$ are invertible elements of a ring of the form indicated with h_1, h_2 nilpotent (with orders of nilpotency n_1, n_2 say) the

the product u_1u_2 is of course invertible; is it again of the special form, "infinitesimally near 1" in the sense that

$$u_1u_2 = 1-h$$

for some nilpotent h of some order? Start with the special case $h_1{}^2=0=h_2{}^2$, $h_1h_2 = 0$. What if $h_i = t_i\varepsilon$ where $\varepsilon^2=0$?

REMARK (An Embedding) Any given integral domain R can be realized as subrin of a field F, by constructing F to consist of equivalence classes of fractio $\frac{x}{s}$ where x in R, s in R and $s \neq 0$.

The condition that a ring R has "no" (i.e. no nonzero) nilpotent elements is often referred to in geometry and analysis by saying the R is **"reduced"**. IT IS MORE GENERAL THAN THE CANCELLATION (i.e. integral domain) property, since for example R = $\mathbb{R}^2$ with co-ordinatewise multiplication is reduced (i.e. has no nilpotent element) but not an integral domain since it has non trivial idempotent elements $\langle 0,1\rangle$, $\langle 1,0\rangle$. In logical notation with variables, R is reduced if and only if

$$\exists n[x^{n+1}=0] \vdash x=0$$

holds for all x in R. Since the $\exists$ occurs on the left, this is one of its eliminable cases. But more profoundly (i.e. using something of the quantitat content of the theory of rings and not merely logical form):

EXERCISE If a ring satisfies

$$\forall x \boxed{x^2=0 \Longrightarrow x=0}$$

then it is reduced. Hint: Show that if $x^{n+1} = 0$ then $x^{2n}=0$ and hence using on main assumption then also $x^n = 0$. By induction then the n can be knocked dow

Commutative Rings

Some important examples of commutative Rings which arise in real linear algebra.

The ring $\mathbb{R}$ of real numbers is in fact a field, as is the ring $\mathbb{R}[i]$ of complex numbers. The latter has the property

$$\forall a \exists x [x^2 = a]$$

which $\mathbb{R}$ does not, and in fact $\mathbb{R}[i]$ has solutions to any polynomial equations:

$$\forall a_0 \forall a_1 \ldots \forall a_{n-1} \exists x [x^n = a_0 + a_1 x + a_2 x^2 + \ldots + a_{n-1} x^{n-1}]$$

holds in $\mathbb{R}[i]$, for any n in $\mathbb{N}$. On the other hand $\mathbb{R}$ permits solutions to certain other equations which $\mathbb{R}[i]$ does not, for example

$$\forall x_1 \forall x_2 \exists y \; [\, y^2(1 + x_1^2 + x_2^2) = 1]$$

holds in $\mathbb{R}$ but not in $\mathbb{R}[i]$.

The idea of adjoining to $\mathbb{R}$ an element satisfying a certain equation, for example adjoining i satisfying $i^2=-1$ to get $\mathbb{R}[i]$, can be used for other equations as well. For example

$$\mathbb{R}[\varepsilon]$$

is obtained by adjoining ε with $\varepsilon^2=0$. Since the equation is quadratic, $\mathbb{R}[\varepsilon]$ shares with $\mathbb{R}[i]$ being "two-dimensional over $\mathbb{R}$" (in the linear sense which we will be studying in detail) but $\mathbb{R}[\varepsilon]$ is <u>not</u> a field nor even an integral domain since it contains a non-zero nilpotent element. In more detail every element of $\mathbb{R}[\varepsilon]$ can be uniquely expressed in the form

$$a + a'\varepsilon$$

where a,a' are <u>real</u>. For example

$$1 = 1 + 0 \cdot \varepsilon$$

is the multiplicative identity of $\mathbb{R}[\varepsilon]$ while

$$0 = 0 + 0 \cdot \varepsilon$$

is the additive identity and

$$\varepsilon = 0 + 1 \cdot \varepsilon$$

The addition in $\mathbb{R}[\varepsilon]$ is performed in the obvious manner, and if we work out the multiplication using the distributive law and the speci rule $\varepsilon^2=0$ we find

$$\boxed{(a+a'\varepsilon)(b+b'\varepsilon) = ab + (ab'+a'b)\varepsilon}.$$

where indeed ab , ab'+a'b is a new pair of real numbers obtained fro the two pair s a,a' b,b' which we wanted to multiply. In particula if a = b = 0 and a'=b'=1 we recover the fact that ε^2=0, yet $\varepsilon \neq 0$ since 0 is the element both of whose components are 0.

The significance of the above multiplication law can be better unde stood if we consider first another, "infinite-dimensional", example of a ring

$\mathbb{R}[t]$ = ring of all polynomials $a_0+a_1t+a_2t^2+\ldots+a_nt^n$
of all possible degrees n with all possible
real coefficients a_i.

As the notation suggests the polynomial ring $\mathbb{R}[t]$ can also be viewed as the result of adjoining an element t to $\mathbb{R}$, but this t has no special properties[which, paradoxically, is itself a very special property.] The polynomial ring is an integral domain, but definitely not a field since indeed the only (non-zero) polynomials whose reciprocals exist as polynomials are the constant ones a_0. However, the ring of all rational functions $\mathbb{R}(t)$, which is too big to be of much use, is a field. An important ring between the two is

$$\mathbb{R}[t,\tfrac{1}{t}]$$

sometimes called the ring of "Laurent polynomials", consists of all rational functions which can be expressed in the form

20 $$\frac{a_{-m}}{t^m} + \frac{a_{-m+1}}{t^{m-1}} + \dots + \frac{a_{-1}}{t} + a_0 + a_1 t + a_2 t^2 + \dots a_n t^n$$

for some $n, m \in \mathbb{N}$ with a_i all real. The significance of this ring for linear algebra is that t is invertible (i.e. has a reciprocal) but has no special properties beyond that. Now suppose $\mathcal{R}$ is either $\mathbb{R}[t]$ or $\mathbb{R}[t, \frac{1}{t}]$ and that α is any given real number or any given non-zero real number in the second case. Define a mapping $\mathcal{R} \xrightarrow{p} \mathbb{R}$

by sending any f in R to its value at α.

$$p(f) = f(\alpha) \quad \text{for all } f \text{ in } \mathcal{R}$$

Then it is clear that p is not only a linear transformation, but even a ring homomorphism i.e.

$$p(fg) = p(f)p(g) \quad \text{for all } f, g \text{ in } \mathcal{R}$$

$$\cdot \quad p(1) = 1$$

because of trivial properties of evaluation at α. But now define a mapping $$\mathcal{R} \xrightarrow{v} \mathbb{R}[\varepsilon]$$

as follows

$$v(f) = f(\alpha) + (Df)(\alpha) \cdot \varepsilon \quad \text{for all } f \text{ in } \mathcal{R}$$

where Df denotes the derivative of f (Recall that $\mathcal{R} = \mathbb{R}[t]$ or $\mathcal{R} = \mathbb{R}[t, \frac{1}{t}]$ is a ring of functions in which differentiation makes sense.)

THEOREM v is a ring homomorphism

1) Proof = Exercise showing that Leibniz rule for differentiation is equivalent to the rule for multiplication in the ring $\mathbb{R}[\varepsilon]$ obtained by adjoining a nilpotent quantity to $\mathbb{R}$.

Remark: In differential geometry one considers the ring $\mathbb{R}^X$ of all smooth real functions defined on a domain X in higher dimensional space. Then a homomorphism $\mathbb{R}^X \to \mathbb{R}$ is determined by a point p of X and a homomorphism $\mathbb{R}^X \to \mathbb{R}[\varepsilon]$ is determined by differentiating in the direction of a certain tangent vector to X

To finish this preliminary survey of examples of commutative rings which are basic to linear algebra, consider, for a given n, the set $\mathbb{R}^n$ of all n-tuples of real numbers, and define this time the multiplication too (as usually the addition) in coordinatewise fashion:

$$\langle x_1, \dots x_n\rangle\langle y_1 - y_n\rangle = \langle x_1 y_1, x_2 y_2, -, x_n y_n\rangle$$

Then in particular $1=\langle 1,1,\dots 1\rangle$ is the identity for $\mathbb{R}^n$. This ring is closely connected with the important problem of diagonalizing linear transformations and matrices.

2) **EXERCISE** With multiplication defined this way $\mathbb{R}^n$ is not a field; in fact it has precisely 2^n idempotent elements.

2x2 Real Matrices

They form a four-dimensional linear space over $\mathbb{R}$. The multiplication rule

$$\begin{pmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{pmatrix}\begin{pmatrix} b_{11} & b_{21} \\ b_{12} & b_{22} \end{pmatrix} = \begin{pmatrix} a_{11}b_{11}+a_{21}b_{12} & a_{11}b_{21}+a_{21}b_{2} \\ a_{12}b_{11}+a_{22}b_{12} & a_{12}b_{21}+a_{22}b_{22} \end{pmatrix}$$

This rule is associative and is distributive with respect to addition, while having

$$1 = 1_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

as identity.

EXERCISE The subsystem of all matrices of the form

$$\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$$

is closed under multiplication and addition, and contains 1; any two matrices in this subsystem <u>commute</u>, i.e. AB = BA for the matrix product of any two A,B of this special form. Hence these special matrices form a commutative ring. This ring is in fact isomorphic to $\mathbb{R}^2$ with the latter's co-ordinatewise multiplication. For example

$$\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

are the two non-trivial idempotent elements.

EXERCISE

$$\varepsilon \underset{\text{def}}{=} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

is <u>nilpotent</u> under matrix multiplication. If A is <u>any</u> nilpotent 2 x 2 matrix the order of nilpotency is 1 (or 0). The set of all 2 x 2 real matrices of the form

$$\begin{pmatrix} \lambda & \delta \\ 0 & \lambda \end{pmatrix}$$

(the two diagonal entries being required to be equal in each matrix, but arbitrary for various matrices in the set) is a 2-dimensional linear space over $\mathbb{R}$ is closed under matrix <u>multiplication</u>, i.e.

where the two • entries are the same. Moreover the multiplication among thes special matrices in our set is again commutative, i.e.

$$\begin{pmatrix}\mu & \eta\\ 0 & \mu\end{pmatrix}\begin{pmatrix}\lambda & \delta\\ 0 & \lambda\end{pmatrix} \quad = \text{ same answer as above, not merely same kind}$$

Thus this is a commutative subring of the 2 x 2 matrices which is actually isomorphic to $\mathbb{R}[\varepsilon]$, since

$$\begin{pmatrix}\lambda & \delta\\ 0 & \lambda\end{pmatrix} = \begin{pmatrix}\lambda & 0\\ 0 & \lambda\end{pmatrix} + \begin{pmatrix}0 & \delta\\ 0 & 0\end{pmatrix}$$

$$= \lambda\cdot 1 + \delta\cdot\varepsilon \quad \text{for any } \lambda, \delta \text{ in } \mathbb{R}$$

Exercise .

$$\varepsilon^* \overset{}{\underset{df}{=}} \begin{pmatrix}0 & 0\\ 1 & 0\end{pmatrix}$$

is also nilpotent under matrix multiplication. However (calculate)

$$\varepsilon\varepsilon^* \neq \varepsilon^*\varepsilon$$

and (in contrast to the commutative case where the sum of nilpotents is again nilpotent (albeit of higher order))

$$(\varepsilon + \varepsilon^*)^3 \neq 0$$

nor will any matrix power kill $\varepsilon + \varepsilon^*$

EXERCISE One can find two matrices of the form $\begin{pmatrix}\lambda_1 & \delta\\ 0 & \lambda_2\end{pmatrix}$, diagonal entries not the same, which do not commute under matrix multiplication; but the product is at least again of the same form; these are called "triangular" matrices.

EXERCISE The matrices of the form

$$\begin{pmatrix}a & b\\ -b & a\end{pmatrix}$$

are closed under multiplication and commute among themselves, and two of ther satisfy the matrix equation $\boxed{x^2 + 1 = 0}$

More Exercises on 2 x 2 matrices

6) Assuming $\boxed{\varepsilon^2 = 0, \ \varepsilon^{*2} = 0, \ (\varepsilon+\varepsilon^*)^2 = 1}$, show that

$$(\varepsilon\varepsilon^*)^2 = \varepsilon\varepsilon^* \qquad (\varepsilon^*\varepsilon)^2 = \varepsilon^*\varepsilon$$

$$\varepsilon^*\varepsilon\varepsilon^* = \varepsilon \qquad (\varepsilon\varepsilon^*)(\varepsilon^*\varepsilon) = (\varepsilon^*\varepsilon)(\varepsilon\varepsilon^*)$$

$$\varepsilon\varepsilon^* + \varepsilon^*\varepsilon = 1 \qquad (\varepsilon-\varepsilon^*)^2 = -1$$

and check by matrix maultiplication.

7) Find a 2x2 matrix T which has a reciprocal T^{-1} for which

$$T\varepsilon = \varepsilon^* T$$

Express T^{-1} in terms of $\varepsilon, \varepsilon^*$

8) The <u>trace</u> of a square matrix is the sum of the entries on the main diagonal

$$\operatorname{tr}\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11} + a_{22}$$

Then $\boxed{\operatorname{tr}(ab) = \operatorname{tr}(ba)}$, $\operatorname{tr}(c^{-1}ac) = \operatorname{tr}(a)$ if c is invertible,

$\operatorname{tr}(a \pm b) = \operatorname{tr}(a) \pm \operatorname{tr}(b)$, $\operatorname{tr}(ab - ba) = 0$

9) What is the dimension of the space of all 2x2 matrices whose trace = 0 ?

MATRICES satisfying SIMPLE EQUATIONS

We will show later that any 2x2 matrix t satisfies a quadratic equation ("Cayley-Hamilton") $\boxed{t^2 - tr(t)t + det(t) = 0}$ where tr(t) and det(t) are scalars. In fact tr was defined in <u>exercise 8</u> and the determinant of a 2x2 matrix is defined by $det\begin{pmatrix} t_{11} & t_{12} \\ t_{21} & t_{22} \end{pmatrix} = t_{11}t_{22} - t_{12}t_{21}$

EXERCISE 10 Verify the above Cayley-Hamilton equation for $t =$ $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$.

EXERCISE 11 Assuming the Cayley-Hamilton equation (and that the scalars are a <u>field</u> (such as $\mathbb{R}$)) show that if det(t) ≠ 0, then t is <u>invertible</u> , in fac for scalars λ, d $t^{-1} = \frac{1}{d}(\lambda t - t^2)$

EXERCISE 12 Assuming the Cayley-Hamilton equation, show that if t is a 2x2 matrix having det(t)=0, then either $t^2=0$ <u>or</u> $t^2 = \lambda t$ for λ a non-zero scal

EXERCISE 13 $t = \lambda e,\ e^2 = e \implies t^2 = \lambda t$ (λscalar). <u>Any</u> 2x2 matrix (over $\mathbb{R}$) is either invertible <u>or</u> nilpotent <u>or</u> a non zero scalar multiple of an idempotent.

EXERCISE 14 There is a 3x3 (even upper triangular)* matrix which is neither invertible <u>nor</u> nilpotent <u>nor</u> a scalar multiple of an idempotent.

For example $t = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$. t^2 is an <u>idempotent</u> e, and there is a <u>nilpotent</u> ε such that $t = e + \varepsilon,\ e\varepsilon = \varepsilon e = 0$. Hence $\boxed{t^3 = t^2}$ but t satisfies no simpler equation. Show that this t is neither invertible nor nilpotent nor a scalar multiple of a idempotent.

---

* Recall that a square matrix is called upper triangular if all its nonzero entries are on or above the main diagonal.

15) Let V be a linear space (not necessarily $\mathbb{R}^2$) with two given linear operators $\begin{tikzcd} V \arrow[loop right, "\epsilon"] \arrow[loop below, "\epsilon^*"] \end{tikzcd}$ satisfying the equations of exercise 6 . Can V be one-dimensional? If T is a given operator with $T^2 = 1$ and we define $A^* = T\,A\,T$ for any operator A, show that

$$(AB)^* = A^* B^*$$

$$A^{**} = A$$

If in particular $T = \epsilon + \epsilon^*$, $i = \epsilon - \epsilon^*$

show that $i^* = -i$, $(\epsilon)^* = \epsilon^*$

How must A be related to T, i, ϵ in order that

$$A\,A^* = A^* A \,?$$

How must A,B be related to T, i, ϵ in order that

$$|A|^2 |B|^2 = |AB|^2$$

where we <u>define</u> $|A|^2 = AA^*$?

In the case $V = \mathbb{R}^2$, $\epsilon = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $T = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ show that the set of all A for which $A = A^*$ is a <u>commutative</u> ring and that

$$A = \begin{pmatrix} b & a \\ -a & b \end{pmatrix} \Rightarrow A\,A^* = (a^2 + b^2)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$A = \begin{pmatrix} 0 & a \\ -a & 0 \end{pmatrix} \Rightarrow A^* = -A$$

Linear Spaces

For any given ring R, there is a "linear category" (to be explained presently) of R-linear spaces and R-linear transformations. It is traditiona since Emmy Nöther, to call R-linear spaces "R-modules" when one wants to emphasize that R is not necessarily a field, and to call R-linear spaces "R-vector spaces" when one wants to insist that R is a field. In either case one often says that the quantities in R are being used as *scalars* for "scalar-multiplying" the quantities in the R-linear spaces in the category of R-linear spaces, and also that R is the *ground ring* for this category, which means the same thing. There is nothing wrong with the term "module" since it is appropriately abstract, though it does tend to obscure the fact that all that's involved in it is the widespread and important mathematical phenomenon of linearity ; in weaker or stronger senses depending on R . The term "vector space", though firmly rooted, is from a geometrical or physical point of view misleading in two ways:

① The vectors drawn as arrows denote quantities in a linear space which are acting by translation on the points of physical space (which is not the same space: there is a zero vector but *no* distinguished zero-point in physical space, there is vector addition of translations, but no distinguished addition of points); the linear spaces of linear algebra are slightly more abstract since they are not furnished with an action on points; we can and do consider the richer structure where this action is taken into account, but we no longer obtain a "linear category" but another kind of category.

② Forces, pressures, momenta etc. varying over a region X do form linear spaces of "vectors" in the physical sense which can be multiplied by $\mathbb{R}$ indeed, but can in fact usually also be scalar-multiplied by smooth functions from the ring $R = \mathbb{R}^X$ of smooth functions on X, and R is not a field in the algebraic sense.

The term "linear space" should however occasion no confusion.

DEFINITION OF LINEAR SPACE

The first part of the structure characteristic of a linear space V does not involve the ground ring R of scalars, it is an operation of addition

$$V \times V \xrightarrow{+} V$$

which assigns to each pair of elements of V another element of V .Thus to specify a particular linear space we must first tell what the elements of V are and how to add them. But there are some restrictions (or axioms on how this can be done: we must be sure that there is a zero-element O in V which satisfies

$$v + 0 = v = 0 + v \qquad \text{for all } v \text{ in } V$$

relative to the specified addition, and further the addition must satisfy the commutative/associative law

$$(v_1 + v_2) + (w_1 + w_2) = (v_1 + w_1) + (v_2 + w_2)$$

for all quadruples v_1, v_2, w_1, w_2 in V

Then the scalar multiplication

$$R \times V \longrightarrow V$$

must be made clear; it must assign to each scalar a in R and each v in V a "product" or "multiple" av in V, and this must satisfy several axioms:

$$a(v_1 + v_2) = av_1 + av_2$$
$$a \cdot 0 = 0$$

where both additions are "vector" addition and both O's likewise, as well as

$$1 \cdot v = v$$
$$(ab)v = a(bv)$$

where the multiplication "internal to R" is required to be compatible with the scalar multiplication between R and V, and finally

$$(a+b)v = av+bv$$
$$0v=0$$

where we have "internal" addition in R on the left and "vector" addition in V on the right (and of course scalar multiplications on both sides).

EXERCISE If (as usually assumed for rings) R has an operation of minus R → R satisfying $a+(-a)=0$ for all a in R, then using the above axiom on scalar multiplication we can construct an operation $V \to V$ on any R-linear space (call it again minus) for which we can prove for all v : $v+(-v)=0$

EXAMPLE R itself is a basic example of an R-linear space, for any ring R. If X is any set, then the set

R^X = set of all R-functions on X

= set of all mappings with domain X and codomain R

is an R-linear space where we define

f + g

to be the function whose values are given by

(f+g) (x) = f(x)+g(x) for all x in X

and similarly for a in R, f in R^X

af

is the function given by

(af)(x) = a(f(x)) all x in R

since f(x) in R (for each x) it can be R-multipled by a since R is a ring.

In case X is a finite set, which is moreover ordered, it is usual to specify functions f on X just by listing their values and in view of conventions we will write these lists vertically and call them column vecto

Thus if $X=\{1,2,3\}$ is a three element set and $f: X \to R$ has

$$f(1)=6 \qquad f(2)=\frac{1}{3} \qquad f(3)=10$$

then we can specify f as

$$f = \begin{pmatrix} 6 \\ \frac{1}{3} \\ 10 \end{pmatrix}$$

Then the addition of functions defined above tells to add two column vectors by adding at each level e.g.

$$\begin{pmatrix} 6 \\ \frac{1}{3} \\ 10 \end{pmatrix} + \begin{pmatrix} 2 \\ 0 \\ -4 \end{pmatrix} = \begin{pmatrix} 8 \\ \frac{1}{3} \\ 6 \end{pmatrix}$$

while scalar multiplication a f says to multiply <u>each</u> entry in a column vactor by the same a

$$a\begin{pmatrix} 6 \\ \frac{1}{3} \\ 10 \end{pmatrix} = \begin{pmatrix} 6a \\ \frac{a}{3} \\ 10a \end{pmatrix} \qquad \text{for any a in R}$$

LINEAR TRANSFORMATION

If V,W are two R-linear spaces, then a linear transformation from V to W is any mapping $V \xrightarrow{T} W$

which satisfies these conditions (of "linearity")

$$T(v_1+v_2) = T(v_1) + T(v_2)$$
$$T(av) = aT(v)$$

for any v, v_1, v_2 in V and any a in R; note that the operations of vector addition and scalar multiplication on the left hand side of the equation are those given by V, whereas on the right hand side of the equations they are those given by W, even though they are denoted by the same symbol and on both sides.

2) **EXERCISE** The only R-linear transformations $R \xrightarrow{T} R$ are those given by

$$T(r) = rs \qquad \text{for all } r$$

where S is a given element of R. Conversely, if S is given and T is defined by this equation, then T is R-linear.

) **EXERCISE** If S_1, S_2 are two given scalars in R and if V is an R-linear space, then we can construct an R-linear transformation

$$V \times V \xrightarrow{(S_1, S_2)} V$$

as follows. First we make the set V x V (of all ordered pairs of elements of V) itself into an R-linear space by defining the sum of v and w (written as column vectors)

$$\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} + \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} = \begin{pmatrix} v_1 + w_1 \\ v_2 + w_2 \end{pmatrix} \qquad \text{all } v_i, w_i \text{ in } V$$

and the scalar multiple

$$a\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} a v_1 \\ a v_2 \end{pmatrix}$$

Then we define the mapping (S_1, S_2) on such "column vectors" by

$$(S_1, S_2)\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = S_1 v_1 + S_2 v_2$$

the result coming out in V. Verify that the mapping thus defined satisfies the linearity conditions, and so is an R-linear transformation. In particular addition $V \times V \xrightarrow{+} V$ is an R-linear transformation.

Is multiplication $V \xrightarrow{s} V$ by a fixed scalar s an R-linear transformation?

) **EXERCISE** (In the case V = R) the <u>only</u> R-linear transformations

$$R \times R \xrightarrow{T} R$$

are of the form T = (S_1, S_2) as in the previous exercise, with S_1, S_2 scalars determined (how?) by T.

Exercise The only R-linear transformations $R \rightarrow R \times R$ are also given by pairs of scalars (better written this time as column vectors themselves, rather than rows).How are the two scalars determined by the given R-linear transformation T?

EXERCISE $R^2 = R \times R$

That is , the definition of addition and scalar multiplication comes out the same whether we regard R^2 as a special case of the function space constructic R^X in which X is an abstract set with two elements, or regard it as the special case V = R of the "square" $V \times V$ construction for an arbitrary R-linear space V.

Linear Combinations

If I is a finite index set and $\lambda_i \in R$ for each $i \in I$ and $v_i \in V$ for each $i \in I$, where V is an R-linear space, then

$$\boxed{\sum_{i \in I} \lambda_i v_i}$$

is a single element of V called the linear combination of the v's with coefficients the λ's. For this to make sense the family of coefficients and the family of "vectors" must be indexed by a common set I (but there may be repetitions in one or both of the two families); and of course each coefficient in the family of λ's must be a scalar and each vector in the family of v's must be in the same linear space V over the ring of scalars.

For example if I is the set of all commodities available in a certain supermarket and if λ_i is the price of the i-th commodity in dollars per unit amount of the commodity, a specific operation of linear combination is thus determined. Namely for each visit of a shopper to the supermarket let v_i be the amount the i-th commodity purchased by the shopper. (Thus V is the one-dimensional space of all possible amounts of commodities, e.g. of all possible weights - we could even imagine these to be negative since the shopper may be returning some unsatisfactory previously-purchased items). Then of course the linear combination (with coefficients the given prices) of the amount-family $(v_i)_{i \in I}$ for any shopper's visit is

$$\sum_{i \in I} \lambda_i v_i = \text{amount of dollars which that shopper is expected to pay at the check-out.}$$

In this example, V is one-dimensional, and in that respect only the example is very special compared to the ones we have to consider. But even here V is not the same as R, which latter consists of "pure" quantities, whereas V consists of "weights". The ratio of two weights is a pure quantity; i.e. a pure quantity multiplied by a weight (scalar multiplication) is another weight. If we chose a fixed weight, like one kilogram, then we can define a mapping

$$R \longrightarrow V$$

by sending any pure λ to the weight equivalent to λ kilograms; this (linear) mapping so defined is invertible (hence an "isomorphism"). That is, the choice of unit permits us to "identify" the two spaces R,V. However, this identification is conditional; if we choose a different unit weight $v_0 \in V$, say the weight of one pound, the two spaces R, V remain unchanged (and distinct but the mapping $R \longrightarrow V$ (defined by $\lambda \longmapsto \lambda u_0$) via which we make the "identification" is different from the one based on kilograms.

If $R \rightrightarrows V$, then the composition of one with the inverse of the other is a linear map from R to R, hence (by a previous exercise), determined by multiplication by a fixed (pure) scalar, namely the conversion factor from one unit to the other. To simplify, we could also identify "dollar" as a certain weight of silver.

In most of our examples V will not be one-dimensional, but a multi-dimensional linear space, for example a space of intensional functions defined on a region, or a space of extensional distributions on a region, or a space of translation vectors in physical affine space, or a space of column vectors or a space of matrices, etc. For any R-linear space V and any finite index set I we can form a new space V^I

Elements of V^I = all families of elements of V indexed by I

Addition of Elements of V^I : The sum of two families $(v_i)_{i\in I}$ and $(\bar{v}_i)_{i\in I}$ is the family $v + \bar{v}$ whose i-th entry is the sum of the i-th entries of the two:

$$(v+\bar{v})_i \underset{\text{def}}{=} v_i+\bar{v}_i \qquad i\in I$$

Scalar Multiplication of Elements of V^I by Elements of R : $(\lambda v)_i \underset{\text{def}}{=} \lambda v_i \qquad i\in I$

The dimension of V^I is bigger than the dimension of V by a factor equal to the cardinality (number of elements) of I.

$$\dim(V^I) = |I|\cdot\dim(V)$$

(In our supermarket example V is one-dimensional, but V^I is 341-dimensional if there are 341 commodities available; it is the space of all possible purchases by all possible shoppers at the given store). Then if $(\lambda_i)_{i\in I}$ is any fixed family of scalars indexed by the finite set I, the process of forming linear combination with those scalars as coefficients is a well-defined mapping

$$\begin{tikzcd} V^I \arrow[d, "\sum_{i\in I}\lambda_i(\ )_i"] \\ V \end{tikzcd}$$

since any family (element of V^I) can be substituted into the blank and the result will be a unique element of V itself.(Recall that the definition of "mapping" is any process which satisfies a condition of the type "$\forall\exists!$......."

1) EXERCISE The operation $\sum_{i\in I} \lambda_i (\ )_i$ of forming linear combinations with fixed family of scalars as coefficients, is a linear transformation between R-linear spaces.

2) EXERCISE $(R^2)^I \xrightarrow{\sim} (R^I)^2$ i.e. construct a tautological invertible linear transformation between the two indicated linear space

3) EXERCISE Not every R-linear transformation $V^I \longrightarrow V$ is of the form "linear combination with certain given coefficients" if $\dim(V) > 1$. That is, construct a counterexample with $V = R^2$

$\boxed{\text{THEOREM}}$ Any R-linear transformation $R^I \xrightarrow{T} R$ is of the form $T = \sum \lambda_i (\ )_i$ where the coefficients λ_i can be determined from the given transformation T.

PROOF: Assume I is a given finite index set and T is *any given* R-linear transformation $R^I \longrightarrow R$. Let e_j be the family of elements of R whose i-th entry is 0 for any $i \neq j$, and whose j-th entry is 1. That is

$$(e_j)_i = \begin{cases} 1 & i=j \\ 0 & i\neq j \end{cases} \qquad \text{for all } i$$

Thus we have a family of families, i.e. for each $j \in I$ $e_j \in R^I$. These "unit vectors" e_j are defined once and for all; but now for any given linear transformation T whose domain is R^I we can apply T to these special elements of the codomain of T; since in our present case the codomain of T is the linear space R of scalars itself, we can thus define

$$\lambda_j \overset{\text{def}}{=} T(e_j) \qquad j \in I$$

We now try to show that the operation of forming linear combination with these coefficients is the *same* as the given process T, when either is applie to *any* elements of R^I (not only to the very few elements e_j of R^I used to define λ_j from T): Let x_j be *any* family of elements of R(i.e. a single element of R^I). Then

$$x = \sum_{j\in I} x_j e_j$$

That is, since the x_j *themselves* are actual scalars (in the special case considered in this theorem) they can be used as coefficients to form the linear combination of the fixed vectors e_j, and the result is x *itself*.
[Verify] Hence for the given mapping T

$$\begin{aligned} Tx = T\sum x_j e_j &= \sum x_j T(e_j) && \text{since } T \text{ is linear} \\ &= \sum x_j \lambda_j && \text{by our definition of } \lambda_j \\ &= \sum_{j\in I} {}_j\, x_j && \text{since } R \text{ is commutative} \\ &= \text{the linear combination, with the coefficients} \\ &\quad\ \text{of the arbitrary } x \in R^I \end{aligned}$$

Since this holds for all x, $T = \sum \lambda_j (\ )_j$ as we wanted to show

$\boxed{\text{COROLLARY}}$ If V is a one-dimensional R-linear space, the only R-linear trans formations $V^I \longrightarrow V$ are given by linear combination with a fixed family of coefficients.

Sketch of Proof: Choose a unit u of V (i.e. a non-zero element; multiplying it by arbitrary scalars is an invertible linear transformation $R \xrightarrow{u} V$ with inverse $V \xrightarrow{u^{-1}} R$ (the existence of such u's is the definition of "one dimensional"). Then if T is any linear $V^I \longrightarrow V$, $u^{-1} \circ T \circ u^I$ is a linear transformation $R^I \longrightarrow I$ which by the theorem is determined by coefficients λ which can be transported back to T and be shown to work for T itself.

We have used above the fact that any linear transformation T preserves (an equational condition) any linear combination; this is in fact really what one wants linear transformation to do, and so could very well be taken as the definition of "linearity" for a mapping T, except that it may seem excessive to invoke arbitrarily large finite sets I in a definition if one can avoid it, even though in the applications of the definiton it is precisely the arbitrary finite sets which come up. Let us spell out this more liberal definition; first we need to make explicit a certain tautological construction.

If T is any mapping whose domain is a set V and whose codomain is a set W, and I is any finite index set then we define another mapping

$$V^I \xrightarrow{T^I} W^I$$

with domain the set of all I-indexed families of elements of V, and codomain the set of all I-indexed families of elements of W, by the specific formula

$$\boxed{(T^I(v))_i = T(v_i)}$$

Then

DEFINITION' If V,W are R-linear spaces and $V \xrightarrow{T} W$ is any mapping, v say T is linear if and only if for any finite set I and any family$(\lambda_i)_{i\varepsilon}$ of scalars (i.e. $I \xrightarrow{\lambda} R$ is any mapping) the following diagram is "commut tive", i.e. the indicated equality comes out true).

$$\begin{tikzcd} V^I \arrow[r, "T^I"] \arrow[d, "\sum_{i\in I}\lambda_i()_i"'] & W^I \arrow[d, "\sum_{i\in I}\lambda_i()_i"] \\ V \arrow[r, "T"'] & W \end{tikzcd}$$

$\forall (v_i)_{i\varepsilon I}$ family in V

$$\boxed{T(\Sigma\lambda_i v_i) = \Sigma\lambda_i T(v_i)}$$

"down then across" "across then down"

The equality required is of elements of W, constructed in the two indicated ways; as is customary we have used the same notation $\Sigma\lambda_i(\ )_i$ for the linea combination operation in the two different linear spaces V,W even though the concrete interpretation depends on the specific definitions of addition and scalar multiplication given in the specification of the linear structure on the two spaces.

Now the liberal definition of linearity implies the previously given definition of linearity as follows: Let the index set I be the two-element set whose only elements are 1,2, and let the family of scalars be $\lambda_1=1, \lambda_2=$
Then

$$\sum_{i\varepsilon I}\lambda_i v_i = v_1+v_2$$

i.e. for this particular choice of index set and family of coefficients, "linear combination" reduces to simple "addition" of a pair of vectors: hence the general linearity condition on T implies in particular that

$$\begin{tikzcd} V^{\{1,2\}} \arrow[r, "T^{\{1,2\}}"] \arrow[d, "+"'] & W^{\{1,2\}} \arrow[d, "+"] \\ V \arrow[r, "T"'] & W \end{tikzcd}$$

$$T(v_1+v_2) = T(v_1)+T(v_2)$$

In words, if T preserves general linear combinations, then T preserves addition in particular. Now if I is instead a one element set but $(\lambda_i)_{i\varepsilon I}$ is an arbitrary "family" (one element) λ, then linear "combination" with coefficient λ is just simple scalar multiplication, so if T preserves genera linear combinations then in particular T preserves scalar multiplication by any scalar λ :

$$\begin{tikzcd} V \arrow[r, "T"] \arrow[d, "\lambda()"'] & W \arrow[d, "\lambda\cdot()"] \\ V \arrow[r, "T"'] & W \end{tikzcd}$$

$$T(\lambda v) = \lambda T(v) \quad \text{all } v$$

Note that if I has only one element, $V^I = V$. Similarly, if I is empty, $V^I =$ (the linear space with only the zero element, and we get T(O) = O.

Briefly, the three clauses in the original definition of "linearity" of a mapping are special cases of the liberal definition. Conversely, as needed in the applications, the three-clause definition implies that T preserves general linear combinations; this has the usual sort of proof for such things: (a) by mathematical induction and the associative and distributive laws, one proves it for arbitrarily large ordered sets of the form

$$I_n = \{1,2,3,..,n\}$$

then (b) one uses the fact that for any finite set I there exists an n and a bijection (invertible mapping) $\beta : I_n \longrightarrow I$ which can be used to transport $\beta^{-1}(\ )\beta$ the theorem from I_n to I.

The $\beta : I_n \longrightarrow I$ just mentioned can be a source of confusion in applications, since β is not uniquely determined by I (though n is)—β amount to counting the elements of I in a particular order, so that $\beta(1)$ is the "first" (according to β) element of I, $\beta(2)$ is the "second" element of I according to β , etc. In the case where I is the set of commodities available in the supermarket, a mapping $\beta : I_n \longrightarrow I$ could be somebody's subjective choic of a "shopping list" written down in a particular order (more precisely, this would normally be injective but not bijective, since the shopper would not bother to list the items of which he plans to buy zero amount). If there exists a bijection $\beta : I_n \longrightarrow I$, one says (by definition) that $|I| = n$; but in that case there exist n! (factorial) different choices for β , none of which are to be preferrred in general, for example the objective meaning of the supermarket bill

$$\mathcal{B}_\lambda(v) = \sum_{i \in I} \lambda_i v_i$$

for given price family λ is independent of the subjective ordering in the shopper's shopping list, as well, as of the subjectively chosen ordering in which the check-out person rings up the amounts v_i.

The ordering of the index set of course is involved in the notation for families of quantities as column vectors, row vectors, and matrices; the meaning of matrix multiplication is objective, but the way the matrices are written in a particular case changes when the choice of ordering is changed - what must be maintained to maintain the meaning is the integrity of the index sets and of their elements. Thus if I, J are sets, by a J x I matrix of scalar: is meant a doubly-indexed family

$$b_{ji} \in R \qquad \begin{matrix} j \in J \\ i \in I \end{matrix}$$

and if K is a third set by a KxJ matrix is ment a doubly indexed family

$$a_{kj} \in R \quad \begin{matrix} k \in K \\ j \in J \end{matrix}$$

Since the index set J is the same, the matrix product ab can be defined as the KxI matrix

$$\boxed{(ab)_{ki} \overset{=}{\underset{df}{=}} \sum_{j\in J} a_{kj} b_{ji} \qquad \begin{matrix} k \in K \\ i \in I \end{matrix}}$$

When we are moreover given orderings of the three index sets I,J,K, then a,b, and ab can be displayed rectangularly and the above definition is seen to be equivalent to the " each row of a times each column of b to get all the entries of ab" description where each of the "times" really means a different linear combination with index set J. The matrix multiplication situation coul conveniently be described as follows:

$$\begin{tikzcd} & J \arrow[dr, "a", "+" marking] & \\ I \arrow[ur, "c", "+" marking] \arrow[rr, "ab"', "+" marking] & & K \end{tikzcd}$$

where the mark on the arrows indicates that these are not ordinary mappings from I to J, etc. but matrices with entries in R. However,

THEOREM A JxI matrix of scalars determines an actual mapping $R^I \longrightarrow R^J$ which is R-linear, and all R-linear mappings from R^I to R^J are so determined.

PROOF: Let b be a J x I matrix . We will define a mapping T as follows. For any $x \in R^I$ we can view x as an $I\times I_1$ matrix (a "column vector", if I were ordered), where I_1 is a one element set, $\{I\times I_1 \xrightarrow{\approx} I\}$

$$\begin{tikzcd} I_1 \arrow[r, "x", "+" marking] \arrow[rr, bend right, "bx"', "+" marking] & I \arrow[r, "b", "+" marking] & J \end{tikzcd}$$

The matrix product is a $J\times I_1$- matrix, i.e. a J-indexed family of scalars, i.e an element of R^J. Recalling the definition of matrix product, we see that we have defined

$$\boxed{(T_b(x))_j := \sum_{i\in I} b_{ji} x_i \qquad \text{all } j \in I, \quad x \in R^I}$$

It can then be verified that T_b is R-linear and that, conversely, any R-line mapping $R^I \xrightarrow{T} R^J$ is of this form, where the entries b_{ji} can be recovered from T-as-a-process by feeding in the special 0,1 "column vectors" e as before

40

THEOREM If $I \xrightarrow{b} J \xrightarrow{a} K$ are matrices with scalar entries then in $R^I \xrightarrow[T_b]{} R^J \xrightarrow[T_a]{} R^K$ we have

$$\boxed{T_{ab} = T_a \circ T_b}$$

i.e. composition of linear transformations between these special spaces R^I i represented by matrix multiplication ab

PROOF Calculate!

Linear Categories

In general a specific category is determined by specifying objects and morphisms together with domain, codomain, composition, and identity, subject to associativity and identity laws. The domain and codomain of a morphism are objects, and it is helpful to write

$$X \xrightarrow{f} Y$$

$$\begin{tikzcd} X \arrow[loop right, "end"] \end{tikzcd}$$

to mean that "f is a morphism whose domain is the object X and whose codomain is the object Y". An endomorphism of X is any morphism whose domain and codomain are both X. The composition of two morphisms is meaningful iff the codomain of the first is the same object as the domain of the second, as in

$$X \xrightarrow{f} Y \xrightarrow{g} Z$$

where Y is the common object; in that case the composition is another morphism, denoted by gf, whose domain and codomain are as indicated in

$$\begin{tikzcd} X \arrow[r, "f"] \arrow[drr, bend right, "gf"'] & Y \arrow[dr, "g"] & \\ & & Z \end{tikzcd}$$

The associative law states that if $X \xrightarrow{f} Y \xrightarrow{g} Z \xrightarrow{h} W$, then $h(gf) = (hg)f$, i.e. that in the diagram

$$\begin{tikzcd} X \arrow[r, "f"] \arrow[dr, "gf"'] & Y \arrow[d, "g"] \arrow[dr, bend left, "hg"] & \\ & Z \arrow[r, "h"] & W \end{tikzcd}$$

the two outer paths from X to W are actually the same morphism. Finally, for every object X there is an identity endomorphism 1_X, determined among all the (in general many) endomorphisms of X by the conditions that for any morphism f with domain X and any morphism x with codomain X,

$$f1_X = f, \quad 1_X x = x$$

$$\begin{tikzcd} T \arrow[r, "x"] \arrow[dr, "x"'] & X \arrow[d, "1_X"] \arrow[dr, "f"] & \\ & X \arrow[r, "f"'] & Y \end{tikzcd}$$

(here T,Y are any objects of the category).

As two simple examples of categories we mention 1) the category whose objects are all finite sets and whose morphisms are all possible mappings between finite sets, with composition the usual "substitution" and 2) the category whose objects are again all possible finite sets (thought of as "index sets") but whose morphisms $I \longrightarrow J$ are all possible JxI matrices of real numbers with matrix multiplication as composition

$$\begin{tikzcd} I \arrow[r, "A"] \arrow[dr, "BA"'] & J \arrow[d, "B"] \\ & K \end{tikzcd} \qquad (BA)_{ki} = \sum_{j \in J} B_{kj} A_{ji}$$

This category ② may be considered as "larger" than the category ① since any mapping $I \xrightarrow{f} J$ determines a matrix by

$$f_{ji} = \begin{cases} 1 & \text{if } f(i) = j \\ 0 & \text{if } f(i) \neq j \end{cases}$$

Such matrices may be characterized by the condition that in every "column" there is just one 1, the rest of the entries being 0; of course most matrices do not meet this stringent requirement, which is why we say "② > ①"

① **EXERCISE** If $gf = \gamma$ as mappings, then the matrix corresponding to γ is equal to the matrix product of the matrices corresponding to g, f.

But there is also an important qualitative difference, namely the category whose morphisms are mappings is distributive $A \times (B+C) = A \times B + A \times C$ for a natural sum and product on the objects (to be explained) whereas the category whose morphisms are matrices is linear $A \times B = A + B$ for the sum and product defined in the same way .

If we have in a category four objects and four morphisms for whic

$$\gamma_2 f_1 = f_2 \gamma_1 \qquad \begin{tikzcd} X \arrow[r, "f_1"] \arrow[d, "\gamma_1"'] & Y \arrow[d, "\gamma_2"] \\ Z \arrow[r, "f_2"'] & W \end{tikzcd}$$

i.e. the two composites define the same morphism $X \longrightarrow W$, then one says we have a commutative square. A square as pictured is said to have a diagonal-fill-in if there exists g from Y to Z for which

$$g f_1 = \gamma_1 , \qquad f_2 g = \gamma_2$$

Proposition If a square has a diagonal fill-in, then it is a commutative square.

Proof: This is easily seen to be a just restatement of the associative law.

② **EXERCISE:** Give simple examples in the category of finite sets and mappings of a square that is not commutative and also of a square that is commutative but which does not have a diagonal fill -in.

A morphism $X \xrightarrow{f} Y$ in a category is said to be invertible or to be an isomorphism if there exists in the same category another morphism g for which the following two equations are true 43

$$gf = 1_X$$
$$fg = 1_Y$$

$$\begin{tikzcd} X \arrow[loop left, "1_X"] \arrow[r, shift left, "f"] & Y \arrow[l, shift left, "g"] \arrow[loop right, "1_Y"] \end{tikzcd}$$

This g is unique if it exists or more strongly

③ **EXERCISE** If $g_1 f = 1_X$ and $f g_2 = 1_Y$ then $g_1 = g_2$.
Thus we can name g in terms of f $\boxed{g = f^{-1}}$ where f^{-1} is defined only if f is invertible and is called the inverse of f.

EXERCISE 1_X is invertible
If f is invertible, so is f^{-1} and $(f^{-1})^{-1} = f$
If f, φ are both invertible and composable, then φf is invertible and

$$\boxed{(\varphi f)^{-1} = f^{-1}\varphi^{-1}} \qquad \begin{tikzcd} X \arrow[r, shift left, "f"] & Y \arrow[l, shift left] \arrow[r, shift left, "\varphi"] & Z \arrow[l, shift left] \end{tikzcd}$$

(The socks-and-shoes principle)

An isomorphism in the category of sets and mappings is often called a bijection. Two objects X,Y are said to be isomorphic $X \cong Y$ iff there exists at least one isomorphism between them. In effect two isomorphic objects (though they may be different) are mathematically indistinguishable (unless more is known about them than just that they lie in the category in question), however in that case they are in fact coming from objects in a "richer" category where they may not be isomorphic).

④ **EXERCISE** In the category of sets and mappings, two sets are isomorphic iff they have the same number of elements.

⑤ **EXERCISE** In the category of matrices, there exist isomorphisms $2 \nrightarrow 2$ which are not induced by mappings $2 \longrightarrow 2$ (However, if there exists an invertible matrix $I \nrightarrow J$, it does follow that there exists also an invertible mapping $I \longrightarrow J$ but some calculation is needed to produce it)

But the problem of the existence of an isomorphism between two given objects (and the resulting division of all objects into equivalence classes known as isomorphism classes - e.g. there is one isomorphic class of finite sets for each natural number n = 0,1,2,3...., and the class corresponding to n consists of all finite sets isomorphic to the set $[n] = \{1,2,....,n\}$) by no means finishes the role of isomorphism because where there is one there are usually many:

A morphism which is both an endomorphism and also an isomorphism is called an automorphism, and the set $\mathrm{Aut}_{\mathcal{C}}(X)$ of all automorphisms of the object X in the category $\mathcal{C}$ is called the automorphism group

4 ⑥ **EXERCISE** In the category of sets and mappings, if a set X has n elements, then Aut(X) has n! = n$\cdot$(n-1)(n-2)$\cdot$ .3.2.1 elements. $n! \leq n^n$ because in general the number of mappings from X to Y is m^n if X has n elements and Y has m elements.

⑦ **EXERCISE** In the category whose morphisms are real matrices Aut ([n]) is an n^2-dimensional (non-linear) space consisting of all n x n matrices whose determinant does not vanish.

⑧ **EXERCISE** In any category, if $X \cong Y$, then there is a bijection between Aut(X) and Aut(Y) which preserves composition. (Choose $X \longrightarrow Y$ in order to define this). Further, the number of isomorphisms $X \longrightarrow Y$ is the same as the size of Aut (X).

A very important role is played by pairs of morphisms satisfying the one equation

$$gf = 1_X \qquad\qquad \begin{tikzcd} X \arrow[r, shift left, "f"] & Y \arrow[l, shift left, "g"] \end{tikzcd}$$

⑨ **EXERCISE** In the category of sets and mappings, if we are given two sets X and Y, then there exists a pair f,g as above iff $n \leq m$ where $n = |X|$ and $m = |Y|$.

In general the existence of a "section/retraction" pair f,g as above is a very strong proof that X is "smaller" then Y; in categories richer than sets "smaller than"is more general than this but one still works to use it in "local or" approximate ways because it is so explicit when one can get it. The importance is reflected in the many names that have come into use to describe various aspects:

f is called a split monomorphism iff there exists a retraction g for f [i.e. $\exists g\,[gf = 1_X]$]

g is called a split epimorphism iff there exists a section f for g [i.e. $\exists f\,[gf = 1_X]$]

A pair f,g is called a splitting for an endomorphism e of Y iff fg = e and $gf = 1_X$

An endomorphism e is called splittable iff $\exists X \exists fg$....

⑩ **EXERCISF** Any splittable endomorphism e is idempotent ee = e.

⑪ **EXERCISE** If e is an idempotent endomorphism of Y and if X_1, X_2 occur in two splitting pairs for the same e, then $X_1 \cong X_2$

⑫ **EXERCISE** In the category of linear spaces and linear transformations if $\begin{tikzcd} X \arrow[r, shift left, "f"] & Y \arrow[l, shift left, "g"] \end{tikzcd}$ with $gf = 1_X$, let Y_0 be the kernel (nullspace) of the split epimorphism g. Say (for f understood) that "$y \in X$" iff $\exists x\,[fx=y]$

Show that any element y of Y can be uniquely expressed as a sum

$$y = y_1 + y_0 \qquad y_1 \in X, \quad y_0 \in Y_0$$

(The same is actually true in any linear category).

In the category of linear spaces and linear transformations (as well as in the category of sets and mappings, with the exception of the case X = 0) we have a very explicit way of handling images, (which will not be directly available in more structural categories of linear systems, etc: A morphism f is called quasi-invertible if there exists another morphism g in the same category with $\boxed{fgf = f}$

(13) **EXERCISE** If f is quasi invertible, then there exists a "quasi-inverse" g for which both fgf = f and gfg = g [Hint: given a g satisfying only the one equation, show that $\bar{g} =_{\mathrm{def}} gfg$ is an improved version satisfying both. Moreover $\bar{\bar{g}} = \bar{g}$, so that no further improvement is possible with=out more information.] If g is quasi inverse to f , then both gf and fg are idempotent (endomorphisms of domain and codomain of f, respectively).

The following applies in particular to the category of linear spaces and linear transformations.

Theorem: In a category in which every morphism has a quasi-inverse and every idempotent has a splitting pair, it follows that every morphism f has a factorization

f = jp

$$\begin{tikzcd}
X \arrow[rr, "f"] \arrow[dr, "p"'] & & Y \\
& I \arrow[ur, "j"'] &
\end{tikzcd}$$

where j is a (split) monomorphism and p is a (split) epimorphism.

(14) Proof: **EXERCISE** Let g be a quasi-inverse for f and let X',i,p and Y',j,r be splittings for the idempotents gf and fg respectively. Define $X' \xrightarrow{f'} Y'$ by f' = rfi. Then f'has an inverse g' defined in a dual manner. (Thus we may identify X',Y' and assume f',g' are identities) jf'p = f so f is factored as required.

Actually the mono-epi factorization is essentially determined by f (though the splittings g,i,r are not) and is called the image of f.

In general "epi" and "mono" are not such simple equational properties but involve a universal quantification over all (or at least many) morphisms of the category in question:

DEFINITION $X \xrightarrow{f} Y$ is a monomorphism iff $\forall\ T\ \forall\ T \overset{x_1}{\underset{x_2}{\rightrightarrows}} X\ \left[fx_1 = fx_2 \Rightarrow x_1 = x_2\right]$

(15) **EXERCISE** A morphism in the category of sets is a monomorphism iff it is injective.

(16) **EXERCISE** Any split monomorphism is a monomorphism (has very simple proof, as does the following "dual").

(17) **EXERCISE** Any split epimorphism is an epimorphism, where

DEFINITION: $Y \xrightarrow{g} Z$ is an epimorphism iff

$$\forall W \; \forall Z \underset{w_2}{\overset{w_1}{\rightrightarrows}} W \left[w_1 g = w_2 g \Rightarrow w_1 = w_2 \right]$$

(18) **EXERCISE** A mapping of sets is an epimorphism iff it is surjective [Hint: Take W a two-element set (of "Wahrheitswerte")]

(19) **EXERCISE** If every morphism has a quasi-inverse, then every epimorphism and every monomorphism splits (in the two respective senses).

Unfortunately in most categories the foregoing exercise does not apply so we are forced to consider two distinct ways that an object may be smaller than another; in fact we combine the two as follows "$X \leq Y$" means $\exists \; X \twoheadleftarrow S \rightarrowtail Y$ where heavy head means epi and tail means mono. Fortunately in most categories iterating this idea does not lead to more complications; i.e. $\leq$ so defined is already transitive.

Why are categories like Lin, $\mathrm{Lin}^{\downarrow}$, $\mathrm{Lin}^{\swarrow}$, and $\mathrm{Lin}^{\downarrow\uparrow}$ considered as "linear" categories? There are essentially two answer s, which can be shown to be essentially equivalent. The first is that "maps can be added For example, if $V \overset{A}{\underset{B}{\rightrightarrows}} W$ are two linear transformations between two linear spaces, we can define a mapping

$$V \xrightarrow{A+B} W$$

by the formula

$$(A+B)(v) = A(v) + B(v) \text{ for all } v \in V$$

Then A+B is <u>linear</u> (if A,B are) because

$$\begin{aligned}(A+B)(\Sigma\lambda_i v_i) &= A(\Sigma\lambda_i v_i) + B(\Sigma\lambda_i v_i) &&\text{— definition of A+B}\\ &= \Sigma\lambda_i Av_i + \Sigma\lambda_i Bv_i &&\text{— since A,B linear}\\ &= \Sigma\lambda_i(Av_i+Bv_i) &&\text{— since W linear space}\\ &= \Sigma\lambda_i(A+B)v_i &&\text{— definition of A+B again}\end{aligned}$$

Now if

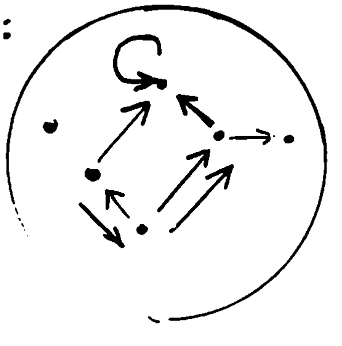
$$\begin{tikzcd} U \arrow[r, "T"] & V \arrow[r, shift left, "A"] \arrow[r, shift right, "B"'] & W \arrow[r, "S"] & X \end{tikzcd}$$

are all linear then

$$\boxed{\begin{aligned}(A+B)T &= AT + BT\\ S(A+B) &= SA + SB\end{aligned}}$$

Together with the associative law of composition (which holds for any category) and the commutativity, associativity, and 0 laws for addition for morphisms between each given pair V,W, the above <u>distributive laws</u> are the essential axioms for "linear category".

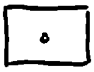
We get a whole class of different examples of linear categories as follows. (See theorem below). Consider any "directed graph" $\mathbf{G}$ which means a given set I of "vertices", a given set G of "edges" and a given pair ∂_0, ∂_1 of maps $G \overset{\partial_0}{\underset{\partial_1}{\rightrightarrows}} I$ thought of as assigning to each edge g its "source vertex" $\partial_0(g)$ and its "target vertex" $\partial_1(g)$. No axioms are imposed in general on the structural maps ∂_0, ∂_1 of a graph $\mathbf{G} = \langle I, G, \partial_0, \partial_1 \rangle$ If I,G are finite sets, then the graph $\mathbf{G}$ can actually be pictured by drawing the vertices as dots and the edges as arrows:

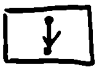


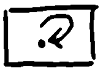
In the pictured example, I has 5 elements, G has 8 elements and the definitions of ∂_0, ∂_1 are forced by the picture.

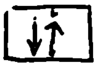
Exercise: Find a pair of edges $g_1 g_2$ in the above picture for which $\partial_0(g_1) = \partial_0(g_2)$ and $\partial_1(g_1) = \partial_1(g_2)$ but $g_1 \neq g_2$ (Parallel edges"). Find also a pair $h_1, h_2 \in \mathbb{G}$ for which $\partial_0(h_1) = \partial_1(h_2)$ and $\partial_1(h_1) = \partial_0(h_2)$ ("feedback edge"). Find also an edge g such that $\partial_0(g) = \partial_1(g)$ ("Loop") as well as a vertex i such that $\forall h \in G[\partial_0(h) \neq i \text{ and } \partial_1(h) \neq i]$ ("Isolated vertex") and finally a vertex j such that $\exists h \in G[\partial_1(h) = j]$ and $\forall h \in G[\partial_0(h) \neq j]$ ("Sink?").

Some of the simplest basic examples of graphs would be the following four.









Now if we are given a graph $\mathbb{G}$ and a category $\mathbb{C}$ (for example the category $\mathbb{C}$ = Lin of real linear spaces and real linear transformations or the category $\mathbb{C} = \mathbb{S}$ of sets and arbitrary mappings) then we can construct an interesting and useful category $\mathbb{C}^{\mathbb{G}}$ as follows (in case $\mathbb{C}$=Lin, we could call Lin$^{\mathbb{G}}$ the category of "linear systems of shape $\mathbb{G}$ and $\mathbb{G}$-linear morphisms between such systems"):

0) An object $\mathbb{V}$ of $\mathbb{C}^{\mathbb{G}}$ is any system of the following type: to every vertex $i \in I$ of $\mathbb{G}$ we choose an object V_i of $\mathbb{C}$, and to every edge $g \in G$ we choose a morphism T_g of $\mathbb{C}$ subject only to the condition that whenever $i \xrightarrow{g} j$ in $\mathbb{G}$ (i.e. whenever $\partial_0(g) = i$ and $\partial_1(g) = j$) we must specify a $\mathbb{C}$-morphism T_g whose domain is the $\mathbb{C}$-object V_i and whose codomain is the $\mathbb{C}$-object V_j (i.e.

$$i \xrightarrow{g} j \text{ in } \mathbb{G} \Longrightarrow V_i \xrightarrow{T_g} V_j \text{ in } \mathbb{C}.$$

1) If $\mathbb{V} = \langle V,T \rangle$, $\mathbb{W} = \langle W,S \rangle$ are two objects of $\mathbb{C}^{\mathbb{G}}$, then by a morphism $\mathbb{V} \xrightarrow{A} \mathbb{W}$ is meant any assignment of a $\mathbb{C}$-morphism A_i to every vertex $i \in I$ of $\mathbb{G}$ which satisfies the "$\mathbb{G}$-homomorphism" or"$\mathbb{G}$-naturalit condition: for all edges $g \in G$ in $\mathbb{G}$ the equation below must hold (where $i = \partial_0(g), j = \partial_1(g)$)

$$\boxed{A_j T_g = S_g A_i}$$

$$\begin{tikzcd}
V_i \arrow[r, "A_i"] \arrow[d, "T_g"'] & W_i \arrow[d, "S_g"] \\
V_j \arrow[r, "A_j"] & W_j
\end{tikzcd}$$

Remark: We often just write g alone for T_g, S_g etc. when the context is understood.

2) If $\mathbb{U} \xrightarrow{B} \mathbb{V} \xrightarrow{A} \mathbb{V}$ are $\mathcal{C}^{\mathbb{G}}$ morphisms then the <u>composition</u> AB is defined as follows

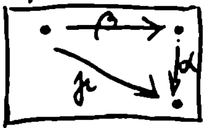
$$\boxed{(AB)_i \overset{=}{\text{def}} A_i B_i} \quad \begin{array}{l}\text{composition in } \mathcal{C} \\ \text{for each vertex } i \in I \text{ of } \mathbb{G}\end{array}$$

Exercise AB again satisfies the $\mathbb{G}$-naturality conditions for each edge $g \in G$ of $\mathbb{G}$, and is hence again a $\mathcal{C}^{\mathbb{G}}$-morphism.

Exercise: A is invertible (i.e. is an isomorphism) in $\mathcal{C}^{\mathbb{G}}$ if and only if A_i is invertible in $\mathcal{C}$ for each vertex $i \in I$ of $\mathbb{G}$. (Hint: the main point is to show that the A_i^{-1} (taken together) is $\mathbb{G}$-natural if A itself is (and if the individual inverses exist)).

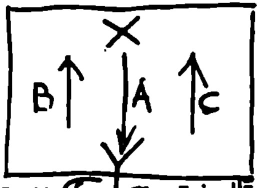
Exercise If we consider the particular graph $\boxed{\downdownarrows}$ with two vertices and two parallel edges, and take $\mathcal{C} = \mathcal{S}$ = the category of all sets and all mappings, then $\mathcal{S}^{\downdownarrows}$ is the category of all graphs and graph-morphisms where the latter means a pair of mappings, working on edges and on vertices respectively, which "preserve" the source and target relations of two given graphs.

Now nearly all categories of mathematical interest can be viewed as full subcategories $\mathfrak{D} \subseteq \mathcal{C}^{\mathbb{G}}$ where $\mathcal{C} = \mathcal{S}$ or $\mathcal{C}$=Lin and $\mathbb{G}$ is an appropriately chosen graph. Here "full" means that we take the same definition (of $\mathbb{G}$-natural) for the morphisms of $\mathfrak{D}$ as for the morphisms of $\mathcal{C}^{\mathbb{G}}$, but restrict the objects, often by equations. For example, if $\mathbb{G}$ =



it is often reasonable to restrict to the full subcategory $\mathfrak{D}$ of $\mathcal{C}^{\mathbb{G}}$ which consist of all $\mathbb{V} = \langle V, T \rangle$ for which $\boxed{T_\alpha \quad T_\beta \quad = T_\gamma}$ in $\mathcal{C}$. Note that in this example there are three $\mathcal{C}$-objects V_i involved in each $\mathcal{C}^{\mathbb{G}}$-object, and the equation defining $\mathfrak{D}$ might just be written "$\alpha\beta = \gamma$", by the T-omission remark above.

If $\mathcal{C}$ itself is a linear category, then there are many more types of equations which would define reasonable $\mathfrak{D}$, since we can use not only composition in $\mathcal{C}$ but also the linear structure of $\mathcal{C}$ in making up equations. For example if $\mathbb{G} = \boxed{\cdot \circlearrowleft d}$ then very important in geometry and analysis is the subcategory $\mathfrak{D} \subset \text{Lin}^{\circlearrowleft}$ consisting of all $V^{\circlearrowleft d}$ for which $\boxed{d \circ d = 0}$ An example fundamental in $\boxed{\text{feedback control}}$ is the graph $\mathbb{F}$ =



and the subcategory $\mathfrak{F} \subset \text{Lin}^{\mathbb{F}}$ defined by the equation

$$\boxed{C = B(1_Y - AC)}$$

More exactly, note that there is a subgraph $\mathbb{F}_0 \subset \mathbb{F}$ consisting of only A, B, and that such a subgraph induces a trivial "forgetting"

$$\mathrm{Lin}^{\mathbb{F}} \longrightarrow \mathrm{Lin}^{\mathbb{F}_0}$$

which in this case just omits C from any $\mathbb{F}$-system to get the underlying $\mathbb{F}_0$-system. Composing this with the inclusion we get a trivial process

$$\mathcal{F} \longrightarrow \mathrm{Lin}^{\mathbb{F}_0}$$

The first non-trivial problem of feedback control theory is, given any linear $\mathbb{F}_0$-system $X \overset{A}{\underset{B}{\rightleftarrows}} Y$ in $\mathrm{Lin}^{\mathbb{F}_0}$, to find all $\mathbb{F}$-systems <u>in $\mathcal{F}$</u> which restrict to it, i.e. to find all $Y \xrightarrow{C} X$ which are solutions of the equation $C = B(1_Y - AC)$.

Exercise In case the Lin-endomorphism 1_Y+BA of Y is <u>invertible</u> (for example if BA is nilpotent), then the feedback-control problem <u>has a unique</u> solution C.

** [redacted] On a first reading one could skip to the theorem on the next page

Exercise* Lin itself can be considered as a full subcategory of the (non linear) category $\mathcal{S}^{\mathbb{M}}$ where $\mathbb{M}$ is the (rather large) graph whose vertices are the natural numbers 0,1,2,3.... and whose edges are all rectangular real matrices of all possible sizes, and $\partial_0(A) = m$, $\partial_1(A) = n$ if A is n x m. Hint: To a linear space V associate the $\mathbb{M}$-system in $\mathcal{S}$ whose n-th vertex is V^n = the set of all n-tuples of vectors from V, with the matrices acting in a sensible way, and show that an $\mathbb{M}$-natural morphism between <u>such</u> objects of $\mathcal{S}^{\mathbb{M}}$ is essentially just a linear transformation between the linear spaces V (in particular if A is a $\mathbb{M}$-natural morphism, then all the mappings A_n are actually determined by the one mapping A_1).

Exercise * (Introduction to topology, advanced calculus, functional analysis). By a (symmetric) metric space X is meant a set X of points equipped with a "distance function", which means a function $X \times X \xrightarrow{d} [0,\infty)$ (where $[0,\infty) = \{r \in \mathbb{R} \mid r \geq 0\}$) satisfying the four axioms

$$d(x,y) + d(y,z) \geq d(x,z) \text{ for all } x,y,z \text{ in } X$$

$$d(x,y) = d(y,x)$$

$$d(x,x) = 0$$

$$d(x,y) = 0 \Longrightarrow x = y.$$

If $X \xrightarrow{f} Y$ is a mapping between two metric spaces, f is called a <u>continuous</u> map iff $\forall x \forall \epsilon > 0 \exists \delta > 0$

$$\forall x' [d_X(x,x') < \delta \Rightarrow d_Y(f(x),f(x')) < \epsilon]$$

The composition $\begin{tikzcd} X \arrow[r, "f"] \arrow[rr, bend right] & Y \arrow[r, "g"] & Z \end{tikzcd}$ of two continuous maps is again continuous so we get a category $\mathcal{T}$ with a full inclusion of categories

$$\mathcal{T}_{met} \subset \mathcal{S}^{\mathbb{T}}$$

if we define the graph '$\mathbb{T}$' as follows. There are only two vertices $\mathbb{1}$' and $\bar{\mathbb{N}}^{op}$, but many edges, as follows: There is exactly one edge $\mathbb{1} \xrightarrow[c]{} \bar{\mathbb{N}}^{o}$ but there are edges $\bar{\mathbb{N}}^{op} \xrightarrow[n]{} \mathbb{1}^{op}$ for each n=0,1,2,3.... and for n = ∞ as well; there is exactly one edge $\mathbb{1}^{op} \to \mathbb{1}^{op}$ ("identity"), but between $\bar{\mathbb{N}}^{op} \xrightarrow{s_n} \bar{\mathbb{N}}^{op}$ there is an edge for each natural number n = 0,1,2,...

The inclusion $\mathcal{T}_{met} \hookrightarrow \mathcal{S}^{\mathbb{T}}$ is defined by associating to each metric space X the $\mathbb{T}$-system of sets $V_1 = X$, $V_{\bar{\mathbb{N}}} = Cgt(X)$ where Cgt(X) is the set of "convergent sequences in X", i.e. the set of all continuous maps $\bar{\mathbb{N}} \to X$, where $\bar{\mathbb{N}}$ is the metric space whose points are 0,1...,∞. the natural numbers together with one more point called ∞ and whose distance function is $d(n,m) = \left| \frac{1}{n} - \frac{1}{m} \right|$, where $\frac{1}{\infty} \underset{def}{=} 0$. The interpretation of the abstract edges of $\mathbb{T}$ in the particular $\mathbb{T}$-system corresponding to a given metric space X is

$$\begin{tikzcd} Cgt(X) \arrow[loop left, "s_n"] \arrow[r, shift right, "n"'] & X \arrow[l, shift right, "C"'] \end{tikzcd}$$

where n evaluates a given sequence at n(in particular, the value at ∞ is the "limit" of the sequence), C assigns to any point the sequence which is constantly that point, and S_n shifts any sequence $x_0, x_1, ..x_\infty$ by n, i.e.

$$s_n(X)_k = x_{k+n}$$

$$s_n(X)_\infty = x_\infty$$

Then the proof of fullness $\mathcal{T}_{met} \hookrightarrow \mathcal{S}^{\mathbb{T}}$ is first a simple calculation showing that any $\mathbb{T}$-natural f

$$\begin{tikzcd} Cgt(X) \arrow[loop left] \arrow[r, "f_{\bar{\mathbb{N}}}"] \arrow[d, shift left] & Cgt(Y) \arrow[loop right] \arrow[d, shift left] \\ X \arrow[u, shift left] \arrow[r, "f_1"'] & Y \arrow[u, shift left] \end{tikzcd}$$

is actually determined by f_1 via $(f_{\bar{\mathbb{N}}}(x))_n = f_1(x_n)$ for any $x \in Cgt(X)$ and that this is true in particular for $n = \infty$, i.e. that

$$\lim_n f(x_n) = f(\lim_n x_n)$$

which can be proved to be equivalent to f_1 being continuous.

**

THEOREM For any graph $\mathbb{G}$ and for $\mathbb{C}$=Lin (or more generally any linear category $\mathbb{C}$), the category $Lin^{\mathbb{G}}$ of $\mathbb{G}$-linear systems is again a linear category

Proof: If $\mathbb{V} \overset{A}{\underset{B}{\rightrightarrows}} \mathbb{W}$ in Lin$\mathbb{G}$, then define $(A+B)_i = A_i + B_i$ in Lin (previously defined) for each vertex $i \in I$ in $\mathbb{G}$. Then A+B is again $\mathbb{G}$-natural, for if $i \xrightarrow[g]{} j$ is any edge in $\mathbb{G}$, then

$$\begin{aligned} (A+B_j)T_g = (A_j+B_j)T_g &= A_jT_g + B_jT_g \quad \text{since Lin itself is a Lin Cat} \\ &= S_gA_i + S_gB_i \quad \text{since A,B assumed } \mathbb{G}\text{-natural} \\ &= S_g(A+B)_i \quad \text{def} \end{aligned}$$

$$\begin{tikzcd} V_i \arrow[r, "(A+B)_i"] \arrow[d, "T_g"'] & W_i \arrow[d, "S_g"] \\ V_j \arrow[r, "(A+B)_j"'] & W_j \end{tikzcd}$$

Thus the addition of morphism in Lin $\mathbb{G}$ (with given domain and codomain) is <u>well defined</u> and fairly obviously satisfies the commutative, associative, and zero laws (for given domain $\mathbb{V}$ and codomain $\mathbb{W}$) since additi in Lin itself does and we have already proved many times that "products' $\prod_{i\in I} \mathrm{Lin}(V_i, W_i)$ of linear spaces again satisfy the CAZlaws; we are dealing with the subspace $\mathrm{Lin}_{\mathbb{G}}(\mathbb{V},\mathbb{W})$ of this product consisting of <u>$\mathbb{G}$-natural</u> families $(A_i)_{i\in I}$, but identities like the C A Z laws are clearly inherited by subspaces, if only they are closed under addition, and that is what we proved above. To finish the proof we need only show that Lin $\mathbb{G}$ satisfies the two <u>distributive</u> axioms characteristic of linear categories. So assume the

$$\begin{tikzcd} \mathbb{U} \arrow[r, "C"] & \mathbb{V} \arrow[r, shift left, "A"] \arrow[r, shift right, "B"'] & \mathbb{W} \arrow[r, "D"] & \mathbb{X} \end{tikzcd}$$

are all $\mathbb{G}$-linear morphisms. Then for i a vertex of $\mathbb{G}$,

$((A+B)\ C)_i = (A + B)_i C_i$ definition of composition

$= (A_i + B_i)C_i$ definition of addition

$= A_iC_i + B_iC_i$ since Lin satisfies distributive

$= (AC)_i + (BC)_i$ definition of composition

$= (AC + BC)_i$ definition of addition

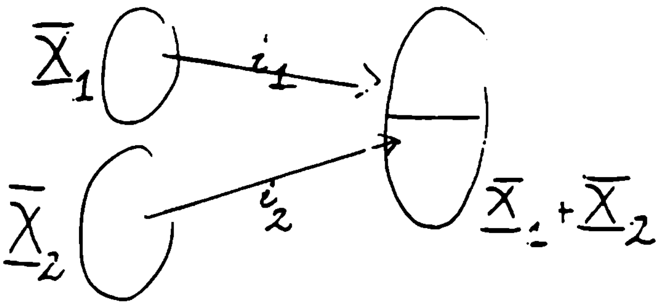
Hence $(A + B)C = AC + BC$ since above holds for all i.

Exercise Complete the proof by showing that

$D(A + C) = DA + DC$

Exercise The subcategories of systems defined by equations like $\alpha\beta = 1^{\nu}$, $d^2 = 0$, $C = B(1 - AC)$ are all <u>also</u> linear categories. Hint: The main point here is to understand that there is nothing to prove since these were all <u>full</u> subcategories.

Now we come to the second form of the condition that a category be "linear", which involves the construction of (cartesian) <u>products</u> and of <u>coproducts</u>. Let us start with the (<u>non</u> linear) example of <u>co</u>products $X_1 + X_2$ in the category $\mathbb{S}$ of sets and mappings; it is simply the ("disjoint") union of the two sets X_1, X_2 and hence comes with two <u>injection</u> mappings $X_k \xrightarrow{i_k} X_1 + X_2$, $k = 1,2$



where $i_k(X) = X$ considered as an element of the (disjoint) union X_1+X_2 for $x \in X_k$, $k = 1,2$. This pair of injection mappings has the following "universal mapping property" (UMP): if Y is any set and if $X_k \xrightarrow{f_k} Y$ are any two mappings then there exists a unique (single) mapping f for which $f_1 = f\iota_1$ and $f_2 = f\iota_2$ (compositions)

$$\begin{tikzcd}
X_1 \arrow[dr, "i_1"] \arrow[drr, bend left, "f_1"] & & \\
 & X_1 \oplus X_2 \arrow[r, dashed] & Y \\
X_2 \arrow[ur, "i_2"'] \arrow[urr, bend right, "f_2"'] & &
\end{tikzcd}$$

Namely, we can define f for any $x \in X_1 + X_2$ by cases as follows:

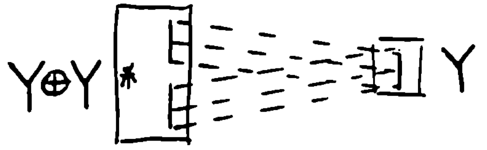
$$f(x) = \begin{cases} f_1(x) \text{ if x comes from } X_1 \\ f_2(x) \text{ if x comes from } X_2 \end{cases}$$

Since all x in $X_1 + X_2$ come from either X_1 or X_2, f(x) is thus defined for all x, and without ambiguity since no x comes from both (i.e. the union is taken in the "disjoint" sense - in fact no other sense *has* sense since we did *not* assume that X_1,X_2 lie inside any common larger set in which they could "overlap"). By construction f satisfies the two equations $f_k = f\iota_k$ for k = 1,2 and moreover these equations with f_1,f_2 given, *force* this construction of f_1 so it is "unique", subject to those equations, as well. Since f is determined by f_k's we may denote it by $f = (f_1,f_2)$.

As a very special case of the above, we could take $X_1 = X_2 = Y$, $f_1 = f_2 = 1_Y$. Then we get (as the resulting f) a standard map

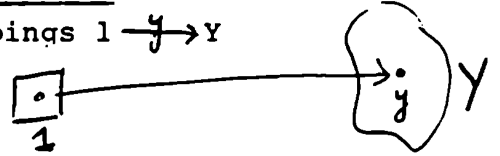
$$Y \oplus Y \xrightarrow{\Delta^*} Y$$

sometimes called the "codiagonal" map of Y: for each element y of Y, there are *exactly two* elements x of $Y\oplus Y$ for which $\Delta^*(x) = y$.



suggest the equation $Y \oplus Y = 2 \times Y$ which we will prove in a moment.

If 1 denotes any *one-element* set, then elements of a general set Y may be identified with mappings $1 \xrightarrow{y} Y$



and hence by the UMP of coproducts, mappings $2 \longrightarrow Y$ may be identified with *ordered pairs* of elements of Y: $y = (y_1,y_2)$

$$\begin{tikzcd}
1 \arrow[dr, "i_1"] \arrow[drr, bend left, "y_1"] & & \\
 & 2 \overset{def}{=} 1 \oplus 1 \arrow[r, dashed, "y"] & Y \\
1 \arrow[ur, "i_2"'] \arrow[urr, bend right, "y_2"'] & &
\end{tikzcd}$$

We will use the UMP as the definition of the concept of coproduct $\oplus$ with "injections" in any category; it will tend to have radically different concrete interpretations in each particular category.

The empty set $\emptyset$ is characterized by the fact that for any set Y, there is exactly one mapping $\emptyset \longrightarrow Y$ (called "the empty mapping with codomain Y"). In any category this "UMP" characterizes a particular object called the coterminal object; depending on the category, it may be far from "empty")

Proposition: In any category having coproducts $\oplus$ and coterminal object $\emptyset$, we have that $\emptyset \oplus X \cong X$

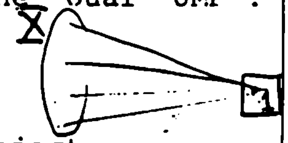
Proof: Consider any Y and any f_1, f_2

$$\begin{tikzcd}
\emptyset \arrow[dr, "e"] \arrow[drr, bend left, "f_1"] & & \\
& X \arrow[r, dashed] & Y \\
X \arrow[ur, "1_X"] \arrow[urr, bend right, "f_2"'] & &
\end{tikzcd}$$

then we can define $f = f_2$; $f_2 \cdot 1_X = f_2$ by the identity property, and $f_2 e = f_1$ since there is only one $\emptyset \longrightarrow Y$ no matter how we represent it. Thus the UMP for $\emptyset \oplus X$ is satisfied by X if we take the two injections to be the empty map e and the identity 1_X.

Now the one element set $\mathbb{1}$ is characterized by the "dual" UMP": for any set X there is exactly one mapping $X \longrightarrow 1$

In any category this



characterizes the kind of object known as a terminal object.

Exercise Any two terminal objects in the same category are isomorphic.

Similarly we can "dualize" the UMP for coproducts to obtain the important notion of (cartesian) product of two objects Y_1, Y_2 in a category: it should be an object $Y_1 \times Y_2$ equipped with "projections" p_1, p_i such that for any object X and any pair of morphisms $X \xrightarrow{f_1} Y_1$, $X \xrightarrow{f_2} Y_2$ there exists a unique morphism f for which both $f_k = p_k f$ $k = 1,2$

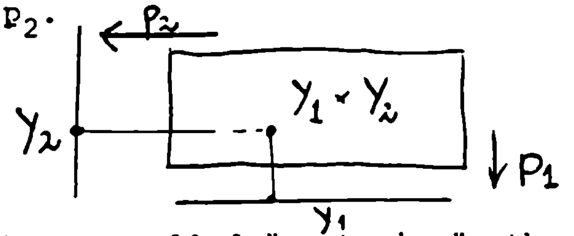
$$\begin{tikzcd}
& & Y_1 \\
X \arrow[r, dashed, "f"] \arrow[urr, bend left, "f_1"] \arrow[drr, bend right, "f_2"'] & Y_1 \times Y_2 \arrow[ur, "p_1"'] \arrow[dr, "p_2"] & \\
& & Y_2
\end{tikzcd}$$

Since f is uniquely determined by f_1, f_2, we may denote it $f = \begin{pmatrix} f_1 \\ f_2 \end{pmatrix}$

If there is an object E in our category (such as $E = \mathbb{1}$ in the category of sets) for which we consider morphisms $E \xrightarrow{x} X$ to be "elements" of X for all X, then the above UMP for the projections just says (taking X = E) that the S

elements of $Y_1 \times Y_2$ "are" ordered pairs of elements of Y_1, Y_2

That is, any element $E \xrightarrow{y} Y_1 \times Y_2$ of course defines an ordered pair of elements $E \xrightarrow{p_k y} Y_k$ $k = 1,2$, but the UMP says that conversely every such pair comes from a unique y. Thus we may picture $Y_1 \times Y_2$ as "rectangularized" by p_1, p_2.



showing why these products are called "cartesian", though the Y_1, Y_2 are arbitrary sets and are only very schematically shown above as lines

"Dual" to the isomorphism $\emptyset \oplus X \xrightarrow{\simeq} X$, we have

Exercise In any category with products x and a terminal object 1, $X \xrightarrow{\approx} 1 \times X$.

Exercise There is a diagonal morphism $X \xrightarrow{\Delta} X \times X$.

Exercise The meaning of $f = \begin{pmatrix} f_1 \\ f_2 \end{pmatrix}$ in general can be understood on arbitrary elements by using the diagram

$$\begin{tikzcd}
 & & Y_1 \\
E \arrow[r, "x"'] \arrow[urr, bend left, "f_1 x"] \arrow[drr, bend right, "f_2 x"'] & \bar{X} \arrow[r, dashed, "f"] \arrow[ur, bend left, "f_1"] \arrow[dr, bend right, "f_2"'] & Y_1 \times Y_2 \arrow[u, "p_1"'] \arrow[d, "p_2"] \\
 & & Y_2
\end{tikzcd}$$

and the uniqueness in the UMP to get

$$\begin{pmatrix} f_1 \\ f_2 \end{pmatrix}(x) = \begin{pmatrix} f_1 x \\ f_2 x \end{pmatrix}$$

Hint: Do both sides have the same composite with each projection?

Combining the UMP's of coproducts and products we obtain the following in any category that has both: Any morphism $X_1 \oplus X_2 \xrightarrow{f} Y_1 \times Y_2$ is uniquely represented by a "matrix" whose entries are themselves morphisms (between "smaller" objects).

$$\boxed{\begin{array}{l} f_{kj} = p_k f i_j \\ k,j = 1,2 \end{array}}$$

$$\begin{tikzcd}
\bar{X}_1 \arrow[dr, "\bar{i}_1"] & & & Y_1 \\
 & X_1 \oplus X_2 \arrow[r, "f"] & Y_1 \times Y_2 \arrow[ur, "p_1"] \arrow[dr, "p_2"'] & \\
\bar{X}_2 \arrow[ur, "\bar{i}_2"'] & & & Y_2
\end{tikzcd}$$

Caution: These matrices cannot in general be "multiplied".

In particular, there is a "canonical" morphism

$$\boxed{(V \times X_1) \oplus (V \times X_2) \xrightarrow{d} V \times (X_1 \oplus X_2)}$$

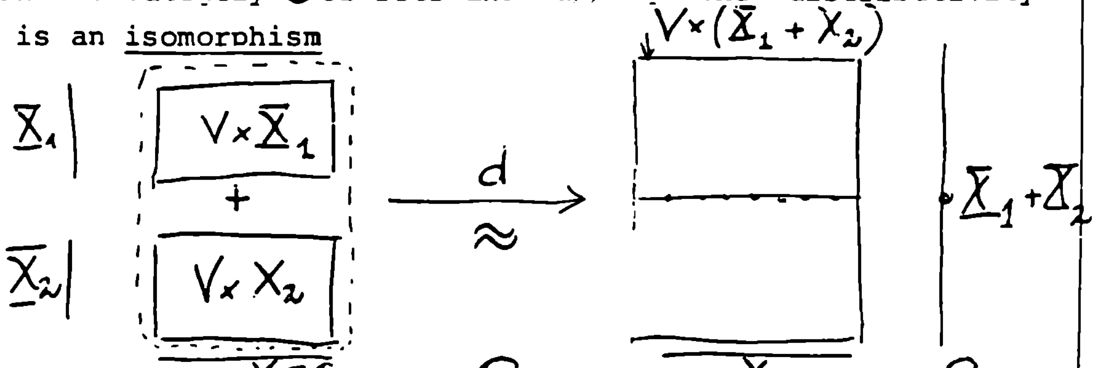
(for any three objects in any category having both coproducts and products) defined by

$$p_V d \bar{i}_K = \bar{p}_K \qquad k = 1,2$$

$$p_\oplus d \bar{i}_K = i_K q_K \qquad k = 1,2$$

where $V \times (X_1 \oplus X_2) \xrightarrow{p_V} V \times (X_1 \oplus X_2) \xrightarrow{p_{\oplus}} X_1 \oplus X_2$, $V \times X_k \xrightarrow{\bar{p}_k} V$, $V \times X_k \xrightarrow{q_k} V$ are projections, and $X_k \xrightarrow{i_k} X_1 \oplus X_2$, $V \times X_k \xrightarrow{\bar{i}_k} V \times X_1 \oplus V \times X_2$ are injections (draw a mapping diagram).

Exercise In the category $\mathcal{S}$ of sets and mappings the "distributivity" morphism d is an isomorphism



Exercise Any category $\mathcal{S}^{\mathcal{G}}$ where $\mathcal{G}$ is any given graph and $\mathcal{S}$ is the category of sets, satisfies distributivity of products over coproducts in the sense that all the canonical morphisms d are isomorphisms.

Hint: Coproducts and products not only exist in $\mathcal{S}^{\mathcal{G}}$ but may be compute in the naive manner, i.e. for each vertex i of $\mathcal{G}$ at a time.

Caution "Distributivity" (i.e. d isomorphism) will often not hold in subcategories $\mathcal{D}$ of $\mathcal{S}^{\mathcal{G}}$ because the meaning of coproducts may change to a highly non-naive interpretation; the "U" in UMP refers to all Y in $\mathcal{D}$ only, and moreover the $X_1 + X_2$ which works in $\mathcal{S}^{\mathcal{G}}$ may not lie in $\mathcal{D}$.

The distributive law d does not hold true in the category Lin of linear spaces and linear transformations. To understand this we must calculate what coproducts and products in Lin mean. For the case of products this is easy: if Y_1, Y_2 are linear spaces, then the set $Y_1 \times Y_2$ equipped with co-ordinate-wise addition and scalar multiplication is again a linear space, the projections p_k are linear transformations by construction, and if f_1, f_2 are any two linear transformations with a common linear space X as domain, then f defined by $fx = \begin{pmatrix} f_1x \\ f_2x \end{pmatrix}$ is again linear. But for coproducts something very different from the $\mathcal{S}$ case happens.

THEOREM For any two linear spaces V_1, V_2 there is a canonical linear isomorphism

$$\boxed{V_1 \oplus V_2 \xrightarrow{\sim} V_1 \times V_2}$$

PROOF: First we note that the single-element linear space 0 is both coterminal and terminal in Lin. This implies (we knew it already) that between any two linear spaces V,W, there is a zero linear transformation

$$\begin{tikzcd} & 0 \arrow[dr] & \\ V \arrow[ur] \arrow[rr, "0_{W,V}"'] & & W \end{tikzcd}$$

Since morphisms from coproducts to products are always defined by "matrices" of smaller morphism, we can in our case define a morphism by means of the "identity matrix" whose entries are

$$\begin{pmatrix} 1_{V_1} & 0_{V_2 V_1} \\ 0_{V_1 V_2} & 1_{V_2} \end{pmatrix}$$

We finish the proof by showing that the familiar $V_1 \times V_2$ satisfies <u>also</u> the UMP of <u>co</u>product in Lin if we define appropriate injections

$$V_k \xrightarrow{i_k} V_1 \times V_2 \qquad k = 1,2$$

obviously these have to be

$$i_1(v) = \begin{pmatrix} v \\ 0 \end{pmatrix} \quad \text{for any } v \in V_1$$

$$i_2(u) = \begin{pmatrix} 0 \\ u \end{pmatrix} \quad \text{for any } u \in V_2$$

Now to verify the UMP we have to consider any linear space W and any two linear transformations $V_k \xrightarrow{A_k} W$ $k = 1,2$; then we define

$$V_1 \times V_2 \xrightarrow{A} W$$

by using the formula

$$\boxed{A\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = A_1 v_1 + A_2 v_2}$$

Then note that

$$(A i_1)(v) = A\begin{pmatrix} v \\ 0 \end{pmatrix} = A_1 v \qquad \text{all } v \in V_1$$

$$(A i_2)(u) = A\begin{pmatrix} 0 \\ u \end{pmatrix} = A_2 u \qquad \text{all } u \in V_2$$

so that $\boxed{A i_k = A_k}$ $k=1,2$ as required, and on the other hand <u>since A is required to be a morphism in Lin</u>, i.e. to preserve sums in particular, the definition $\boxed{}$ is forced by the agreement with A_k on i_k, so that A is <u>unique</u>, as also required by the $\oplus$ UMP.

With a little more care, using the properties of the "identity matrix" above instead of elements u,v. etc, the above theorem is proved to hold in <u>any</u> linear category which has products (or coproducts, dualizing..). Either using that, or calculating directly, we get

Exercise $\mathbb{V}_1 \oplus \mathbb{V}_2 \xrightarrow{\sim} \mathbb{V}_1 \times \mathbb{V}_2$ in Lin^{G}.

Exercise In a <u>linear</u> category, block matrices can be multiplied (of course taking care of the order when "multiplying" (i.e. composing) the components

$$\begin{tikzcd} U_1 \oplus U_2 \arrow[r, "\begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}"'] & V_1 \times V_2 & V_1 \oplus V_2 \arrow[l, "\sim"'] \arrow[r, "\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}"'] & W_1 \times W_2 \end{tikzcd}$$

$$(AB)_{ik} = \sum^{2} A_{ij} B_{jk}$$

where

$$U_k \xrightarrow{B_{jk}} V_j \xrightarrow{A_{ij}} W_j$$

In practice, in a linear category we just identify $V_1 \times V_2 = V_1 \oplus V_2$. But it is well to keep in mind how different they are in a non-linear category, which at the opposite extreme often satisfies the "distributive law" for objects: that the canonical map d is an isomorphism. Since in a non-linear category there is no "addition" for morphisms in general, there should be no confusion between the distributive law for addition of linear transformations which holds in a linear category on the one hand, and the distributive law for "addition" (i.e. coproducts of objects in some typical non-linear categories like $\mathcal{S}$, $\mathcal{S}^{\downarrow}$, $\mathcal{S}^{\circlearrowleft}$, $\mathcal{S}^{\downarrow\uparrow}$; $\mathcal{S}$ on the other hand.

Exercise The canonical map $V \times X_1 \oplus V \times X_2 \xrightarrow{d} V \times (X_1 + X_2)$ is usually not an isomorphism in a linear category. Hint: Take $V = X_1 = X_2 = \mathbb{R}$ in Lin and compare the dimensions of the two sides of d.

Since the codiagonal map in a linear category is (up to the "identity" matrix) just the internal addition operation of the given object (e.g. linear space) W:

$$\boxed{a\begin{pmatrix}w_1\\ w_2\end{pmatrix} = w_1 + w_2 \qquad \begin{tikzcd} W \times W \arrow[rr, bend right, "a"'] & W \oplus W \arrow[l, "\sim"'] \arrow[r, "\Delta^*"] & W \end{tikzcd}}$$

we see that the addition of morphisms $V \underset{T_2}{\overset{T_1}{\rightrightarrows}} W$ can be reconstructed as the composite

$$(*) \quad \boxed{\begin{tikzcd} V \arrow[r, "\Delta"] \arrow[rrrr, bend right, "T_1 + T_2"'] & V \times V \arrow[r, "T_1 \times T_2"] & W \times W & W \oplus W \arrow[l, "\sim"'] \arrow[r, "\Delta^*"] & W \end{tikzcd}}$$

using only the fact that the "identity matrix" for W is an isomorphism. Since in turn the very existence of the "identity matrix" for all objects W depends on the fact that the unique morphism from the coterminal object to the terminal object is invertible (thus the coterminal object has a unique element (the inverse of the canonical map) which may be called 0), and since the concepts of coterminal and terminal objects, coproduct and product objects of pairs of objects, and codiagonal and diagonal maps ($\Delta(x) = \begin{pmatrix}x\\ x\end{pmatrix}$) of an object, depend only on the composition of morphisms in the given category, we see that our second characterization of "linear categories", namely

$$V \oplus W \xrightarrow{\sim} V \times W,$$

implies the first, namely the existence of a good notion of addition for morphisms, provided only that we prove the

Exercise In any category with coproducts, products etc. in which there are zero maps and $V\oplus W \xrightarrow{\sim} V\times W$ is always an isomorphism, <u>define</u> addition of morphisms by $\circledast$ above. Then this addition satisfies the commutativity and associative laws for $V \;\substack{\longrightarrow\\ \longrightarrow\\ \longrightarrow}\; W$ and the distributive laws for $U \longrightarrow V \rightrightarrows W \longrightarrow Y$.

Conversely in any category equipped with a good notion of adding morphisms, we can characterize the injections and projections of an object $V_1\oplus V_2$, which serves simultaneously as coproduct and product, by purely "algebraic" equations as follows

$$\boxed{\begin{aligned} p_2i_1 &= 0\\ p_1i_2 &= 0\\ p_2i_2 &= 1_{V_2}\\ p_1i_1 &= 1_{V_1}\\ i_1p_1 + i_2p_2 &= 1_{V_1\oplus V_2}\end{aligned}}$$

$$\begin{tikzcd} V_1 \arrow[dr, "i_1"] & & V_1 \\ & V_1\oplus V_2 \arrow[ur, "p_1"] \arrow[dr, "p_2"] & \\ V_2 \arrow[ur, "i_2"'] & & V_2 \end{tikzcd}$$

Exercise Recalling that $i_1(x) = \begin{pmatrix}x\\0\end{pmatrix}$, $i_2(y) = \begin{pmatrix}0\\y\end{pmatrix}$, $p_1\begin{pmatrix}x\\y\end{pmatrix} = x$, $p_2\begin{pmatrix}x\\y\end{pmatrix} = y$ on elements, prove the above equations for the morphisms. The equations imply that $e_k \underset{\text{def}}{=} i_kp_k$ is an <u>idempotent</u> endomorphism of X for $k=1,2$, and that $e_1 + e_2 = 1_X$, where $X = V_1\oplus V_2$, while $e_1e_2 = 0 = e_2e_1$. Such a pair of idempotent endomorphisms of an object X in a linear category (i.e. with products $= 0$ and sum $= 1_X$) is often called a "decomposition of the identity into disjoint idempotents" and plays an important role in all kinds of spectral analysis. We can recover the component spaces from a space X equipped with such a decomposition of the identity, since

$$\begin{aligned} V_1 &= \mathrm{Ker}(e_2)\\ V_2 &= \mathrm{Ker}(e_1)\end{aligned}$$

with i_1, i_2 as the inclusions, and that p_k is constructed by using the same rule as e_k but proving that all values of e_k actually lie in the subspace $V_k\subseteq X$.

Exercise In a linear category the coproduct and product universal mapping properties follow from the algebraic equations discussed above. For example if $V_k\xrightarrow{A_k}W$ are any two morphisms, we can <u>define</u> $V_1\oplus V_2\xrightarrow{A}W$ using as the sum $A = A_1p_1 + A_2p_2$ of two morphisms obtained by composing with the p_k's. Then the algebraic equations governing the p's and i's imply that the A so defined is the unique morphism satisfying both $A_1 = Ai_1$ and $A_2 = Ai_2$. Dually, if we are given $U\xrightarrow{B_k}V_k$ we can define a single B using addition and the i's, then use the governing equations to prove B is the unique morphism allowing recovery of the B_1, B_2 using the p's.

There is a process of construction $\mathcal{S} \xrightarrow{\mathbb{R}[\,]} \mathrm{Lin}$

as follows. For each set S, $\mathbb{R}[S]$ is the linear space of all those real valued functions on S which have finite support. (So that $\mathbb{R}[S] \hookrightarrow \mathbb{R}^S$ is actually an equality if S is finite). Moreover if $S \xrightarrow{f} T$ is any mapping, then a linear transformation

$$\mathbb{R}[S] \xrightarrow{\mathbb{R}[f]} \mathbb{R}[T]$$

is defined by

$$(\mathbb{R}[f](v))_t = \sum_{\substack{s\in S\\ f(s)=t}} v_s \qquad \begin{matrix} t\in T \\ v\in\mathbb{R}[S]\end{matrix}$$

Exercise The process $\mathbb{R}[\,]$ is "functorial", i.e. if $S \xrightarrow{f} T \xrightarrow{g} U$ are any mappings of sets, then

$$\mathbb{R}[gf] = \mathbb{R}[g]\,\mathbb{R}[f]$$

as linear transformations $\mathbb{R}[S] \longrightarrow \mathbb{R}[U]$.

Proposition Consider the mapping

$$S \xrightarrow{\varsigma} \mathbb{R}[S]$$

defined by $\varsigma(s)(s') = \begin{cases} 1 \text{ if } s = s' \\ 0 \text{ if } s \neq s' \end{cases}$

Then for any linear space W and any mapping $S \xrightarrow{f} W$, there is a unique linear transformation $\mathbb{R}[S] \xrightarrow{\bar{f}} W$ for which $\boxed{\bar{f}\cdot\varsigma = f}$

Proof: Define $\bar{f}(v) = \sum_{s\in S} v_s \cdot f(s)$

The sum is well-defined since v has finite support and clearly has the three stated properties.

DEFINITION A "family of vectors" $S \xrightarrow{f} W$ is

1) linearly independent iff its linear extension $\mathbb{R}[S] \xrightarrow{\bar{f}} W$ is injective
2) linearly spanning iff f is surjective
3) a linear basis iff f is an isomorphism

A linear space W is said to have dimension = S iff there exists a basis family f for V with domain S.

THEOREM ($\mathbb{R}$ is a field) Every W in Lin has a dimension. $\mathbb{R}g_1 = \mathbb{R}g_2 \Rightarrow g_1 = g_2$, for any mappings $T \overset{g_1}{\underset{g_2}{\rightrightarrows}} U$ in $\mathcal{S}$. If there exists an isomorphism $\mathbb{R}[S] \xrightarrow{\sim} \mathbb{R}[T]$ in Lin, there exists one induced by an isomorphism $S \xrightarrow{\sim} T$ in $\mathcal{S}$.

THEOREM (Rank) The above theorem remains almost true if $\mathcal{S}$ is replaced by $\mathcal{S}^{\downarrow}$ and Lin by $\mathrm{Lin}^{\downarrow}$. However this does not generalize to graphs $\mathcal{G}$ with loops.

CATEGORY	$\mathcal{S}$	LIN	LIN^{2}	$\mathrm{LIN}^{\mathbb{N}}$
Objects	sets	linear spaces	linear transformations	linear operators
Morphisms	mappings between sets	linear transformations between linear spaces	$\begin{tikzcd} \cdot \arrow[r, "T_0"] \arrow[d, "A"'] & \cdot \arrow[d, "B"] \\ \cdot \arrow[r, "T_1"'] & \cdot \end{tikzcd}$ $\boxed{T_1A = B T_0}$	$\begin{tikzcd} \cdot \arrow[loop left, "A"] \arrow[r] & \cdot \arrow[loop right, "B"] \end{tikzcd}$ $\boxed{T A = B T}$
Endomorphisms	self-maps on sets	linear operators on linear spaces	$\begin{tikzcd} \cdot \arrow[loop right, "T_0"] \arrow[d, "A"] \\ \cdot \arrow[loop right, "T_1"] \end{tikzcd}$ $\boxed{T_1A = A T_0}$	$\begin{tikzcd} \cdot \arrow[loop left, "A"] \arrow[loop right, "T"] \end{tikzcd}$ $\boxed{AT = TA}$ (Commuting pairs of linear operators)
Property of Objects invariant under isomorphism	Cardinality	$\boxed{\text{Dimension}}$	$\boxed{\text{RANK}}$ of linear transformations and dimensions of domain and codomain	$\boxed{\text{Spectrum}}$ including multiplicities of linear operators
"Elements" represented by	1	R	$\begin{tikzcd} R \arrow[d, "1"] & O \arrow[d] \\ R & R \end{tikzcd}$	$R[t]$

In these
four
diagrams,
← the dots
are
arbitrary
linear
spaces an
the arrow
← are
arbitrary
linear
trans-
formation

Rank of a Linear Transformation

The spectral analysis of the diagram $\begin{smallmatrix}\bullet\\ \downarrow\\ \bullet\end{smallmatrix}$ is much simpler than that of the diagram $\begin{tikzcd}\bullet \arrow[loop right]\end{tikzcd}$

The latter involves the continuous infinity of eigenvalues as "colors", whereas the former involves only a whole number called "rank", at least when the ground rig of scalars is a field. This is basically because the maps in the category Lin $\begin{smallmatrix}\bullet\\ \downarrow\\ \bullet\end{smallmatrix}$ (whose <u>objects</u> are linear transformations A,B), are arbitrary commutative squares

$$\begin{tikzcd}\bullet \arrow[r, "T_0"] \arrow[d, "A"'] & \bullet \arrow[d, "B"] \\ \bullet \arrow[r, "T_1"'] & \bullet\end{tikzcd} \qquad T_1A = BT_0;$$

even if A,B happen to be endomaps (linear "operators") we do not require that the T_0 and T_1 be the same; this means in particular that there are "many" isomorphisms in our category, hence "few" invariants. In fact, besides the dimensions of the domain and codomain of A , the only invariant is the dimension of the <u>image</u> of A (assuming that Lin itself is the category of linear spaces over a <u>field</u> of scalars).

Exercise 1) Given any three whole numbers $\langle n, r, m\rangle$ for which $r \leq \min(n,m)$, construct a map $n \longrightarrow m$ of finite sets of the indicated cardinalities whose image has cardinality r, and show that the induced linear map $\begin{tikzcd} R^n \arrow[d, "B"'] \\ R^m \end{tikzcd}$ has image of dimension r, in fact factorize it into epic and monic linear maps

$$\begin{tikzcd} R^n \arrow[dd, "B"'] \arrow[dr, two heads] & \\ & R^r \arrow[dl, tail] \\ R^m & \end{tikzcd}$$

Exercise 2) If R is a field, show that for any linear transformation $\begin{tikzcd} V \arrow[d, "A"] \\ W \end{tikzcd}$ with finite dimensional V, W

there exists a unique triple $\langle n,r,m\rangle$ for which A is isomorphic in $\mathrm{Lin}^{\downarrow}_{R}$ to your example B above.

Exercise 3) Knowing the invariant $\langle n,r,m\rangle$ of a linear transformation A we can detect two properties which A may have: show that

"A is epi " iff $r = m \leq n$

"A is mono" iff $r = n \leq m$

(just as for finite sets)

Exercise 4) The dimension of the kernel of A is $n - r$, (the "nullity" of A ; the other difference $m - r$ might be called the "co-nullity").

Spectral Analysis

Recall that Newton analyzed sunlight (by passing it through a prism) into a "Spectrum" consisting of various colors λ each of which has a certain intensity m_λ. A basic program for analyzing an object $\mathbb{V}$ in a linear category such as $\mathrm{Lin}^{\mathbb{G}}$ is to show that it is isomorphic to a big sum (in the $\oplus$ sense) of relatively simple objects Λ (which Λ themselves cannot be further broken down as a ("parallel connection") $\oplus$ of still simpler objects; the number m_Λ of summands of each given type Λ in the decompositon of $\mathbb{V}$ is usually called the multiplicity(rather than intensity) of the "color" Λ in $\mathbb{V}$.

In each linear category, we have to determine first what the irreducible "color" Λ should be (see below for some of the simpler examples) and then to analyze an "arbitrary" object $\mathbb{V}$, we study for each Λ the "eigenvectors" of type Λ in $\mathbb{V}$ in the sense of

DEFINITION: In a given category, an eigenvector of type Λ in $\mathbb{V}$ is just any morphism (in the category in question) $\Lambda \longrightarrow \mathbb{V}$. That is, an eigenvector is really nothing but a morphism in the appropriate category, except that we imagine that the domain is some kind of extremely special object (the "eigenvalue" in question should be thought of as the special information required to specify Λ in particular among all possible objects). We say that Λ is an eigenobject for $\mathbb{V}$ if there exists an injective eigenvector $\Lambda \hookrightarrow \mathbb{V}$; thus an "eigenvalue" of $\mathbb{V}$ would be special information determining an eigenobject of $\mathbb{V}$, i.e. one which will actually occur at least once in the expression of $\mathbb{V}$ as a $\oplus$ of "colors" (in those categories where the program succeeds.

Linguistic Remark: Terms like "eigenvalue", etc. are hybrids as follows
German: Eigenwerte
British: Characteristic Value

The hybrid is widely used by American chemists, engineers, mathematicians, economists, etc. perhaps partly because it helps discriminate from other kinds of values that might be "characteristic" of a given situation.

For example, if we consider Lin itself, then the basis theorem says that every $\mathbb{V}$ is isomorphic to some $\mathbb{R}^n$ which is actually $\mathbb{R}\oplus\mathbb{R}\oplus\ldots\oplus\mathbb{R}$(n terms). Thus we could say that Lin is "monochromatic" with $\mathbb{R}$ as the only "color", which can occur with various "intensity" (=dimension) as we consider different objects $\mathbb{V}$; in fact knowing the one intensity is equivalent, up to isomorphism, to finding the object in the case of Lin. There is no need for "eigenvalues" since no information is required to distinguish between colors since there is only one.

The category of central importance in this connection is Lin$^{\mathbb{G}}$ where $\mathbb{G} = \boxed{\circlearrowleft}$ is a single loop, i.e. the category whose objects are (single) linear operators, however, since the answer will be simpler we will consider first the case $\mathbb{G} = \left(\begin{smallmatrix}\bullet\\ \downarrow \\ \bullet\end{smallmatrix}\right)$, i.e. the category Lin$^{\begin{smallmatrix}\bullet\\ \downarrow \\ \bullet\end{smallmatrix}}$ whose objects are arbitrary linear transformations (whose domain and codomain dimensions may differ) which are to be morphically compared by commutative squares involving two comparison arrows from Lin (since $\mathbb{G}$ has two vertices). Any object in this category is actually isomorphic in this category to one of the form

$$\begin{tikzcd} \mathbb{R}^m \arrow[d] \\ \mathbb{R}^n \end{tikzcd} \begin{pmatrix} 1 & & & & & \\ & 1 & & & \text{\Large 0} & \\ & & 1 & & & \\ & & & 0 & & \\ & \text{\Large 0} & & & 0 & \\ & & & & & 0 \end{pmatrix}$$

where the number of 1's is the (isomorphism-invariant) rank, which is of course $\leq$ min(n,m)

Exercise: There exist no isomorphisms L

$$\begin{tikzcd} \mathbb{R}^4 \arrow[r, "L_0", "\sim"'] \arrow[d, "\begin{pmatrix} 1&0&0&0 \\ 0&1&0&0 \\ 0&0&0&0 \end{pmatrix}"'] & \mathbb{R}^4 \arrow[d, "\begin{pmatrix} 1&0&0&0 \\ 0&1&0&0 \\ 0&0&1&0 \end{pmatrix}"] \\ \mathbb{R}^3 \arrow[r, "\sim", "L_1"'] & \mathbb{R}^3 \end{tikzcd}$$

in Lin$^{\begin{smallmatrix}\bullet\\ \downarrow \\ \bullet\end{smallmatrix}}$.

Recall that $\oplus$ in Lin$^{\mathbb{G}}$, for example Lin$^{\circlearrowleft}$ is calculated "vertex by vertex" so that

$$\left(\begin{tikzcd} V_0 \arrow[d, "T"'] \\ V_1 \end{tikzcd}\right) \oplus \left(\begin{tikzcd} W_0 \arrow[d, "S"] \\ W_1 \end{tikzcd}\right) = \left(\begin{tikzcd} V_0 \oplus W_0 \arrow[d, "T\oplus S"] \\ V_1 \oplus W_1 \end{tikzcd}\right)$$

with the internal structural map being just "parallel connection, no interaction":

$$(T\oplus S)\begin{pmatrix} v \\ w \end{pmatrix} = \begin{pmatrix} Tv \\ Tw \end{pmatrix} \text{ in } V_1 \oplus W_1 \text{ for } \begin{pmatrix} v \\ w \end{pmatrix} \in V_0 \oplus W_0.$$

As in arithmetic we use natural numbers as coefficients when large sums with repeated summands are involved, for example

$$2\mathbb{V} \oplus 3\mathbb{W} = \mathbb{V} \oplus \mathbb{V} \oplus \mathbb{W} \oplus \mathbb{W} \oplus \mathbb{W}$$

ⓐ

$$\begin{tikzcd} \mathbb{R} \arrow[dd, "0"] & & 0 \arrow[dd, "0"] & & \mathbb{R} \arrow[dd, "0"] \\ & \oplus & & = & \\ 0 & & \mathbb{R} & & \mathbb{R} \end{tikzcd}$$

ⓑ

$$\begin{tikzcd} & \mathbb{R} \arrow[dd] & & \mathbb{R} \arrow[dd, "1"] & & 0 \arrow[dd] & & V_0 \arrow[dd, "T"] \\ 3 & & \oplus 2 & & \oplus 7 & & = & \\ & 0 & & \mathbb{R} & & \mathbb{R} & & V_1 \end{tikzcd}$$

where $\dim(V_0) = 5$, $\dim(V_1) = 9$, $\operatorname{rank}(T) = 2$.

ⓒ

$$\begin{tikzcd} \mathbb{R} \arrow[dd, "\lambda"'] & & \mathbb{R} \arrow[dd, "\mu"] \\ & \cong & \\ \mathbb{R} & & \mathbb{R} \end{tikzcd} \iff$$

both λ, μ are <u>non</u>-zero

<u>or</u>

both $\lambda = \mu = 0$

Thus if we denote the <u>three "colors"</u> for Lin as follows

$$\begin{tikzcd} 0 \arrow[d] & \mathbb{R} \arrow[d, "1"] & \mathbb{R} \arrow[d] \\ \mathbb{R} \arrow[d, equal] & \mathbb{R} \arrow[d, equal] & 0 \arrow[d, equal] \\ \mathbb{D}^* & \mathbb{D} & \mathbb{K} \end{tikzcd}$$

then every object of Lin$^{\downarrow}$ is isomorphic to a unique object of the form

$$\begin{tikzcd} V_0 \arrow[dd, "T"'] & \\ & \cong m_0 \cdot \mathbb{K} \oplus r \cdot \mathbb{D} \oplus m_1 \cdot \mathbb{D}^* \\ V_1 & \end{tikzcd}$$

where

$m_0 = \dim(V_0) - \operatorname{rank}(T)$

$r = \operatorname{rank}(T)$

$m_1 = \dim(V_1) - \operatorname{rank}(T)$

EXERCISE: The three colors of $\mathrm{Lin}^{\dot\downarrow}$ are not completely independent since there are morphisms $\mathbb{D}^* \hookrightarrow \mathbb{D} \longrightarrow \mathbb{K}$ where the first is injective (a monomorphism which is however <u>not</u> a <u>split</u> monomorphism) and the second is surjective (an epimorphism which is however <u>not</u> a <u>split</u> epimorphism) and the composite of the two in 0.

EXERCISE: If

$$0 \longrightarrow U \longrightarrow V \longrightarrow W \longrightarrow 0$$

is an exact sequence in any $\mathrm{Lin}^{\mathbb{G}}$ and if either $U \to W$ is <u>split</u> (in $\mathrm{Lin}^{\mathbb{G}}$) as a monomorphism or $V \to W$ is split as an epimorphism, then $W \oplus U \xrightarrow{\sim} V$ (Hint: use the algebraic equations which characterize $\oplus$ in any linear category). This shows "why" neither morphism in $\mathbb{D}^* \hookrightarrow \mathbb{D} \twoheadrightarrow \mathbb{K}$ splits, since $\mathbb{K} \oplus \mathbb{D}^*$ is <u>not</u> isomorphic to $\mathbb{D}$.

EXERCISE: If $\mathbb{V} = \begin{matrix} V_0 \\ {\scriptstyle T}\downarrow \\ V_1 \end{matrix}$ is any object of $\mathrm{Lin}^{\dot\downarrow}$, then in it an eigenvector of type $\mathbb{D}$ is essentially any vector in V_0, an eigenvector of type $\mathbb{D}^*$ is essentially any vector in V_1, $\mathbb{K} \to \mathbb{V}$ is essentially any vector in the <u>kernel of T</u>. Thus T is <u>injective</u> iff $\mathbb{K}$ is <u>not</u> an "eigenvalue" of $\mathbb{V}$. If we consider any

$$\mathbb{D} \twoheadrightarrow \mathbb{K} \xrightarrow{v} \mathbb{V}$$

then the composite "is" v, but considered as a vector of V_0, rather than (as given) as a vector of Ker(T). If we consider any $v \in V_0$ then in

$$\mathbb{D}^* \hookrightarrow \mathbb{D} \xrightarrow{v} \mathbb{V}$$

the composite "is" $\boxed{Tv}$ in V_1. (All statements in this exercise are essentially obvious if one draws the rectangular diagrams in Lin which they describe).

EXERCISE: In $\mathrm{Lin}^{\dot\downarrow}$, the three endomorphism rings (i.e. consisting of endomorphisms <u>in the sense of $\mathrm{Lin}^{\downarrow}$</u>) of all three distinct colors $\mathbb{D}^*, \mathbb{D}, \mathbb{K}$ are actually isomorphic to the same ring, namely $\mathbb{R}$ itself.

The most striking feature of Lin^{2} from the spectral point of view is that there is a continuous infinity of possible "colors" Λ (the ones which are usually called ordinary possible eigenvalues) and which have <u>one</u>-dimensional underlying spaces) as well as a further necessary infinite family of colors which is partly continuous and partly discrete (these are connected with nilpotency phenomena and have all possible dimensions, although they are much simpler than the general operator (object of Lin^{2}) which we need to analyze) and finally another continuous family of <u>two</u>-dimensional "colors" which are usually referred to as "complex".

Let us consider first an example of the last-mentioned complex "color" $\mathbb{C}$: It is nothing but the rotation through a right angle in two-dimensional space, considered as an object of $\mathrm{Lin}^{\circlearrowleft}$:

$$\mathbb{C} = \boxed{\mathbb{R}^2 \circlearrowleft \begin{pmatrix}0&-1\\1&0\end{pmatrix}} \qquad \begin{pmatrix}0&-1\\1&0\end{pmatrix}$$

EXERCISE: $\mathbb{C}$ cannot be decomposed in $\mathrm{Lin}^{\circlearrowleft}$. Since a non-trivial decomposition would have to involve one-dimensional summands, it must be shown that there is no isomorphism

$$^{a}\circlearrowright\mathbb{R} \oplus \mathbb{R}\circlearrowleft^{b} \longrightarrow \mathbb{R}^2 \circlearrowleft \begin{pmatrix}0&-1\\1&0\end{pmatrix}$$

i.e. that the indicated equations for a 2 x 2 matrix α and a pair of 1 x 1 matrices a,b has no solution with α invertible, i.e. that there is no choice of basis for $\mathbb{R}^2$ with respect which the right-angle rotation is expressed by a matrix of the form $\begin{pmatrix}a&0\\0&b\end{pmatrix}$ (which intuitively would just stretch $\mathbb{R}^2$ without rotating it). Determine all $\mathrm{Lin}^{\circlearrowleft}$ - endomorphisms of $\mathbb{C}$, i.e. all 2 x 2 matrices A for which

$$A\begin{pmatrix}0&1\\-1&0\end{pmatrix} = \begin{pmatrix}0&1\\-1&0\end{pmatrix}A$$

If $V \circlearrowleft^{T}$ is any object of $\mathrm{Lin}^{\circlearrowleft}$, then in it an "eigenvector" of type $\mathbb{C}$ is any choice of family of two vectors in V for which $Tv_1 = -v_2, Tv_2 = -v_1$, i.e. to a choice of a single vector v_1 for which $T^2 v_2 = -v_1$. $\mathbb{C}$ is an "eigenvalue" of $V \circlearrowleft^{T}$ iff such a family of two which is linearly independent exists, i.e. iff there exists a plane in V which is closed under the operation T and on which T looks like (up to $\mathbb{R}$-linear isomorphism) rotation through a right angle. What does linear indenendence of such a family of two mean in terms of the single v which generates it?

Now we notice that all operators on one-dimensional space must be considered as distinct possible colors or eigenvalues in analyzing objects in $\mathrm{Lin}^{\circlearrowleft}$.

EXERCISE: If $\mathbb{R}\circlearrowleft^{\lambda}$, $\mathbb{R}\circlearrowleft^{\mu}$ are any two real numbers for which there exists an isomorphism $^{\lambda}\circlearrowright\mathbb{R} \longrightarrow \mathbb{R}\circlearrowleft^{\mu}$ of $\mathrm{Lin}^{\circlearrowleft}$, then $\lambda = \mu$. If on an arbitrary linear space V of dimension n, we consider the very special operator of multiplication by λ , then there exist isomorphisms

$$^{\lambda}\circlearrowright\mathbb{R}^n \xrightarrow[\alpha]{\sim} V\circlearrowleft^{\lambda}$$

in $\mathrm{Lin}^{\circlearrowleft}$ (In fact in this very special case, any isomorphism α in Lin will be one in $\mathrm{Lin}^{\circlearrowleft}$). Of course

$$\lambda = \begin{pmatrix}\lambda & 0\\ 0 & \lambda\end{pmatrix}$$ when operating on $\mathbb{R}^n$.

So $\boxed{^{\lambda}\circlearrowright V \cong n\cdot\mathbb{R}^{\circlearrowleft}}$, a $\oplus$ of n copies. Thus our object has only one eigenvalue λ , but of multiplicity n. But we could have, for example, on an n + m dimensional space just two eigenvalues

$$n\cdot\mathbb{R}^{\circlearrowleft\lambda}\oplus m\cdot\mathbb{R}^{\circlearrowleft\mu}\cong\mathbb{R}^{n+m}\circlearrowleft\begin{pmatrix}\lambda & & & & \bigcirc\\ & \lambda & & & \\ & & \ddots & & \\ & & & \mu & \\ \bigcirc & & & & \mu\mu\end{pmatrix}$$

where the matrix has n λ's and m μ's on the diagonal.

EXERCISE: An eigenvector of type ${}^{\lambda}\circlearrowright\mathbb{R}$ in $V^{\circlearrowleft T}$ is essentially just a vector v of V for which $\boxed{Tv=\lambda v}$ In particular an eigenvector in $V^{\circlearrowleft T}$ of type $\circlearrowright\mathbb{R}$ is just a vector in Ker(T) and indeed an eigenvector of type ${}^{\lambda}\circlearrowright\mathbb{R}$ may also be considered as a vector in the kernel of the operator $T-\lambda$ (which leads to a method for calculating which eigenvalues λ actually occur in $V^{\circlearrowleft T}$, namely by solving the n-th degree polynomial equation det($T-\lambda$) = 0; the latter is often called the "characteristic equation" of T, but to be consistent in language-mixing one should call it the "eigenequation of T ".

EXERCISE: Is the complex number i isomorphic to the complex number -1 in the category Lin$^{\circlearrowleft}$ of real linear operators? In other words, is there any invertible solution a to the matrix equation

$$\begin{pmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{pmatrix}\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}\underset{?}{=}\begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix}\begin{pmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{pmatrix}$$

As the last exercise suggests, it is not always the most effective procedure to insist on the spectral philosophy in its purest form but to choose for example a class of identify-able objects which includes all the colors but which may contain many pairs of "different" objects which are actually isomorphic. For example any nilpotent operator in Lin$^{\circlearrowleft}$ is isomorphic to a strictly upper-triangular matrix

$$\begin{pmatrix}0 & & c & c & c & & \\ & 0 & & c & c & c & c\\ & & 0 & c & c & c & c\\ & & & 0 & c & c & c\\ \bigcirc & & & & 0 & & \\ & & & & & 0 & \\ & & & & & & 0\end{pmatrix}$$

where there are zeroes below and on the main diagonal, but a more precise analysis would show that any such is in turn isomorphic to one having only 1's (and 0's) above the diagonal, with the precise arrangement of 1's reflecting the way in which the various Lin -$\oplus$- ands are either coupled through T or actually split off as Lin$^{\circlearrowleft}$ -$\oplus$- ands.

EXERCISE: An eigenvector of type $\begin{pmatrix}0&1\\0&0\end{pmatrix}\circlearrowright\mathbb{R}^2$ in $\mathcal{V}\circlearrowleft T$ is essentially a vector v of $\mathcal{V}$ for which $T^2v = 0$. The object $\begin{pmatrix}0&1\\0&0\end{pmatrix}$ is an eigenvalue in $\mathcal{V}\circlearrowleft T$ iff such a v exists for which the family v, Tv of two vectors in $\mathcal{V}$ is linearly independent.

EXERCISE: The object $\begin{pmatrix}0&1\\0&0\end{pmatrix}$ of $\mathrm{Lin}^{\circlearrowleft}$ has to be accepted as a "color" in its own right, since it cannot be represented as a $\oplus$ of two one-dimensional objects in $\mathrm{Lin}^{\circlearrowleft}$. On the other hand "colors" are again not completely unrelated, as there exists an exact sequence

$$0\circlearrowright\mathbb{R}\hookrightarrow\mathbb{R}^2\circlearrowleft\begin{pmatrix}0&1\\0&0\end{pmatrix}\twoheadrightarrow\mathbb{R}\circlearrowleft 0$$

in $\mathrm{Lin}^{\circlearrowleft}$ (which however by the first part of the exercise does not split.)

EXERCISE: Determine explicitly which 2 x 2 matrices are isomorphic in $\mathrm{Lin}^{\circlearrowleft}_{\mathbb{R}}$ to an upper-triangular matrix, and that any upper-triangular matrix $\begin{pmatrix}\lambda&b\\0&\mu\end{pmatrix}$ is either a color (i.e. $\oplus$ -irreducible) in its own right (if $b \neq 0$) or is isomorphic to the $\oplus$ of two one-dimensional eigenvalues, i.e. to $\begin{pmatrix}\lambda&0\\0&\mu\end{pmatrix}$. Hint: If $\begin{pmatrix}\alpha&\beta\\\delta&\varkappa\end{pmatrix}$ is the given arbitrary matrix, then since we know that determinants, trace, and eigenvalues remain unchanged under isomorphism, the λ, μ must be determined by the two equations

$$\lambda\mu' = \alpha\varkappa - \beta\delta$$
$$\lambda+\mu = \alpha+\varkappa$$

Knowing this, isomorphisms can be found. Or alternatively, an undetermined putative isomorphism can be considered

$$\mathbb{R}^2\xrightarrow[L]{}\mathbb{R}^2\circlearrowleft\begin{pmatrix}\alpha&\beta\\\delta&\varkappa\end{pmatrix}$$

then the one equation stating that

$$L^{-1}\begin{pmatrix}\alpha&\beta\\\delta&\varkappa\end{pmatrix}L \text{ has lower left corner } 0$$

can be shown to have a solution with non-zero determinant, provided a suitable discriminant involving trace and determinant is positive.

EXERCISE $\begin{pmatrix}\lambda&b_1\\0&\mu\end{pmatrix}\cong\begin{pmatrix}\lambda&b_2\\0&\mu\end{pmatrix}$

if both b_1, b_2 are non-zero, but if $b_1 = 1$, $b_2 = 0$ they are not isomorphic.

EXERCISE: Find a simple example of a 2 x 2 matrix A which is not upper-triangular but which is isomorphic LT = AL to an upper triangular T by means of an (invertible) lower triangular L. Show that this could not possibly be done with an upper-triangular L, since LTL^{-1} is upper-triangular whenever L and T are. Show that the lower-triangular L can be taken to be in the manifestly invertible form $1 - H$ where H is nilpotent, or a composition of such.

Some useful properties of operators (even of a "spectral" nature) can be proved more easily without explicit spectral analysis

EXERCISE: If we have an exact sequence

$$\begin{tikzcd} U \arrow[loop above, "a"] \arrow[r, hook, "i"'] & V \arrow[loop above, "T"] \arrow[r, "p"'] & W \arrow[loop above, "b"] \end{tikzcd}$$

in $\mathrm{Lin}^{\circlearrowleft}$ and if $a^2 = 0$, $b^2 = 0$, then $T^4 = 0$. (Though not necessarily $T^2 = 0$). Use the fact that $U \oplus W \xrightarrow{\approx} V$ in Lin (though not in $\mathrm{Lin}^{\circlearrowleft}$) to show first that

$$T = \begin{pmatrix} a & \alpha \\ 0 & b \end{pmatrix}$$

where $W \xrightarrow{\alpha} U$ is an arbitrary linear transformation.

The spectral approach is extremely important for solving linear differential equations. We will consider instead an example of linear difference equations, which are both of importance in their own right and as a method of computerized solution of differential equations, as well as analogous in many respects. Whereas for (ordinary) differential equations we would consider the operator $\frac{d}{dt}$ on an infinite-dimensional space of smooth functions, for difference equations we may consider a different object of $\mathrm{Lin}^{\circlearrowleft}$ as follows: Let $\mathbb{R}^{\mathbb{N}}$ denote the infinite-dimensional space of sequences of real numbers, so that a typical vector x in $\mathbb{R}^{\mathbb{N}}$ is an arbitrary infinite sequance $x_0, x_1, x_2, \ldots$ of real numbers, which $(x + y)_n = x_n + y_n$, etc. for all n. On this space we consider the shift operator S

$$\mathbb{R}^{\mathbb{N}} \circlearrowleft S \qquad (Sx)_n = x_{n+1} \text{ all } n = 0, 1, 2, \text{ all } x \in \mathbb{R}^{\mathbb{N}}$$

which assigns to each sequence the new one obtained by shifting; S is clearly $\mathbb{R}$-linear. In order to consider "second-order" equations in particular, we consider the $\mathbb{R}$-linear map

$$\mathbb{R}^{\mathbb{N}} \xrightarrow{\pi} \mathbb{R}^2 \qquad \pi(x) = \begin{pmatrix} x_0 \\ x_1 \end{pmatrix}$$

We want to relate this to certain operators on $\mathbb{R}^2$, but not in such a way that π is a $\mathrm{Lin}^{\circlearrowleft}$ morphism. Rather it will be of interest to consider operators T on $\mathbb{R}^2$ together with $\mathrm{Lin}^{\circlearrowleft}$-morphisms in the direction opposite to $\widetilde{\pi}$

$$\begin{tikzcd} \mathbb{R}^{\mathbb{N}} \arrow[loop left, "S"] & \mathbb{R}^2 \arrow[l, "G"'] \arrow[loop right, "T"] \end{tikzcd} \qquad \boxed{GT = SG}$$

i.e. G is an "eigenvector" of S with "eigenvalue" T , which moreover satisfies the <u>initial condition</u>

$$\boxed{\widetilde{\pi} \circ G = \text{identity on } \mathbb{R}^2}$$

For example consider the equation studied by Fibonacci (=Leonardo of Pisa 1250, who participated in mathematical contests sponsored by the Emperor Frederick II of Sicily):

$$x_{n+2} = x_{n+1} + x_n .$$

It is clear that if we start with any given pair of initial values x_0, x_1, then recursively applying Fibonacci's equation, we generate a uniquely determined complete sequence x. Let G be the mapping from pairs to sequences thus generated; by construction we have $\widetilde{\pi} \circ G$=identity, and G may be referred to as the solution operator for Fibonacci's equation. However, when we speak of "solving" an equation we usually intend that the solution be expressed in terms of generally-understood mathematical operations; thus if we started with say $x_0 = 1$, $x_1 = 1$, but need to know the 1013th value of the resulting sequence, can we calculate it by means of operations known from high-school algebra without actually going through the 1011 steps suggested by the equation itself? Note that the equation can be expressed in terms of our shift operator S as

$$S^2 x = Sx + x$$

or

$$(S^2 - S - 1)\, x = 0$$

where the combination in parentheses is a single new operator on $\mathbb{R}^{\mathbb{N}}$. Centuries of experience with difference equations suggest looking for special solutions (analogous to the special solutions $e^{\lambda t}$ appropriate for differential equations)in the form of "geometrical progressions":

$$x_n = \lambda^n \qquad \text{all } n = 0, 1, 2, \ldots$$

where λ is to be determined. In fact if we substitute this into the equation we find that λ must be one of the two numbers

$$\lambda_\pm = \frac{1 \pm \sqrt{5}}{2}$$

of the "golden mean". Thus there are two special solutions

$$\left(\frac{1 + \sqrt{5}}{2}\right)^n , \quad \left(\frac{1 - \sqrt{5}}{2}\right)^n$$

of Fibonacci's equation, and we will see that every solution can be <u>uniquely expressed as a linear combination of these two</u>

$$x_n = a\left(\frac{1+\sqrt{5}}{2}\right)^n + b\left(\frac{1-\sqrt{5}}{2}\right)^n$$

where the coefficients a, b can be determined from the initial values a, b.

EXERCISE: Given x_0, x_1 solve the equation $a + b = x_0$

$$a\lambda_+ + b\lambda_- = x_1 \quad \text{for } a,b$$

Thus the solutions constitute a two-dimensional subspace of the infinite-dimensional space $\mathbb{R}^{\mathbb{N}}$, with G as the inclusion. (A different second-order equation from Fibonacci's would involve a different pair of eigenvalues and a different G). Now notice that if x is a solution of Fibonacci's equation

$$S^2x = Sx + x$$

then its shift y = Sx is another solution:

$$S^2y = S^3x = S(Sx+x) = S(Sx)+(Sx) = Sy+y$$

In other words the inclusion $\mathbb{R}^2 \xrightarrow{G} \mathbb{R}^{\mathbb{N}}$ is actually a sub-object in $\mathrm{Lin}^{\circlearrowleft}$ (not only in Lin). This means that G is a $\mathrm{Lin}^{\circlearrowleft}$-morphism, if we only make explicit the operator T on $\mathbb{R}^2$ for which $\boxed{GT = SG}$

Since

$$G\begin{pmatrix}x_0\\x_1\end{pmatrix} = \begin{pmatrix}x_0\\x_1\\x_0+x_1\\x_0+2x_1\\2x_0+3x_1\\3x_0+5x_1\\5x_0+8x_1\\\vdots\end{pmatrix}$$

we have $SG\begin{pmatrix}1\\0\end{pmatrix} = \begin{pmatrix}0\\1\\1\\2\\3\\5\\\vdots\end{pmatrix}$ but $GT\begin{pmatrix}1\\0\end{pmatrix} = \begin{pmatrix}t_{11}\\t_{21}\\t_{11}+t_{21}\end{pmatrix}$, hence $t_{11} = 0$, $t_{21} = 1$.

on the other hand $SG\begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}1\\1\\2\\\vdots\end{pmatrix}$ but $GT\begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}t_{12}\\t_{22}\\t_{12}+t_{21}\end{pmatrix}$, hence $t_{12} = 1$, $t_{22} = 1$

Thus it appears that

$$T = \begin{pmatrix}0 & 1\\1 & 1\end{pmatrix}$$

EXERCISE: The statement that all values of G satisfy Fibonacci's equation is equivalent to the equation $S^2G = SG+G$ for linear transformation. If $GT=SG$, then using the latter twice we get

$GT^2 = GT+G$, or $G(T^2-T-1)=0$. Since also $\pi G = 1$, we get $T^2-T-1=0$ or $T^2 = T+1$.

EXERCISE: Verify that for the above T, we have indeed that $GT = SG$ where G is the map obtained by interating the Fibonacci equation with given initial values, and also where G is expressed in the "closed form" with powers of $\lambda_{\pm} = \frac{1 \pm \sqrt{5}}{2}$ and coefficients a,b obtained from x_o, x_1 with help of $\frac{1}{\sqrt{5}}$.

The initial values emphasized by Fibonacci were whole numbers $x_o = 1$, $x_1 = 1$, which leads to the "Fibonacci numbers" 1,1, 2, 3, 5, 8, 13, 21, 34,... all of which are whole; it is thus striking that the "closed form" expressing all of them as a fixed linear combination of two powers λ^n must involve not only denominator $\frac{1}{2^n}$ but even the irrational number $\sqrt{5}$. On the other hand, what is considered "explicit" depends on the problem at hand: if we want rational approximation to a solution (like $\frac{1+\sqrt{5}}{2}$) of a polynomial equation (like $\lambda^2 - \lambda - 1 = 0$), one way to obtain them is by just iterating the corresponding difference equation, as follows

EXERCISE: Investigate the claim: If x is "any" solution of $S^K x = aS^{K-1}x + \cdots + b\lambda + c$ then $\lim_{n\to\infty} \frac{x_n + 1}{x_n}$ = the largest solution of $\lambda^K = a\lambda^{K-1} + \cdots + b\lambda + c$.

Determinants computed by Linear Categories and Nilpotents

While certain invariants such as the trace and the multiplicity of all non zero eigenvalues are (not only invariant under isomorphism in Lin$^{\circlearrowleft}$ but also) the same for T and S which are "weakly equivalent" in the sense of the definition

$$T \sim S \Leftarrow \exists A, B \left[BA = T \wedge AB = S \right],$$

some other invariants such as the determinant can be changed by weak equivalence, so other methods are needed to compute them.

Exercise Prove that isomorphism implies weak equivalence, i.e. that if for given $T \circlearrowright V$, $W \circlearrowleft S$ there is known to exist at least one $V \xrightarrow{L} W$ for which $LT = SL$ and L^{-1} exists, then it is possible to construct $V \underset{B}{\overset{A}{\rightleftarrows}} W$ for which $BA = T$ and $AB = S$.

Exercise There exist $R \underset{B}{\overset{A}{\rightleftarrows}} R^2$ such that $BA = 1$ (hence has determinant 1) but $S = AB$ has determinant 0.

For calculating determinants, the following Axioms for determinants (of operators on finite dimensional spaces) are useful:

1. If $T \circlearrowright V$, $W \circlearrowleft S$ are such that $T \cong S$, i.e.

$$\exists V \xrightarrow{L} W \left[L^{-1} \text{ exists and } LT = SL \right]$$

then $\det(T) = \det(S)$

2. For any $W \circlearrowleft S$ with $\dim(V) = n$, there exists an upper triangular $n \times n$ matrix T and an isomorphism in Lin$^{\circlearrowleft}$ $T \circlearrowright \mathbb{R}^n \underset{L}{\xrightarrow{\sim}} W \circlearrowleft S$

3. If $T = \begin{pmatrix} t_1 & \cdots & \cdots \\ & \ddots & \vdots \\ 0 & & t_n \end{pmatrix}$ is any upper triangular $n \times n$ matrix, then $\boxed{\det(T) = t_1 t_2 \cdots\cdots t_n}$ the product of the n diagonal elements [Caution: for a square matrix which has non-zero elements both above and below the main diagonal, the determinant is not given by such a simple formula, but by a much more complicated formula.]

4. If $\mathbb{R}^n \circlearrowleft S_0$ is a given matrix [and recall that given any $V \circlearrowleft S$ we can find a matrix isomorphic to it by choosing any basis $\mathbb{R}^n \underset{\beta}{\xrightarrow{\sim}} V$ for V and defining $S_0 = \beta^{-1} S \beta$], then an isomorphism $T \circlearrowright \mathbb{R}^n \underset{L}{\xrightarrow{\sim}} \mathbb{R}^n \circlearrowleft S_0$ can be constructed as a composition of isomorphisms $L = L_0 L_1 \ldots L_p$ of the extremely simple kind known as "elementary row and column operations":

$$\begin{tikzcd}
\mathbb{R} \arrow[loop left, "T"] \arrow[r, "\sim", "L_p"'] & \mathbb{R}^n \arrow[loop above, "S_{p-1}"] \arrow[r, "\sim", "L_{p-1}"'] & \mathbb{R}^n \arrow[loop above, "S_{p-2}"] \arrow[r, "\sim", "L_{p-2}"'] & \mathbb{R}^n \arrow[loop above, "S_{p-3}"] \arrow[r, dashed] & \mathbb{R}^n \arrow[loop above, "S_2"] \arrow[r, "L_1"'] & \mathbb{R}^n \arrow[loop above, "S_1"] \arrow[r, "L_0"'] & \mathbb{R}^n \arrow[loop above, "S_0"]
\end{tikzcd}$$

$$\boxed{S_{k+1} = L_k^{-1} S_k L_k} \qquad \boxed{T = L_p^{-1} S_{p-1} L_p \text{ and } T \text{ is upper triangular}}$$

where "elementary" means in particular that each L_k (while <u>not</u> an <u>endomorphism</u> in Lin$\mathcal{Q}$) is an endomorphism in Lin of $\mathbb{R}^n$ of the special form

$$\boxed{L_k = c_k \cdot I_n + H_k\,,\ c_k \neq 0,\ H_k \text{ nilpotent endo of } \mathbb{R}^n}$$

so that it is trivial to see that L_k^{-1} exists!

Exercise In fact, if

$$L = \begin{pmatrix} c & 0 & 0 & 0 & 0 \\ & c & & & \\ & & c & 0\ 1 & 0 \\ & 0 & & c & 0\ 0 \\ & & & & c\ 0 \end{pmatrix} = c + \begin{pmatrix} 0 & 0 & c & & 0 \\ & 0 & & & \\ & & c & 0\ 1 & 0 \\ & 0 & & 0\ c & 0 \\ & & & 0 & \end{pmatrix}$$

with the 1 in the i-th row and j-th column only, then for any square matrix A of the same size

$$\bar{A} \underset{\text{def}}{=} L^{-1} A L$$

is just like A except that the -th row of A has been replaced by the sum of the original -th row plus c times the -th row. Thus by choosing $c \neq 0$ and the 'pivotal" position i,j correctly we can arrange that the entry in the second row and first column of $\bar{A}$ is 0. Then by a second choice of $\bar{c}$ and i,j (hence an $\underline{L}$), we can arrange the $\bar{\bar{A}} = \underline{L}^{-1}\ \bar{A}\ \underline{L}$ has still more zeroes, and continue until the result has <u>all</u> entries below the main diagonal 0. Since A, $\bar{A}$, $\bar{\bar{A}}$,...etc. are also isomorphic in Lin$\mathcal{Q}$, all isomorphism-invariant properties, such as the value of the determinant, remain the same for each of A, $\bar{A}$, $\bar{\bar{A}}$,..., we can compute the determinant of A by computing the determinant of $\bar{\bar{\bar{A}}}$, but if that is <u>upper triangular</u>, then its determinant is just simply the product of the n <u>diagonal</u> entries of $\bar{\bar{\bar{A}}}$. This is very often the best way to calculate the determinant of a large square matrix A (unless perhaps we know a lot a priori about the structure of A).

Exercise (Following problem 1 on the "last" test) If $V \underset{B}{\overset{A}{\rightleftarrows}} W$ are any linear transformations (with V,W not necessarily $\mathbb{R}^n$, then tr(BA) = tr(AB). If ${}^{T}\circlearrowright V$ is an <u>idempotent</u> operator$(T^2 = T)$, then tr(T) is a <u>whole</u> number and tr(T) ≤ dim(V).

<u>Hint:</u> If $W \underset{\text{def}}{=} \{v \in T \mid Tv = v\}$, and B is the inclusion map of W into V, then we can construct an A (essentially the "rule" of A is the same as that of T, but the idempotence of T must be used to show that it goes into W) for which

$$\begin{tikzcd} V \arrow[loop left, "T"] \arrow[r, shift left, "A"] & W \arrow[l, shift left, "B"] \arrow[loop right, "1_W"] \end{tikzcd} \quad BA=T,\ AB=1_W.$$

But the $\boxed{tr(1_W) = dim(W)}$ can be proved for any <u>identity</u> endomorphism by using any basis.

Test

DEFINITIONS TO BE USED ON TEST

If A is an nxn square matrix, its *trace* is defined as the sum of the diagonal entries

$$\mathrm{tr}(A) = \sum_{i=1}^{n} a_{ii}$$

where a_{ij} are entries of A. If T is a linear endomorphism (operator) on a linear space V, and if $R^n \xrightarrow{\alpha} V$ is a linear isomorphism (or coordinate system for V) then (provisionally) the *trace of T relative to α* may be defined as the trace of the matrix which represents T relative to

$$\mathrm{tr}_{\alpha}(T) = \mathrm{tr}\,(\alpha^{-1}T\alpha)$$

If λ is any scalar, then

$$V_{\lambda}(T) = \{v \in V \mid T v = \lambda v\}$$

is a linear subspace of V called the *λ-th eigenspace* of T, whose dimension $\dim V_{\lambda}(T)$ is called the *multiplicity* of λ as an eigenvalue of T; if the multiplicity $\dim V_{\lambda}(T) > 0$, i.e. if $\exists v \, [v \neq 0 \text{ and } Tv = \lambda v]$ then one says λ *is* an eigenvalue of T.

Two linear operators $V \circlearrowleft^{T}$, $W \circlearrowleft^{S}$ are *weakly equivalent* if there exist linear transformations $V \xrightarrow{A} W$, $W \xrightarrow{B} V$ for which $\boxed{BA = T \text{ and } AB = S}$. A *map* (in the category $\mathrm{Lin}^{\circlearrowleft}$) from T to S is a linear transformation $V \xrightarrow{A} W$ for which $\boxed{AT = SA}$; T and S are *isomorphic* (as linear operators) if there exists an invertible map A from T to S (then A^{-1} will be a map from S to T).

A <u>map</u> (in the category $\mathrm{Lin}^{\circlearrowleft}$) from one linear transformation $V_0 \xrightarrow{T} V_1$ to another $W_0 \xrightarrow{S} W_1$ is a <u>pair</u> $V_i \xrightarrow{A_i} W_i$ of linear transformations for which $A_1 T = S A_0$; the map is an isomorphism iff both A_i are.

1. If A is an nxm matrix and B is an mxn matrix, show that $\mathrm{tr}(AB) = \mathrm{tr}(BA)$.

2. Show that the trace of a linear operator T is well-defined independently of a coordinate system, i.e. that if α, β are two coordinate systems on the space V on which T acts, then (although the matrices $\alpha^{-1}T\alpha$ and $\beta^{-1}T\beta$ are usually <u>different</u>) the numbers $\mathrm{Tr}_{\alpha}(T) = \mathrm{tr}_{\beta}(T)$ are always equal.

3. Prove that if two linear operators T, S are weakly equivalent, then $\mathrm{tr}(T) = \mathrm{tr}(S)$.

4. If T,S are weakly equivalent (witnessed by A,B) show that T^2, S^2 are weakly equivalent (by constructing $\bar{A}$, so that $\bar{A}$, B witness this new equivalence).
How are $\mathrm{tr}(T^2)$, $\mathrm{tr}(S^2)$ related? What about T^3, S^3 ?

5. Show that if A,B witness the weak equivalence of T, S then A,B are in particular maps in $\mathrm{Lin}^{\circlearrowleft}$. If A is a map from T to S <u>and if A is invertible</u>, construct B so that A,B witness the weak equivalence of T,S. But construct a simple example of $V^{\circlearrowleft T}$, $W^{\circlearrowleft S}$ which are weakly equivalent, but for which $\dim V = 1$, $\dim W = 2$ (so that no invertible A could exist.)

6. If T, S are weakly equivalent linear operators and if λ is an invertible scalar (i.e. a non-zero scalar if the scalars form a field), show that the multiplicity of λ as an eigenvalue of T equals the multiplicity of λ as an eigenvalue of S ; do this by constructing a linear isomorphism $V_\lambda(T) \xrightarrow{A_\lambda} W_\lambda(S)$ and constructing its inverse.

7. Give a simple example of a pair of linear transformations $V \underset{B}{\overset{A}{\rightleftarrows}} W$ for which $\lambda = 0$ is an eigenvalue of $S = AB$, but $\lambda = 1$ is the only eigenvalue of $T = B\,A$. [In fact, one example correctly chosen will work both for problem 7 and for the last part of problem 5.]

Conclusion: The hypothesis $\lambda \neq 0$ is needed in problem 6.

8) Given two scalars λ, μ, multiplication defines linear transformations $R \xrightarrow{\lambda} R$, $R \xrightarrow{\mu} R$ between one-dimensional spaces $V_0 = V_1 = R$, $W_0 = W_1 = R$. How must λ, μ be related in order that these two transformations are isomorphic as objects of the category $\mathrm{Lin}^{\substack{\bullet \\ \downarrow \\ \bullet}}$?

9. But considering two scalars as operators $R \circlearrowleft^{\lambda}$, $R \circlearrowleft^{\mu}$ how must they be related to be isomorphic in $\mathrm{Lin}^{\circlearrowleft}$?

10. Consider an arbitrary 2x2 matrix as an object

$$R^2 \circlearrowleft \begin{pmatrix} t_{11} & t_{12} \\ t_{21} & t_{22} \end{pmatrix} = T$$

of the category $\mathrm{Lin}^{\circlearrowleft}$ of linear operators. Try to find an upper triangular matrix

$$R^2 \circlearrowleft \begin{pmatrix} s_{11} & s_{12} \\ 0 & s_{22} \end{pmatrix} = S$$

and an *isomorphism* A in $\mathrm{Lin}\,\wp$ between S and the given T . Find conditions on the four scalar entries of T for which such an S and A exist, explaining why the condition is different in the two cases R = real numbers and R = complex numbers, giving a simple example in the real case for which the condition is not satisfied.

Hint: S would have the *same* determinant as T ,
and determinant $A \neq 0$.

Application of the "Geometric Series" Formula $\frac{1}{1-D} = \sum_{k=0}^{\infty} D^k$ to Differential Equations.

The above formula can be valid for "nilpotency reasons" even when D is not nilpotent, in situations where D is acting on a big linear space in "locally nilpotent" fashion. For example consider the space of all polynomials in one variable t and interpret D as ordinary differentiation:

$\mathbb{R} \subset$ linear polynomials $\subset$ quadratic polynomials $\subset$ cubic polynomials $\subset \dots \subset$ all polynomials $\circlearrowleft D$

While D is not nilpotent in acting on all polynomials (since there exist polynomials of degree $>$ any given n), for each given polynomial g,

$$D^{n+1}g=0$$

if we take n $\geq$ deg(g), because applying D decreases the degree by one and hence interating D enough times kills the given polynomial g. That is

"$\exists n \forall g \; D^{n+1}g=0$" is <u>false</u>, but $\forall g \exists n \; D^{n+1}g=0$ is <u>true</u>

(This of course would not be true if we allowed g to be a rational function or an exponential function).

Suppose we need to find all solutions u of the differential equation

$$u-u'=g$$

where g is a given polynomial "forcing term". Now a basic general principle of linear algebra is that, if T is any <u>linear</u> transformation, then the set of all solutions u of

$$Tu=g$$

can be parameterized by the linear space of all solutions f of the "homogeneous equation" $Tf=0$

provided we can find <u>one</u> "particular" solution $\bar{u}$ to $T\bar{u}=g$.

The parameterization is just $f \mapsto u \underset{\text{def}}{=} f+\bar{u}$, i.e. $T\bar{u}=g \Rightarrow \exists$ "*bijection*" (or one-to-one correspondence) $\{f \mid Tf=0\} \rightleftarrows \{u \mid Tu=g\}$

where the inverse of the bijection is $f \underset{\text{def}}{=} u-\bar{u}$.

EXERCISE Prove this basic principle of linear algebra.

Now in applying this basic principle to the above differenti equation we take $T = 1 - D$, so that the homogeneous equation is just $f - f' = 0$ whose solutions are well known from elementary calculus to be parameterize by integration constants A : $f(t) = Ae^{t}$

Thus we need only find a particular solution $\bar{u}$ to our differential equation $\bar{u} - \bar{u}' = g$

But since g is a polynomial, our "local nilpotency" geometric series tells us that $\boxed{\bar{u} = \sum_{k=0}^{\infty} D^{k} g}$ is the unique* polynomial solution; in fact the degree of $\bar{u}$ is the same as the degree of g.

EXERCISE: Verify this for $g(t) = 3t^{2} - t + 7$,

Thus the general solution of $u - u' = g$ is $u = Ae^{t} + \sum_{k=0}^{\infty} D^{k} g$ where the series is actually finite for any given polynomial g.

While we have taken $D = (\ )'$ to keep the solution of the homogeneous equation simple and thus emphasize the method of finding the particular solution, exactly the same method could be used if $D = 10(\ )' + 3(\ )'' - 2(\ )'''$ Then the solution f of $(1-D)f = 0$ would involve periods of oscillation, damping and growth constants, etc. but so long as the forcing term g is a polynomial the particular polynomial solution $\bar{u}$ could still be found by iterating (this more complicated) D on g and adding the results, since even this more complicated D is still "locally nilpotent" when acting on polynomials Further (going beyond pure algebra into analysis) even if g is not a polynomial it can be approximated by a polynomial if we are interested in studying the fine structure of a solution u over a brief period of time (say over a time span of a few days when the shortest period of g is annual

---

* Since $u - u' = 0$ has no non-zero polynomial solutions.

Multi - Dimensional Calculus

If f is a (smooth) real-valued function defined on an open interva in the line such that the values of f lie in an open interval Y, and if g is another smooth function defined on Y, then the chain rule of elementary calculus states that

$$(g\circ f)'(x) = g'(f(x))\cdot f'(x)$$

For a multitude of applications in geometry, physics, economics, etc. it is necessary to give the chain rule a meaning and validity also in higher-dimensional situations. If $X \xrightarrow{f} Y \xrightarrow{g} Z$ are mappings of three higher-dimensional non-linear spaces, the meaning of $g\circ f$ is clear:

$$X \xrightarrow{g\circ f} Z \qquad (g\circ f)(x) = g(f(x)) \text{ all } x \text{ in } X$$

But what of ( )' and what of the "multiplication"• on the right-hand side of the chain rule? Note that even in the one-dimensional case, even though x is restricted to X, if $a = f'(x)$ for a certain x then the product av makes sense for any v ; similarly if $b = g'(y)$ for a certain y, b can be multiplied by any w (not necessarily in Y). Moreover

$$(ba)v = b(av) \quad \text{for all } v$$

The chain rule gives another meaning to the product ba in case $y = f(x)$ namely if $c = (g\circ f)'(x)$ at the same x, then

$$c = ba$$

i.e. for all v, $cv = (ba)v = b(av)$

Now if X,Y,Z are not necessarily 1-dimensional (nor linear) we can still attach to each point x a linear space V_x called the tangent space to X at (for example X might be a sphere) and similarly W_y a linear space attached to each point y of Y. Then the derivative of f at x will be a linear transformation $V_x \xrightarrow{f'(x)} W_{f(x)}$

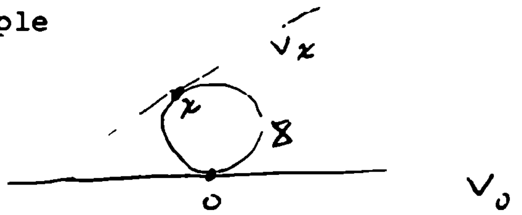
i.e. $f'(x)v \varepsilon W_{f(x)}$ for all $v \varepsilon V_x$, and $f'(x)(\lambda_1 v_1 + \lambda_2 v_2) = \lambda_1 f'(x)v_1 + \lambda_2 f'(x)v_2$.

Now composition of linear transformations may be considered as a generaliz "multiplication". If we have another non-linear map $g: Y \rightarrow Z$ and if we hav chosen $x \in X$, then we can chose not only $y = f(x)$ but also $z = g(y) = g(f(x))$ and if we denote the tangent spaces to Z by U_z for $z \in Z$ then

$$\begin{tikzcd} X \arrow[rr, "g \circ f"] \arrow[dr, "f"'] & & Z \\ & Y \arrow[ur, "g"'] & \end{tikzcd} \qquad \begin{tikzcd} V_x \arrow[rr, "(g \circ f)'(x)"] \arrow[dr, "f'(x)"'] & & U_z \\ & W_y \arrow[ur, "g'(y)"'] & \end{tikzcd} \qquad \begin{matrix} z = g(y) \\ y = f(x) \end{matrix}$$

The chain rule says that the last diagram is commutative, as composition of linear transformations (i.e. "multiplication")

In case all three X,Y,Z of the non linear spaces are included in (affine) linear spaces $X \subset V, Y \subset W, Z \subset U$ it is often possible to identify the various tangent spaces with a common linear subspace $V_0 \subset V$, $W_0 \subset W$, $U_0 \subset U$ of the containing space by rotation and subtractio of vectors, for example



all inside V

so that x corresponds to the origin. If we omit all these identifications, f'(x) becomes (somewhat confusingly) a linear transformation $f'(x): V_0 \to W_0$ between fixed linear spaces, but in general a different linear transformati for each x. If X,Y,Z are open (for example balls) in V, W, U, then $U_0 = U$, $V_0 = V$, $W_0 = W$ but in general they will be lower-dimensional, as in our picture.

While the "co-ordinate free" description is necessary for describing the objective motion of bodies in everyday life and in conceptual physics, equally necessary for numerical calculations is the introduction of (subjectively chosen) "coordinate systems"; such induce among other things the choice of a basis in each tangent space (which in turn induces an identification of the various tangent spaces with each other which may have to be untangled by subtracting and rotating etc.) Thus if $X\subset\mathbb{R}^m$, $Y\subset\mathbb{R}^n$, are open, then for each x, $y=f(x)$ we have linear transformations:

$$\mathbb{R}^m \xrightarrow{f'(x)} \mathbb{R}^n \xrightarrow{g'(y)} \mathbb{R}^p$$

Then $f'(x)v$ is a matrix product for each column vector $v \varepsilon \mathbb{R}^m$ where f'(x) "is" the matrix whose entries are

$$f'(x)_{ji} = \frac{\partial f_j}{\partial x_i} \qquad \begin{array}{l} j=1,\dots n \\ i=1,\dots,m \end{array}$$

where the f_j are the components of the nonlinear map f

$$f(x) = \langle f_1(x), f_2(x) \dots, f_n(x)\rangle$$

and similarly the chosen bases for the tangent spaces to Y and Z will give g'(y) the matrix entries

$$g'(y)_{kj} = \frac{\partial g_k}{\partial y_j} \qquad \begin{array}{l} k=1,\dots,p \\ j=1,\dots n \end{array}$$

Since composition of linear transformations is represented, relative to given coordinates, by matrix multiplication, the chain rule becomes

$$(g\circ f)'(x)_{ki} = \sum_{j=1}^{n} \frac{\partial g_k}{\partial y_j} \frac{\partial f_j}{\partial x_i} \qquad \begin{array}{l} k=1,\dots,p \\ i=1,\dots,m \end{array}$$

where it must be understood that the $\frac{\partial g_k}{\partial y_j}$ are to be evaluated at y = f(x) and not some other y.

Another important construction in calculus is Newton's method, both in the proof of theorems like the implicit function theorem as well as in a great variety of numerical approximations. Here one needs to find x for which f(x) = y, where f is a given non-linear map and y is a given point in the codomain Y of f. Newton's method is to iterate the (even more nonlinear) map $\mathcal{N}_{f,y}(x) = x + \frac{y - f(x)}{f'(x)}$ starting from a point $\bar{x}$ which one believes is "nearly" a solution and repeating $\mathcal{N}$ again and again hoping that the resulting sequence

$$\bar{x},\ \mathcal{N}(\bar{x}) = \bar{\bar{x}},\ \mathcal{N}(\bar{\bar{x}}) = \bar{\bar{\bar{x}}},\ \dots$$

of points in X will "converge" to an actual solution x of f(x) = y. While this hope is amazingly often justified, the discussion of that is given in courses on analysis. Here we just point out that since, if X,Y are more than onedimensional, f'(x) is not a number but a linear transformation, to even get the Newton procedure $\mathcal{N}$ going one has to deal with the inverse $\frac{1}{f'(x)}$ of a linear transformation (and indeed deal with it in a way that can be repeated when one changes from $\bar{x}$ to a new $\bar{\bar{x}}$, thus getting a whole new linear transformation) and this is one of the important problems of linear algebra. (More precisely, if dim(X) $>$ dim(Y), it is only a <u>one sided</u> inverse $S(x)$ for f'(x), i.e. $f'(x) \circ S(x) = 1_{W_x}$, rather than an actual 2-side inverse, which is <u>possible</u> because of the geometry and <u>required</u> for Newton' method.)

The operators grad, curl. div, lap, are linear transformations betwee <u>infinite</u> -dimensional linear spaces of functions. For example if X is open in $\mathbb{R}^n$, then $\mathbb{R}^X$ = the linear space of all smooth real-valued functions on admits the <u>linear</u> self-transformation

$$\mathbb{R}^X \circlearrowleft \mathrm{lap} \qquad \mathrm{lap}(f) = \sum_{i=1}^{n} \frac{\partial^2 f}{\partial x_i^2}$$

named after Laplace, who was instrumental in using it to develop the theory of gravitational and electrical potential.

→ Exercise: Use lap to define a new (non associative) product on functions $X \to \mathbb{R}$ as follows (Here we use the fact that $\mathbb{R}^X$ is a ring as well as a linear space).

$$g \bullet f \underset{\mathrm{def}}{=} \frac{1}{2}\left[\mathrm{lap}(gf) - g\,\mathrm{lap}(f) - \mathrm{lap}(g) f\right]$$

then $g \bullet (uv) = (g \bullet u)v + u(g \bullet v)$ (Leibniz rule!)

$x_i \bullet u = \frac{\partial u}{\partial x_i}$ [note: x_i is a particular g]