Learning by Gestalt

Karl Hayo Siemsen, INK, Hochschule Emden/Leer siemsenKH@gmail.com

A team was selected, some people with an understanding of mathematics, some engineers and myself – maybe a gestalt psychologist. This team brings the ideas and cooperation between Eino Kaila, the Nevanlinna family and their PhD students Ahlfors and Kurki-Suonio to mind. All members of the team agreed that the main task was to optimize our mathematics lessons for engineering students. This is normally a hard task with a high dropout rate. We had to admit we would fail, if we continued the conventional strategy.

Ahlfors had made a general proposal, influenced by Nevanlinna, Kaila, Wittenberg, Polya, Kline and Wertheimer (Fig. 01):

> Hadamard: "The object of mathematical rigor is to sanction and legitimize the conquests of intuition, and there never was any other object for it."
>
> **5. Genetic method**. "It is of great advantage to the student of any subject to read the original memoirs on that subject, for science is always most completely assimilated when it is in then ascent state." wrote James Clerk Maxwell. There were some inspired teachers, such as Ernst Mach, who in order to explain an idea referred to its genesis and retraced the historical formation of the idea. This may suggest a general principle: The best way to guide the mental development of the individual is to let him retrace the mental development of its *great lines*, of course, and *not the thousand errors of detail*.
> This genetic principle may safeguard us from a common confusion: If A is logically prior to B in a certain system, B may still justifiably precede A in teaching, especially if B has preceded A in history. On the whole, we may expect greater success by following suggestions from the genetic principle than from the purely formal approach to mathematics. (Italics highlighting by author)

Fig. 01: Some excerpts from Ahlfors: On the mathematics curriculum

The proposal was influenced by Ernst Mach (not by Ernst Mach the physicist, but by other, less known aspects of Mach "in the same body"). We traced his approach and ideas, carefully transferring it to the present time. As engineers we analyzed the traditional linear model of learning, found many strategic errors and decided to follow Ahlfors' proposal of a genetic model. These are two quite different models. Let us start with the linear one (see Fig. 02).

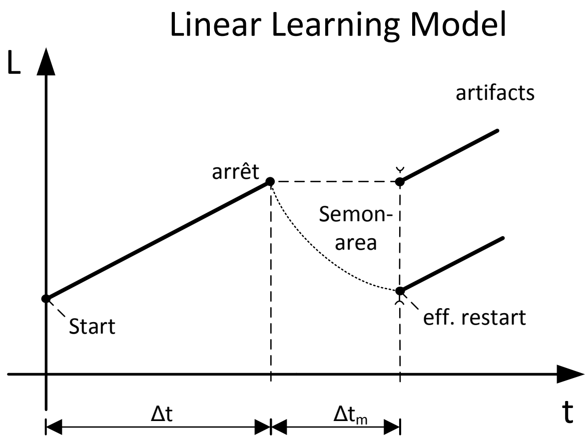


Fig. 02: Learning and building statistics with a linear model.

In Fig. 01 we have a start of the learning process and a stop. The time consumed is Δt. If we want to restart learning, we can take two different points: after Δt_m some contents is forgotten (Semon area). When one ignores this (the logically straightforward way), one starts at the old learning level Δt_m, thereby leaving some laggards behind (students who cannot close the memory gap by themselves). But this is not the biggest error. The undefined starting point makes a higher contribution. No care is taken of testing the kind of pre-knowledge or intuition. This start makes the tacit assumption that nothing is forgotten and that no pre-knowledge is required for axioms. But as Hadamard had observed, axiomatic generalizations, like for instance Hilbert made, require a high experience and intuition of pre-concepts.

No wonder that as a result, learning is reduced to *consuming* a well defined lessen contents with no relation to Gestalt psychology, Psychophysics, *Erkenntnis*theory or transfer to other contents. As science produces more and more contents, there is a contents "overflow" defined by curricula with more and more precise defined static contents. The relation between these contents in different areas is thereby neglected by the curricula.

This problem is resolved by an exponential model (Fig. 03).

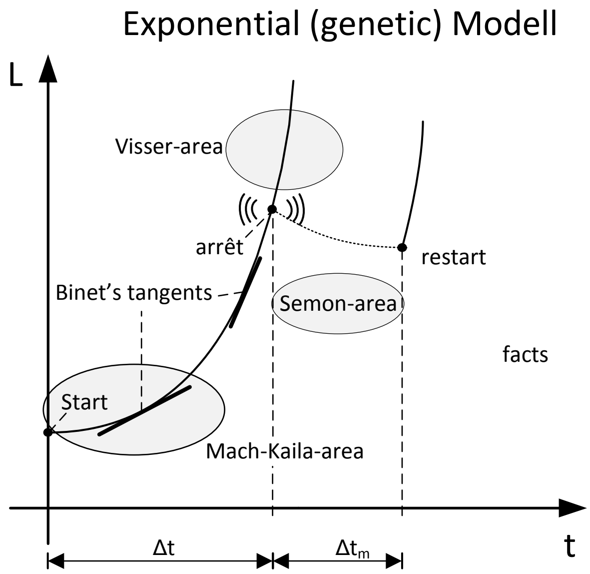


Fig. 03: Learning and building statistics with an exponential (genetic) model

An exponential learning curve has (maximum) two intersections with a linear learning curve. If we choose a learning time long enough, the exponential function *will* win (overtake the linear curve). But our gut feeling will tell: exponential learning consumes much more time. When we implement learning the Mach-Kaila way (which gives room for Gestalt processes, Psychophysics and *Erkenntnis*theory and *seems* to "crawl" along extremely slow), the gut feeling tends to be dominant in time perception. In an experimental arrangement we give exactly the same time Δt as in the linear model. The impact of the Mach-Kaila model and the exponential slope will correct out gut feeling. At the arrêt, we can even measure transfer to other contents. The slope of the tangents to the exponential function of learning was found experimentally by Alfred Binet (the inventor of the intelligence scale) during experiments in the school of Vaney between 1907 and 1910 (see Fig. 04).

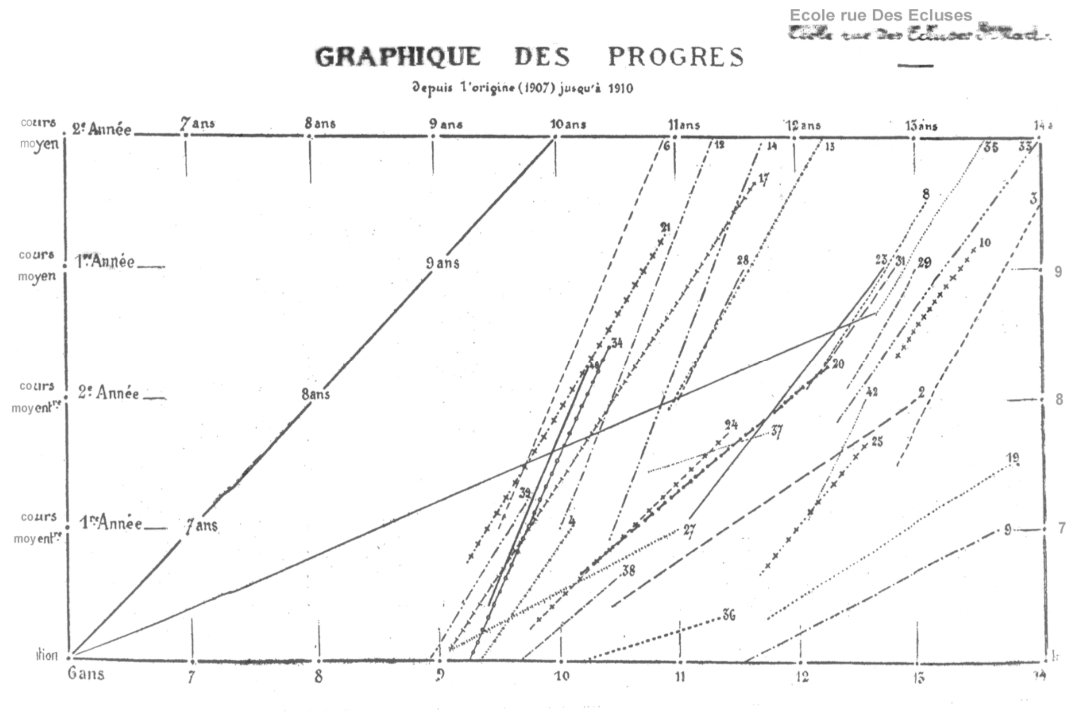


Fig. 04: Alfred Binet’s tangents; the slope of laggards with special education is higher then the slope of students in normal education. The Drawing from Vaney is slightly modified in the caption. Special thanks to the Alfred Binet Archive, Nancy for the kind permission of reproduction.

After the same amount of time Δt_m as given in the linear model, the extent of what is forgotten is less than in the linear model, because the Gestalts are present (and the slope is now exponential). The Visser area is for stabilizing and intuitivizing memory (i.e. making it faster and neuronally better connected by repetition), if needed.

In experimental arrangements, the exponential model is highly competitive. It does not matter that there are 60000 experiments referenced in the metastudy of Hattie with a linear model behind, opposed to less than 10 experiments with the genetic model. The linear model seems to produce artifacts, because of unrecognized starting conditions. The exponential model produces interpretable and non-contradictory facts. When the linear model is chosen to compare very short lessons, maybe the linear model has advantages. But one should have in mind that these short lessons are not typical for school learning and this form of learning might have similarity to "mental dressage" instead of understanding. If one mixes the exponential model with linear parts, one will produce Chimaeras, the antique, today scientifically inconsistent form of genesis, not the Darwinian form.

In the experimental learning model at the start there should be an archetype of the contents, a Gestalt showing the typical property of the contents at a glance. Because of the statistical properties of learning, one has to ensure that the Gestalt is perceived at once by each learner. If not, those laggards should be given opportunity to correct their perspective immediately.

These Gestalts should be as simple and memorable as possible. During the phase of early settling them, logical contradictions (typically used during academic expertise learning) or new different Gestalts should be avoided. While reordering the network brain, the Gestalts are pre-terms, terms in *status nascendi*. They do not already work as terms, so one cannot use them in the normal context of logical thinking of the learner.

One can look at this recipe from a different view called economy of thought (Denkökonomie, as Mach formulated[1]). To reach with a small input of thinking the highest possible output. Or to get the optimum effectivity from small resources. One has to keep consistent thoughts while avoiding inefficiency at the same time. The topic of economy of thought has to be as simple as possible at its start and simultaneously this should not be seen as an abstraction.

Let us have a look at a well known field experiment in Finland, compared with the OECD means (see fig. 05).

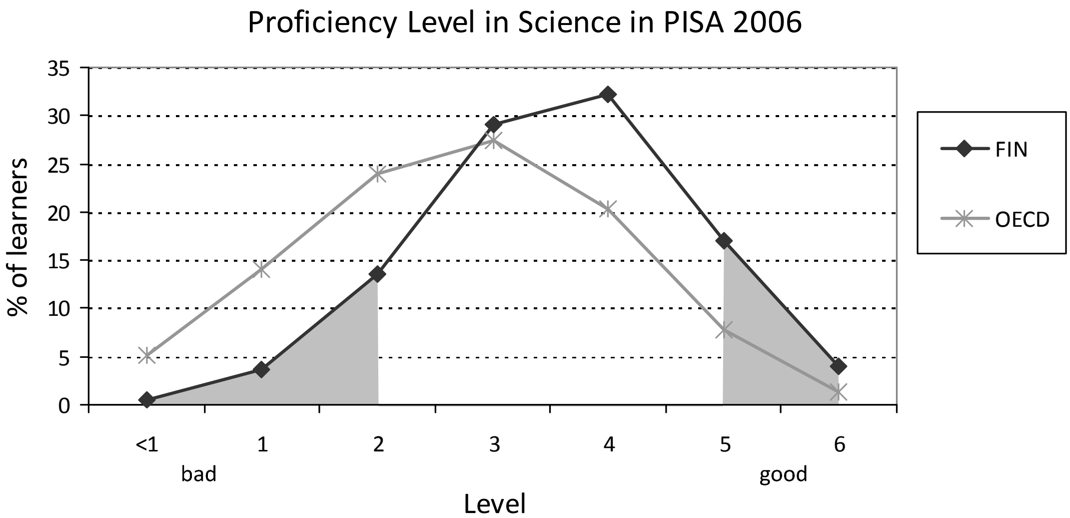


Fig. 05: Finland versus OECD mean in PISA study. The Finland curve is shifted to the right.

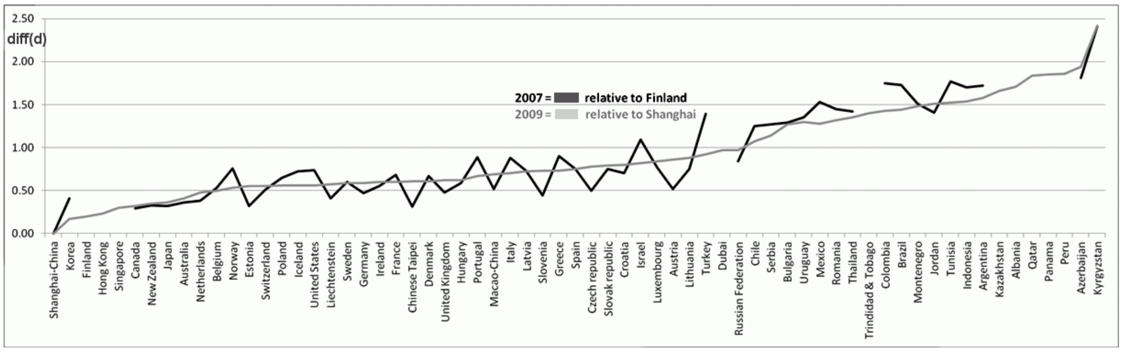


Fig. 06: Comparison worldwide relative to Finland (from an Email from Hattie to us).

We compute a d=1.0 for Finland (Fig. 05). (Hattie (2009) made a comparison with d of 800 metastudies on education.) We compare the Finland d with other countries worldwide building the difference (see Fig. 06). (Some of the Asian countries have double the number of school lessons in science as in Finland, so these countries are not directly comparable.)

[1] A Gestalt for Mach is a thought-economical function. Economizing is an optimization process of input and output (not a minimizing or maximizing process as it is often assumed). A thought process has to be effective (externally) as well as efficient (internally). Both aspects need to be optimized in dependency of a number of variables. So what is optimal might change depending on the importance one attaches to the different variables (for instance easiness to learn vs. logical elegance) or the time-frame one uses (time is one of the variables; one might view the result for an exam tomorrow as relevant or the value for the whole life of the student; the economy can thereby change quite dramatically).

Now we change our view to a lab experiment (see fig. 07).

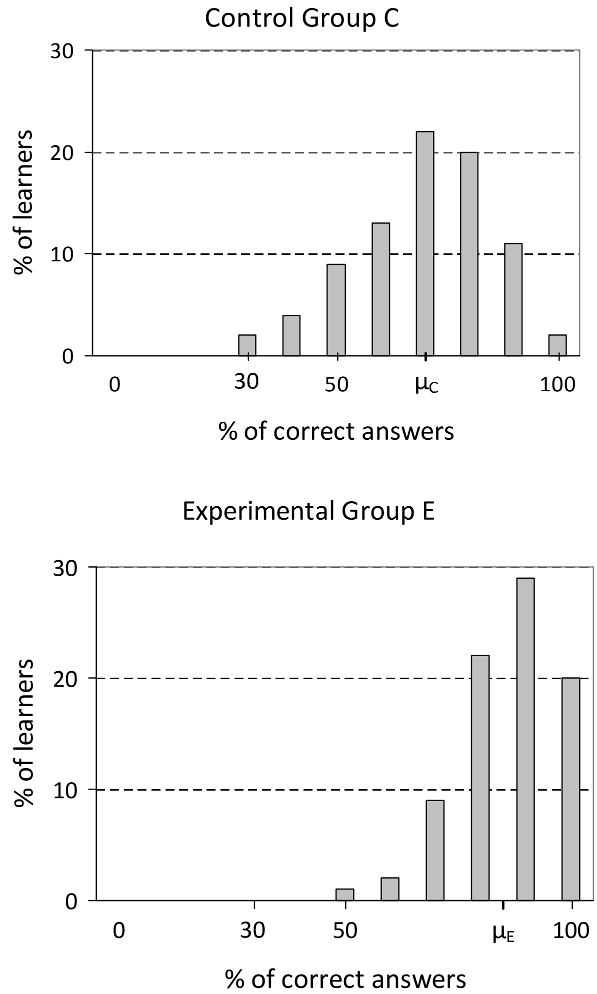


Fig. 07: Genetic-adaptative experiment. Control group and experimental group are educated. The shift in the statistics from left to right (control- to experimental group) brings a higher d.

In a field experiment, some variables are not controllable. In the example of Finland, the number of teachers educated the Mach-Kaila way was about one third of the full population of science teachers there. This seemed to be the biggest singular influence. In Fig. 07 (experimental group), the laggards have disappeared by teaching the Machian sensualistic way.

In principle it is possible to enhance the d value of a teacher (or a team of teachers) by enhancing their strategy to base all teaching on Gestalts, on Psychophysics and on *Erkenntnis*theory without ignoring the time constraint. The result of the enhancement can already be seen after two hours of teaching in an experimental design. This enhancement is stable if (and only if) the teaching style is stable. Each new contents needs new Gestalts independent of the age of the learner and each learner should perceive the Gestalt.

Now let us take another example (see fig. 08).

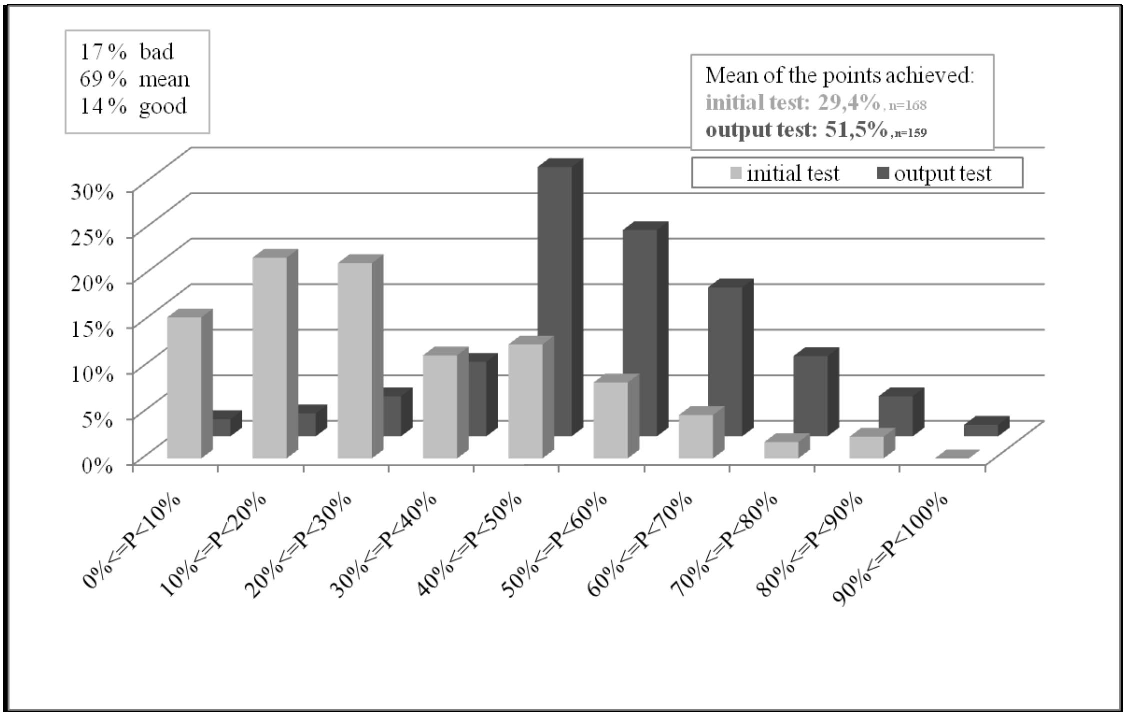


Fig. 08: Statistics of mathematics lesson for students in the first semester (2011). The same group is at start (gray) and at the end of educational treatment (black). $d_{\Delta} = 0.86$ (Δ is the reference to preliminary insights of the students).

The task is to optimize the statistics: "no laggards in the left half" (in the output test). This requires in this case a d_{Δ} much higher than actually supported. In the statistics example we do “some trials with a rake” to shift the laggards to the right. This is just like using a rake to rake the leaves from the left side of a court to the right half. We have the “rake” by the Kaila perceptional approach or very similar by the Machian sensualism. Teaching can be optimized this way by teaching the teachers to do the right things starting with a Gestalt, thereby modifying the learners' world view by himself and ignoring all the contents of the lesson which is forgotten within 14 days by students (let us call this strategy "Wagenschein's razor").

To reach a "d_{Δ}" value as needed much higher than available in the situation, we have to analyze the Gestalts to their simplest form in a way Ahlfors suggested (Fig. 01): reduce the Gestalts to the great lines, avoid the thousand errors of detail in history, ignore the logical sequence of B preceding A in history (the type of error George Sarton made following the historical line carefully while ignoring to reduce the Gestalts to their simplest forms, see Siemsen 2011). One has to be careful while restructuring the contents this way: listen to the proximate empirical answer of experiments to find the simplest elementary forms. Do not trust the "gut feeling", i.e. your intuition without further research.

I want to thank the members of the team, Dirk Rabe, Joachim Wiebe, Walter Schumacher, Luz Ezcurra and Hayo Siemsen for discussion of the exponential genetic learning model, Lisa Martin-Hansen for a hint and Hayo Siemsen especially for his hints to Binet and Vaney.

References

Ahlfors. On The Mathematics Curriculum Of The High School. The Mathematics Teacher of March 1962; American Mathematical Monthly of March 1962

Binet, A. (1911/1975). *Modern Ideas about Children*. Suzanne Heisler, Albi.

Vaney, V. (1911). Les classes pour arriérés. Recruitement, organisation, exercises d'orthopédie mentale. *Bulletin de la Société libre pour l'Etude psychologique de l'Enfant*, 73: 293-300. Archiv Binet, Nancy.

Hattie, J.A.C. (2009) Visible Learning. Oxon: Routledge.

OECD (2007). *PISA 2006 – Science Competencies for Tomorrow's World, Volume I and II*, http://www.oecd.org. Cited 12 Jun 2008.

Mach, E. (1890). Über das psychologische und logische Moment im Naturwissenschaftlichen Unterricht. *Zeitschrift für den physikalischen und chemischen Unterricht*, 4/1, October 1890: 1-5.

Mach, E.: 1897. Über Gedankenexperimente. In: *Zeitschrift für den physikalischen und chemischen Unterricht*, 10, Heft 1, Springer, Berlin

Mach, E. (1905/1926/2002). *Erkenntnis und Irrtum: Skizzen zur Psychologie der Forschung*.5th Edition, Leipzig, Berlin: reprint by rePRINT.

Mach, E. (1883/1888). Transformation and Adaptation in Scientific Thought. *The Open Court*: Jul 12 1888, 2/46, APS Online IIA.

Mach, E. (1892). Facts and Mental Symbols. The Monist;2;2. APS Online pg. 198

Mach, E. (1900). The Concept. *The Open Court* Jun 1900, 14.

Mach, E. (1906b). Sur le rapport de la physique avec la psychologie. *L'Année Psychologique*, Vol. 12: 303-318.

Montessori, M. (2012) Gesammelte Werke, Bd. 11. Psychoarithmetik. Herder, Freiburg.

Polya (Hrsg.) Wittenberg, Alexander Israel. The Prime Imperatives. Priorities in Education, Clarke, Irwin&Company Limited. Toronto, Vancouver 1968

Siemsen, K. H. (1981). *Genetisch-adaptativ aufgebauter rechnergestützter Kleingruppenunterricht. Begründungen für einen genetischen Unterricht*. Frankfurt: Lang.

Siemsen, H. (2011). Ernst Mach, George Sarton And The Empiry of Teaching Science. *Science & Education*, Online First.

Wagenschein, M. Was bleibt? In Rumpf, H. (Hrsg.) ...zäh am Staunen. Kallmeyer. Seelze-Velber 2002. ISBN 378004196-0

Wertheimer, M. (1912/1925). Über das Denken der Naturvölker, Zahlen und Zahlgebilde. In: *Drei Abhandlungen zur Gestalttheorie* (106-163). Erlangen: Verlag der Philosophischen Akademie.

Wertheimer, M. (1924/1938) Gestalt Theory. In: Ellis, W. D. (Ed.) *A Source Book of Gestalt Psychology*. London: Kegan, Trench, Trubner, 1-11.

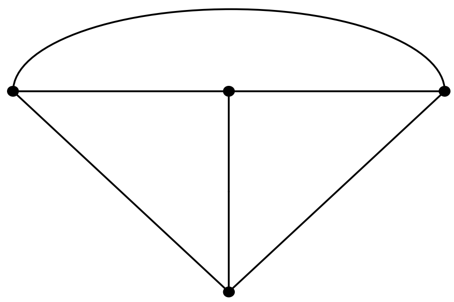
Wittenberg, A. (1957). Vom Denken in Begriffen. Mathematik als Experiment des reinen Denkens. Diss. ETH Zürich. eth-33688-02.pdf.

Wittenberg, A. I. (1957). *Vom Denken in Begriffen: Mathematik als Experiment des reinen Denkens*. Basel: Birkhäuser Verlag.

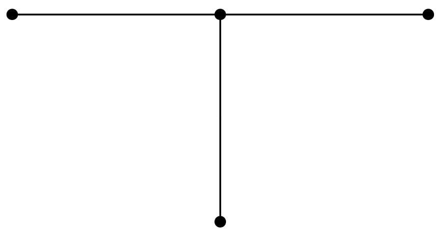
Wittenberg, A. I. (1963/1990). *Bildung und Mathematik*. Stuttgart: Klett.

Appendix: A Graph (an example)

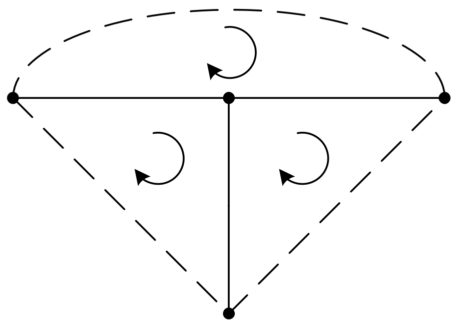
1 A tree, branches of a tree and nodes



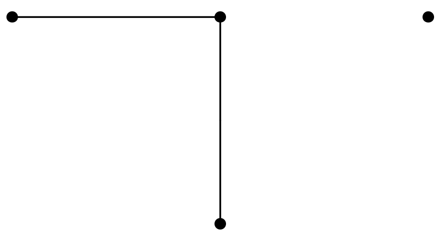
2 *One* complete tree (of many)



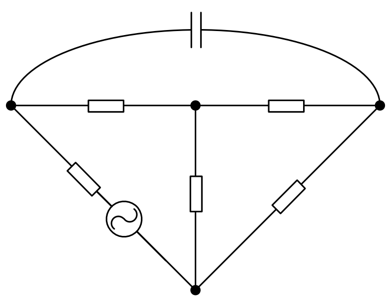
3 Boughs; loop analysis with boughs (Kirchhoff's law in mathematical form)



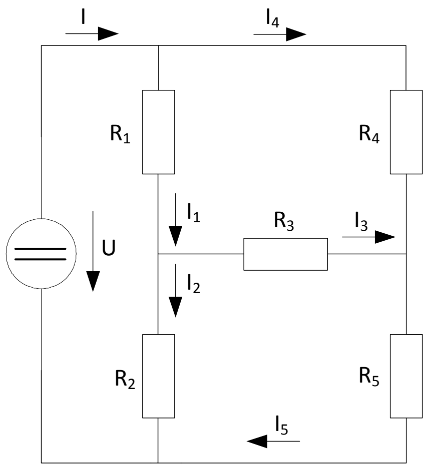
2.1 Breakup of a tree



4 Inflating the tree (Ohm's law, shown later)



5 Wheatstone bridge as an electric circuit. (Turn this picture 90° clockwise and the Gestalt will change to a coherent perception of the previous tree!)



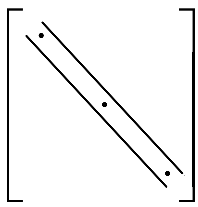
Spezielle Werte der Schaltelemente:

U = 10V	R_1 = 200
R_2 = 300	R_1 = 500
R_4 = 400	R_1 = 400

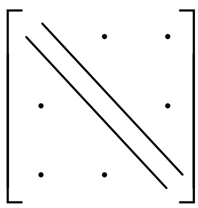
6 A two dimensional table in matrix notation is derived from the graph #4 ((x_i) and (u_i) are the Elements of a vector [ ] each):

$$\begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \cdot \begin{bmatrix} (x_i) \end{bmatrix} = \begin{bmatrix} (u_i) \end{bmatrix}$$

7 The main diagonal is filled with placeholders of those components, which are circulated by a loop variable.



8 The other elements in the table are dummy placeholders of the electronic components which are circulated by *two* loop variables. The indexes of the loop variables define the position in the matrix.



9 If the two loop variables circulate in opposite direction, the placeholder of the components becomes negative leading sign.

10 "u" are sources, which point in the direction of a loop variable, get a negative leading sign (and versed). Now all features of *Kirchhoff's law* are brought to the table.

11 Ohm's law ("the inflating"): in the table electronic devices (or other, for example mechanical devices) instead of the placeholders are positioned following the rules of 7 to 10. The u_i's are voltage sources, the i_i's are loop currents.

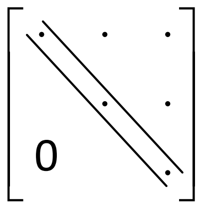
$$\underline{Z}\cdot\underline{I}=\begin{bmatrix}(u_i)\end{bmatrix}$$

12 Table 6 filled in this way is now called the "generalized Ohm's law".

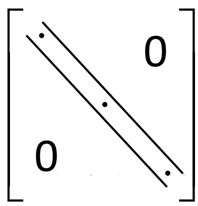
$$\underline{I}=\underline{Z}^{-1}\cdot\begin{bmatrix}(u_i)\end{bmatrix}$$

This Ohm's law is solved with respect to $\underline{I}$ by matrix calculus, for example building the upper triangular matrix first, then the diagonal matrix. The solution of the vector $[(i_i)]$ can be read directly from the matrix.

This matrix Z^{-1} is transformed into



and further to



(The same way of solution can be used for current controlled voltage sources. An alternative for the whole procedure is to use the node voltages (including voltage controlled current sources) instead of loop currents.)