From rules of grammar to laws of nature

Joachim Lambek
McGill University, Montreal

December 3, 2013

Abstract

We look at questions debated by pre-Socratic philosophers, emphasize their linguistic motivation and point out attempted answers by modern physics. We pursue a mathematical treatment of the syntax and morphology of natural languages, investigate the foundations of mathematics and speculate about a quaternionic approach to special relativistic quantum mechanics.

0 Introduction

Many questions about the nature of the world are concealed in our language, even if we are not always aware of the hidden linguistic assumptions. I believe that most, if not all philosophical investigations are motivated by reflections on the structure of language. For example, the intense debate by pre-Socratic philosophers whether matter is continuous or discrete is anticipated by grammatical distinctions, still revealed by those between mass nouns and count nouns in English. Modern physics seems to have resolved this question in favour of the latter, although there still is some disagreement about the nature of space and time, as already discussed by the Greek philosopher Zeno.

As far as we know, language was invented between fifty and a hundred thousand years ago by a small group of humanoid primates in Africa, allowing them to spread over the whole globe and grow to several billions in number. It is believed by some experts that all present day languages have a common origin.

Modern English is a member of the Germanic branch of the Indo-European family of languages, which stretch all the way from Iceland to Ceylon and are still recognizably related. Indo-European itself has been assigned a place in the Eurasiatic super-family that includes also many North-asiatic languages and even supposedly Korean, Japanese and the language of the Eskimos (aka Inuit). A parallel super-family called Afro-asiatic includes many North-African and Middle-Eastern languages formerly called Hamito-Semitic, even such remote members as Hausa and Somali.

Both these super-families have been included in what Russian linguists call Nostratic, together with some other languages,e.g. those in the Caucasus. What is not included in Nostratic are the languages of South-East Asia, of Sub-Saharan Africa and the aboriginal languages of Australia and America.

Since the invention of language, humans have made considerable progress. Here is a rough timetable:

Language: 50,000 years ago,
Writing: 5,000 years ago,
Printing: 500 years ago,
Word processing: 50 years ago.

The medieval undergraduate curriculum comprised *seven liberal arts*. These were made up from the *quadrivium*, consisting of pure and applied mathematics, as already envisaged by the ancient Pythagorean school: geometry, arithmetic, astronomy and music, and preceded by the *trivium*: logic, grammar and rhetoric. The first two of these subjects have by now also undergone mathematical treatment.

Originally, *grammar* was meant to denote the art of reading and writing; but, in medieval Western Europe, it was restricted to the study of Latin, a necessary prerequisite to intellectual advancement. Not surprisingly therefore, knowledge of grammar was seen to provide a person with magical power, to be described by the word "glamour", derived from the word "grammar" and now applied more to fashion models than to intellectuals. Well, this is one etymological interpretation. Dictionaries interpret the word "grammary" to denote a book of magical spells, transformed into "glamour" by the novelist Walter Scott.

Nowadays grammar is largely conceived as consisting of *inflectional morphology*, which tells us how naked words stored in the mental dictionary are transformed into word-forms, and of *syntax*, which tells how these may be strung together to form sentences. In some languages, e.g. in Chinese, the lexical words are not transformed at all, and in others, e.g. in Turkish, the word-forms are constructed by linking together a number of so-called *morphemes* subject only to a phonological process such as *vowel harmony*. In some languages, e.g. in Latin, the mental dictionary must list many clues how the lexical words are to be transformed into word-forms, while in others, e.g. in some Semitic languages, the word-forms can be calculated mathematically, subject only to a few phonological rules.

In the so-called *polysynthetic* languages, e.g. Mohawk, the distinction between word-forms and sentences is not so clear. Not being familiar with any of these, I will point out only one example from Latin, where the single word *fugabantur* translates into English "they were put to flight".

In English we find traces of all the language types discussed above. The inflectional morphology has partly disappeared, yet the verbs still may be transformed by adding the morphemes + *es*, + *ed* and + *ing*, subject to some phonological rules which may transform + *es* into + *ez* or + *ed* into + *et*. However, there are many irregular verbs; for example, there is no way we can calculate *went* and *gone* from *go*, and these forms must be listed in the mental dictionary.

Traditional grammars, influenced by the study of Latin, concentrate on inflectional morphology, but nowadays the rules of syntax have become just as important. Not all grammatical rules can be found in official textbooks or dictionaries, and I will discuss a few such rules in the present text.

Languages tend to differ both in vocabulary and grammar and often seem to conceive the world in different ways. Some linguists believe with Noam Chomsky that there is a *universal grammar* that different languages share and others believe in the *Whorf-Sapir hypothesis* which asserts that the way people see the world is dictated by the structure of their language. Both theories may be correct, although I am not convinced by the evidence presented for either. In my experience, speakers of a language are not necessarily consciously aware of the peculiarities of their language, but outsiders often are.

The present text will discuss questions in the foundations and philosophy of linguistics, mathematics and physics and postpone applications of more technical mathematics as much as possible.

Part I will deal with words in the mental dictionary with some consideration of their origin and meaning.

In Part II we will illustrate how these lexical entries are transformed by inflections, in particular in certain circum-mediterranean languages.

In Part III we investigate how strings of inflected word-forms yield grammatical sentences. Apart from plain rewriting rules this can be done with the help of *categorial* grammars, where each inflected word-form is assigned a grammatical *type*, namely an element of a logical or algebraic system, in which a calculation will ensure that a string of types reduces to the basic type of a well-formed sentence.

In Part IV we review some minimal mathematical prerequisites, introducing some basic algebraic concepts, ultimately the notion of an additive category.

In Part V we attempt to formulate modern physics in terms of Hamiltonian quaternions and are led to speculate that time, like space, has three dimensions.

Finally, in Part VI we turn to a different train of thoughts concerning the foundations of mathematics. We claim that, contrary to a widely held belief, the different philosophies of mathematics called *logicism, formalism, Platonism* and *intuitionism*, when moderately formulated may be reconciled to form what has been called *constructive nominalism.*

Part I. The mental dictionary.

1 Kinship terms

Among the oldest words of a language are kinship terms. For example, compare English

father, mother, brother, sister, son, daughter

with Sanskrit

pitr, matr, bhrater, svasr, sunu, duhitr,

and there can be no doubt about a common origin. Not all words have been preserved without a change in meaning. For example, the English word *nephew* is related to Sanskrit *neptr*, which originally meant *grandson*. How did this change in meaning come about?

Old English originally had two words for *uncle*: the father's brother and the mother's brother. At some point in history it was considered less important to maintain this distinction and the word *uncle* was derived from Latin avunculus, meaning "little grandfather". The reason for this was that a woman's guardian was her father, upon whose death his brother took his place. Conversely, the rôle of grandson now became attached to the word "nephew". In modern Italian, *nepote* still means both grandson and nephew.

Another curiosity is associated with the word *cousin*, meaning "first cousin" to most people, who have no idea what a third cousin twice removed is. In India, first cousins are considered to be siblings, and in some languages what we call "cousins" can refer to almost any relative. For example, in the Trobriand Islands of the Western Pacific, the child of a parent's sibling can be referred to by words whose primary meaning is one of the following, as observed by Malinowski in 1932:

grandparent or *grandchild, mother, father, maternal uncle* or a *man's nephew* or *niece, sibling* of the opposite sex, *younger sibling* of the same sex, *child*.

Not impressed by Malinowski's interpretation, Leach declared in 1958 that these meanings were "utterly incomprehensible". Yet Lownsbury showed in 1965 that these meanings could be derived mathematically if one adopted a small number of simple *reduction rules*, e.g. that a father's sister is to be regarded as a paternal grandmother.

More formally, we apply the mathematical calculus of binary relations to describe kinship. Thus we have

MSFP $\to$ *avunculus*

where the left side refers to a

male sibling of a female parent.

In English we can also express this in reverse order as

a mother's brother

and in some articles I have followed this convention. Here I prefer to stick to the above order, since it is easier to refer to a mother as a *female parent* than as a *parent who is female.* Note that F may be viewed as the identity relation between females.

In the Trobriand language we have

$$G_4CG_3SG_2PG_1 \to X,$$

where $G_i = M$ or F and X is to be calculated: If $G_2 = M$ and $G_3 = F$, $X = tama$, whose primary meaning is "grandparent or grandchild". If $G_2 = F = G_3$, $X = tama$ or *bwada*, whose primary meaning is "older or younger sibling of the same sex", provided $G_1 = G_4$, but $X = latu$, whose primary meaning is "sibling of the opposite sex", provided $G_1 = F$. If $G_2 = M = G_3$, there is no term X in the language.

To account for these and some other data, Lownsbury [1965] postulated a number of rewrite rules, here presented in the notation of [BL1995], except for the order of multiplication:

(1) FSMP $\to$ FPMP,
(1′) CMSF $\to$ CMCF,
(2) GSG′P $\to$ GP,
(2′) CG'SG $\to$ CG,
(3) CFP $\to$ S,
(4) PGP $\to$ P,
(4′) CGC $\to$ C,
(5) CMP undefined.

Here is a sample calculation:

$$\begin{array}{rcll} FC\underline{\mathrm{FSMP}}M & \to & F\underline{\mathrm{CFP}}MPM & \text{by (1)} \\ & \to & \underline{\mathrm{FSMP}}M & \text{by (3)} \\ & \to & F\underline{\mathrm{PMP}}M & \text{by (1)} \\ & \to & FPM & \text{by (4)} \\ & \to & tama & \end{array}$$

Note that CMP is undefined. Some anthropologists have suspected that this is so because Trobrianders did not understand the rôle of the father in procreation. A more likely explanation may be inferred from the observation by the travel writer Theroux in 1992 that during the annual month-long yam festival women are permitted to seduce any man they meet.

As Chomsky famously pointed out; "whereas the Greeks made up number theory, others made up kinship systems".

Compared to the elaborate kinship systems of many primitive tribes, the English one appears to be quite pedestrian, since all genders, except those at the beginning of a kinship description, are irrelevant. However, even in some English dialects there are some surprising complications. For example, Goodenough [1965] observed that some American speakers apply

MSPP $\to$ *great uncle*

but

MCCS $\to$ *grand + nephew*.

The distinction between *grand+* and *great* may usually be explained by postulating

$$\begin{array}{rcl} \mathrm{GCC} & \to & \textit{great}\ \mathrm{GC\ before\ C} \\ & \to & \textit{grand}\ +\ \mathrm{GC\ otherwise,} \\ \mathrm{GPP} & \to & \textit{great}\ \mathrm{GP\ before\ P} \\ & \to & \textit{grand}\ +\ \mathrm{GP\ otherwise,} \end{array}$$

We may thus calculate at the end of a kinship description:

MPPP $\to$ *great* MPP $\to$ *great grand* + MP $\to$ *great grandfather*,

MCCC $\to$ *great* MCC $\to$ *great grand* + MC $\to$ *great grandson*,

MCCCS $\to$ *great* MCCS $\to$ *great grand* + MCS $\to$ *great grandnephew*

However, when it comes to calculate MSPP, none of the above rules may apply. It is therefore not surprising that Goodenough's Yankees should have picked the rule

GSPP $\to$ *great* SP,

thus leading to the calculation

$$\begin{array}{rcl} \text{MSPPP} & \to & \textit{great}\ \text{MSPP} \\ & \to & \textit{great great}\ \text{MSP} \\ & \to & \textit{great great uncle.} \end{array}$$

as opposed to

MCCCS $\to$ *great grandnephew*.

2 Number words

As old as kinship terms are words denoting natural numbers, although they are usually not as well preserved. Consider for example the English *five* with the French *cinq*. On the face of it, they have no sound in common. Yet English *five* is evidently related to German *fünf* and French *cinq* to Latin *quinque*, and one still recognizes a resemblance between *fünf* and *quinque*. In fact, philologists derive both these words from a proto-Indo-European **pengue*, which seems to be cognate to English *finger*.

Another old word for finger is **tik*, presumably going back to the origin of all human languages and still surviving in Latin *decem* (= ten) and English *ten* and *digit*, as well as German *zehn* and *zeigen* (= to point).

Although Indo-European and Hamito-Semitic are not particularly closely related, the words for *six* and *seven* bear a surprising similarity in both language families. This similarity had not been ascribed to a common origin, but to an early contact between the two families in Anatolia. Other words that have similarly been borrowed are *star* and *wine*.

Sometimes it is not individual words that resemble each other in different languages but the similarity between two words in one language that resembles that in another. For example, a similarity between *new* and *nine* can be observed in all Indo-European languages; in French even the same word *neuf* serves for both. This similarity may be explained by assuming that at one time the thumb was not considered to be a finger and people counted up to *eight*, making *nine* a new number.

Such an explanation won't work to explain the similarity between *two* and *tooth*, which prevails in both Indo-European and Semitic languages. For example, the masculine word for *two* in Hebrew can also denote a pair of *teeth*.

Since Indo-European and Semitic are not very closely related, one may wonder whether the similarity between *two* and *tooth* has an even older origin. The proto-world-language postulated by Greenberg and Ruhlen contains the word * *pal* for two, but records no word for *tooth*. On the other hand, the rival Nostratic proto-language postulated by the Russian school does contain the word $^*p^hal$ for *tooth*, still surviving in Telugu *palu*. I wonder whether this is just a coincidence, or whether the similarity might indeed go back fifty thousand years.

Since first writing the above, it was brought to my attention by Claudia Casadio that even in Mohawk, a native American language still spoken in two suburbs of Montreal, the words for *two* (*tekuni*) and *tooth* (*tekeuno*) resemble one another. While the picture of an elephant or mammoth comes to mind, there is probably no meaningful explanation for the similarity between the two words.

The meaning of number words ought to be clear. All but *one* serve as plural adjectives. Yet, mathematicians often employ a rather confusing language, for example when declaring that the two roots of an equation are equal. Evidently, if they are equal they were not two in the first place.

Ignoring such abuse of language, we all know the meaning of the adjective *two*. But what about the noun *two*, supposedly an entity living in a Platonic universe? A naive explanation claims that the number 2 consists of two platinum balls kept in a drawer in Paris. Not much more seriously, I recall the famous intuitionist Brouwer saying that 2 is the principle of *twoity*, to which the famous geometer Coxeter responded by claiming that “twoity” ought to be replaced by either “binity” or “twoness”.

A more serious interpretation by the logicists Frege and Russell claimed that 2 is the collection of all pairs of entities. But this leads to the question whether such a collection is itself an entity. As an alternative, von Neumann defined 2 as a special unordered pair, namely the set consisting of 0 and 1. In the same spirit, he defined 1 as the set consisting of 0 alone and 0 as the empty set.

A more dynamic view conceives 2 not as a set, but as a function. Thus, according to Wittgenstein and Church, 2 is the process of iteration which assigns to every function $f : x \mapsto x$ its iterate $f^2 = f \circ f$, where

$$(f \circ f)(x) = f(f(x)).$$

This leaves us with the question: What is x?

Perhaps the most undisputed view goes back to Emanual Kant, and might be translated into modern terminology as saying that 2 is the equivalence class of all expressions in the language of mathematics which are provably equal to $1 + 1$. Of course, the definition presupposes the acceptance of a formal language, such as that of Giuseppe Peano, in which 2 is defined as the successor of 1. We shall return to this is a later chapter.

3 Nouns

Let us now turn to the wider class of nouns and what they denote. We begin with a couple of quotations from Lewis Carroll:

> ‘The time has come’, the Walrus said
> ‘to talk of many things,
> Of shoes - and ships - and sealing wax -,
> Of cabbages-and-kings.......’
> ‘A loaf of bread’ the Walrus said,
> ‘is what we chiefly need:

pepper and vinegar besides
are very good indeed - ....'

Note that *shoes, ships, cabbages* and *kings* are plurals of count nouns, while *wax, bread, pepper* and *vinegar* are mass nouns, which have no plurals and do not admit an indefinite article.

For example, beans may be counted, but rice may not. Two hundred years ago, it was believed that *pease* was a mass noun like *rice*, but nowadays *pea* is viewed as a new singular with the plural *peas*. On occasion, I have wondered whether *rice* may someday be viewed as the plural of *rouse*.

A similar distinction in classical Greek may have led pre-Socratic philosophers to debate whether the world is continuous or discrete. The former view was entertained by Thales, who claimed that all matter is made out of one substance, namely water, while Democritus asserted that all matter consists of indivisible entities he called "atoms".

The single substance of Thales was later extended to four, namely earth, water, air and fire, reminding us of the modern solid, liquid, gas and energy, the last only recognized as a kind of matter in the early twentieth century.

In parallel to the philosophical debate there was one in mathematics: what is more fundamental, geometry or arithmetic? The latter view was favoured by the legendary Pythagoras, whose followers only reluctantly admitted that the diagonal of a unit square could not be measured by the ratio of two whole numbers. The philosophical problem was resolved by Aristotle, who pointed out that, even if water cannot be counted, cups of water can. The mathematical problem had already been resolved in Plato's Academy, when it was realized that ratios of geometric quantities could be measured by what we now call "continued fractions" (Theaetetus) or "Dedekind cuts" (Eudoxus).

The philosopher Zeno had shown, with the help of his famous paradoxes, that space and time cannot be discrete (and perhaps also that there are problems with continuity). Modern physicists, with some exceptions, still believe that space and time are continuous, but have adopted the view that matter and energy are discrete. Yet, the probability that a particle appears at a certain position in space-time is still represented by a real number between 0 and 1. Indeed, quantum mechanics requires for its description also complex numbers, or even quaternions, as we will see in a later chapter.

Nouns in many familiar languages display three features: number, gender and case. Number is relatively easy in English, where the plural of regular count nouns is formed by attaching the suffix $+s$, subject to certain phonological rules, e.g. transforming *knife* $+s$ into *knives*. But there are many irregular plurals that have to be memorized, e.g. the plural of *ox* is *oxen*. In other Indo-European languages, such as Latin or German, the plural formation may be much more complicated.

Nouns in Semitic languages admit three so-called numbers: singular, plural and dual. The Hebrew plural is even simpler than the English one: a nominally masculine one ends in $+im$ and a feminine one ends in $+ot$. Unfortunately, there are exceptions, as any man entering a room labelled *nashim* will discover. On the other hand, Arabic plurals are quite irregular and place a considerable burden on one's memory.

Many languages attach a *gender* to each noun. In European languages this may be masculine, feminine and sometimes neutral. In some sub-Saharan languages there appear many other classes of nouns, for instance those denoting vegetables or flat objects.

The meaning of genders should not be taken too seriously. I doubt whether German speakers think of the table as male and of the sun as female, and that French speakers think the opposite. Although English nouns carry no visible gender, some English speakers think of ships as being female.

Let me take a closer look at the word *person* of Etruscan origin. In Latin it had the meaning of "mask", ultimately referring to a character in a play. The modern meaning of "person" is roughly the same in German, French and English, although it is feminine in German and French, even when referring to a male. Yet, speakers of those languages may feel uneasy when they have to refer to a male person in a subsequent sentence as *she*. Curiously, in French the word *personne* has also come to mean *nobody*.

Paradoxically, in English the word *person* has acquired the meaning of *woman* in attempted political correctness. Trying to avoid words ending in +*man* , people might say *chairperson* in place of *chairman*,but only think of doing so when referring to a woman.

The Latin noun had six *cases*, so named after the Latin word *casus* meaning "fall", reminding us of the six ways a die can fall. In many modern languages, the ancient cases have been eliminated in favour of prepositional phrases. Thus English has replaced the original dative and genetive cases by propositional phrases involving *to* and *of*. It is widely believed that the old genetive case has survived by attaching the ending +'*s* to a noun, but actually this ending is attached to a noun phrase, as in

the queen of England's hat.

The original cases still survive in the pronouns *him* or *her* and *his* or *her*. For example,

he gave her roses

is an optional alternative to

he gave roses to her.

Comparing the Semitic languages Arabic and Hebrew, we notice that Arabic has retained cases like German, but Hebrew has replaced them by prepositional phrases like English.

4 Verbs

There are several types of verbs in English, among which we mention intransitive verbs like *sleep* that require no object, transitive verbs like *kill* that require one object and ditransitive verbs like *give* that require two objects called "indirect" and "direct", as illustrated in the following sentences:

she sleeps; she loves him; she gives him cookies.

There are also auxiliary verbs like *may* as in

she may sleep.

Traditionally, European languages are expected to express *tense*, supposedly denoting time. The so-called *simple* tenses are formed by a modification of the verb called "conjugation". Thus Latin has five simple tenses: present, past, future, present subjunctive and past subjunctive. These appear in three patterns, here called "aspects", as in *amo, amavi* and *amor*, which translate as "I love", "I have loved" and "I am loved" respectively.

English grammar is usually described in terms inherited from Latin grammar and supposedly has three tenses, not counting the almost obsolete subjunction. However, the English future is not a simple tense and requires an auxiliary verb to express it. In Victorian times one distinguished between *shall* in the first person and *will* in the third. There was a joke concerning a foreigner who fell into the Thames and shouted

I will drown and no one shall save me.

Nowadays the future is usually expressed by *will* or, more recently, by *gonna*. Not counting the almost obsolete subjunctive, this leaves two simple tenses, traditionally called "present" and "past", although the so-called present tense does not necessarily denote the present time. It would be more reasonable to follow the example of Japanese or Arabic grammarians and call these tenses "incomplete" and "complete" respectively. Concerning the compound future tense, I see no reason why we should call it a tense at all, unless we also admit a "permissive tense" expressed by *may* and a "possibility tense" as expressed by *can*. Similar compound tenses exist in Turkish, where they are formed by agglutination.

A question arises why we spell the negation *cannot* in one word, but not *must not*? There is a rational explanation for this: *she cannot sleep* is the negation of *she can sleep*, while *she must not sleep* is not the negation of *she must sleep*, which is better rendered as *she may not sleep*.

I sometimes wonder whether English recognizes a unique future at all. In fact, I have come to speculate that even nature admits three dimensions of time, as we shall see in a later chapter.

Transitive verbs, namely verbs requiring an object, are often interpreted as representing a causal action. Thus *X kills Y* means that *X causes Y to die*. In languages without strict word order, X and Y may be augmented by different case endings. In English, cases are only indicated in pronouns. Thus *she kills him* makes it clear that *he* is the victim. Yet, there is no reason why *he* should be in the same case as in *he dies*. In fact, in the so-called "ergative" languages, *kills* requires an object in the nominative case.

I recall telling a Georgian friend that his language was of interest to linguists because of this property. After asking me to explain the term "ergative" to him, he said: "I have studied Latin grammar in school and am convinced that in our language, as in Latin, the subject is in the nominative case and the object in the accusative one." I had to ask him to compare the subjects of "she sleeps" and "she kills him" in his language to make him see that things were not quite as he thought. I mention this example to illustrate the fact that speakers of a language make use of internalized rules that they may not be aware of.

Nowadays we have become accustomed to see many, if not most, transitive verbs as expressing a causal action, with the subject as agent and the object as patient, although, from a logical point of view, a transitive verb just represents a binary relation. For some verbs this relation is symmetric, as in

she resembles him, she marries him,

so the question of causality does not arise. Even where it does, different languages may not agree as to what is the subject and what is the object. Thus Francophones would say *elle me manque*, where Anglophones say *I miss her*. Even in English, the rôles of subject and object may have become interchanged over time. Thus *I like her* would have been rendered as *she likes me* a few centuries ago, like modern *she pleases me*.

It seems that *ditransitive* verbs, logically ternary relations, always express causality in English. Thus, *he gives her roses* and, equivalently, *he gives roses to her* mean "he causes her to have roses", and similarly if *roses* is replaced by *a cold*. The indirect object *her* originally appeared in the so-called *dative* case, indistinguishable from the *accusative* case in modern English, and may be replaced by the prepositional phrase *to her*, though in a different position in the sentence. Similarly

show means "cause to see",

tell means "cause to hear",

teach means "cause to know",

explain means "cause to understand".

The importance of time and change was the subject of intense debate between pre-Socratic philosophers, in particular between Heraclitus and Parmenides, and so was the nature of causality. For example, when I see a flower, am I or the flower the active cause? Empedocles believed it was the seeing person, a belief shared more recently by the philosopher Berkeley. Of course,physics has settled the problem in favour of *light* rays as against *sight* rays. Still, one language often treats causation and perception analogously. For example, the sentences

I make her sleep, *I see her come*

share the absence of *to* in the infinitive.

5 Adjectives and participles

Adjectives in European languages may serve two distinct rôles in sentence formation. In *predicative* position, as in English

the wine is good

they help to complete the verb phrase involving the copula *is*, a conjugational form of the auxiliary verb *be*; but in *attributive* position, as in

the good wine,

they help to modify the noun within a larger noun-phrase.

The Latin adjective is subject to modifications due to the features gender, number and case, all of which must agree with the noun both in predicative and attributive position.

In French, the case has disappeared, but the features number and gender must still agree with the noun in both positions.

In German, the attributive adjective is modified by the features number, gender and case, all of which must agree with the noun, but the predicative adjective is not. Thus we have

der gute Mann, *der Mann ist gut.*

An additional complication in German distinguishes between noun-phrases beginning with the definite and indefinite article:

der gute Mann, *ein guter Mann.*

Being a native German speaker, I don't recall ever being taught a rule for how this is done and only became aware of the distinction when asked to teach a course on scientific German.

Fortunately, the English adjective is completely invariant and carries no features whatsoever, perhaps contributing to turn English into a universal lingua franca.

Genuine English adjectives may be modified by certain adverbs such as *very* and *more*, as in

the very good wine, *the wine is more tasty*,

thus ruling out the number words discussed in an earlier chapter as being adjectives. The fact that *more good* is usually replaced by *better* is another story we will not go into here.

The present participles of intransitive verbs and the past participles of transitive verbs also resemble adjectives, but fail our test. We can say

the sleeping dog, *the killed rabbit*,

but not

*the *very sleeping dog*, *the *very killed rabbit*.

Yet there are transitive verbs whose present and past participles are both genuine adjectives, e.g. the verb *interest*, since we can say

the very interesting book, *the very interested reader*.

In fact, there is a surprisingly large class of transitive verbs with this property. Searching my own subconscious memory, I came up with the following partial list:

alarm, amuse,annoy, astonish, bewilder, bother, charm, confuse, convince, depress, disappoint, discourage, disillusion, displease, distress, disturb, edify, encourage, excite, exhaust, fascinate, flabbergast, frighten, frustrate, infuriate, interest, intimidate, intoxicate, oblige, perturb,please, puzzle, relieve, satisfy, scare, surprise, tantalize, tempt, threaten, torment, trouble, worry.

Thus we might say

I was very puzzled to hear this very amusing joke.

How do English speakers know which transitive verbs belong to this class? Consulting a printed dictionary turns out to be only partially reliable. But introspection reveals that the verbs listed above have something in common: they all refer to the *causation of a mental or emotional state*. For example,

she frightens him

means

she causes him to be afraid.

To test this explanation, I submit the verbs *move* and *touch* that can have two distinct meanings. We cannot say

*the engine was *very moving*, *the train was *very moved*

but we can say

her performance was very moving, *the audience was very moved*

and similarly when *move* is replaced by *touch*. For both verbs it is only the metaphorical meaning that refers to an emotional state.

As another test, may I suggest a *psycho*-linguistic experiment? People suffering from Asperger's syndrome are supposed to have difficulty in realizing other people's emotional states. I wonder whether they will come up with the same list of transitive verbs?

6 Adverbs and prepositions

In French one can form an adverb from an adjective by attaching the ending +*ment*, evidently derived from Latin *mens*, meaning "mind". Curiously, in English the same can be done with the ending +*ly*, ultimately derived from a German root meaning "body", still surviving with the meaning "corpse" in the first part of English *lychgate* and in the German *leiche*. Does this suggest that French and English find themselves on opposite sides of the mind-body debate?

Prepositions may be viewed as transitive adverbs, inasmuch as they require an object to turn into an adverbial phrase. For example, English *of* and *to* serve to replace earlier genetive and dative cases, which still survive in German.

Similarly, Hebrew has used prepositions to replace earlier cases still surviving in Arabic. Biblical Hebrew even employs a preposition to replace the accusative case by an adverbial phrase.

Other prepositions may serve a different function and may exhibit certain peculiarities. Consider for instance the English sentence

Alice saw a spider behind her.

Here the pronoun *her* could refer to Alice or to her friend Julia, who might have been mentioned in a previous sentence. But this is not so in the sentence

Alice bought a purse for her.

Clearly *behind* refers to a spacial location and *for* does not.

There is a common belief in a connection between the structure of a language and the way native speakers of that language see the world. A strong form of this opinion is known as the *Sapir-Whorf hypothesis*,which asserts that language shapes the speaker's world-view.

Yet, some arguments that have been proposed in favour of this hypothesis seem to be contradicted by actual observation. Several authors subscribe to the belief that Aymara, a language spoken in Peru and Columbia, stands out as seeing the future behind and the past in front, whereas familiar languages have been asserted to do the opposite. In fact, familiar European languages such as English and Latin resemble Aymara in this respect. The English preposition *before* and the Latin preposition *ante* refer both to the past and the front, whereas English *after* and Latin *post* may refer both to the future and the back. Think of nautical *aft* and cognate German *Afler*, which usually denotes an orifice in the posterior.

Even biblical Hebrew starts the book of Genesis with the word *bereshit*, meaning "in the beginning", the ultimate past, where the root *resh* or *rosh* means "head".

7 Capulet semantics

Semantics is supposed to derive the meaning of sentences from the meaning of words, as stated in the dictionary, mental or printed. Since the meaning of sentences has been famously discussed by Montague, I will call the latter *Capulet semantics*.

In languages such as English, which has more or less abolished inflectional morphology, a word listed in the dictionary may have several distinct syntactic types and several distinct meanings. For example, *sound* can be a noun, with an acoustic or nautical meaning, a verb or an adjective. Strictly speaking, we should regard these lexical entries as representing separate words.

The meanings of many words are widely misunderstood. For example

Midsummer night is not in the middle but at the beginning of summer.

The *north magnetic pole* of the earth is not near the geographic north-pole, but in Antarctica. For, if earth and compass needle are both regarded as magnets, it is the south pole of the former which attracts the north pole of the latter.

To *decimate* does not mean to reduce *to* 10%, but *by* 10% .

Rule of thumb is listed in the Oxford dictionary as referring to a "rough practical method", but it originally referred to the thickness of a stick with which a man was permitted to beat his wife.

Of course, many words change their meaning over time. For example, the word *gay* changed its meaning from "joyful" to "homosexual" only quite recently.

This becomes most obvious when we compare two languages, such as English and German, with a common origin less than 2000 years ago. We notice that English

clean is cognate with German *klein* (= small),

gate is cognate with German *Gasse* (= alley),

knight is cognate with German *Knecht* (= servant),

knave is cognate with German *Knabe* (= boy),

ordeal is cognate with German *Urteil* (= judgement),

small is cognate with German *schmal* (= slender),

stove is cognate with German *Stube* (= parlour),

timber is cognate with German *Zimmar* (= room),

town is cognate with German *Zaun* (= fence).

My favourite English example of a historical change in meaning, is the transformation of *long beard* into *lumber*. This involved the long bearded Teutonic tribe, the *Langobardi*, who invaded Northern Italy and settled in *Lombardi*. The prosperous *Lombards* founded banks elsewhere in Europe, which later deteriorated into pawnshops and ultimately into junk shops, the junk consisting mostly of *lumber*.

For amusement I have been collecting a number of pairs of words that look alike but have opposite meaning.

cleave can mean "stick together" or "rend apart";

witch can mean "ugly old woman" or "charming young woman";

sanction can imply approval or disapproval;

homely means "humble" in Britain but "ugly" in America;

moot is listed as "debatable" in many dictionaries, but seems to mean "no longer debatable";

to table a motion can mean "to submit a motion for later discussion" or "to dismiss it".

That two words pronounced similarly can have different meanings is often exploited as a pun. Usually puns are intended as jokes. For example, when a jester had been reprieved from the gallows on condition that he would stop punning, he declared (in an American accent): "no noose is good news". Some puns convey useful information. For example, when Shakespeare declares "there are more reasons than blueberries", we may infer that in his time *reason* was pronounced like *raisin*. Some puns are bilingual. For example, after conquering a province in what is now Pakistan, the British military commander sent a one-word telegram saying *peccavi*, a Latin expression meaning "I have sinned".

More serious are puns intended to be instructive. For example, the first book of the Hebrew bible contains an elaborate double pun based on the fact that the Hebrew words for "knowledge" and "intercourse" are the same, as are the words for "cunning" and "naked". When the usual translation tells us that Adam and Eve, after eating from the tree of knowledge, discover that they are naked, it would make more sense if *knowledge* was replaced by *intercourse* or *naked* by *cunning*. On the other hand, the puzzling assertion that the serpent is the most cunning of all animals makes more sense if *cunning* is replaced by *naked*.

Some puns may not have been intended at all. When the Canadian treasury issued a new two-dollar coin, one of my sons drew my attention to the fact that it bore the picture of the queen with a *bear* behind.

Part II. Verbal inflections.

8 Mediterranean conjugation

Those of us who were brought up on the study of Latin were left with the impression that the essence of grammar is not syntax (how sentences are formed from words) but inflectional morphology (how words must be modified to prepare them for their rôle in sentence formation). In particular, nouns and adjectives are modified by what is called "declension", and verbs are modified by what is called "conjugation". Having touched upon declension in an earlier chapter, we will now turn our attention to conjugation in three circum-mediterranean languages: Latin, Turkish and Hebrew.

Although the Eurasiatic language family, to which both Latin and Turkish belong, is not particularly closely related to the Afro-asiatic family, which embraces both Hebrew and Arabic, we do notice a common way of classifying inflected verb forms according to aspect, tense and person, which are completely absent in some languages such as Chinese. Thus Latin, Turkish and Hebrew admit "finite" conjugational verb-forms $C^k_{ji}(V)$, where V refers to the verb, k to the aspect, j to the tense and i to the person. However, the details differ in the three languages under consideration: the number of aspects, tenses and persons is $(3, 5, 6)$ in Latin, $(4, 8, 6)$ in Turkish and $(7, 2, 10)$ in Hebrew.

In each of these languages the so-called finite verb-forms are derived from a verb stem. In Latin there are two stems, one for aspects 1 and 3, and one for aspect 2. In Turkish there is just one stem, obtained from the infinitive by chopping off the ending + *mak*. In Hebrew, as in Arabic and perhaps other Semitic languages, the stem depends on the aspect and tense, according to a formula made up from the verb, which is completely described by 3 consonants and 2 vowels, from which all verb-forms can be calculated mathematically, all apparent exceptions being explicable by certain phonological rules.

In Latin, the finite verb-forms belonging to each of the aspects *amo, amor* and *amavi* may be calculated by somewhat diverging mechanisms that have to be memorized. In Turkish, the different verb-forms are produced by stringing together a number of morphemes, subject only to certain phonological rules, such as vowel harmony which insists that each vowel is influenced by the vowel in the preceding syllable. In Hebrew and Arabic, all finite verb-forms can be calculated from the verb presented by three consonants and two vowels.

The following formulas will account for the *finite verb-forms*, excluding infinitives and participles, in the languages under consideration. They have been slightly altered from the articles in which they were first proposed, in order to facilitate comparison between the three languages. From now on, V stands for verb, S for stem, T for tense and P for person.

In Latin we postulate the formula

$$C^k_{ji}(V) \to S^k(V) T^k_j P^k_i, \tag{8.1}$$

where k ranges from 1 to 3, i from 1 to 6 and j from 1 to 5, to account for a total of 90 different finite verb-forms.

In Turkish we postulate

$$C^k_{ji}(V) \to S(V) T_j A^k P_i, \tag{8.2}$$

where k ranges from 1 to 3, j from 1 to 8 and i from 1 to 6, to account for a total of 384 finite verb-forms. Now A stands for aspect.

In Hebrew we postulate

$$C^k_{ji}(V) \to P_{ji} S^k_j(V) Q_{ji}, \tag{8.3}$$

where k ranges from 1 to 7, $j = 1$ or 2 and

$i =$ 1, 2m, 2f, 3m, 3f, 4, 5m, 5f 6m, 6f

to account for a total of 140 finite verb-forms. Here both P and Q stand for person, m for masculine and f for feminine.

Let us now turn to some detailed calculations for the three languages in examples chosen for the sake of illustration. For a more complete description, the reader is referred to the literature: Lambek [1975], Bargelli and Lambek [2004] and Lambek and Yanofsky [2008].

9 Latin conjugation

In the sample calculations below we follow [L1975], but ignore the distinction between long and short vowels and make some other simplifications.

For the translation from Latin to English, we bear in mind that the three Latin aspects are what we call *normal, passive* and *perfect* respectively. The five tenses are the *present* and *past indicative*, the *present* and *past subjunctive* and the *future*. The six persons are the three persons *singular* followed by the three persons *plural*. Note that the stems $S^1(V)$ and $S^3(V)$ agree, but that $S^2(V)$ has to be memorized separately.

Some Latin verbs are *defective*, inasmuch as they do not admit all three aspects. For example *memini* (I remember) exists only in the third (perfect) aspect and *nascor* (I am born) only in the second (passive) aspect.

Even if almost all the stems $S^k(V)$ have been stored in the mental dictionary, there still remain some irregularities that have to be memorized, namely for the common verbs *sum, edo, possum* and *volo*, the last together with *nolo* and *malo*.

The formula (8.1) gives rise to the following sample calculations:

$$\begin{array}{rcl} C^1_{36}(amo) & \to & S^1(amo)\ T^1_3 P^1_6, \\ & \to & ama\ T^1_3 P^1_6, \\ & \to & amaba\ P^1_6, \\ & \to & amabant, \end{array}$$

meaning "they loved", using $T^1_3 \to$ *be* after *a* or *e* and $P^1_6 \to nt$.

$$\begin{array}{rcl} C^3_{51}(do) & \to & S^3(do)\ T^3_5 P^3_1, \\ & \to & ded\ T^3_5 P^3_1, \\ & \to & deder\ P^3_1, \\ & \to & dedero, \end{array}$$

meaning "I will have given", using $T_5^3 \to er$ and $P_1^3 \to o$ after a consonant.

$$\begin{aligned} C_{13}^2(lego) &\to S^2(lego)\, T_1^2 P_3^2, \\ &\to leg\, T_1^2 P_3^2, \\ &\to legi\, P_3^2, \\ &\to legitur, \end{aligned}$$

meaning "it was read", using $S^2(lego) \;\to\; leg$,

$$\begin{aligned} T_1^2 &\to i \text{ except after } a \text{ or } e, \\ P_3^2 &\to tur. \end{aligned}$$

10 Turkish conjugation

We will now turn to the conjugation of the Turkish verb, following [BL2001 and 2007], but first we take a look at the phonetic rule of vowel harmony, which ensures that certain vowels are chosen to be in accordance with vowels appearing in a preceding syllable. Similar rules exist in other members of the Ural-Altaic subfamily of the Eurasiatic superfamily.

Turkish vowels are classified by *height, low* or *high, shape, flat* or *round*, and *depth, back* or *front*. Each vowel V_{hsd} is completely described by its height h, shape s and depth d, where h, s and d are all equal to 1 or 2.

Thus

$$\begin{array}{ll} V_{111} = \imath & V_{211} = a \\ V_{112} = i & V_{212} = e \\ V_{121} = u & V_{221} = o \\ V_{122} = \ddot{u} & V_{222} = \ddot{o}. \end{array}$$

Here the Greek letter $\imath$ stands for a short i and $\ddot{u}$ and $\ddot{o}$ are pronounced as in German. In addition, there are two variable vowels which are here denoted by the Greek letters ε and ω and which have been called *amphibian* and *chameleon* respectively.

The rule of vowel harmony states that

$$\varepsilon \to V_{11d} \text{ after } V_{hsd}C,$$

C being any cluster of consonants, and

$$\begin{aligned} \omega \to\ & V_{2sd} \text{ after } V_{hsd} \\ \to\ & \emptyset \text{ otherwise.} \end{aligned}$$

In other words,

$\varepsilon \to a$ or e, depending on whether the vowel in the preceding syllable is back or front,

$\omega \to$ the high vowel of the same shape and depth as the preceding syllable, but it is swallowed if there is no intermediate consonant.

In addition there are a few other phonological rules. It would be tedious to state them all here, since, in the illustrations below, we will use only two:

$$\begin{array}{rcl} (d) & \to t & \text{after a voiceless consonant,} \\ & \to d & \text{otherwise,} \\ (y) & \to y & \text{between two vowels or between} \\ & & \text{a vowel and a voiced } d \text{ or } m, \\ & \to & \emptyset \text{ otherwise.} \end{array}$$

The different finite verb-forms in Turkish may be calculated with the help of formula (8.2), where $S(V)$ is the verb stem, obtained by deleting the ending $+mak$ of the infinitive. There are six persons P_i as in Latin, eight tenses T_j called present, aorist, future, preterit, dubitative past, optative, subjunctive and necessitative, and four aspects A^k called simple, narrative, reportative and conditional.

In the two illustrations below, we will make use of the following computational rules for T_i, A^k and P_j:

$$\begin{aligned}
T_4 \to &\quad (d)\omega y \text{ before } A^2 \text{ or } A^3,\\
T_7 \to &\quad (y)\varepsilon(y),\\
A^2 \to &\quad (d)\omega n \text{ before } P_2,\\
\to &\quad (d)\omega \text{ before } P_i \text{ when } i \neq 2,\\
A^3 \to &\quad m\omega\ \d{s},\\
P_1 \to &\quad m,\\
P_2 \to &\quad \emptyset \text{ after } n\\
P_6 \to &\quad \ell\varepsilon r.
\end{aligned}$$

We are now ready to carry out two sample calculations:

$$\begin{aligned}
C^2_{41}(yazmak) &\to \quad yaz\ T_4A^2P_1,\\
&\to \quad yaz\ (d)\omega yA^2P_1,\\
&\to \quad yaz\ d\omega yA^2P_1,\\
&\to \quad yaz\ diyA^2P_1,\\
&\to \quad yazdiyd\omega P_1,\\
&\to \quad yazdiydiP_1,\\
&\to \quad yazdiydim \text{ (I wrote).}
\end{aligned}$$

Note that the first person ending $+m$ is shared by several Eurasiatic languages. It prevails in Polish, survives in the Latin past and *sum*, as well as in English *am*.

$$\begin{aligned}
C^3_{76}(kalmak) &\to \quad kal\ T_7A^3P_6\\
&\to \quad kal\ (y)\varepsilon(y)A^3P_6\\
&\to \quad kala\ (y)A^3P_6\\
&\to \quad kala\ (y)m\omega\d{s}P_6\\
&\to \quad kala\ ym\omega\d{s}P_6\\
&\to \quad kalayml\d{s}P_6\\
&\to \quad kalaymi\d{s}\ell\varepsilon r\\
&\to \quad kalaym\ell\d{s}\ell ar
\end{aligned}$$

([that] they be said to stay).

I hope the above two examples will convey an idea how all the finite Turkish verb-forms can be calculated. It should be pointed out that Turkish nouns can also be treated like verbs as regards conjugation. Thus *doktorum* means "I am a doctor" and *adamdi* means "it was a man".

11 Hebrew conjugation

Conjugation in biblical, as well as in modern Hebrew, is given by the formula (8.3), see [LY2008]. As in Arabic [BL2001], the verb V is completely described by three consonants F (first), M (middle) and L (last) and two vowels α_1 and α_2. We write

$$V = (FML, \alpha_1\alpha_2).$$

There are seven aspects, numbered $k = 1$ to 7:

normal active, normal passive, causative active, causative passive, intensive active, intensive passive, reflexive.

There are two tenses, numbered $j = 1, 2$, usually called "past" and "future", although the so-called future rather corresponds to the Arabic "perfect", a term that is better avoided since it may be confused with the Latin perfect, which corresponds to the English compound past.

There are ten persons numbered

$$i = 1, 2m, 2f, 3m, 3f, 4, 5m, 5f, 6m, 6f,$$

referring to five persons singular followed by five persons plural. The gender distinction is indicated by m (masculine) and f (feminine), wherever it is required.

To determine the stem $S_i^k(V)$ of the verb $V = (FML, \alpha_1\alpha_2)$ we require a list of *stem rewrite rules*, only two of which will be made use of here:

$$\begin{aligned} S_1^1(V) &\to iFM\alpha_1 L, \\ S_2^5(V) &\to FiMM(e)L, \end{aligned}$$

subject to the phonological rule

$$\begin{aligned} (e) \to &\quad a \text{ before two consonants,} \\ \to &\quad e \text{ before a single consonant at the end.} \end{aligned}$$

The vowels α_1 and α_2 have to be a, e or o. Originally, biblical Hebrew had only three vowel *phonemes*, as modern standard Arabic still has. Unfortunately, about 800 AD the Masorites insisted on preserving the vowel *allophones* in their spelling. This makes it difficult for non-experts to determine the original phonemes in many contexts.

To compute the prefixes P_{ji} and the suffixes Q_{ji}, we require another list of rewrite rules, only four of which will be made use of here:

$$\begin{aligned} P_{2,1} &\to \emptyset, \quad & Q_{2,1} &\to tiy, \\ P_{1,5m} &\to t, \quad & Q_{1,5m} &\to uw. \end{aligned}$$

We look at two sample calculations for the regular verb $V = (q_1 \underset{.}{t}\ell, oa)$ meaning "to kill":

$$\begin{aligned} C_{2,1}^5(V) \to &\quad P_{2,1} S_2^5(V) Q_{2,1}, \\ \to &\quad S_2^5(V) Q_{2,1}, \\ \to &\quad qi\underset{.}{t}\underset{.}{t}(e)\ell Q_{2,1}, \\ \to &\quad qi\underset{.}{t}\underset{.}{t}a\ell Q_{2,1}, \\ \to &\quad qi\underset{.}{t}\underset{.}{t}a\ell tiy, \end{aligned}$$

translated into "I was murdered".

$$\begin{aligned} C_{1,5m}^1(V) \to &\quad P_{1,5m} S_1^1(V) Q_{1,5m}, \\ \to &\quad t S_1^1(V) Q_{1,5m}, \\ \to &\quad tiFM\alpha_1 L Q_{1,5m}, \\ \to &\quad tiq\underset{.}{t}o\ell Q_{1,5m}, \\ \to &\quad tiq\underset{.}{t}o\ell uw, \end{aligned}$$

translated into "you (males) killed".

The above illustrations were chosen to avoid bringing in phonological rules other than those for (e). With the help of a few other phonological rules one may easily calculate the finite forms of all *regular* verbs, by which I mean verbs not containing the semivowels w and y, any guttural consonants, nor the letter n. To handle these so-called *weak* verbs, further phonological rules are required, e.g.

$$h \to \emptyset \text{ after a consonant.}$$

Hebrew and Arabic are still remarkably similar, just like Spanish and Portuguese. It seems that native Arabic speakers can understand Hebrew, just as Portuguese speakers can understand Spanish, but the converse does not appear to be the case.

As it turns out, three consonants will not suffice for late biblical and modern Hebrew, and the same is true for Arabic. Both Arabic and Hebrew have employed the same strategy to accommodate a fourth consonant, by allowing MM in aspects 5, 6 and 7 to be replaced by M_1M_2, where $M_1 \neq M_2$. Not surprisingly, the four-consonant verbs exist only in those three aspects.

Part III. From word to sentence.

12 Production grammars

Traditional grammarians of classical languages were mostly concerned with inflectional morphology of words, such as declension of nouns, adjectives and pronouns, and conjugation of verbs. However, influenced by the prevailing Chomskian orthodoxy, recent grammarians seem to be more interested in *syntax*, which deals with how words are put together to form sentences. I myself have pursued three mathematical approaches to syntax over the years: *production grammars*, the *syntactic calculus* and *pregroup grammars*, which will be discussed briefly in this and subsequent chapters.

Production grammars are usually presented in a form in which the sentences of a natural language are generated by certain *rewrite* rules from a symbol S, standing for "sentencehood", in a metalanguage that incorporates certain auxiliary grammatical terms. The most popular production grammars are *contextfree*, where all rewrite rules have the form $A \to B$ or $A \to BC$. For example, we may adopt the rule

$$\begin{aligned} S \to &\quad \text{Subj Pred}, \\ \text{Subj} \to &\quad \textit{John}, \\ \text{Pred} \to &\quad \text{V}_{\text{trans}}\text{Obj}, \\ \text{V}_{\text{trans}} \to &\quad \textit{likes}, \\ \text{Obj} \to &\quad \textit{her}, \end{aligned}$$

with the obvious abbreviations for subject, predicate, transitive verb and object, yielding

$$\begin{aligned} S \to &\quad \text{Subj Pred} \\ \to &\quad \textit{John}\ \text{Pred} \\ \to &\quad \textit{John}\ \text{V}_{\text{tr}}\text{Obj} \\ \to &\quad \textit{John likes}\ \text{Obj} \\ \to &\quad \textit{John likes her}. \end{aligned}$$

More realistically, we may forget about context-freeness and incorporate some inflectional morphology into the grammar, as in the following rewrite rules:

$$\begin{aligned} S \to &\quad \text{Subj Pred}, \\ \text{Subj} \to &\quad \text{NP}\ P_3, \\ \text{NP} \to &\quad \textit{John}, \\ \text{Pred} \to &\quad T_1\ \text{VP}, \\ P_3 T_1 \to &\quad C_{13}, \\ \text{VP} \to &\quad V_{\text{trans}}\ \text{NP}, \\ \text{V}_{\text{trans}} \to &\quad \textit{like}\ \text{Acc}, \\ C_{13}\ \textit{like} \to &\quad \textit{likes}, \\ \text{NP} \to &\quad \textit{she}, \\ \text{Acc}\ \textit{she} \to &\quad \textit{her}, \end{aligned}$$

where

NP = noun phrase

P_3 = third person singular,

T_1 = present tense

C_{13} = present tense third person singular,

Acc = accusative.

These rules yield

$$\begin{aligned}
S \to\ & \text{Subj Pred} \\
\to\ & \text{NP } P_3 \text{ Pred} \\
\to\ & \textit{John } P_3 \text{ Pred} \\
\to\ & \textit{John } P_3 T_1 \text{ VP} \\
\to\ & \textit{John } C_{13} \text{ VP} \\
\to\ & \textit{John } C_{13} \text{ V}_{\text{trans}}\text{NP} \\
\to\ & \textit{John } C_{13}\ \textit{like} \text{ Acc NP} \\
\to\ & \textit{John likes} \text{ Acc } \textit{she} \\
\to\ & \textit{John likes her.}
\end{aligned}$$

This example will help to illustrate the insight by Miller [1956] that the number of items in the short term temporary memory must be restricted. Thus, having uttered *John*, we may have to store four items C_{13} *like* $A_{cc}NP$ in the memory, and uttering *likes* these reduce to three: *likes* A_{cc} *she*. Miller asserted that the maximum number of items that can be stored in the temporary memory are about seven and he called them "chunks of information". This may be verified by anyone who is trying to dial a ten-digit telephone number.

Miller's limitation alone does not suffice to avoid all sentences that are difficult to analyze. Who would suspect that

(1) $$fish\ fish\ fish\ fish\ fish$$

u can be analyzed as a grammatical English sentence in two different ways, one of which is even a tautology?

Readers will not find it difficult to construct a production grammar that will generate the tautology

(1) $$seals\ hunt\ fish\ [that]\ seals\ hunt$$

and the more informative sentence

(2) $$seals\ [that]\ whales\ hunt\ hunt\ fish.$$

Now replace the plural nouns *seals* and *whales* by the plural *fish* and take advantage of the fact that the rules of English grammar allow us to omit the relative pronoun [*that*] when introducing a sentential object. Then both (2) and (3) will be transformed into (1).

An amusing mathematical puzzle asks to show that $fish^{2m+1}$ may be analyzed as a grammatical English sentence in $f(m+1) = \frac{1}{m+1}\binom{2m}{m}$ different ways, when $f(m+1)$ is the so-called *Catalan* number, as was first proved by Telyn Kusalik.

A production grammar may be viewed as a special kind of partially ordered algebraic system, namely a *freely generated partially ordered semigroup*, also known as a "semi-Tue system".

Let me remind the reader that a (multiplicative) *semigroup* is a set endowed with a binary operation denoted by $\cdot$ or just by juxtaposition, satisfying the *associative* law

$$(ab)c = a(bc).$$

A *partially ordered* semigroup also possesses a binary relation $\to$ satisfying

$$a \to a, \quad \frac{a \to b \;\; b \to c}{a \to c}, \quad \frac{a \to b \;\; b \to a}{a = b}, \quad \frac{a \to b \;\; c \to d}{ac \to bd}.$$

The semigroup is called a *monoid* if it contains an element 1 such that

$$a1 = a = 1a.$$

In an *additive* semigroup or monoid, $\cdot$ is replaced by $+$ and 1 by 0. A *group* is a monoid in which each element a has an inverse, written a^{-1} or $-a$:

$$a^{-1}a = 1 = aa^{-1}, \;\; -a + a = 0 = a + 0.$$

13 Anaphora

A basic syntactical question ignored by traditional grammarians is the following:

(0) When can a pronoun represent a noun phrase A in the same sentence?

For example, in the sentence

(1) *although John*$_1$ *did not like her*$_2$, *he*$_1$ *kissed the girl*$_2$

her$_2$ can represent *the girl*$_2$ and *he*$_1$ can represent *John*$_1$, as is indicated by the numerical subscripts. On the other hand in the reverse order sentence

(2) *he kissed the girl*$_2$, *although John did not like her*$_2$,

her$_2$ can still represent *the girl*$_2$, but *he* can no longer represent *John*.

The answer to question (0) can be phrased as a necessary and sufficient condition in a rudimentary production grammar which allows us to recognize certain *key* constituents of a sentence, by which we will understand the following:

(DS) a direct sentence,

(NP) a noun-phrase,

(SC) a subordinate clause,

(QS) a quasi-sentence,

(STP) a spatio-temporal prepositional phrase.

We also ought to consider

(IS) an indirect sentence

like *that he snores*, which may be viewed as a subordinate clause or as a sentential noun phrase.

We will now state a set of conditions (C0) to (C3) which I believe to be necessary and sufficient for a pronoun to represent a noun-phrase A in the same sentence:

(C0) The pronoun agrees with A in the features person, number and gender.

(C1) The pronoun is properly contained in the key constituent $B \neq DS$ not containing A.

(C2) A is properly contained in the key constituent $C \neq STP$ and the pronoun comes *after* C.

(C3) The pronoun and A are not contained in the same minimal key constituent.

The necessity of (C0) is obvious and need not be discussed further. We proceed to look at some illustrations of (C1):

$John_1$ *kissed* (*the girl* [*that*] he_1 *admired*$)_{NP}$;

$(his_1$ $girlfriend)_{NP}$ *kissed* $John_1$;

$(John_1$ *said* ([*that*] he_1 *liked Jane*$)_{IS}$;

(*that* he_1 *liked Jane*$)_{IS}$ *was clear to* $John_1$;

$John_1$ *drinks tea* (*when* he_1 *is thirsty*$)_{SC}$;

(*when* he_1 *is tired*$)_{SC}$ $John_1$ *goes to bed*;

$Alice_1$ *heard a rabbit* (*behind* $her_1)_{STP}$;

(*in front of* $her_1)_{STP}$ *Alice saw a rabbit*;

$Elizabeth_1$ *distrusted the queen* (*before* $her_1)_{STP}$

However, the exception $B \neq DS$ must be observed, as in the following example:

$$^*She_1 \textit{ likes John and } (Jane_1 \textit{ kissed him})_{DS}$$

where *she* cannot refer to *Jane*.

Recall that *with* may be considered as a spatial preposition, but not *for*, since we may have

$$Alice_1 \textit{ brought a purse with } her_1$$

but not

$$Alice_1 \textit{ bought a purse for } {}^*her_1.$$

Here *her* may refer to *Alice* in the first example, but not in the second.

We now turn to some illustrations of (C2):

$(John_1$ *met* $Jane_2)_{DS}$ *and* she_2 *kissed* him_1;

$(John's_1$ $girlfriend)_{NP}$ *kissed* him_1;

(*when* $John_1$ *sleeps*$)_{SC}$ he_1 *snores.*

Note that in condition (C2) the pronoun must occur after C. For example, the following is not permitted:

$$^*he_1 \textit{ snores when } (John_1 \textit{ sleeps})_{SC}.$$

What if $C = STP$? The following example shows that *she* cannot refer to Alice:

$$(\textit{behind } Alice_1)_{STP} \ {}^*she_1 \textit{ saw a rabbit.}$$

Finally, let us look at some illustrations of (C3):

I claim that $({}^*he_1$ *saw* $John_1)_{DS}$;

$I\ said\ that\ (John_1\ saw\ {}^*him_1)_{DS}$;

$({}^*he_1\ and\ John_1's\ brother)_{NP}\ snore$;

$({}^*his_1\ and\ John's\ brother)_{NP}\ snore$.

Here *he, him* and *his* cannot refer to *John*.

Special consideration must be given to the cases when B and C are what I have called quasi-sentences (QS), as in the following illustrations:

$John_1\ let\ Jane_2\ (\sharp_2\ kiss\ him_1)_{QS}$;

$John_1\ wanted\ [for]\ Jane_2\ (\sharp_2\ to\ kiss\ him_1)_{QS}$;

$John_1\ wanted\ ([for\ him_1]\ \sharp_1\ to\ kiss\ Jane)_{QS}$;

but

$John_1\ let\ {}^*him\ (\sharp_1\ kiss\ Jane)_{QS}$;

$John_1\ let\ Jane_2\ (\sharp_2\ kiss\ {}^*her_2)_{QS}$;

$John_1\ wanted\ (\sharp_1\ to\ kiss\ {}^*him)_{QS}$.

Here the symbol $\sharp$ serves as the subject of the quasi-sentence, roughly corresponding to Chomsky's PRO.

14 Syntactic calculus

An algebraic system with linguistic applications goes back to K. Ajdukiewicz and Y. Bar-Hillel, but is now best described as a *residuated monoid*: it is a partially ordered monoid with a partial order $\to$, a multiplication $\otimes$ and two additional binary operations / (over) and \ (under) satisfying

$$a \otimes b \to c \text{ iff } a \to c/b \text{ iff } b \to a\backslash c.$$

For linguistic purposes one is interested in the residuated monoid of *types* freely generated by a partially ordered set of *basic types*. Originally this was the discrete set $\{n, s\}$, where s stood for declarative sentence and n for name or noun-phrase. Here is a simple illustration:

$$\begin{gathered} John\ likes\ her \\ n \otimes (n\backslash(s/n)) \otimes ((s/n)\backslash s) \to s, \end{gathered}$$

which is easily proved by noting the *contractions*

$$n \otimes (n\backslash x) \to x,\ \ x \otimes (x\backslash s) \to s,$$

where $x = s/n$.

Of course, we can replace *her* by *Jane*, and this is justified by the *expansion*

$$n \to (s/n)\backslash s.$$

When I first proposed this system, the referee Bar-Hillel pointed out that the presence of such expansions necessitated a search for a decision procedure. Fortunately such a procedure had already been found for the intuitionistic propositional calculus by Gentzen, based on his famous "cut elimination theorem".

The proof of cut elimination was even easier for the present system, which could be thought of as a kind of "substructural" logic, with $\otimes$ standing for conjunction and / and \ for implications $\Leftarrow$ and $\Rightarrow$. Gentzen's original system had contained three *structural* rules, interchange, weakening and contraction, based on the usual logical properties of conjunction:

$$a \otimes b \to b \otimes a,\ a \otimes b \to a,\ a \to a \otimes a.$$

The absence of these structural rules simplified the proof of cut elimination considerably.

Note that the arrow in the syntactic calculus is the reverse of the arrow in a production grammar, indicating that we are dealing with *recognition* rather than *generation* of sentences.

Gentzen's argument involved deductions of the form $\Gamma \to b$, where $\Gamma = a_1 \cdots a_n$ and b might be replaced by the empty string $\emptyset$. Moreover, the operations $\otimes, /$ and $\backslash$ may be introduced by the following *introduction* rules:

$$\frac{\Gamma \to a\ \ \Delta \to b}{\Gamma\Delta \to a \otimes b}, \quad \frac{\Gamma a b \Delta \to c}{\Gamma(a \otimes b)\Delta \to c},$$

$$\frac{\Lambda b \to a}{\Lambda \to a/b}, \quad \frac{\Lambda \to b\ \ \Gamma a \Delta \to c}{\Gamma(a/b)\Lambda\Delta \to c}$$

and a symmetric rule for \. To these we must add the *identity* rules

$$\frac{\Lambda \to \emptyset}{\Lambda \to I}, \quad \frac{\Gamma\Delta \to c}{\Gamma I \Delta \to c}$$

when I is the unity element of the monoid.

The so-called *cut rule*

$$\frac{\Lambda \to a\ \ \Gamma a \Delta \to b}{\Gamma\Lambda\Delta \to b}$$

can now be shown to be redundant: it need not be postulated but can be proved. This is also so for the identity rules, except for basic types. Here is a sample cut-free proof to show that *John likes her* is a well-formed sentence:

$$\dfrac{n \to n \quad \dfrac{\dfrac{s/n \to s/n \quad \dfrac{\dfrac{n \to n\ \ s \to s}{(s/n)n \to s}}{s \to s}}{(s/n)(n(s/n)\backslash s) \to s}}{}}{s(n\backslash(s/n))(s/n\backslash s) \to s}$$

The symbols $\otimes$ and I have been avoided in this illustration.

The problem with such "proof trees" is that they occupy much space on pages or blackboards and are somewhat difficult to reproduce by people engaged in conversation. This is why I prefer the pregroup grammars developed in a later chapter to the above syntactic calculus, although this is still preferred by a number of mathematical linguists, possibly with additional modifications, because of its usefulness in semantics. Its set-theoretical interpretation is usually ascribed to Richard Montague, but can be traced back to Haskell Curry.

The syntactic calculus, as improved by Moortgat and others, has remained popular among a small coterie of computational linguists, largely motivated by its closeness to the set-theoretical interpretation that goes by the name of *Montague semantics*.

Recall that *Capulet semantics* deals with the meaning of words one finds in dictionaries. For example, a noun-phrase of type n may represent an element of the set N of entities, and an inflected transitive verb like *eats* or *kills* may represent a binary relation between entities, which may be thought of as functions $N \times N \to S$ or, equivalently, as elements of the set $S^{N \times N}$. More specifically, *kills* should be defined as the causative of *dies* and *eats* should somehow require that the entity to be eaten is solid. A word of warning to readers on both sides of the Atlantic: in America soup is eaten and tea is drunk, while in Britain soup is drunk and tea may be eaten, when it consists of cucumber sandwiches.

Montague semantics deals with the meaning of sentences. The main idea is very simple. Let us assume that n and s are the only basic types, that n denotes entities which are elements of a set N and that s denotes truth-values, which are elements of a set S. Suppose that types a, b and c have already been interpreted by sets A, B and C respectively, then $a \otimes b$ will be interpreted as the direct product $A \times B$ and c/b and $a \backslash c$ will be interpreted as sets of functions C^B and C^A respectively.

This is not the whole story, because Montague realized that, in order to accommodate ordinary discourse, one ought to abandon one of the rules of logic that permits substitution of equals for equals. For example, from the fact that "John was the murderer", we cannot infer that "the police know that John was the murderer".

Another problem with the syntactic calculus is that one ought to admit other basic types than n and s. To illustrate this, look at the simple sentence *Napoleon liked wine*. Here *Napoleon* is a name of type n, and *wine* is the name of a substance, say of type m, *liked* is an inflected form of the transitive verb *like* listed in the dictionary. Strictly speaking *liked* will have type $\pi_3 \backslash s_2 / o$, where

$\pi_3 =$ subject in the third person singular,
$s_2 =$ declarative sentence in the past tense
and $o =$ object.

Moreover, we require the ordering of basic types,

$$n \to \pi_3, \quad m \to o.$$

Then we may calculate the sentencehood of

$$\begin{array}{c} Napoleon\ liked\ wine \\ n\ (\pi_3 \backslash s_2 / o)\ m \end{array}$$

as follows:

$$\begin{array}{rcl} n(\pi_3 \backslash s_1 / o)m & \to & \pi_3(\pi_3 \backslash s_2 / o)m \\ & \to & (s_2/o)m \\ & \to & (s_2/o)o \\ & \to & s_2 \end{array}$$

15 Pregroup grammars

To accommodate Miller's restriction on the short term memory, we look at another algebraic system. We define a *pregroup* as a partially ordered monoid in which each element a has a *left* *adjoint* a^ℓ and a *right adjoint* a^r such that

$$a^\ell a \to 1 \to a a^\ell, \;\; a a^r \to 1 \to a^r a.$$

In a *pregroup grammar* we work with the pregroup of *types* freely generated from a partially ordered set of *basic types*. For the purpose of illustration, consider the following set of basic types:

$\pi_2 =$ second person subject,

$s_1 =$ statement in the present tense,

$s =$ declarative sentence,

$p_1 =$ present participle,

$p_2 =$ past participle,

$o =$ object,

$q =$ yes-or-no question,

$q' =$ question.

For the moment we also require that $q_1 \to q \to q', \;\; s_1 \to s$.

To each inflected word form one associates one or more types, enabling one to perform a calculation on the string of types underlying a string of words.

For example, consider the following three typed strings:

(1)
$$\begin{array}{l} \textit{you have been seeing her.} \\ \pi_2\ (\pi_2^r s_1 p_2^\ell)(p_2 p_1^\ell)(p_1 o^\ell)o \qquad \to s_1 \to s \end{array}$$

(2)
$$\begin{array}{l} \textit{have you seen her ?} \\ (q_1 p_2^\ell \pi_2^\ell)\pi_2(p_2 o^\ell)o \qquad \to q \to q' \end{array}$$

(3)
$$\begin{array}{l} \textit{whom have you seen} - ? \\ (q' o^{\ell\ell} q^\ell)(q_1 p_2^\ell \pi_2^\ell)\pi_2(p_2 o^\ell) \qquad \to q' \end{array}$$

Note that the word *have* has been assigned a different type in (1) than in (2) and (3). The double ℓ in (3) signals the presence of a Chomskyan trace, here indicated by a dash at the end of (3).

Here is how a mental calculation might proceed step by step for example (3):

$$\begin{array}{ccl} \textit{whom} \quad \textit{have} & & \\ (q' o^{\ell\ell} q^\ell) \;\; (q_1 p_2^\ell \pi_2^\ell) & \longrightarrow & q' o^{\ell\ell} p_2^\ell \pi_2^\ell, \\ \textit{whom have you} & & \\ (q' o^{\ell\ell} p_2^\ell \pi_2^\ell)\pi_2 & \longrightarrow & q' o^{\ell\ell} p_2^\ell, \\ \textit{whom have you seen} & & \\ (q' o^{\ell\ell} p_2^\ell)\ (p_2 o^\ell) & \longrightarrow & q' o^{\ell\ell} o^\ell \longrightarrow q' \end{array}$$

Note that in successive stages of the calculation the short term memory must store 6, 4, 5, 3, 5, 3, 1 types of length 1, which I am tempted to identify with Miller's *chunks of information.* Recall that Miller had famously declared that our short term memory cannot hold more than 7 ± 2 such chunks, as anyone trying to dial a long telephone number can verify.

It turns out that pregroup grammars admit a very simple decision procedure, allowing easy calculations proceeding from left to right, as in elementary algebra. This is based on the following observation, the so-called *Switching Lemma*, which Buszkowski has shown to be equivalent to a cut-elimination theorem.

In any calculation one may assume, without loss of generality, that all *contractions*

$$a^\ell a \to 1,\ aa^r \to 1$$

come before any *expansions*,

$$1 \to aa^\ell,\ 1 \to a^r a$$

It follows that as long as one wishes to reduce a string of types to a type of length 1, no expansions are needed at all. Yet, expansions may be required for theoretical discussions.

As it turns out, a kind of pregroup grammar without expansions had been anticipated by Zellig Harris. He did not even require repeated adjoints, which I found useful to account for Chomskyan traces, although Anne Preller later showed that repeated adjoints could be avoided by clever retyping.

Evidently a residuated monoid may be converted into a pregroup by putting

$$a/b = ab^\ell,\ b\backslash a = b^r a.$$

The converse requires ambiguity. For example, abc^ℓ may be converted to $a(b/c)$ or to $(ab)/c$. This allows us to transfer the semantic interpretation of the syntactic calculus to pregroup grammars, again with an inbuilt ambiguity. Thus abc^ℓ may denote $A \times C^B$ or $C^{A\times B}$, if a, b and c denote A, B and C respectively.

A more original interpretation of pregroup grammars has been proposed by Mehrnoosh Sadrzadeh and her collaborators. Their idea is to target the interpretations in a finite dimensional vector space and interpreting abc^ℓ as $A + B + C^*$, where

$$C^* = Hom(C, F),$$

F being the underlying field.

16 Multiple pregroups

Since Buszkowski (2001) proved that pregroup grammars are context-free, a problem arose with languages that are known not to admit a context-free grammar, such as Dutch and Swiss German. Now the best known formal language with this property happens to be the intersection of two context-free ones. It was therefore proposed by Ed Stabler (2008) and Brendan Gillon that it might be worthwhile to look at two or more pregroup grammars simultaneously and this idea was taken up by their students Kobele [2008] and Kusalek [2008]. I too found it useful to take another look at French by making two pregroup calculations simultaneously, the first with syntactic types as before and the second restricted to feature types.

The following illustrations will give an idea of this approach:

$$\begin{array}{ccccccc}
elle & peut & etre & heureuse & & \\
\pi_3 & (\pi_3^r s_1 j^\ell) & (i a^\ell) & a & \longrightarrow & s_1 \\
\pi_{3f} & & & \pi_f^r & \longrightarrow & 1
\end{array}$$

Here the bottom line is concerned with the genders of *elle* and *heureuse*;

$$\begin{array}{cccccccc}
nous & pouvons & avoir & embrass\acute{e} & la\ fille, & & \\
\pi_3 & (\pi_3^r s_1 j^\ell) & (i p_2^\ell) & (p_2 o^\ell) & o & \longrightarrow & s_1 \\
\pi_{4m} & & (\pi^r \pi_{3m}) & \pi_{3m}^r & & \longrightarrow & 1
\end{array}$$

where *embrassé* ignores the gender of *fille*. Note that $i \to j$, $\pi_{3f} \to \pi$, and $\pi_{4m} \to \pi$.

Part IV. Mathematical background.

17 Arithmetic versus geometry

The ancient Greeks employed positive integers, positive rationals and positive real numbers, although the last were originally conceived as ratios of geometric quantities. In Plato's Academy two methods were proposed for explaining the real numbers in terms of integers. One by Eudoxus, as rendered in Euclid's Elements essentially by a method now ascribed to Dedekind, where a real number is identified with a subdivision of the ordered set of rationals. The other method, ascribed to Theaetetus, introduced positive reals as continued fractions, which may now be considered as a special case of the Cauchy reals, namely convergent series of rationals. Continued fractions were forgotten for almost two millenia and have fallen out of favour again in recent revisions of elementary education.

The Greeks don't seem to have paid any attention to zero and negative numbers, which were only introduced in India a thousand years later. At first, the Pythagoreans thought that only rational numbers deserved to be studied, but reluctantly admitted irrationals when they realized that the square-root of two could not be written in the form p/q with integers p and $q \neq 0$. The amazingly simple proof of this result is found in the works of Aristotle.

The number system might have stopped with the real numbers, positive and negative, when the renaissance Italian scientist Cardono, better known for his contribution to medicine, realized that there were cubic equations with real coefficients and real solutions which could only be arrived at by intermediate steps involving the square-root of -1. This observation gave rise to the admission of complex numbers of the form $a + bi$, where a and b are real and $i^2 = -1$. For several centuries, this is where the number system stopped, since now all polynomial equations could be solved.

Let us return to the philosophical distinction between continuous and discrete entities. In mathematics this is reflected by the difference in emphasis placed on geometry and arithmetic. The ancient Greeks obtained remarkable results in geometry. There was the theorem attributed to Thales that the diameter of a circle subtends a right angle at the circumference and the theorem supposedly due to Pythagoras that the sum of the squares on two sides of a right-angled triangle equals the square on the hypothenuse. Moreover, Thales was able to exploit the properties of *similar* triangles, namely triangles having the same angles, to calculate the height of distant objects or the distance of objects of known height.

More spectacular were two applications of geometry to the universe later in Alexandria. Thus, Eratosthenes was able to calculate the circumference of the earth to within 5% accuracy by measuring the distance between Alexandria and Assuan due south and by comparing the angle between the sun and a vertical pole at Alexandria at the exact time when it was straight overhead at Assuan. Even more surprising was the realization by Aristarchus that the angle subtended at the earth between the sun and the half-moon must be a right angle, which enabled him to calculate the ratio of the distances of these two heavenly bodies from the earth, although not with great accuracy.

The Greeks obtained equally spectacular results in arithmetic, although at their time without any practical applications. They could prove that there were infinitely many *prime* numbers, that is positive integers with exactly two divisors. They were also able to construct all even *perfect* numbers, that is positive integers which are equal to the sum of their proper divisors, although the proof that they had done so was only discovered much later by Euler. The perfect numbers 6 and 28 were considered to be of mystical significance, since the world was reputedly created in 6 days and a four-week month had approximately 28 days. More recently, the discovery of new perfect numbers had proved to be a challenge to modern computers.

The positive integers can be added and multiplied but not subtracted and divided. The positive rationals and reals can also be divided. If we abandon positivity, subtraction can be carried out, but division by zero is forbidden, and we obtain what mathematicians call a *field*. If we also abandon commutativity of multiplication, we obtain a *skew field*, aka a *division ring*. We will look at more precise definitions of these terms in the next chapter.

18 Algebraic preliminaries

At one time, an *algebraist* was considered to be a bonesetter and some Spanish dictionaries still say so. However, the medieval Persian mathematician Al Khowarismi introduced the term *algebra* in the title of a book and it came to refer to a manipulation of symbols, originally in the process of solving equations.

A thousand years later, algebra came to deal with sets with structure, consisting of operations and sometimes also relations. Of special interest were *groups*, namely monoids $(A, 1, \cdot)$ in which every element a had an inverse a^{-1} such that $a^{-1}a = 1 = aa^{-1}$ or, if you prefer additive monoids $(A, 0, +)$, where each element a had an inverse $-a$ such that $-a + a = 0 = a + -a$. In particular, *Abelian* groups were groups satisfying the commutative law $a \cdot b = b \cdot a$, or $a+b = b+a$.

A *ring* $(A, 0, +, -, 1, \cdot)$ consists of an Abelian group $(A, 0, +, -)$ together with a monoid $(A, 1, \cdot)$ satisfying the distribution law

$$a(b + c)d = abd + acd.$$

A *division ring*, aka *skew-field* is a ring in which every nonzero element a has a multiplicative inverse a^{-1} such that $a^{-1}a = 1 = aa^{-1}$. It is called a *field* if multiplication is commutative.

It soon became clear that algebraists were interested not only in sets with structure but also in mappings between them that preserved the structure. In 1945, Samuel Eilenberg and Saunders Mac Lane introduced the notion of a *category* which consisted of objects and morphisms, namely arrows between objects.

The word *category* had been introduced by Aristotle with a different meaning and was a compound of the Greek words *cata* and *agora* meaning "down" and "forum" respectively. According to Eilenberg and Mac Lane, a *category* consists of *objects* $A, B, C, ...$ and *arrows* $f : A \to B,\ g : B \to C,\ h : C \to D, ...$ Moreover, arrows may be compared as follows:

$$\frac{f : A \to B \ g : B \to C}{gf : A \to C}$$

satisfying the associative law $(hg)f = h(gf)$. There were also assumed to be *identity* arrows $1_A : A \to A,\ !_B : B \to B, ...$ such that

$$f 1_A = f = 1_B f.$$

Among the categories Eilenberg and Mac Lane originally had in mind were the following:

the category of sets, whose objects are sets and whose arrows are functions;

the category of partially ordered sets, where arrows are order preserving functions;

the category of groups, whose arrows are homeomorphisms;

the category of topological spaces, whose arrows are continuous functions;

the category of small categories, whose arrows are functors.

A category is said to be *small*, if the class of objects is a set. A *functor* $F : \mathcal{A} \to \mathcal{B}$ between two categories sends objects to objects and arrows to arrows, such that

$$\frac{f : A \to B}{F(f) : F(A) \to F(B)},\ F(gf) = F(g)F(f),\ F(1_A) = 1_{F(A)}.$$

Largely due to the insight of Bill Lawvere it soon became clear that the objects of many mathematical categories are already categories in disguise. Among these degenerate categories we find partially ordered sets $A, B, \ldots$ which allow at most one arrow between A and B. We also recognize monoids as categories with just one object, so that the source and target of an arrow need not be named. Conversely, we may think of a category as a monoid with many objects.

In a later chapter we will be mainly interested in additive categories, where the set of arrows from A to B forms an additive semigroup. Additive categories were also called rings with many objects by Barry Mitchell. In particular, a ring may be viewed as an additive category with one object and a ring with two objects is known as a *Morita context.* In a later chapter we will look at a ring with three objects.

19 Quaternions

Inasmuch as any polynomial equation with complex coefficients has a complex solution, the number system now seemed to be complete. It therefore came as quite a surpise when Hamilton discovered the division ring of quaternions, sacrificing only the commutative law of multiplication. Quaternions had the form

$$a = a_0 + i_1 a_1 + i_2 a_2 + i_3 a_3,$$

when a_0 to a_3 were real numbers and the so-called *units* i_1 to i_3 satisfied

$$i_1^2 = i_2^2 = i_3^2 = i_1 i_2 i_3 = -1.$$

Any quaternion a has a *conjugate*

$$\overline{a} = a_2 - i_1 a_1 - i_2 a_2 - i_3 a_3$$

such that

$$a\overline{a} = a_0^2 + a_1^2 + a_2^2 + a_3^2$$

allowing us to form the inverse

$$a^{-1} = a(a\overline{a})^{-1}$$

if $a \neq 0$.

Unfortunately, the quadratic form obtained from $a\overline{a}$ turned out to be unsuitable for physical applications, which required

$$a_0^2 - a_1^2 - a_2^2 - a_3^2$$

instead, as was shown in Einstein's special theory of relativity in 1905. Nonetheless, only a few years later, applications of quaternions to special relativity were carried out by Conway [1911,1912] and Silberstein [1912,1924]. However they had to allow some of the components of a quaternion to be imaginary and were thus working with *biquaternions.*

The idea which we will pursue here follows a different approach, based on the observation that quaternions are more elegantly equipped to describe a six-dimensional space-time, hinting that time has three dimensions like space. This hypothesis has indeed been proposed by Peter Brendan, whose motivation was however not based on quaternions but on the attempt to derive quantum mechanics from general relativity.

We begin by looking at the *regular* matrix representations of quaternions. With any quaternions ψ there is associated a 4-vector $(\psi_0, \psi_1, \psi_2, \psi_3)$, whose transposed column vector will be denoted by$[\psi]$.

The regular *left* and *right* regular representations of any quaternion a will be denoted by $L(a)$ and $R(a)$ respectively and are defined by

$$L(a)[\psi] = [a\psi], \;\; R(a)[\psi] = [\psi a].$$

We note that

$$\begin{aligned} L(a)L(a') =& \;\; L(aa'), \;\; R(a)R(a') = R(a'a), \\ L(a)R(a') =& \;\; R(a')L(a). \end{aligned}$$

The two matrix representations can be related with the help of the diagonal matrix

$$\Gamma = (1, -1, -1, -1)$$

by noting that $\Gamma[\psi] = [\overline{\psi}]$ and by calculating

$$R(a) = \Gamma L(\overline{a})\Gamma, \;\; L(a) = \Gamma R(\overline{a})\Gamma.$$

A well known theorem about central simple algebras implies that, if $\mathcal{H}$ is the skew-field of quaternions, then $\mathcal{H} \otimes \mathcal{H}^{op}$ is isomorphic to the ring of all 4×4 real matrices. Computationally, this means that every such matrix A can be written uniquely in the form

$$A = \sum_{\alpha,\beta=0}^{3} L(i_\alpha)R(i_\beta)x_{\alpha\beta},$$

where $i_0 = 1$ and $x_{\alpha\beta} \in \mathbb{R}$. This expression becomes more manageable if A is *skew-symmetric*, that is if A coincides with its negative *transposed.* For then every skew-symmetric matrix is uniquely of the form

$$A = L(\mathbf{a}) + R(\mathbf{a}'),$$

when $\mathbf{a}$ and $\mathbf{a}'$ are 3-*vectors*, that is, quaternions whose scalar parts are zero.

If we define the *conjugate* A^* of A as

$$A^* = L(\mathbf{a}) - R(\mathbf{a}')$$

we easily verify that

$$\begin{aligned} AA^* = &\ L(\mathbf{a}\mathbf{a}) - R(\mathbf{a}'\mathbf{a}') \\ = &\ -I(\mathbf{a} \circ \mathbf{a}) + I(\mathbf{a}' \circ \mathbf{a}'), \end{aligned}$$

where the small circle denotes the Heaviside scalar product and I is usually omitted. This suggests writing

$$\begin{aligned} A \odot A = &\ -(\mathbf{a} \circ \mathbf{a}) + (\mathbf{a}' \circ \mathbf{a}') \\ = &\ -(a_1^2 + a_2^2 + a_3^2) + (a_1'^2 + a_2'^2 + a_3'^2) \end{aligned}$$

as a kind of six-dimensional scalar product.

According to Einstein and Minkowski there are four dimensions of space-time, three of space and one of time. According to a theory proposed by Kaluza and Klein, there ought to be five dimensions, four of space and one of time. More recently, string theorists have proposed ten dimensions, nine of space and one of time. Here we will explore the possibility of a six-dimensional universe, where both space and time have three dimensions. In fact, we will represent location in space-time by the matrix

$$L(i_1x_1 + i_2x_2 + i_3x_3) \ + \ R(i_1t_1 + i_2t_2 + i_3t_3),$$

where the old time t is now replaced by the 3-vector $\mathbf{t} = i_1t_1 + i_2t_2 + i_3t_3$, but we can still recapture t except for its sign, by

$$t^2 = \mathbf{t} \circ \mathbf{t} = t_1^2 + t_2^2 + t_3^2.$$

20 Infinitesimals

Among the famous paradoxes of Zeno there is one that says "the flying arrow is at rest". One way to read this is saying that for an object to move it must first pass through an infinitely small distance, but how can such a distance be different from zero?

In the 18th century mathematicians did employ *differentials*, not to be assumed to be zero, for developing the differential calculus, only to be severely criticized by the philosopher Berkeley. In the 19th century mathematicians got around Zeno's difficulty by introducing a rigorous definition of limits that avoided infinitesimals. Mathematicians of the 20th century realized that the existence of differentials could be justified after all.

Thus Abraham Robinson used formal logic to show that the existence of infinitesimals was consistent, and hence could be assumed without leading to a contradiction. On the other hand, Kock and Lawvere considered differentials to be real numbers of square zero. To do so they had to assume the existence of numbers not unequal to zero, other than zero itself, thus assuming an underlying logic in which two negations do not necessarily cancel each other. Such a logic had indeed been proposed by the philosopher mathematician Brouwer, who called his mathematics "intuitionistic". Indeed, Penon was led to define infinitesimals as numbers not unequal to zero. We will pursue this definition and justify it.

The rule that two negations cancel each other or, equivalently, that for every proposition p either p or $\neg p$ must hold is usually attributed to Aristotle. But even he had some doubts about this and mentioned as an example the proposition "there will be a sea-battle tomorrow". The so-called Aristotelian rule was abandoned by Brouwer and his followers, even by many mathematicians that do not otherwise agree with his philosophy.

As it turns out, in a field or skew-field the elements not unequal to zero are precisely the elements of the Jacobson radical, sometimes also attributed Perlis or Chevalley. The *radical* consists of all those elements h for which $1 - xh$ is invertible for all elements x. In a division ring this means that $1 - xh \neq 0$, i.e. that h is *not* invertible, that is, that h is not unequal to zero.

We will now justify Penon's definition by showing that in a *normed* division ring these are precisely the elements which are infinitely small. A ring A is said to have a *norm* N if to each element a of A there is associated a non-negative real number $N(a) = x^2$ such that

$$N(n1) = n^2,\ N(ab) = N(a)N(b)$$

for all integers n. Evidently, the field of reals is normed with $N(a) = a^2$, the field of complex numbers and the skew field of quaternions with $N(a) = a\bar{a}$, $\bar{a}$ being the conjugate of a. We can call an element h of A *infinitely small* if

$$\forall_{n\in\mathbb{N}}(N(nh) < 1).$$

Theorem. In a normed division ring A, the Penon infinitesimals are the same as the infinitely small elements.

Before proving this, we make some reasonable assumptions about the field $\mathbb{R}$ of real numbers. It should be endowed with a binary relation $<$, which we may think of as the negation of a partial order $\geq$, hence which should satisfy $0 < 1$ and

$$\neg\neg(x < y) \Rightarrow x < y \Rightarrow \neg(y < x).$$

Moreover, we assume all real numbers are bounded:

$$\forall_{x\in\mathbb{R}}\exists_{n\in\mathbb{N}}(x^2 < n^2).$$

We are now ready to prove the theorem:

Proof. On the one hand, $h = 0$ implies that for any $n \in \mathbb{N}$,

$$N(nh) = N(n)N(0) = n^2 0^2 = 0 < 1,$$

hence

$$\neg\neg(h = 0) \Rightarrow \neg\neg(N(nh) < 1) \Rightarrow N(nh) < 1.$$

On the other hand, $\neg(h = 0)$ implies the existence of an element x such that $xh = 1$ and $N(x) < n^2$ for some $n \in \mathbb{N}$, hence

$$1 = N(1) = N(xh) = N(x)N(h) < n^2 N(h) = N(nh)$$

and so $\neg(N(nh) < 1)$. Thus

$$N(nh) < 1 \Rightarrow \neg\neg((Nh) < 1) \Rightarrow \neg\neg(h = 0).$$

QED

21 What is a calculation?

Gödel needed a special kind of calculable function, the so-called primitive recursive ones, to be defined below. There are others as we will see.

A subset A of $\mathbb{N}^k$ is said to be *recursively enumerable* if there is a primitive recursive function $f : \mathbb{N} \to \mathbb{N}^k$ such that

$$A = \{f(n)|n \in \mathbb{N}\}.$$

A function $h : \mathbb{N}^k \to \mathbb{N}$ is said to be *recursive* if its graph

$$\{\langle a, h(a)\rangle | a \in \mathbb{N}^k\}$$

is recursively enumerable. This means that

$$n = h(a) \text{ iff } \exists_{m:\mathbb{N}}(n = f(m) \wedge a = g(m))$$

for some primitive recursive function f and g. In other words that $h = fg^{\vee}$, where $g^{\vee}$ is the converse of the relation g and juxtaposition denotes the relative product.

It turns out that A is recursive if and only if both A and its complement are recursively enumerable. Once the terms of the language of higher order arithmetic have been replaced by their Gödel numbers,one realizes that the set of proofs is recursive, hence the set of theorems is recursively enumerable, but not recursive in view of Gödel's incompleteness theorem.

The word "calculus" is Latin for "pebble" and calculations originally meant to be achieved by moving pebbles. This is not just a figure of speech, since everything a modern computer can compute can already be done more slowly by a stone-age computer that is simpler than the better known Turing machine and will now be described.

We require an unlimited number of locations, say holes in the ground. (If all holes have been used up just dig another.) Each location is supposed to have an unlimited capacity. (If a hole is not deep enough, just dig a little more deeply.) We also require an unlimited number of pebbles. (If all pebbles have been used up, go down to the beach and pick some more.) Note that a modern computer may suffer from similar limitations: it may run out of storage space or electricity.

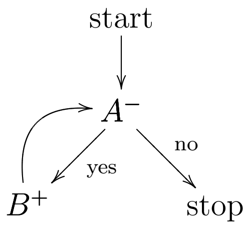
Programs are made up from the following basic instructions:

$$\text{start} \ \to\ ,\ \to\ \text{stop}, \qquad \begin{tikzcd} & & {} \\ {} \arrow[r] & B^- \arrow[ur, "\text{yes}"] \arrow[dr, "\text{no}"'] & \\ & & C \end{tikzcd}$$

For example, the program

$$\begin{tikzcd} \text{start} \arrow[d] & \\ A^- \arrow[loop left, "\text{yes}"] \arrow[dr, "\text{no}"] & \\ & \text{stop} \end{tikzcd}$$

allows us to empty location A, and the program

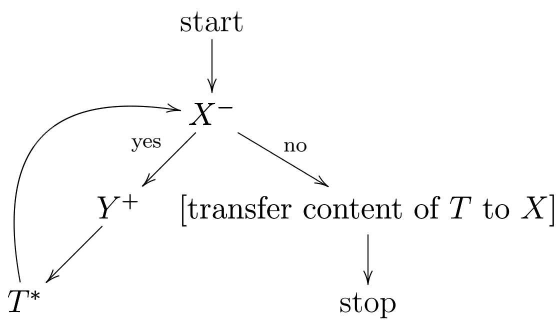


allows us to transfer the content of A to B and suggests how addition of numbers can be performed. Surprisingly, every recursive function can be calculated in a similar fashion.

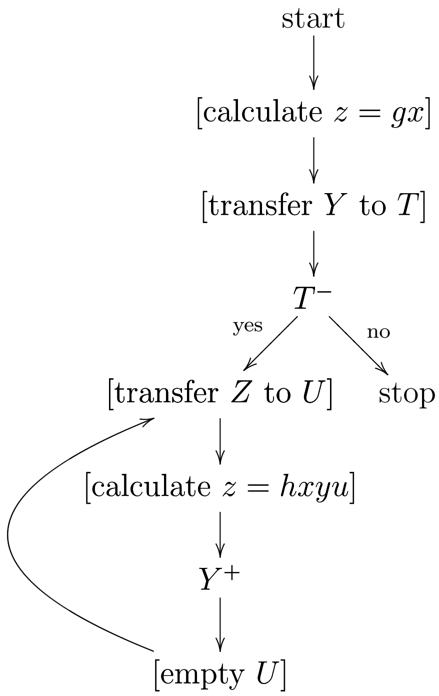
To calculate a numerical function, say $z = fxy$, of two variables, we begin with x pebbles at location X, y pebbles at location Y and no pebbles at certain other locations that are reserved to serve as temporary storage locations.

The answer is to appear at location Z, and the input x and y should re-appear at locations X and Y respectively in order not to be forgotten.

Here, for example, is a program that will calculate the successor function $y = Sx$, containing the transfer program as a subroutine and T as a temporary storage location:



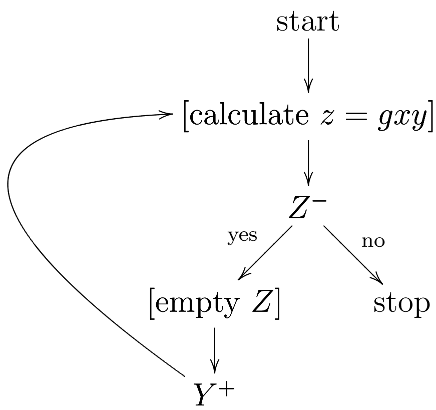
Aside from a few trivial subroutines, the calculation of all primitive recursive functions requires the *recursion scheme* to obtain fxy from subroutines for gx and $hxyz$:



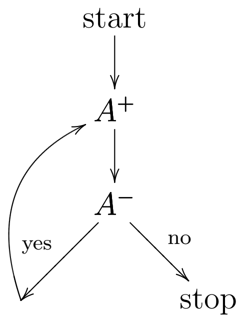
To calculate all recursive functions, not just the primitive recursive ones, we also require the *minimization scheme*, say

$fx =$ smallest y such that $gxy = 0$

for producing a function of one variable from a known function of two variables. This can be done by the program



As it turns out, a numerical function can now be calculated if and only if it is recursive. Not every conceivable program will help to calculate a numerical function. For example, the following *bad* program clearly leads to a calculation that will never stop:



However, in general, there is no way of verifying whether a flow-diagram leads to a good or bad program.

In the spirit of Gödel numbering, each flow-diagram can be represented by a natural number. We can then say that the set of all flow-diagrams is recursive, but the set of good ones, leading to calculations that will ultimately terminate, is not.

Part V. Speculations about six dimensions.

22 Classifying fundamental particles

The pre-Socratic debate whether material objects were made from infinitely divisible substances or from indivisible discrete atoms was resolved in the nineteenth century in favour of the latter, even though what were then called *atoms* have now turned out to be divisible after all.

In the early twentieth century it was realized that the so-called atoms of the periodic table consisted of yet smaller particles, *electrons, protons* and *neutrons.* Later in the twentieth century *neutrinos* were added to these and the protons and neutrons were seen to consist of still smaller particles called *quarks.* Concerning the related question, whether space and time are continuous or discrete, the jury is still out.

Influenced by an idea of Harari and Shupe, I was led to describe the surviving particles, all fermions of spin $1/2$, by three-vectors with entries $1, -1$ or 0, each representing one third of the charge. A similar description applied to the forces of nature, all represented by bosons of spin 1, except for the graviton and the hypothetical Higgs particle. I found it convenient to introduce a fourth component into the vector, the *fermion* number: it was 1 for fermions, -1 for anti-fermions and 0 for bosons. More recently, I expanded the fermion number into a three-vector as well. The *fermion vector* was to account for the higher generation fermions and to distinguish between left-handed and right-handed particles.

Here is what the charge vector for fermions looked like: $(-1,-1,-1)$ for electrons, $(+1,+1,+1)$ for positrons, $(0,0,0)$ for neutrinos, $(-1,0,0),(0,-1,0)$ and $(0,0,-1)$ for the three so-called *colours* of down-quarks and $(0,+1,+1),(+1,0,+1)$ and $(+1,+1,0)$ for the three colours of up-quarks. The signs were reversed for anti-quarks. For bosons the charge vector was $(0,0,0)$ for photons and $Z^0,(-1,-1,-1)$ for $W^-,(+1,+1,+1)$ for W^+ and for colour-changing gluons it was $(-1,+1,0),(0,-1,+1),(+1,0,-1)$ and those negatives. Note that the sum of the three entries of the charge vector is three times the charge of the electron, if this is taken to be -1.

The fermion vector is $(+1,0,0),(0,+1,0)$ and $(0,0,+1)$ for the three generations of left-handed fermions and $(0,-1,-1),(-1,0,-1)$ and $(-1,-1,0)$ for the three generations of right-handed fermions. Note that the sum of the entries of the fermion vector is equal to the old fermion number modulo 3, e.g.

$$0-1-1=1 \text{ mod } 3,$$

which means that the sum differs from the old fermion number by a multiple of 3.

The fermion vector of any boson could now be $x+x+x$, since this is 0 modulo 3, the old fermion number. Some decision has to be made about x. Since photons come in two varieties, in clockwise or counter-clockwise rotation about the axis of spin, we ought to pick $x=+1$ or -1 accordingly. For gluons, I see no reason against picking $x=0$. But, for weak bosons Z^0, W^- and W^+, I reluctantly chose $x=-1$, to ensure that they will interact with left-handed fermions only, although I may be wrong about this.

Juxtaposing the charge vector and the fermion vector we obtain a six-vector

$$a=(a_1,a_2,a_3;\ a_1',a_2',a_3').$$

Under addition two such six-vectors will combine to another with components $1,-1$, and 0 if and only if the corresponding fundamental particles will combine to a third fundamental particle, as shown in a Feynman diagram. To illustrate such six-vectors we adopt a simplified notation, replacing $+1$ and -1 by $+$ and $-$ respectively and restrict attention to first generation particles.

$$
\begin{array}{rcll}
{}_Le^- & = & - - - + 0\ 0 & \text{left handed electron} \\
\gamma & = & 0\ 0\ 0 - - - & \text{clockwise photon} \\
\hline
{}_Re & = & - - -\ 0 - - - & \text{right-handed electron} \\
 & & & \\
{}_Le^- & = & - - - + 0\ 0 & \text{left-handed electron} \\
{}_{\overline{k}}e^+ & = & + + +\ 0 + + & \text{antiright-handed positron} \\
\hline
\overline{\gamma} & = & 0\ 0\ 0 + + + & \text{counterclockwise photons} \\
 & & & \\
{}_Lu_b & = & + 0 + + 0\ 0 & \text{left-handed blue up-quark} \\
g_{by} & = & 0 + -\ 0\ 0\ 0 & \text{blue-yellow gluon} \\
\hline
{}_Lu_y & = & + + 0 + 0\ 0 & \text{left-handed yellow up-quark} \\
 & & & \\
{}_Lu_r & = & 0 + + + 0\ 0 & \text{left-handed red up-quark} \\
Z^0 & = & 0\ 0\ 0 - - - & \text{neutral weak boson} \\
\hline
{}_Ru_r & = & 0 + + 0 - - & \text{right-handed red up-quark} \\
 & & & \\
{}_L\nu & = & 0\ 0\ 0 + 0\ 0 & \text{left-handed neutrino} \\
W^- & = & - - - - - - & \text{negative weak boson} \\
\hline
{}_Re^- & = & - - -\ 0 - - - & \text{right-handed electron}
\end{array}
$$

Recalling that second and third generation electrons are called *muons* and *tauons* respectively and are represented by the Greek letters μ and τ, we also obtain

$$
\begin{array}{rcll}
{}_L\mu^- & = & - - -\ 0 + 0 & \text{left-handed muon} \\
{}_L\nu_e & = & 0\ 0 + + 0\ 0 & \text{left-handed electron neutrino} \\
\hline
{}_{\overline{R}}\tau^- & = & - - - + + 0 & \text{antiright-handed tauon}
\end{array}
$$

Our method of classification also leaves room for three generations of dark matter, with fermion vectors $(- + +), (+ - +)$ and $(+ + -)$ respectively, which do not allow addition with (xxx) when $x = \pm$, to yield another fermion vector. Unfortunately, our method of labelling does not distinguish between the neutral weak boson Z^0 and the clockwise photon γ.

Since our classification does not apply to the graviton and the Higgs boson, it raises the question whether these bosons are fundamental. Surprisingly, two separate articles in a recent issue of Scientific American [306, May 2012] have suggested that these are compound particles.

23 The Lorentz category

I have argued in chapter 19 that mathematical elegance required that 4-dimensional space-time quaternions be replaced by 6-dimensional skew-symmetric matrices. This argument may be extended to momentum-energy, both kinetic and potential, and to the Maxwellian charge-current density. Thus we ought to consider the skew-symmetric matrices

$$\begin{array}{ll} X = L(\mathbf{x}) + R(\mathbf{t}) & \text{(space-time)} \\ P = L(\mathbf{p}) + R(\mathbf{m}) & \text{(kinetic energy-momentum)} \\ \Phi = L(\mathbf{A}) + \mathbf{R}(\varphi) & \text{(electro-magnetic potential)} \\ J = L(\mathbf{J}) + \mathbf{R}(\rho) & \text{(charge-current density).} \end{array}$$

To account for partial differentiation, we also require

$$D = L(\nabla) - R(\nabla'),$$

when $\nabla = i_1\dfrac{\partial}{\partial x_1} + i_2\dfrac{\partial}{\partial x_2} + i_3\dfrac{\partial}{\partial x_3}$ and $\nabla' = i_1\dfrac{\partial}{\partial t_1} + i_2\dfrac{\partial}{\partial t_2} + i_3\dfrac{\partial}{\partial t_3}$, the minus sign being due to contravariance of differentiation.

Given any skew-symmetric 4×4 matrix S and any 4×4 matrix Q, QSQ^T will also be skew-symmetric, since

$$(QSQ^T)^T = Q^{TT}S^TQ^T = -QSQ^T.$$

Evidently, the transformations $S \mapsto QSQ^T$ are closed under composition and form a group if $\det Q = 1$. Writing

$$Q^\sharp = Q^{-T} = (Q^T)^{-1},$$

we call

$$X \mapsto QXQ^T = QXQ^{-\sharp}$$

a *Lorentz* transformation provided

$$XX^* = -X \odot X$$

is invariant.

To ensure that XX^* itself will be invariant, a necessary and sufficient condition is that

$$(*) \qquad X^* \mapsto Q^\sharp X^* Q^{-1}.$$

To prove the sufficiency of $(*)$, note that $(*)$ implies

$$XX^* \mapsto QXQ^TQ^\sharp X^*Q^{-1} = QXX^*Q^{-1} = XX^*.$$

To prove the necessity of $(*)$, assume the invariance of XX^*. For the moment also suppose that $XX^* \neq 0$, then

$$X^* = X^{-1}(XX^*) \mapsto Q^\sharp X^{-1}Q^{-1}(XX^*) = Q^\sharp X^*Q^{-1}.$$

If $XX^* = 0$, a continuity argument should salvage the result.

It appears that all interesting physical entities A in special relativity can be represented by four by four real matrices accompanied by instructions on how they are transformed under a coordinate transformation $X \mapsto QXQ^T = QXQ^{-\sharp}$. These instructions take the form $A \mapsto Q^uAQ^{-v}$, where $Q^{-v} = (Q^v)^{-1}$ and u and v are elements of the commutative monoid $\{0, 1, \sharp\}$, where $Q^0 = 1$, $Q^1 = Q$ and $Q^\sharp = Q^{-T} = (Q^T)^{-1}$. Note that $u0 = 0$, $u1 = u$ and $\sharp\sharp = 1$. We write $A : u \to v$ and consider A to be an arrow in the *Lorentz* category. In particular, the basic entities $S = X, P, \Phi, J$ and D may all be viewed as arrows $1 \to \sharp$ and S^* as arrows $\sharp \to 1$, rendering SS^*: $1 \to 1$.

The Lorentz category is an additive category with three objects, which Barry Mitchell would have called a ring with three objects. When $A : u \to v$ and $B : v \to w$, we write $AB : u \to w$, opposite to the order of composition preferred by categorists, so that $AB \mapsto Q^u ABQ^{-w}$. In what follows, assume that A, B and C are *basic* physical entities, say

$$A = L(\mathbf{a}) + R(\mathbf{a}'),\ B^* = L(\mathbf{b}) - R(\mathbf{b}'),$$

thus $AB^* : 1 \to 1$ and $AB^*C : 1 \to \sharp$.

The trace of AB^* is

$$\frac{1}{2}(AB^* + BA^*) : 1 \to 1$$

and may be called the *scalar* part of AB^*. Its negative is

$$A \odot B = \mathbf{a} \circ \mathbf{b} - \mathbf{a}' \circ \mathbf{b}' : 1 \to 1$$

and extends the Heaviside scalar product to skew-symmetric matrices. On the other hand,

$$\begin{aligned} \frac{1}{2}(AB^* - BA^*) &= L(\mathbf{a} \times \mathbf{b}) + R(\mathbf{a}' \times \mathbf{b}') \\ &\quad + L(\mathbf{b})R(\mathbf{a}) - L(\mathbf{a}')R(\mathbf{b}') : 1 \to 1 \end{aligned}$$

extends the Heaviside vector product.

We also have $AB^*C : 1 \to \sharp$, but this need not be skew-symmetric. Twice its skew-symmetric part is

$$\begin{aligned} AB^*C + BA^*C = &\ A(B^*C + C^*B) + (AB^* + BA^*)C \\ &- (CA^* + AC^*)B, \end{aligned}$$

hence the skew-symmetric part is

$$\text{skew}(AB^*C) = -A(B \circ C) + B(C \odot A) - C(A \odot B) : 1 \to \sharp. \tag{23.1}$$

24 Six-dimensional special relativity

Recalling that $S \mapsto QSQ^T$ and $S^* \mapsto Q^\sharp S^* Q^{-1}$ when $S = X, P, \mathbf{\Phi}$ or J, we now have arrows

$$S : 1 \to \sharp,\ S^* : \sharp \to 1,\ SS^* : 1 \to 1,\ S^*S : \sharp \to \sharp,$$

although as matrices the last two are equal to $-(S \odot S)I$.

Physicists require that for real particles (virtual ones excepted) that $S \odot S$ is negative, e.g.,

$$X \odot X = -S^2,\ P \odot P = -\boldsymbol{\mu}^2,$$

where S is called the *interval* and $\boldsymbol{\mu}$ the *rest mass*. If $\boldsymbol{\mu} \neq 0,\ P = \boldsymbol{\mu}(dX/ds)$, hence

$$(dX/ds) \odot (dX/ds) = -1,\ (dX/ds) \odot P = -\boldsymbol{\mu}.$$

Force $d\mathbf{p}/ds$ and power $d\boldsymbol{\mu}/ds$ may be combined into dP/ds, the rate of change of *kinetic* energy-momentum. For the electron, the *potential* energy-momentum is $-e\boldsymbol{\Phi}$.

We now turn to the six-dimensional generalization of the Maxwell-Lorentz theory of the electron. We calculate

$$\begin{aligned} D^*\boldsymbol{\Phi} = & \; (L(\nabla) + R(\nabla'))((L(\mathbf{A}) + R(\boldsymbol{\phi})) \\ = & \; -D \odot \Phi + B - E, \end{aligned}$$

where

$$-D \odot \boldsymbol{\Phi} = L(\nabla \circ \mathbf{A}) + R(\nabla' \circ \boldsymbol{\phi})$$

is the scalar part of $D^*\Phi$ and

$$\begin{aligned} B = & \; L(\nabla \times \mathbf{A}) + R(\nabla' \times \boldsymbol{\phi}), \\ -E = & \; L(\nabla)R(\boldsymbol{\phi}) + R(\nabla')L(A) \end{aligned}$$

generalize the magnetic and electric fields $\mathbf{B}$ and $\mathbf{E}$ to six dimensions. Hence the six-dimensional analogue of the vector part of $D^*\boldsymbol{\Phi}$ is

$$\frac{1}{2}(D^*\boldsymbol{\Phi} - \boldsymbol{\Phi}^*D) = D^*\boldsymbol{\Phi} + D \odot A = B - E = F,$$

generalizing the usual electro-magnetic field. Hopefully it incorporates a pre-quantum treatment of the weak and strong forces.

Maxwell's equation may now be written

$$DF - J = 0,$$

that is

$$L(\nabla) - R(\nabla')(B - E) = L(\mathbf{J}) + R(\boldsymbol{\phi})$$

where B and E are defined as above and generalize the usual magnetic field $\mathbf{B}$ and electric field $\mathbf{E}$. Note that in the familiar four-dimensional universe

$$R(\boldsymbol{\phi}) = i\boldsymbol{\phi}, \;\; R(\nabla') = i\partial/\partial t,$$

hence we obtain the usual Maxwell equations, containing those of Faraday, Gauss and Ampère, as well as the non-existence of a magnetic monopole.

The six-dimensional equation of *continuity* takes the form $D \odot J = 0$. It must hold since the scalar part of F, hence of $D^*J = D^*DF$, is zero.

As far as F is concerned, $\boldsymbol{\Phi}$ is determined only up to a *gauge* transformation

$$\boldsymbol{\Phi} \mapsto \boldsymbol{\Phi} - D\boldsymbol{\Theta} = \boldsymbol{\Phi}_{\boldsymbol{\Theta}}$$

where $\boldsymbol{\Theta} : 1 \to 1$ is a scalar with $D^*D\boldsymbol{\Theta} = 0$. Physicists usually assume that $D \odot \boldsymbol{\Phi}_{\boldsymbol{\Theta}} = 0$, but I prefer instead

(24.1) $$(dX/ds) \odot \boldsymbol{\Phi}_{\boldsymbol{\Theta}} = 0,$$

which may be achieved by postulating that

$$d\boldsymbol{\Theta}/ds = (dx/ds) \odot \boldsymbol{\Phi},$$

To obtain the special relativistic Lorentz force, we apply (23.1) with

$$A = -e(dX/ds), B = D, C = \mathbf{\Phi_\Theta}$$

and obtain

$$\begin{aligned}\mathrm{skew}(-eD^*\mathbf{\Phi_\Theta}) = \ & e(dX/ds) - eD(\mathbf{\Phi_\Theta} \odot (dX/ds)) \\ & +e\mathbf{\Phi_\Theta}((dX/ds) \odot \overleftarrow{D}) : 1 \to \sharp\end{aligned}$$

Here the second term on the right hand side is zero by (24.1), the third term is $e(d\mathbf{\Phi_\Theta}/ds)$ and the first term may be transferred to the left hand side to yield

$$\mathrm{skew}\left(-e\frac{dX}{ds}F\right) = \ \mathrm{skew}(-eD^*\mathbf{\Phi_\Theta} + D \odot \mathbf{\Phi_\Theta}) = e(d\mathbf{\Phi_\Theta}/ds)$$

that is

$$\frac{d}{ds}(P - e\mathbf{\Phi_\Theta}) = 0.$$

25 Six-dimensional relativistic quantum mechanics

The special relativistic Schroedinger equation, also called the Klein-Gordon equation, takes the form

$$DD^*[\boldsymbol{\psi}] = -\boldsymbol{\mu}^2[\boldsymbol{\psi}],$$

where we allow $\boldsymbol{\psi} = \boldsymbol{\psi}(X)$ to be a quaternion. This may evidently be replaced by two first order differential equations, provided the rest-mass $\boldsymbol{\mu} \neq 0$:

$$D^*[\boldsymbol{\psi_1}] = \boldsymbol{\mu}[\overline{\boldsymbol{\psi_2}}],\ D[\overline{\boldsymbol{\psi_2}}] = -\mu[\boldsymbol{\psi_1}],$$

where $[\boldsymbol{\psi_1}] : 1 \to 0$ and $[\overline{\boldsymbol{\psi_2}}] : \sharp \to 0$, $[\boldsymbol{\psi_1}]$ and $[\boldsymbol{\psi_2}]$ being solutions of the Klein-Gordon equation.

If we accept Penrose's zigzag description of the electron, we may be satisfied with this pair of equations; but, following Dirac, we may wish to replace it by a single first order equation. An easy way to do this is to introduce an auxiliary entity $K : 1 \to \sharp$ which possesses an inverse $K^{-1} : \sharp \to 1$ such that

$$D^* = -K^{-1}DK^{-1}$$

and put

$$[\boldsymbol{\psi}] \ = \ [\boldsymbol{\psi_1}] - K[\overline{\boldsymbol{\psi_2}}] : 1 \to 0$$

so that

$$D^*[\boldsymbol{\psi}] = -\boldsymbol{\mu}K^{-1}[\boldsymbol{\psi_2}], \tag{25.1}$$

as a straight-forward calculation shows.

Does such an auxiliary entity K exist? Yes, if we make the assumption that one of the three time dimensions is redundant, say t_3 in a suitable coordinate system, that is to say, the action takes place in a five-dimensional subspace of six-dimensional space-time. For example, let $K = R(i_3)$ in a suitable frame of reference and allow all other coordinate systems to be introduced by Lorentz transformations. Thus we have not only $X^* = -K^{-1}XK^{-1}$ but also $S^* = -K^{-1}SK^{-1}$ when $S = P$ or D, as will be exploited below.

Putting $K = L(\mathbf{k}) + R(\mathbf{k}')$, we may rewrite (25.1) in pure quaternionic form:

$$\overrightarrow{\nabla}\psi + \psi\overleftarrow{\nabla}' + \mu(\mathbf{k}\boldsymbol{\psi} - \boldsymbol{\psi}\mathbf{k}') = 0. \tag{25.2}$$

Introducing $\boldsymbol{\eta} = \boldsymbol{\eta}(P)$ by

$$\boldsymbol{\mu\eta} = KP^* = -PK^{-1},$$

we have $\boldsymbol{\eta}^{\mathbf{2}} = -1$ and obtain the following explicit solution of (21.1):

$$[\boldsymbol{\psi}(X)] = (\exp \boldsymbol{\eta}(X \odot P))[\boldsymbol{\psi}(0)].$$

Moreover, since

$$\mu\eta^T = P^*K^{-1} = -KP,$$

we have $\eta^T = \mu^{-1}K^{-1}P$ and $(\eta^T)^2 = -1$, hence (21.1) may be rewritten thus:

$$\eta^T D^*[\psi] = P^*[\psi],$$

hence $\eta^T(P)$ replaces the traditional constant complex root of -1.

From the Dirac equation (25.1) we infer that

$$[\psi]^T\overrightarrow{D}^*[\psi] = -\mu[\psi]^T K^*[\psi],$$

where the right-hand side is skew-symmetric. Hence so must be the left-hand side and therefore

$$[\psi]^T\overleftrightarrow{D^*}[\psi] = [\psi]^T\overrightarrow{D}^*[\psi] + [\psi]^T\overleftarrow{D}[\psi] = 0.$$

Multiplying this by the constant skew-symmetric matrix S on the right, assuming that $S : 0 \to 0$ is Lorentz invariant, and recalling that the trace of a product is independent of the order of the factors, we obtain

$$\operatorname{trace}(\overrightarrow{D}^*[\boldsymbol{\psi}]S[\boldsymbol{\psi}]^T) = \operatorname{trace}([\boldsymbol{\psi}]^T\overleftrightarrow{D^*}[\boldsymbol{\psi}]S) = 0.$$

Now recall that, for skew-symmetric matrices A and B, $\operatorname{trace}(AB^*) = -A \odot B$ and write

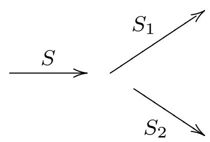
$$J_S = [\boldsymbol{\psi}]S[\boldsymbol{\psi}]^T, \tag{25.3}$$

then $D \odot J_S = 0$ looks like Maxwell's equation of continuity, suggesting that J_S resembles Maxwell's charge-current density and so yields a probabilistic interpretation of $[\boldsymbol{\psi}]$.

But what is $S : 0 \to 0$ in (21.3)? I am tempted to put

$$S = L(\mathbf{a}) + R(\mathbf{b})$$

where $(\mathbf{a}, \mathbf{b})$ is the six-vector representing a fundamental particle, then the equation $S = S_1 + S_2$ associated with a Feynman diagram



implies the corresponding equation $J_S = J_{S_1} + J_{S_2}$ for charge-current densities. This looks promising, but fails to account for the coupling constant.

26 The spread of time

A Latin proverb asserts that "time flees". For some reason this is rendered in English as "time flies". The contention of the present text is that "time is spread out" and this deserves a few words of explanation.

Some natural languages admit only two simple tenses: past and non-past. Many languages, even Victorian English, refuse to acknowledge that time proceeds in one direction only and allow additional simple tenses, such as the subjunctive, and many compound tenses which describe modalities.

While orthodox physics recognizes three dimensions of space and one of time, some recent theories have introduced additional dimensions, mostly of space. Thus, string theory requires six extra spatial dimensions; but, at least one physicist, Peter Gillan, has proposed a total of three temporal dimensions. His purpose was to derive quantum mechanics from general relativity in six dimensions. In the present text I haven't gone quite as far, restricting attention to special relativity, and have discussed three temporal dimensions mainly for the sake of mathematical elegance. However, if the derivation of Dirac's equation in chapter 25 is taken seriously, one of these three dimensions is redundant and all the action takes place in a five-dimensional subspace of six-dimensional space-time. The question arises how the extra temporal dimensions can be interpreted and whether they serve any purpose.

At one time I suspected that these extra dimensions ought to account for the direction of the axes of spin, but I was unable to make this idea work. Modern quantum theory assumes that a particle can be at different spacial locations at the same time with different probabilities, as is often illustrated by Schrödinger's cat. This seems to be easier to visualize if time has more than one dimension. I suspect that the extra temporal dimension (or dimensions) can be exploited to account for the weak force (or perhaps also the strong force) even before proceeding from classical physics to quantum theory.

Even if we cannot explain the additional dimensions of time, we may refer to Plato's famous metaphor of the cave and say that, when real time on the outside is projected onto the wall of the cave, the temporal dimensions coalesce into one.

Part VI. Foundations of mathematics (with Phil Scott).

27 Philosophy of mathematics

While most practicing mathematicians are too busy doing mathematics to care about the foundations of their subject, the few who do care are traditionally divided into four schools:

logicism, formalism, Platonism, intuitionism.

I have argued repeatedly, sometimes in collaboration with Phil Scott and Jocelyne Couture and even in an article in the Encyclopedia Britannica, that moderate forms of these four schools are in fact compatible. Their present adherents are more or less in agreement about the language of mathematics, which I take to be higher order arithmetic, although intuitionists will reject the axiom of the excluded third and some extreme intuitionists might downplay the importance of any formal language.

More than a century ago, Peirce, Frege, Russell and Peano tried to express the language of mathematics in logical notation, using the propositional connectives $\wedge, \vee, \Rightarrow$ and $\neg$ and the quantifiers $\forall$ and $\exists$ (in present notation). These were subject to the usual rules of logic, although with a difference between classical mathematicians and intuitionists. The former accepted the Aristotelian rules which asserted the truth of $p \vee \neg p$ and of $\neg\neg p \Rightarrow p$ for all propositions p, which the latter rejected.

It was also necessary to accept the symbols $=$ for equality and $\in$ for membership, subject to the axioms of *extensionality*

$$a = b \Leftrightarrow \forall_{x:A}(x \in a \Leftrightarrow x \in b)$$

and *comprehension*

$$a \in \{x : A|\phi(x)\} \Leftrightarrow \phi(a)$$

where x, a and b are of type A, the types having been introduced by Russell and Whitehead to avoid the famous Russell paradox.

In particular, one now needs a type Ω of *truth-values* and two operations on types: $A \times B$ the type of all pairs (a, b) when a and b are of types A and B respectively, and C^A the type of all functions from A to C. The type Ω^A is usually written as PA so that $\alpha = \{x : A|\phi(x)\}$ is of type PA.

Pure mathematicians are not interested in types distinguishing between animals and vegetables or between males and females, but they do require a type N of natural numbers. At first, logicists attempted to *define* the natural numbers semantically, e.g. 2 as the set of all two-element sets, but such definitions led to problems and were ultimately replaced by the postulational approach of Peano, who introduced symbols 0 of type N and S, the successor function, of type N^N, from which one could construct all numerals of type N

$$0,\ S0,\ SS0,\ SSS0,\ \ldots\ldots,$$

here summarized by S^n0, where the superscript n belongs to the metalanguage and S^n stands for n-fold repetition of the successor function.

Peano's approach relied on the principle of *mathematical induction*: from $\phi(0)$ and $\forall_{x:N}(\phi(x) \Rightarrow \phi(Sx))$ one may infer $\phi(S^n 0)$ for each $n : N$, or even $\forall_{x:N}\phi(x)$.

There now arises a philosophical question: do the symbols $S^n 0$ denote entities living in a Platonic universe or is there nothing else than the symbolism? The former position was taken by many mathematicians, including Gödel, while the latter has been defended by Hilbert and Mac Lane. Hilbert, a formalist, expected all true mathematical propositions to be provable, even if the proof has not yet been discovered. Mac Lane repeatedly asserted that "Platonism is clearly wrong".

The Platonist Gödel had shown that there are *true* mathematical propositions in what he believed to be a Platonic universe which are not provable, at least in classical mathematics. His argument depended on the so-called ω-rule: if $\phi(S^n 0)$ is true for all n, then so is $\forall_{x:N}\phi(x)$. Some present day intuitionists accept this rule as well, but it is rejected by the so-called *Brouwer-Heyting-Kolmogorov interpretation*, which admits this implication only if the proofs of $\phi(S^n 0)$ for each n have been obtained in a *uniform* way. After all, the individual proofs might increase in length with n.

According to Brouwer, a proposition is true if and only if it can be known. To us, as to many mathematicians, this means "can be proved".

Classically, but not intuitionistically, the ω-rule is equivalent to the ω^*-rule: if $\exists_{x:N}\phi(x)$ is true, then so is $\phi(S^n 0)$ for some n. This rule is not only acceptable to intuitionists, it has even been proved to hold if "true" is replaced by "provable". It is equivalent to the following two assertions, when $A = N$:

(27.1) If $\exists_{x:A}\phi(x)$ is true then so is $\phi(a)$ for some closed term a of type A, "closed" meaning that it contains no free variables. This is known as the *existence property*.

(27.2) Every closed term of type N is provably equal to a term of the form $S^n 0$. This is sometimes rendered as *all numerals are standard*.

Both (27.1) and (27.2), hence the ω^*-rule, have been proved for pure intuitionistic higher order arithmetic, by Lambek and Scott [1988], although Lewis Carroll's turtle might raise some questions about the validity of the proof [1895].

28 Gödel's argument re-examined

We would like to offer the seriously interested reader a short summary of Gödel's argument, although many readers of the present expository account may prefer to skip this chapter as being too technical.

Gödel began by expressing mathematical expressions in formal higher order arithmetic by natural numbers. The details of how he did this are not important, as he might as well have put all expressions of the formal language in alphabetic order and replaced the n^{th} expression by the number n, hence by the numeral $S^n 0$ in the formal language itself.

His argument invoked the notion of a special kind of numerical function

$$f(m_1, \cdots, m_k) = m_{k+1}$$

called *primitive recursive*. Aside from trivial constructions, these may be formalized in Peano's arithmetic with the help of the *recursion scheme*

$$f(..., 0) = g(...), \;\; f(..., Sy) = h(..., y, f(..., y))$$

where g and h already express primitive recursive functions. This scheme may be illustrated by Peano's definition of addition:

$$x + 0 = x,\ x + Sy = S(x + y).$$

Gödel realized that the functions

$$f(m) = S^m 0 \notin \alpha_m,\ g(n) = n^{th} \text{ theorem},$$

where α_m is the m^{th} comprehension term, are primitive recursive and proved that every primitive recursive function h is expressible in higher order arithmetic by a function symbol χ such that

$$S^{h(m)}0 = \chi(S^m 0)$$

is a theorem, i.e.

$$\vdash S^{h(m)}0 = \chi(s^m 0).$$

It follows that there are function symbols ϕ and ψ such that

$$\vdash S^{f(m)}0 = \phi(S^m 0),\ \vdash S^{g(n)}0 = \psi(S^n 0).$$

Now let

$$\alpha_k = \{x : N | \exists_{y:N} | \phi(x) = \psi(y)\},$$

then $S^k 0 \notin \alpha_k$ is true iff $\neg \exists_{y:N}(\phi(S^k 0) = \psi(y))$ is true

iff $\forall_{y:N}(\psi(S^k 0) \neq \phi(y))$ is true

if $\phi(S^k 0) \neq \psi(S^n 0)$ is true for all $n \in \mathbb{N}$ (by the ω-rule) (*)

iff $S^{f(k)}0 \neq S^{g(n)}0$ for all n

iff $f(k) \neq g(n)$ for all n

iff not $\vdash S^k 0 \notin \alpha_k$.

Being a Platonist, Gödel believed that there is a real world of mathematics in which all natural numbers live and all theorems are true. He also believed in the ω-rule for *classical* mathematics. It then follows from the above argument that classical higher order arithmetic contains a proposition, to wit $S^k 0 \notin \alpha_k$, which is true but not provable.

The argument fails for intuitionistic mathematics if one rejects the ω-rule in view of the Brouwer-Heyting-Kolmogorov interpretation, although I have met at least one intuitionist who still believed in the ω-rule.

Classically, though not intuitionistically, the ω-rule for truth is equivalent to the ω^*-rule and this holds even if truth is replaced by provability [LS2011]. Therefore, the word "if" in $(*)$ above can be replaced by "iff", if the word "true" is replaced by "theorem", hence $S^k 0 \in \alpha_k$ is not a theorem either. Thus neither $S^k 0 \in \alpha_k$ nor $S^k 0 \notin \alpha_k$ are provable and neither of them has been shown to be true.

29 Constructive nominalism

Gödel was widely believed to have put an end to Hilbert's program, which aimed to show that all true propositions of formal arithmetic are provable. An examination of Gödel's argument shows that this is indeed the case for *classical* arithmetic. On the other hand, in *intuitionistic* arithmetic all true propositions are assumed to be *knowable*, a property which many mathematicians, including the present authors, tend to interpret as meaning *provable*.

Let us remind the reader that a *proof* of p is a finite sequence of propositions, each of which is an *axiom* or follows from preceding elements of the sequence by a *rule of inference*, the last element of the sequence being p. It is crucial that there be a decision procedure for determining whether a sequence of propositions is a proof, hence that the set of provable propositions, aka *theorems*, is *recursively enumerable*.

If we identify true and provable propositions in intuitionistic arithmetic, we may justifiably claim that Hilbert's program applies to intuitionistic arithmetic after all, in spite of the alleged antagonism between Hilbert and Brouwer.

Gödel's argument that produced a certain proposition which can be neither proved nor disproved is still valid for intuitionistic arithmetic. However, the assertion that this proposition is true was based on the ω-property and should be rejected in view of the Brouwer-Heyting-Kolmogorov interpretation.

Applying the identification of truth with provability allows us to construct a Platonic universe from pure intuitionistic type theory, taking as its objects all provably equivalence classes of terms of the same type. This nominalistic construction is technically known as the *free topos*. It also happens to be the initial object in the category of all small toposes. The abstract notion of a topos had been introduced by Bill Lawvere as a cartesian closed category with an object of truth-values and an object of natural numbers.

In summary, we claim that moderate forms of the four conventional schools of mathematical philosophy are compatible: This compromise goes under the name *constructive nominalism*.

(27.1) *logicism*, if one admits as logical the symbols for equality and membership and abandons the attempt to define natural numbers semantically, introducing them by Peano's axioms instead;

(27.2) *formalism*, if we abandon Hilbert's attempt to show that all true statements are provable classically, restricting provability to statements acceptable to moderate intuitionists as knowable;

(27.3) *Platonism*, by identifying the real world of mathematics with the nominalistic universe constructed from pure intuitionistic type theory, in which mathematical entities are sets of provably equivalent terms of the language;

(27.4) *intuitionism*, although rejecting some of its more extreme claims.

While moderate intuitionists might accept the free topos as a Platonic universe of mathematics, Gödel's incompleteness theorem implies that an analogous nominalistic construction does not produce a Platonic universe for classical mathematicians, because they would not admit a proposition p for which neither p nor $\neg p$ is true, since $p \vee \neg p$ is a classical axiom.

So what do we mean when we say that a proposition is *true*? Here we ought to invoke the notion of a *model*, this being a possible universe of mathematics, and we ought to replace "true" by “true in a model”. Numbers, sets of numbers, sets of sets of numbers etc all must lead a precarious existence in a model.

The Gödel-Henkin completeness theorem asserts that a mathematical proposition is *provable* if and only if it is *true in every model.* This result is also valid for intuitionistic arithmetic, although there a single model, the free topos, suffices for the completeness proof, and may involve a non-constructive metalanguage. Whether a distinguished model can be constructed for classical mathematics has not yet been shown, but appears to be unlikely.

As far as we know, classical mathematics must, and constructive mathematics may accept many possible universes, aka models, for which it is required that

(1) $0 \neq 0$ is not true,

(2) $p \vee q$ is true only if p or q is true,

(3) $\exists_{x:A}\phi(x)$ is true only if $\phi(a)$ is true for some closed term **a** of type A.

The completeness theorem then shows that it does not suffice to look only at models in which all numerals are standard.

Extra material

30 Epilogue

This expository monograph has attempted to present some insights leading from linguistics to modern physics, as indicated by its title, as well as some ideas on the philosophy of mathematics which had previously been elaborated with Phil Scott [LS2011]. If asked to single out conceivably important contributions to knowledge that have been ignored in the literature, I will mention the following in parts I to VII.

Part I. Among transitive English verbs there are some whose present and past participles are genuine adjectives, namely those that describe causation of a mental or emotional state, such as *interest, amuse* and *frighten.* One may ask whether people with Asperger's symptoms will come up with the same list.

Part II. There is a common assumption that certain Andean languages are unusual in seeing the future behind and the past in front of the speaker. As I wish to point out again, the same is true in European languages as seen by the English prepositions *before* and *after* or the Latin prepositions *ante* and *post*, both referring to spatial as well as temporal location.

Part III. There is a surprising regularity in the inflectional morphology of Arabic and Hebrew verbs. Both languages allow the speaker to calculate the many finite conjugational forms of the verb as presented by three consonants and two vowels. All apparent irregularities can be explained by certain phonological rules.

Part IV. Among three mathematical formulations of syntax I now prefer *pregroups.* This algebraic or logical system allows the construction of sentences from the *types* assigned to inflected word forms, while respecting Miller's [1956] restriction on the short term memory to seven (plus or minus two) chunks of information. If one wishes to retain a Montague kind of semantics, one must abandon uniqueness of interpretation or replace sets by finite vector spaces.

Part V. The mathematical background required here is rather restricted and there is nothing new, except the emphasis on quaternions and the realization that infinitesimals of a division ring can be described algebraically as elements of the Jacobson radical. Whereas other publications have suggested the application of compact monoidal categories to both linguistics and physics, these are avoided here for the sake of readability.

Part VI. An attempt is made to describe the Maxwell-Lorentz theory of the electron in terms of Hamiltonian quaternions and to extend this description to the Dirac equation as well as to other fundamental particles. It appears that this presentation becomes more transparent and elegant if time is allowed to spread over three dimensions like space.

Part VII (with Phil Scott). We wish to contradict the widely held belief that Gödel had shown that Platonism is incompatible with Hilbert's formalist program. His impressive argument ceases to be convincing when classical mathematics is replaced by a moderate form of Brouwer's intuitionism that has been called *constructive nominalism* in [CL1991]. Hilbert's proposal that all true mathematical statements should be provable is automatically satisfied intuitionistically if Brouwer's *knowable* is interpreted as *provable*.

31 Acknowledgements

I am indebted to the following:

1. Danièle Bargelli, for teaching me Arabic and Turkish;
2. Yehoshua Bar-Hillel, for asking me to find a decision procedure for the syntactic calculus;
3. Wojciech Buszkowski, for producing rigorous proofs to settle several basic problems in categorial grammar;
4. Claudia Casadio, for initiating the application of noncommutative linear logic to linguistics;
5. Jocelyne Couture, for her collaboration on the philosophy of mathematics;
6. Chandler Davis, for encouraging me to publish popular expositions of some ideas on Linguistics, Philosophy and Physics in The Mathematical Intelligencer;
7. George Findlay, for collaborating on the original investigation of residuated monoids;
8. Brendan Gillon, for his active interest in my program and for encouraging his students to develop it further;
9. Nawal Kamel, for urging me to study French conjugation;
10. Tely Kusalik, for pointing out the relevance of Catalan numbers;
11. Larry Lambek for checking the text;
12. Michael Lambek, for arousing my interest in kinship terms;

13. Alex Melzak, for getting me to study the stone-age computer;
14. Gordon Pall, for getting me interested in quaternions;
15. Marcia Rodriguez, for helping with the index;
16. Phil Scott, for collaborating on the categorical foundations of mathematics;
17. Elaine Swan, for typing this book and earlier manuscripts;
18. Noson Yanofsky, for collaborating on the computational aspects of the biblical Hebrew verb;
19. Hans Zassenhaus, for showing me the elegant formulations of regular matrix representations.

References

[1] Abrusci, V.M. and C. Casadio, eds. 1996. *Proofs and Linguistic Categories, Applications of Logic to the Analysis and Implementation of Natural Language*, Proceedings of the 3rd Roma Workshop, Bologna: CLUEB.

[2] Abrusci, V.M. and C. Casadio, eds. 1999. *Dynamic Perspectives in Logic and Linguistics: Proof Theoretical Dimensions of Communication Processes*. Proceedings of the 4th Roma Workshop. Rome: Bulzoni.

[3] Abrusci, V.M. and C. Casadio, eds. 2002. *New Perspectives in Logic and Formal Linguistics*. Proceedings of the 5th Roma Workshop. Rome: Bulzoni.

[4] Abrusci, V.M., C. Casadio, and M. Moortgat, eds. 1994. *Linear Logic and Lambek Calculus*. Proceedings of the 1st Roma Workshop. Utrecht, Amsterdam: OTS/Dyana.

[5] Ajdukiewicz, K. 1935. Die syntaktische Konnexität, *Studia Philosophica*, 1:1-27. Eng. trans., Syntactic connexion. In *Polish Logic*, ed. S. McCall. Oxford: Clarendon Press, 1967.

[6] Bargelli, D. and J. Lambek, 2001a. An algebraic approach to Arabic sentence structure. In [42], 301-315.

[7] Bargelli, D. and J. Lambek, 2001b. An algebraic approach to French sentence structure. In [42], 62-78.

[8] Bargelli, D. and J. Lambek, 2007. A computational view of Turkish syntax and morphology, *Linguistic Analysis*, **34** 2004, 66-84.

[9] Bar-Hillel, Y. 1950. On syntactical categories, *Journal of Symbolic Logic*, 15:1-16.

[10] Bar-Hillel, Y. 1953. A quasi-arithmetical notation for syntactic description, *Language*, 29:47-58.

[11] Bar-Hillel, Y. 1962. Some recent results in theoretical linguistics. In *Logic, Methodology and Philosophy of Science*, eds. E. Nagle et al., 551-557. Stanford: Standord University Press.

[12] Bar-Hillel, Y., C. Gaifman, and E. Shamir. 1960. On categorial and phrase structure grammars, *Bulletin of the Research Council of Israel*, 9:1-16.

[13] Bengston, J.D. and M. Ruhlen. 1994. Global etymologies. In [84], 277-336.

[14] Bhargava, M. and J. Lambek. 1992. A production grammar for Sanscrit kinship terminology, *Theoretical Linguistics*, 18:45-66.

[15] Bhargava, M. and J. Lambek. 1995. A rewrite system of the Western Pacific: Lounsbury's analysis of Trobriand kinship terminology, *Theoretical Linguistics*, 21:241-253.

[16] Brame, M. and Youn Gon Kim. 1998. Directed types: an algebraic theory of production and recognition. *Linguistic Analysis*, 28:117-184.

[17] Buszkowski, W. 1986. Completeness results for Lambek syntactic calculus, *Zeitschr. f. math. Logik und Grund. d. Math.*, 32:13-28.

[18] Buszkowski, W. 1988. Generative power of categorial grammars. In [79], 69-94.

[19] Buszkowski, W. 1997. Mathematical linguistics and proof theory. In [90], 683-736.

[20] Buszkowski, W. 1998. Algebraic structures in categorial grammar, *Theoretical Computer Science*, 199:5-24.

[21] Buszkowski, W. 1999. Embeddability of residuated semigroups into algebras of relations. In [2], 19-22.

[22] Buszkowski, W. 2001. Lambek grammars based on pregroups. In [42], 95-109.

[23] Buszkowski, W. 2002a. Cut elimination for the Lambek calculus of adjoints. In [3], 95-109.

[24] Buszkowski, W. 2002b. Pregroups and grammars. In [3], 85-93.

[25] Buszkowski, W. 2002c. Pregroups, models and grammars. In *Relational Methods in Computer Science*, ed. H.M. de Swart, 35-49. LNCS, 2561. Berlin: Springer.

[26] Buszkowski, W. 2006. Type logics and pregroups, *Studia Logica*, 82: 1-25.

[27] Buszkowski, W., W. Marciszewski, and J. van Benthem, eds. 1990. *Categorial Grammar*. Amsterdam: J. Benjamin.

[28] Casadio, C. 1988. Semantic categories and the development of categorial grammars. In [79], 95-123.

[29] Casadio, C. 2001. Non-commutative linear logic in linguistics, *Grammars*, 4/3:1-19.

[30] Casadio, C. and J. Lambek. 2002. A tale of four grammars, *Studia Logica*, 71(2):315-329. Special issue: *The Lambek Calculus in Logic and Linguistics*, ed. W. Buszkowski.

[31] Casadio, C. and J. Lambek, eds. 2008. *Computation and algebraic approaches to natural languages*, Polimetrica. Milan.

[32] Casadio, C., P.J. Scott and R.A.G. Seely, eds. 2005. *Language and grammar: Studies in Mathematical Linguistics and Natural Language*. CSLI Lecture Notes, no. 168. Stanford: CSLI Publications.

[33] Chomsky, N. 1956. Three models for the description of language, IRE Symposium on Information Theory, MIT, Cambridge, Mass. In *Readings in Mathematical Psychology*, 11:105-124. New York: John Wiley and Sons, 1965.

[34] Chomsky, N. 1957. *Syntactic Structures*. The Hague: Mouton.

[35] Chomsky, N. 1959. On certain formal properties of grammars, *Information and Control*, 2:137-167.

[36] Chomsky, N. 1965. *Aspects of the Theory of Syntax*. Cambridge: The MIT Press.

[37] Chomsky, N. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.

[38] Coecke, B., M. Sadrzadek and S. Clark. 2010, Mathematical foundation for a compositional distributive model of meaning, *Linguistic Analysis*. 36: 345-384.

[39] Curry, H.B. 1961. Some logical aspects of grammatical structure, *Amer. Math. Soc. Proceedings Symposia Applied Math.*, 12:56-68. In [51].

[40] Curry, H.B., R. Feys and W. Craig. 1958. *Combinatory Logic*. Amsterdam: North Holland.

[41] Davis, S. and B.S. Gillon, eds. 2004. *Semantics: A Reader*. Oxford: Oxford University Press.

[42] De Groote, P., G. Morrill and C. Retoré, eds. 2001. *Logical Aspects of Computational Linguistics*. Proceedings 4th International Conference, LACI 2001. LNAI, 2099. Berlin: Springer.

[43] Dowty, D. 1982. Grammatical relations and Montague grammar. In *The Nature of Syntactic Representation*, eds. P. Jacobson and G.K. Pullam, 79-130. Dordrecht: Reidel.

[44] Dowty, D., R. Wall and S. Peters, eds. 1981. *Introduction to Montague Semantics*. Dordrecht: Reidel.

[45] Francez, N. and M. Kaminsky. 2007a. Commutation-extended pregroup grammars and mildly context-sensitive languages. *Studia Logica*, 87:295-321.

[46] Gazdar, G., E. Klein, G. Pullum and I. Sag. 1985. *Generalized Phrase Structure Grammar*. Cambridge: Harvard University Press.

[47] Goodenough, W.H. 1965. Yankee kinship terminology. In E.A. Hammel, ed., Formal semantics, *American Anthropologist*, 67: 259-287.

[48] Greenberg, J. 1996. *Indo-European and Its Closest Relatives: The Eurasiatic Language Family*, vol. 1, *Grammar*. Stanford: Stanford University Press.

[49] Harris, Z.S. 1966. A cyclic cancellation-automaton for sentence well-formedness, *International Computation Centre Bulletin*, 5:69-94.

[50] Harris, Z.S. 1968. *Mathematical Structure of Language*. New York: John Wiley and Sons.

[51] Jakobson, R., ed. 1961. *Structure of Language and Its Mathematical Aspects*. Providence: AMS.

[52] Jesperson, O. 1924. *The Philosophy of Grammar*. London: Allen and Unwin.

[53] Keenan, E.L. and L.M. Faltz. 1985. *Boolean Semantics for Natural Language*. Dordrecht: Reidel.

[54] Keenan, E.L. and E.P. Stabler. 2003. *Bare Grammar: Lectures on Linguistic Invariants*. Stanford Monographs in Linguistics. Stanford: CSLI Publications.

[55] Klein, E. 1971. *A Comprehensive Etymological Dictionary of the English Language*. Amsterdam: Elsevier.

[56] Kusalik, T. Product pregroups as an alternative to inflection. In: Casadio, C. and J. Lambek. 2005. *Computational Approaches to Natural Languages*. Polimetrica, Milan.

[57] Lambek, J. 1958. The mathematics of sentence structure, *American Mathematical Monthly*, 65:154-170.

[58] Lambek, J. 1959. Contribution to a mathematical analysis of the English verb-phrase, *Journal Canadian Linguistic Association*, 5:83-89.

[59] Lambek, J. 1975. A mathematical look at French conjugation, *Theoretical Linguistics* 2:203-214.

[60] Lambek, J. 1979. A mathematical look at Latin conjugation, *Theoretical Linguistics* 6:221-233.

[61] Lambek, J. 1986. A production grammar for English kinship terminology, *Theoretical Linguistics*, 13:19-36.

[62] Lambek, J. 1995c. On the nominalistic interpretation of natural languages. In *Quebec Studies in the Philosophy of Science*, Pt. 1, *Logic, Mathematics, Physics and History of Science: Essays in Honour of Hugues Leblanc*, eds. M. Marion and R.S. Cohen, 69-78. Dordrecht: Kluwer.

[63] Lambek, J. 1996. Number words and language origins, *The Mathematical Intelligencer*, 18:69-72.

[64] Lambek, J. 1999c. Type grammars revisited. In [71], 1-27.

[65] Lambek, J. 2000. Type grammars as pregroups, *Grammars* 4:21-39.

[66] Lambek, J. 2008. *From word to sentence*, Polimetrica, Milan.

[67] Lambek, J. 2009. The syntax-semantics interface and the origins of philosophy, *The mental lexicon* 4:121-128.

[68] Lambek, J. 2012. The mental lexicon and subconscious calculation, *Linguistic Analysis* 38: 57-68.

[69] Lambek, J. Logic and grammar, *Studia Logica* 100:667-681.

[70] Leach, E. 1958. Concerning Trobriand classes and the kinship category *tabu*. In *The Development Cycle of Domestic Groups*, ed. J. Goody. Cambridge: Cambridge University Press.

[71] Lecomte, A., F. Lamarche and G. Perrier, eds. 1999. *Logical Aspects of Computational Linguistics*, LNCS/LNAI, 1582. LACL 1997. Berlin: Springer.

[72] Lounsbury, F.G. 1965. Another view of Trobriand Kinship categories. In E.A. Hammond, ed., Formal semantics, *American Anthropologist*, 67:142-185.

[73] Malinowski, B. 1932. *Sexual Life of Savages*. London: Routledge and Kegan Paul.

[74] Miller, G.A. 1956. The magical number seven plus or minus two: some limits on our capacity for processing information, *Psychological Review*, 63:81-97.

[75] Montague, R. 1974. *Formal Philosophy: Selected Papers*. New Haven: Yale University Press.

[76] Moortgat, M. 1988. *Categorial Investigations: Logical and Linguistic Aspects of the Lambek Calculus*. Dordrecht: Foris.

[77] Moortgat, M. and R.T. Oehrle. 2005. Pregroups and type-logical grammar: searching for convergence. In Casadio et al. eds., *Logic and grammar*, CSLI Lecture Notes 168, Stanford.

[78] Moot. R. and C. Retoré. 2012. *The Logic of categorial grammars*, Springer-Verlag, Berlin.

[79] Oehrle, R.T., E. Bach, and D. Wheeler, eds. 1988. *Categorial grammars and Natural Language Structures*. Dordrecht: Reidel.

[80] Pentus, M. 1993. Lambek grammars are context free. In *Proceedings of the 8th Annual Symposium on Logic in Computer Science*, 429-433.

[81] Pinker, S. 1999. *Words and Rules: The Ingredients of Language*. New York: Basic Books.

[82] Preller, A. 2007. Towards discourse representation via pregroup grammars, *J. Logic, Language and Information*, **16**, 173-194.

[83] Preller, A. and V. Prince. 2008. Pregroup grammars and linear parsing of the French verb phrase. In [31], 53-84.

[84] Ruhlen, M. 1994. *On the origin of Languages*. Stanford: Stanford University Press.

[85] Savitch, W.J. et al., eds. 1987. *The Formal Complexity of Natural Language*. Dordrecht: Reidel.

[86] Sheiber, S. 1985. Evidence against the context-freeness of natural languages, *Linguistics and Philosophy*, 8:333-343.

[87] Theroux, P. 1993. *The Happy Isles of Oceania: Paddling the Pacific*. New York: Ballantine Books.

[88] van Benthem, J. 1983. The semantics of variety in categorial grammar. Report 83-29, Simon Fraser University, Burnaby, B.C. In Buskowski et al. 1990.

[89] van Benthem, J. 1988. The Lambek calculus. In Oehrle et al: 35-68.

[90] van Benthem, J. and A. ter Meulen, eds. 1997. *Handbook of Logic and Language*. Amsterdam: Elsevier.

References

[1] Altmann, S.L. 1986, *Rotations, quaternions and double groups*, Clarendon Press, Oxford.

[2] Altmann, S.L. 1989, Hamilton, Rodrigues and the quaternion scandal, *Math. Magazine* 62:291-308.

[3] Anglin, W.S. and J. Lambek 1995, *The heritage of Thales*, Springer-Verlag, New York.

[4] Bell, J. 2007, The continuous and the infinitesimal in mathematics and philosophy, *Polimetrica*, Milano.

[5] Bishop, E. and D. Bridges 1985, *Constructive Analysis*, Springer-Verlag, Berlin.

[6] Church, A. 1940, The foundation of the simple theory of types, *J. Symbolic Logic* 5:56-68.

[7] Došen, K. and P. Schroeder-Heister eds. 1993, *Substructural Logic*, Oxford University Press.

[8] Eilenberg, S. and S. Mac Lane 1945, General theory of natural equivalences, *Transactions of the American Mathematical Society* 58: 231-294.

[9] Fowler, D.H. 1987, The mathematics of Plato's Academy, Clarendon Press, Oxford.

[10] Gentzen, G. 1935, Untersuchungen über das logische Schliessen, Math. Zeitschrift 39, English translation in M.E. Szabo (ed), The collected papers of Gerhard Gentzen, North Holland, Amsterdam, 68-131.

[11] Girard, J.-Y. 1987, Linear logic, Theoretical Computer Science 50:1-102.

[12] Gödel, K. 1931, Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme, 1931. Monotshefte der Mathematischen Physik 38:173-198.

[13] Gödel, K. 1958, Über eine bisher noch nicht benutzte Erweiterung des finiten Standpunktes, Dialectica 12:281-287.

[14] Henkin, L.A. 1950, Completeness in the theory of types, J. Symbolic Logic 15:81-91.

[15] Jacobson, N. 1980, *Basic Algebra II*, Freeman, San Francisco.

[16] Kleene, S.L. 1952, *Introduction to metamathematics*, Van Nostrand, New York.

[17] Kock, A. 1981, *Synthetic differential geometry*, London Math. Soc. Lecture Notes Series 3333, Cambridge University Press.

[18] Lambek, J. and P.J. Scott, Intuitionist theory and the free topos 1980, *J. Pure and Applied Algebra* 19:215-257.

[19] Lambek, J. and P.J. Scott, 1986, *Introduction to higher order categorical logic*, Cambridge studies in advanced mathematics 7, Cambridge University Press.

[20] Lambek J. and P.J. Scott, 2011, Reflections on the foundations of mathematics, in: G. Sommaragu (ed), *Foundational theories of classical and constructive mathematics*, Western Ontario Series in Philosophy, Springer-Verlag: 171-186.

[21] Lawvere, F.W. 1972, *Introduction to toposes, algebraic geometry and logic*, Springer LNM 274:1-12.

[22] McLarty, C. 1992, Elementary categories, elementary toposes, Clarendon Press, Oxford.

[23] Mines, R., R. Richman and W. Ruitenberg, 1988, A course in constructive algebra, Springer-Verlag, New York.

[24] Penon, J. 1981, Infinitésimaux et intuitionisme, Cahiers topologie géometrie différentielle 22:67-72.

[25] Seldin, J.P. and J.R. Hindley (eds) 1980, To H.B. Curry: essays on combinatory logic, lambda calculus and formalism, Academic Press, London.

[26] Tarski, A. 1996, Logic, semantics, metamathematics, Clarendon Press, Oxford.

References

[1] Barnes, J. 1987/2002 *Early Greek Philosophy*, Penguin Books, London.

[2] Carroll, L. 1985 What the tortoise said to Achilles, *Mind* 4:278-280.

[3] Couture, J. 1997 Analyse logique et analycité: de Carnap à Gödel, *Dialectica* 51:95-117.

[4] Couture, J. and J. Lambek 1991 Philosophical reflections on the foundations of mathematics, *Erkeuntuis* 34:187-209.

[5] Lambek, J. 1981 The influence of Heraclitus on modern mathematics, in: J. Agassi and R.S. Cohen (eds) *Scientific philosophy today*, Reidel Publishing Co., Dordrecht: 111-121.

[6] Lambek, J. 1994 Are the traditional philosophies of mathematics really incompatible? *The Mathematical Intelligencer* 16:56-62.

[7] Lambek, J. 2004, What is the world of mathematics, *Annals Pure and Applied Logic* 126:149-158.

[8] McCall, S. (ed) 1967, *Polish Logic*, Oxford, Clarendon Press.

[9] Negrepontis, S. 2001, *The anthypharetic nature of Plato's dialectics*, Manuscript, Athens University.

[10] Peirce, C.S. 1887, The logic of relations, *The atomist* 7:161-217.

[11] Russell, B. 1945, *A history of western philosophy*, Simon and Schuster Inc. New York.

[12] Tait, W.W. 1986, Truth and proof: the Platonism of mathematics, *Synthèse* 68:341-370.

References

[1] Adler, S.L. 1994 Composite leptons and quarks constructed as triply occupied quasiparticles in quaternionic quantum mechanics, *Physics Letters*. B332, 358-365.

[2] Adler, S.L. 1995 *Quaternions, quantum mechanics and quantum fields*. Oxford University Press.

[3] Conway, A.W. 1911 On the application of quaternions to some recent developments of electrical theory, *Proc. Royal Irish Acad.* A29, 80.

[4] Conway, A.W. 1912 The quaternionic form of relativity, *Phil. Mag.* 24, 208.

[5] Conway, A.W. 1948 Quaternions and quantum mechanics, *Pontificia Academia Scientiarum* 12, 204-277.

[6] Conway, A.W. 1951 His life work and influence, *Proc. Second Canadian Math. Congress*, Toronto: University of Toronto Press.

[7] Davis, W.H. et al. (eds) 1998 *Cornelius Lanczos' collected papers with commentaries*, Volume iii, North Caroline State University, Raleigh.

[8] Dirac, P.A.M. 1945 Applications of quaternions to Lorentz transformations, *Proc. Roy. Irish Acad.* A50, 261-270.

[9] Feynman, R.P. 1985 *QED, the strange theory of light and matter*, Princeton University press, Princeton NJ.

[10] Feynman, R.P., R.B. Leighton and M. Sands, 1963/1977 *The Feynman Lectures on Physics*. Addison Wesley Publ. Co., Reading, Mass.

[11] Finkelstein, D., J.M. Jauch and D. Speiser, 1979 Notes on quantum mechanics, in: Hooker (ed.), *Logico-algebraic approach to quantum mechanics* II, 367-421, Reidel Publ. Co., Dordrecht.

[12] Gell-Mann, M. 1994 *The quark and the jaguar*, W.H. Freeman, New York.

[13] Gillan, P. 2010 How general relativity contains the Standard Model, arχiv:hep-th/011029603.

[14] Gsponer, A. 2005 *Integer-quaternion formulation of Lambek's representation of fundamental particles and their interaction*, ISRI Report, Geneva.

[15] Gsponer, A. and J.-P. Hurni, 1998 *Lanczos-Einstein-Petiau: from Dirac's equation to non-linear wave mechanics*, in: W.H. Davis et al.

[16] Gürsey, F. 1955 Contributions to the quaternionic formalism in special relativity, *Rev. Faculté Sci. Univ. Istanbul* A20, 149-171.

[17] Gürsey, F. 1958 Correspondence between quaternions and four-spinors, *Rev. Faculté Sci. Univ. Istanbul* A21, 33-54.

[18] Harari, H. 1979 A schematic model of quarks and leptons, *Physics Letters* 86B, 83-86.

[19] Harari, H. 1983 The structure of quarks and leptons, *Scientific American* 248, 56-68.

[20] Lambek, J. 1995 If Hamilton had prevailed; quaternions in physics, *The Mathematical Intelligencer* 17, 7-15.

[21] Lambek, J. 2000 Four-vector representations of fundamental particles, *International J. of Theoretical Physics* 39, 2253-2258.

[22] Lambek, J. 2013 In praise of quaternions, C.R. Math. Rep. Acad. Sci. Canada 35, 1-16.

[23] Lambek, J. 2013 A six-dimensional classification of fundamental particles.

[24] Lambek, J. 2013 Quaternions and three temporal dimensions.

[25] Lambek, J. 2013 Six-dimensional Lorentz category.

[26] Lanczos, C. 1998 Three articles on Dirac's equation, Zeitschrift für Physik 57(1929), 447-493 (reprinted in W.H. Davis et al).

[27] Maxwell, J.C. 1869 Remarks on the mathematical classification of physical quantities, *Proc. London Math. Soc.* 3, 224-232.

[28] Moore, W. 1989 *Schrödinger, Life and Thought*, Cambridge, Cambridge University Press.

[29] Nahin, P.J. 1990 Oliver Heaviside, *Sci. American*, 122-129.

[30] Penrose, R. 2004 *The road to reality: a complete guide to the laws of the universe*, Jonathan Cape, London.

[31] Ryder, L.H. 1985/1997 *Quantum field theory*, Cambridge University Press.

[32] Silberstein, L. 1912 Quaternionic form of relativity, *Phil. Mag.* 23, 790.

[33] Silberstein, L. 1924 *Theory of relativity*, MacMillan, London.

[34] Shupe, M.A. 1979 A composite model of leptons and quarks, *Physics Letters* 86B, 87-92.

[35] Sudbery, A. 1986 *Quantum mechanics and the particles of nature*. Cambridge University Press, Cambridge.

[36] Synge, J.L. 1972 Quaternions, Lorentz transformations and the Conway-Dirac-Eddington matrices, *Commun. Dublin Inst. Adv. Studies* A21, 1-67.

[37] Weiss, P. 1941 On some applications of quaternions to restricted relativity and classical radiation theory, *Proc. Roy. Irish Acad.* A46, 129-168.

[38] Weyl, H. 1950 *The theory of groups and quantum mechanics*, Dover Publication, New York.

Table of Contents

Prologue

Description for marketing purpose

The book "From rules of grammar to laws of nature" ought to be suitable for advanced undergraduate and beginning graduate students in the liberal arts, as well as for interested readers who are not put off by some rudimentary mathematical notation.

It discusses language inspired pre-Socratic philosophy and some elementary computational linguistics. It concentrates on certain aspects of *morphology*, the structure of words, and some recent mathematical approaches to *syntax*, which deals with the composition of words to form sentences.

In morphology, we restrict attention to kinship grammars, of interest to anthropologists, and to the construction of finite verb-forms in three mediterranean languages. In Latin there are ninety finite verb-forms, which may be calculated only after some special verb-stems have been memorized. In Turkish all verb-forms are constructed by concatination of certain *morphemes*, subject only to the so-called *vowel harmony*, which requires that the pronunciation of any vowel depends on that in the preceding syllable. In Arabic, as in biblical Hebrew (chosen for detailed illustration), the verb-forms can be calculated, provided the verbs are presented by three consonants and two vowels, subject only to certain phonological rewrite rules.

In syntax we concentrate on *production grammars*, made up from rewrite rules, and two logical or algebraic systems: the *syntactic calculus*, inspired by Ajdukiewicz and Bar-Hillel, and the more recent *pregroup grammars*, which may be traced back to Zellig Harris.

Turning to Natural Philosophy, nowadays called *Physics*, we look at special relativity and relativistic quantum mechanics, in particular at the electro-magnetic theories of Maxwell-Lorentz and Dirac. In place of the now fashionable representations of groups and Lie algebras, we go back to Hamiltonian quaternions, which were first applied to special relativity shortly before World War I, and whose *regular representations* came into prominence soon after World War II.

In place of the usual four-dimensional space-time, we admit six dimensions, allowing three dimensions for time. While the extra dimensions were introduced for the sake of mathematical elegance, they ought to help us to understand how Schrodinger's cat can be alive and dead at the same "time". We also present a classification of elementary particles with the help of six-vectors made up from the entries $0, 1$ and -1.

The book also discusses the notion of computability and presents a stone-age calculator that is much simpler than the original Turing machine.

The foundations of mathematics are of interest to philosophers, though not to most practicing mathematicians. They are here discussed in collaboration with Phil Scott. In particular, we compare ideas of the *formalist* Hilbert, the *Platonist* Gödel and the *intuitionist* Brouwer. We show that, when moderately formulated, they do not contradict each other.

In tribute to Brouwer, we employ intuitionistic higher order arithmetic, in which the rule of the excluded third, asserting that every proposition is either true or false, no longer holds. The language models, aka *possible worlds*, are the elementary toposes of Lawvere. It turns out that the initial object in the category of models may be constructed linguistically and may serve as a Platonic universe acceptable to Gödel, in which all true statements are provable as expected by Hilbert.