Werner Fenchel, a pioneer in convexity theory and a migrant scientist

Christer O. Kiselman

Abstract. Werner Fenchel was a pioneer in introducing duality in convexity theory. He got his PhD in Berlin, was forced to leave Germany, moved to Denmark, and lived during almost two years in Sweden. In a private letter he explained his views on the development of duality in convexity theory and the many terms that have been used to describe it. The background for his move from Germany is sketched.

1. Introduction

Werner Fenchel (1905–1988) was a pioneer in convexity theory and in particular the use of duality there. When asked about his views on the many terms used to express this duality he described in a private letter (1977) the whole development from Legendre and onwards, as well as his preferences concerning the choice of terms. The background for his leaving Germany and moving to Denmark and later to Sweden is sketched in Section 12.

2. Convex sets and convex functions

In a vector space $E$ over the field $\mathbb{R}$ of real numbers—think of $E$ as the two-dimensional plane, identified with $\mathbb{R}^2$, or three-dimensional space, identified with $\mathbb{R}^3$—we define, given two points $a$ and $b$, the segment with endpoints $a$ and $b$:

$$[a, b] = \{(1-t)a + tb; \ t \in \mathbb{R}, \ 0 \leq t \leq 1\}.$$

A subset $A$ of $E$ is said to be convex if

$$\{a, b\} \subset A \implies [a, b] \subset A.$$

The modern theory of convex sets starts with the work of Hermann Minkowski (1864–1909); see his book (1910), a large part of which was published in 1896.

When calculating with functions, it will be convenient to allow infinite values, thus to let functions take values in $\mathbb{R} \cup \{-\infty\} \cup \{+\infty\}$, adding two infinities to $\mathbb{R}$. This is the set of extended real numbers, and will be denoted by $\mathbb{R}_e$ or $[-\infty, +\infty]$. It is ordered so that $-\infty < x < +\infty$ for all $x \in \mathbb{R}$.

The operation of addition,

$$\mathbb{R} \times \mathbb{R} \ni (x, y) \mapsto x + y \in \mathbb{R},$$
will cause difficulties when we try to define sums like $(+\infty) + (-\infty)$. A convenient solution, pioneered by Moreau (1963; 1966–1967:9) is to extend it in two different ways to operations $\mathbb{R} \times \mathbb{R} \to \mathbb{R}$, viz. as upper addition $x + y$, defined as $+\infty$ if one of the terms is equal to $+\infty$, and lower addition $x + y$, defined as $-\infty$ if one of the terms is equal to $-\infty$. (Some authors allow only values in $[-\infty, +\infty]$, but then the rules for avoiding $-\infty$ become complicated: if $f$ is admissible, maybe $-f$ is not; the infimum of an admissible function may not be admissible, etc.)

It is convenient to define convex functions with the help of convex sets. We define the finite epigraph of a function $f: E \to \mathbb{R}$ by

$$\text{epi}^F(f) = \{(x, t) \in E \times \mathbb{R}; \ t \geq f(x)\} \subset E \times \mathbb{R}.$$  

Then a function $f: E \to \mathbb{R}$ is defined to be convex when $\text{epi}^F(f)$ is convex.

In the other direction we note that a set $A$ is convex if and only its indicator function $\text{ind}_A$ is convex. Here $\text{ind}_A: E \to \mathbb{R}$ is defined to take the value 0 in $A$ and $+\infty$ in its complement. We have $\text{epi}^F(\text{ind}_A) = A \times [0, +\infty]$.

So far, there is no duality.

The book by Bonnesen and Fenchel (1934) is mainly concerned with convex bodies; it mentions the supporting function but does not go into the duality theory. Also—remarkably—Fenchel’s survey article (1983), a translation of a talk given in 1973, does not mention duality.

It is tempting to believe that convex sets possess some kind of regularity. But this is not so. Let $W$ be any subset of $\mathbb{R}^n$, as irregular as you like, for instance a set which is not measurable in the sense of Lebesgue (if you believe in the axiom of choice). Then the set

$$\{x \in \mathbb{R}^n; \|x\|_2 < 1\} \cup \{x \in W; \|x\|_2 = 1\},$$

where we use the Euclidean norm $\| \cdot \|_2$, is convex.

### 3. Duality in convexity theory

The definition of a convex set is done from the inside and does not need duality, but there is also a definition from the outside, which requires an introduction of duality.

We introduce the algebraic dual of $E$, which is the set of all linear forms $\xi: E \to \mathbb{R}$ and is denoted by $E^\ast$. An element of $E^\ast$ is thus a function $\xi: E \to \mathbb{R}$ such that $\xi(x + ty) = \xi(x) + t\xi(y)$ for all elements $x, y$ of $E$ and all real numbers $t$. We can now speak of hyperplanes in $E$: they are given by an equation $\xi(x) = c$ for a nonzero element $\xi$ of $E^\ast$ and a constant $c \in \mathbb{R}$, as well as affine functions on $E$; they are of the form $f(x) = \xi(x) + c$.

If $E$ has a topology, we also speak about the dual of $E$, meaning the set of all continuous linear forms on $E$ and denoted by $E'$. If $E$ is equal to $\mathbb{R}^n$ with the usual topology, we have $E' = E^\ast$. Every linear form can then be written as $x \mapsto \xi \cdot x = \xi_1 x_1 + \cdots + \xi_n x_n$ for some vector $(\xi_1, \ldots, \xi_n) \in \mathbb{R}^n$.

By duality in convexity theory I mean any consideration involving the dual or the algebraic dual to a given space. A typical example is that of a norm $\| \cdot \|'$ dual
to a given norm $\| \cdot \|$ and defined by

$$\|\xi\|' = \sup_{x \in E} \xi(x), \quad \xi \in E'.$$

If $A$ is a closed convex subset of $\mathbb{R}^n$, then it can be described as the intersection of a family of closed half spaces. Here a closed half space is of the form $\{ x \in \mathbb{R}^n; \xi \cdot x \leq t \}$ for some $\xi \in \mathbb{R}^n \setminus \{0\}$ and some real number $t$. This result is essentially the Hahn–Banach theorem in the finite-dimensional case. It follows that if $A$ is an open convex subset of $\mathbb{R}^n$ it can be described as the intersection of a family of open half spaces. An open half space is of the form $\{ x \in \mathbb{R}^n; \xi \cdot x < t \}$.

In general we need half spaces of a more general kind. Let us say that $Y$ is a refined half space if it is convex and satisfies

$$\{ x \in \mathbb{R}^n; \xi \cdot x < t \} \subset Y \subset \{ x \in \mathbb{R}^n; \xi \cdot x \leq t \}.$$

This means that $Y$ is the union of the open half space and a convex subset $A_\xi$ of the hyperplane given by the equation $\xi \cdot x = t$. Then any convex subset of $\mathbb{R}^n$ can be described as the intersection of a family of refined half spaces $Y_\xi$, $\xi$ of norm 1, taking $A_\xi$ as the set of points $y$ in $A$ which satisfy $\xi \cdot y = t$. (Cf. the notion of refined digital hyperplane in my paper (2004:456, Definition 6.2).)

4. The Legendre transformation

The Legendre transformation is named for Adrien-Marie Legendre (1752–1833), who introduced it in classical mechanics to go from the Lagrangian\footnote{Refers to Joseph-Louis Lagrange (1736–1813).} to the Hamiltonian\footnote{Refers to William Rowan Hamilton (1805–1865).} formulation: the Hamiltonian is, in modern language, the Legendre transform of the Lagrangian (see, e.g., Goldstein 1950:217).

If $f: \mathbb{R}^n \to \mathbb{R}$ is a differentiable mapping with gradient $\text{grad } f: \mathbb{R}^n \to \mathbb{R}^n$, we may take $\xi = (\text{grad } f)(x)$ as a new independent variable, thus looking at the inverse

$$(\text{grad } f)(x) = \xi \mapsto x = (\text{grad } f)^{-1}(\xi).$$

This is in general a multivalued mapping, which, in Legendre’s time, was no big deal. The Legendre transform of $f$ is the mapping

$$\mathbb{R}^n \ni \xi \mapsto \xi \cdot ((\text{grad } f)^{-1}(\xi)) - f((\text{grad } f)^{-1}(\xi)) = \xi \cdot x - f(x).$$

(See Hiriart-Urruty & Lemaréchal 2001:209.) If $f$ is strictly convex, then $x$ is uniquely determined by $\xi$, and we arrive at the Fenchel transform (see Section 6 below). If $f$ is convex, the point $x$ is not uniquely determined by $\xi$: the expression $\xi \cdot x - f(x)$ can attain its maximum for several choices of $x$. 

Refers to Joseph-Louis Lagrange (1736–1813).

Refers to William Rowan Hamilton (1805–1865).
5. Tropicalization

Tropicalization means, roughly speaking, to replace a sum or an integral by a supremum. A simple example is the $l^p$-norm,

$$\|x\|_p = \left( \sum_{j=1}^n |x_j|^p \right)^{1/p}, \quad x \in \mathbb{R}^n, \quad 1 \leq p < +\infty,$$

which becomes

$$||x||_\infty = \left( \sup_{j=1,...,n} |x_j|^p \right)^{1/p} = \sup_{j=1,...,n} |x_j|$$

when the sum is replaced by the supremum. We shall see that the Fenchel transformation is a case in point.

Tropical geometry is now a field of growing interest; see, e.g., Richter-Gebert et al. (2005).

6. The Fenchel transformation

The Fenchel transform $\tilde{f}$ of a function $f: \mathbb{R}^n \to \mathbb{R}$ is defined as

$$\tilde{f}(\xi) = \sup_{x \in \mathbb{R}^n} (\xi \cdot x - f(x)), \quad \xi \in \mathbb{R}^n.$$ 

Clearly $\xi \cdot x - f(x) \leq \tilde{f}(\xi)$, which can be written as

$$\xi \cdot x \leq f(x) + \tilde{f}(\xi), \quad (\xi, x) \in \mathbb{R}^n \times \mathbb{R}^n,$$

called Fenchel’s inequality. It follows that the second transform $\tilde{\tilde{f}}$ satisfies $\tilde{\tilde{f}} \leq f$.

The operation $f \mapsto \tilde{f}$ is a cleistomorphism (closure operator) if we define the order between functions by inclusion of their finite epigraphs, which is actually a very natural thing to do. This is because $+\infty$ corresponds to vacuum since its finite epigraph is empty, while $-\infty$ corresponds to an infinitely dense black hole—in the model for material objects, the density is $e^{-f}$.

On the other hand, if we order functions by the relation $f(x) \leq g(x)$ for all $x$, then the operation is an anoiktomorphism. In this way, the Fenchel transformation enters the field of mathematical morphology with its complete lattices, ethnomorphisms, cleistomorphisms and anoiktomorphisms.

The equality $\tilde{f} = f$ holds if and only if $f$ is convex, lower semicontinuous, and takes the value $-\infty$ only if it is $-\infty$ everywhere. Let us denote by $\text{CVX}_0(\mathbb{R}^n)$ the family of all functions with these three properties. The restriction of the Fenchel transformation to $\text{CVX}_0(\mathbb{R}^n)$ is thus an involution, i.e., it is equal to its own inverse. Artstein-Avidan & Milman (2009:662) have proved that any decreasing involution $\text{CVX}_0(\mathbb{R}^n) \to \text{CVX}_0(\mathbb{R}^n)$ is, up to some obvious modifications, the restriction to $\text{CVX}_0(\mathbb{R}^n)$ of the Fenchel transformation. Thus the Fenchel transformation is, essentially, unique among all transformations with the indicated properties. While this is, intuitively, hardly surprising, it is a very nice property and a convincing witness to the importance of Fenchel’s work.
If $f$ is defined on an arbitrary vector space $E$, we define the Fenchel transform on its algebraic dual $E^*$:

$$
\tilde{f}(\xi) = \sup_{x \in E} (\xi(x) - f(x)), \quad \xi \in E^*.
$$

Again,

$$
\xi(x) \leq f(x) + \tilde{f}(\xi), \quad (\xi, x) \in E^* \times E.
$$

We can apply the transformation a second time, but then it is convenient to choose first an arbitrary subspace $\Xi$ of $E^*$, for instance $\Xi = E'$, and define

$$
\tilde{\tilde{f}}(x) = \sup_{\xi \in \Xi} (\xi(x) - \tilde{f}(\xi)), \quad x \in E.
$$

The Fenchel transformation $f \mapsto \tilde{f}$ is a tropical counterpart of the Laplace transform. The Laplace transform of a function $g: [0, +\infty[ \to [0, +\infty[$ is

$$
(\mathcal{L}g)(\xi) = \int_0^\infty g(x)e^{-\xi x} dx, \quad \xi \in \mathbb{R}.
$$

If we replace the integral by a supremum and take the logarithm, we get

$$
(\mathcal{L}_{\text{trop}} g)(\xi) = \sup_x (\log g(x) - \xi x) = \tilde{f}(-\xi), \quad \xi \in \mathbb{R}, \quad f = -\log g.
$$

### 7. The supporting function

A special case of the Fenchel transform is obtained when the function is the indicator function of some set: $f = \text{ind}_A$. Then $\tilde{f} = H_A$, the supporting function of $A$,

$$
H_A(\xi) = \sup_{x \in A} \xi \cdot x, \quad \xi \in \mathbb{R}^n.
$$

This function was introduced already by Minkowski (1903:448).

However, conversely, a Fenchel transform $\tilde{f}$ is also a supporting function if we go up in dimension. We equip $\mathbb{R}^n \times \mathbb{R}$ with the inner product

$$
(\xi, \tau) \cdot (x, t) = \xi \cdot x + \tau t, \quad (\xi, \tau), (x, t) \in \mathbb{R}^n \times \mathbb{R}.
$$

Then

$$
\tilde{f}(\xi) = H_{\text{epi}F}(\xi, -1), \quad \xi \in \mathbb{R}^n.
$$

### 8. Galois correspondences

Let $X$ and $Y$ be two ordered sets. A Galois correspondence is a pair $(F, G)$ of decreasing mappings $F: X \to Y$ and $G: Y \to X$ such that $G \circ F$ is larger than the identity in $X$ and $F \circ G$ is larger than the identity in $Y$. Taking both $F$ and $G$ as the Fenchel transformation, we see that the two form a Galois correspondence, provided we order the functions using the inclusion relation between the finite epigraphs. (See, e.g., my paper 2010: Section 4.)

$^3$Named for Évariste Galois (1811–1832).
9. Werner Fenchel


After his PhD, Werner Fenchel became an assistant at Göttingen University. He was dismissed from this position in 1933. He accepted an invitation to Copenhagen and arrived to Denmark in the summer of 1933. (Jessen 1987:89.)

In December 1933 he married Käte Sperling (1905–1983), who was also a mathematician, and who had been ousted from a position as high-school teacher in Berlin. In Denmark, Werner Fenchel was first supported by various foundations until he

---

4Bieberbach joined the National Socialist Lecturers Association in November 1933 and a little later became the representative of this organization at the University of Berlin (Mehrtens 1987:220). He joined the *Sturmabteilung* (SA) on 1934 November 05 and the *Nationalsozialistische Deutsche Arbeiterpartei* (NSDAP) on 1937 May 01 (Bieberbach's personal file in the archive at Humboldt University, signature UK B 220, folio 64; information provided by Reinhard Siegmund-Schultze on 2016 January 26). In a lecture in 1934 he justified the Nazi-organized boycott of Edmund Landau's classes (Mehrtens 1987:227). However, in 1928 he could still serve as advisor for Fenchel: “no one had ever observed antisemitic tendencies in Bieberbach before the summer of 1933” (Schappacher 1998:130). For more on Bieberbach, see Segal (2003).
got a teaching position in 1938 and became a lecturer at Copenhagen University in 1942. (Jessen 1987:89.)

As it turned out, Copenhagen was not sufficiently far away: on 1940 April 09, Germany invaded Denmark. However, during the first years of the occupation, Jews were not deported from Denmark; deportations started in October 1943. A majority of all Jews in Denmark could escape to Sweden, although some were deported to Theresienstadt (Jacoby 2015). Werner Fenchel came to Sweden in October 1943 (Jessen 1987:89).

A Danish school, den danske Skole i Lund, started on 1943 November 15 in Lund in southern Sweden and functioned during 19 months, to the end of the war. Werner Fenchel and many others taught there, for example Harald Bohr. (Siegmund-Schultze 2009:107, 136.) Also his wife Käte Fenchel taught at the Danish school, and Werner Fenchel lectured also at Lund University (Jessen 1987:89).

In 1950/51 Werner Fenchel was at the Institute for Advanced Study in Princeton, NJ, and met Harold W. Kuhn (1925–2014) and Albert W. Tucker (1905–1995) there:

It must therefore have been very exciting for Kuhn and Tucker when they became aware that Werner Fenchel from Copenhagen University, who was the leading expert on convexity at the time, was visiting the Institute for Advanced Study in Princeton as part of a sabbatical year in the USA in 1950/51. Tucker invited Fenchel to give a series of lectures on the theory of convexity at the mathematics department at Princeton University within his ONR project. (Kjeldsen 2010:3250)

Fenchel wrote notes (1951) based on his lectures in Princeton, expanding his short paper (1949):

Fenchel was motivated by problems and connections in pure mathematics, and his lecture notes became highly influential for further developments within the theory of convexity, especially through the work of R. T. Rockafellar [50]. (Kjeldsen 2010:3251)

Werner Fenchel died on 1988 January 24.

10. Werner Fenchel’s letter

I wrote a letter (in Swedish) to Werner Fenchel on 1977 February 24, asking him about the history of duality in convexity theory and the many terms used in that theory. He replied with a letter dated 1977 March 07 (3 pages; in Danish). He reported in detail on the development of convexity theory and the duality which had been known under so many different names, and where he himself had been a pioneer. I translate it into English below.

---

5 Refers to Rockafellar (1970). This book the author “dedicated to Fenchel, as honorary co-author” (Rockafellar (1970:viii)).

6 For the benefit of young readers, let me mention that Fenchel’s letter as well as mine were written on paper using a kind of mechanical device called “typewriter,” which had “arms” or “bars” equipped with profiles in the form of letters, upper case letters being above the corresponding lower case letter, like this: A. A tiny moment before an arm hit the paper, a textile ribbon, slightly moistened with a kind of ink, jumped up in front of the paper and because of the pressure from the arm left a mark on the paper, very much looking like a letter or symbol, although often a little fuzzy.
Dear Kiselman.

Thanks for the letter! It is not easy for me to answer. I have again after many years looked at the relevant literature, which has delayed my answer. I briefly report on the history of the matter.

The classical Legendre transformation was introduced purely formally by Legendre. That it is an involution and closely related to duality was, however, as far as I know, first discovered by Monge somewhat later. One shall of course keep the name Legendre transformation.

W. H. Young has in a small paper from 1912 proved the following. Let \( \varphi(x) \) be positive and differentiable with \( \varphi'(x) > 0 \) for \( x > 0 \), and let \( \psi(y) \) be the inverse of \( \varphi \). Then we have with \( b = \varphi(a) \),

\[
xy - ab \leq \int_x^a \varphi(\xi) d\xi + \int_y^b \psi(\eta) d\eta.
\]

If we put \( f(x) = \int_a^x \varphi(\xi) d\xi \) and \( \tilde{f}(y) = ab + \int_b^y \psi(\eta) d\eta \), we easily see that \( f \mapsto \tilde{f} \) is a special case of the classical Legendre transformation, which Young, however, did not notice, even though this is clear from his proof. He is only interested in the inequality.

The next contribution to the topic is due to Z. W. Birnbaum & W. Orlicz, 1931. They consider convex functions \( f(x) \) which are defined for \( x \in \mathbb{R} \) and satisfy \( f(0) = 0, f(x) > 0 \) for \( x > 0 \), \( f(-x) = f(x) \), \( f(x)/x \to 0 \) as \( x \to 0 \).
Werner Fenchel, in the late sixties or early seventies

\[ \frac{f(x)}{x} \to \infty \text{ as } x \to \infty. \]  

Here the “complementary” function \( \tilde{f} \) to \( f \) is defined by

\[ \tilde{f}(y) = \sup_{x \in \mathbb{R}} (xy - f(x)), \]

and it is shown that \( \tilde{\tilde{f}} = f \). The authors also show that \( f(x) = \int_0^x \varphi(\xi)d\xi \) and \( \tilde{f}(y) = \int_0^y \psi(\eta)d\eta \), where \( \varphi \) is non-decreasing and continuous to the right, and \( \psi \) is its inverse, conveniently defined with the same properties.

In 1939 there appeared a somewhat imprecise note by S. Mandelbrojt\(^7\) where he defined \( \tilde{f} \) by (\( \ast \)) for an arbitrary convex function of one variable, but without investigating under which conditions the transformation \( f \mapsto \tilde{f} \) is an involution. It is clear that he did not know about Birnbaum & Orlicz.

When I wrote my paper “On conjugate convex functions” in 1949, I also did not know about Birnbaum & Orlicz. I define there \( \tilde{f} \) for an arbitrary convex function defined in a convex subset \( C \) of \( \mathbb{R}^n \) by

\[ \tilde{f}(y) = \sup_{x \in C} (\langle x, y \rangle - f(x)) \]

and find a necessary and sufficient condition for the equality \( \tilde{\tilde{f}} = f \) to hold. The inequality

\[ \text{inequality} \]

\(^7\)That his note is imprecise, as Fenchel writes, is due to the fact that Mandelbrojt assumes real values of the functions without noticing that, if \( f \) has real values and does not grow faster than all linear functions, then \( \tilde{f}(y) \) will not be real valued; it will take the value \( +\infty \) when \( y \) or \( -y \) is large.
\[ \langle x, y \rangle \leq f(x) + \tilde{f}(y) \]

is a trivial consequence. I also show that \( f \mapsto \tilde{f} \) under suitable differentiability hypotheses is the Legendre transformation. In Lecture Notes, Princeton 1951, I have given a more systematic treatment of the transformation and several variants. A further development is due to R. T. Rockafellar from 1963.

Generalizations to infinite-dimensional spaces have been undertaken independently by J. J. Moreau, 1962, A. Brøndsted, 1964, and U. Dieter, 1965. In the bibliography in Rockafellar’s book “Convex Analysis” all the mentioned papers are listed.

Now to the question of the name! I admit that the situation is rather chaotic. That I called \( \tilde{f} \) the function conjugated to \( f \) depends on the fact that I considered the relation as a generalization of the “conjugated exponent” of classical analysis \( x^p/p \) and \( y^q/q \) with \( 1/p + 1/q = 1 \). This was also the starting point for Young and Birnbaum–Orlicz. But the connection with that has become very weak and the name is therefore not adequate. Rockafellar and Brøndsted have retained my terminology. Moreau calls \( f \) the function dual to \( f \). But in the special case considered by Birnbaum and Orlicz, he talks about the Young dual. The inequality (**) is called Young’s inequality by several authors. This is OK, but I cannot see any reason for calling \( f \mapsto \tilde{f} \) the Young transformation. If you want to include a personal name, it seems Legendre is the closest at hand, even though it does not create the correct association. The “maximum transformation” I have not seen. That is a possibility. I do not want to add a new name, but nevertheless say, that if I now should give the transformation one, I would let me be guided by the analogy and the connection with polarity between convex sets (in dual vector spaces) and for instance call it parabolic polarity.

With kind regards

Werner Fenchel

I thanked Fenchel in a letter dated 1977 April 02 and told him that Christer Borell and several other persons in Uppsala had discussed it, and that we liked very much the term parabolic polarity, which I had not seen before. However, I remarked that it is rather unwieldy to call \( \tilde{f} \) the function parabolically polar to \( f \).

I later met Werner Fenchel at a conference in Uppsala and could continue the discussion on the history of duality and possible terms.

11. A plethora of terms

As Fenchel wrote concerning names: “I admit that the situation is rather chaotic.” Since the Fenchel transformation is also an example of a Galois correspondence, where the situation concerning concepts and the terms used for them is even more

---

8 Mentioned in my earlier letter to Fenchel; the term was used by Bellman & Karush (1963).
chaotic, this story can be continued: terms like residual mapping, adjunction, upper adjoint, and lower adjoint occur. (See my paper 2010, Section 4). The reason for this state of affairs is pretty clear: the concepts are indeed fundamental in nature, but they are approached from many different directions, often by researchers who are not aware of earlier work.

In Section 6 above I have used the terms that I prefer now. However, several terms have been used for the main concepts involved here: for the supporting function of a set; for a Fenchel transform \( \tilde{f} \); for the Fenchel transformation \( f \mapsto \tilde{f} \); and for Fenchel’s inequality \( \xi \cdot x \leq f(x) + \tilde{f}(\xi) \). I list a few of them below.

- le Gendre [Legendre] (1789; Lú le 1st Septembre 1787): “Au lieu de considérer \( z, p, q \) comme des fonctions de \( x \) et \( y \), rien n’empêche de regarder \( x, y, z \) comme des fonctions de \( p \) et \( q \); […]” (p. 315). This is the very beginning of what we now call the Legendre transformation. Here it is about a function \( z \) of two variables \( x \) and \( y \) with \( p = \partial z/\partial x \) and \( q = \partial z/\partial y \). The symmetry or duality between \( z, p, q \) and \( x, y, z \) is evident. Later in the article the author considers functions of more than two variables.

- Minkowski 1896: in Minkowski (1910) appear the terms “Stützebene” (p. 13); “nirgends concave Fläche” (p. 35), meaning a convex surface; “überall convexe Fläche” (p. 38), meaning a strictly convex surface. The part containing these terms was published already in 1896.

- Minkowski 1897: What is now known as “Minkowski addition” appears for instance in a paper published in 1897, reproduced in Minkowski (1911a:108).

- Minkowski (1903): “konvexer Körper” (p. 447); “Stützebenfunktion” (p. 448).

- Minkowski (1911b): “konvexer Körper” (p. 131); “Stützebene” (p. 136).

- Birnbaum & Orlicz (1931): “konjugierte Potenzen” (p. 2); “Die \( N \)-Funktion \( N(v) \) heißt konjugiert” (p. 6); “komplementäre Funktion” (p. 8).


- Mandelbrojt (1939): “qu’on peut appeler la fonction convexe associée à \( f(x) \)” (p. 977).

- Fenchel (1949): The two functions are called “conjugate” to each other.

- Goldstein (1950): “the Legendre transformation” (p. 215).

- Landau & Lifshitz (1960): “Legendre’s transformation” (p. 131).

- Moreau (1962): “la fonction convexe duale de \( f \)” (p. 2897).

- Bellman & Karush (1963): “maximum convolution” (p. 67); “maximum transform” (p. 68).


• Moreau (1966–1967): “fonction polaire de $h$” (p. 33); “une paire de fonctions duales (ou que chacune est la duale de l’autre) si chacune est la fonction polaire de l’autre” (p. 45).

• Abraham & Marsden (1967): “a Legendre transformation” (p. 122).

• Иоффе & Тихомиров (1968): “конволюция” (for infimal convolution); “преобразование Юнга или сопряженная” ‘Young’s transformation or conjugate’; “неравенство Юнга” ‘Young’s inequality’ (p. 55).

• Rockafellar (1970): “This $f^*$ is called the conjugate of $f$” (p. 104); “support function” (p. 112); “subgradient” (p. 214); “subdifferential” (p. 215); “The classical Legendre transformation for differentiable functions defines a correspondence which, for convex functions, is intimately connected with the conjugacy correspondence” (p. 251); “The Legendre conjugate,” “the Legendre transformation” (p. 256).

• Kiselman (1978): “The Legendre transformation.”


• Fenchel (1983): “support function” (p. 127).

• Hörmander (1983): “supporting function” (p. 105); “The quadratic forms […] are said to be dual” (p. 206); “dual cone” (p. 257).

• Moreau (1993, in a private letter to Hiriart-Urruty): “C’est moi qui ai introduit le terme sous-gradient (et sa traduction anglaise subgradient).”

• Hiriart-Urruty & Lemaréchal (1993): “Theory of Conjugate Functions” (p. 36); “the so-called Fenchel correspondence, and is closely related to the Legendre transform” (p. 38); “Young–Fenchel inequality” (p. 38).

• Hörmander (1994): “Conjugate convex functions (Legendre transforms)” (p. 16); “The Legendre transform (also called the conjugate function)” (p. 17); “Then the Legendre transform (= conjugate function = Fenchel transform)” (p. 67); “supporting function” (p. 69).

• Singer (1997): “subdifferential” (p. 19); “Fenchel–Young inequality” (p. 254); “Fenchel–Rockafellar theorem” (p. 273); “Fenchel–Moreau theorem” (p. 433).

• Gao (2000): “Legendre transformation” (p. 27); “Legendre conjugate transformation” (p. 27); “Legendre Duality Theorem” (p. 29); “Legendre duality pair” (p. 30); “Legendre conjugate functions” (p. 30); “Legendre–Fenchel transformation” (p. 32); “Fenchel transformation” (p. 32); “Fenchel-conjugate function” (p. 32); “Fenchel–Young inequality” (p. 32); “Fenchel-conjugate” (p. 32); “Fenchel conjugate pair” (p. 234); “sub-differential” (p. 236); “sub-gradents” (p. 236).

• Hiriart-Urruty & Lemaréchal (2001): “Legendre transform” (p. 209) for the original Legendre transform; “The conjugate of a function”; “conjugacy operation”; “the Legendre–Fenchel transform” (p. 211).

• Murota (2003): “Legendre–Fenchel transform” (p. 10); “Legendre–Fenchel transformation” (p. 10); “Legendre duality” (p. 12); “subdifferential” (p. 80); “subgradient” (p. 80); “convex conjugate” (p. 81).


In 1977, I still called the transformation the Legendre transformation, following my advisor Lars Hörmander (1931–2012), who, with some emphasis, insisted that the Fenchel transformation is the same as the Legendre transformation.

I later switched to calling it the Fenchel transformation. The motivation for this is that I am now of the opinion that the results of Legendre belong to differential geometry, whereas Fenchel’s seminal work belongs to the theory of ordered sets and complete lattices, which has a very different flavor.

As Fenchel wrote, the name Legendre does not create the correct association. Rockafellar (1970:251–260) distinguishes between the conjugacy correspondence, the Legendre conjugate, and the Legendre transformation. As I mentioned in Section 6 above, the Fenchel transformation is a tropicalization of the Fourier or Laplace transformation and does not invoke differentiability as the Legendre theory does. It is indeed an early example of the tropicalization of mathematics, which is now a very active field of research.

The terms transform for $\tilde{f}$ and transformation for $f \mapsto \tilde{f}$ are very practical and follow a well-known pattern: the Fourier transformation, the Laplace transformation, the Radon transformation, the Fenchel transformation, …

12. Why Werner Fenchel had to leave Germany

Many scientists in Germany lost their jobs after the Nazi Machtübernahme in 1933. Fenchel was one of them. It seems appropriate to describe in some detail this development, which is of interest also because of the general decline of science and culture in Germany. See also Gordin (2015: Chapter 7, Unspeakable).

As described in my paper (2017), German mathematics and physics were world leading in many respects during the nineteenth century and the beginning of the twentieth century. Scientists in many other countries, including Sweden, wrote in German, but after 1945, this became rare in Sweden.

The reasons behind the transition from German to English in scientific writing are often considered to be Germany’s defeat in the Second World War and the subsequent rise of the USA to the world’s leading country in engineering and natural science. This opinion needs to be qualified.

12.1. Science education in Germany 1933–1939

A factor which, according to my opinion, has not been given sufficient attention, is how drastic the loss in German science was already before the start of the Second World War. This can be illustrated by a few figures.
In the year 1932, the number of students in mathematics, insurance mathematics, and physics in Germany was 7,139. Just before the war, in the beginning of 1939, this number had gone down to 1,270, a decrease by 82 per cent in seven years (Mehrtens 1985:85). In Göttingen, where, as Laurent Schwartz wrote in his memoirs (1997:79), the greatest scientists of the world resided, the number of mathematics students went down from 432 till 37, a decrease by 91 per cent during the same period.

We can compare with how Harald Bohr (1887–1951) described Göttingen in an earlier epoch:

While Göttingen was in many ways a provincial town, calm and peaceful, the richest scientific life flourished there. A spirit of genuine international brotherhood of a rare intensity reigned there among the many young mathematicians who came from nearly all over the world to make a pilgrimage to Göttingen, bound together as they were by their common interest in and love for their science. (Bohr 1952:xxi–xxii)

The importance of Göttingen as a center is further discussed by Thomas Schott:

Göttingen and Bohr’s orientation toward this place is not an exceptional instance. It is one of a recurring pattern of how the scientific community fits into the various national societies according to a center and periphery theory […] (Schott 1979:86)

However, the center of a science is not a fixed point. It can move.

12.2. The law of 1933

Among the decisions causing this disaster in higher education and research was the Gesetz zur Wiederherstellung des Berufsbeamtenums ‘Law for the Restoration of the Professional Civil Service’, enacted by the government (not by the Reichstag) on 1933 April 07, little more than two months after the Machtübernahme, and signed by Der Reichskanzler Adolf Hitler, Der Reichsminister des Innern Frick, and Der Reichsminister der Finanzen Graf Schwerin von Krosigk.

As the law was first drafted by the Interior Minister Wilhelm Frick (1877–1946; Reichsminister des Innern 1933–1943), all Beamte nicht arischer Abstammung ‘civil servants of non-Aryan descent’, as well as several other categories, were to be fired immediately at the Reich, Länder, and Municipal levels of government. However, the President of Germany, Paul von Hindenburg (1847–1934; Reichspräsident 1925–1934) objected to the bill until it had been amended to exclude three classes of civil servants from the ban.

The cardinal provision in the law is Article 3:

§ 3

(1) Beamte, die nicht arischer Abstammung sind, sind in den Ruhestand (§§ 8 ff.) zu versetzen; soweit es sich um Ehrenbeamte handelt, sind sie aus dem Amtsvorhaben zu entlassen. (Gesetz zur Wiederherstellung des Berufsbeamtenums)

Civil servants who are not of Aryan descent are to be put into the state of retirement (§§ 8 et seq.); in so far as it is about a commission of trust, they are to be dismissed from service. (My translation)

The law contained provisions for reducing the pensions of these retired civil servants.

The second paragraph of Article 3 contained certain exceptions, imposed by President Hindenburg. However, the first and sixth articles opened for a broader application:
Werner Fenchel, a pioneer in convexity theory

§ 1

(1) Zur Wiederherstellung eines nationalen Berufsbeamten- und zur Vereinfachung der Verwaltung können Beamte nach Maßgabe der folgenden Bestimmungen aus dem Amt entlassen werden, auch wenn die nach dem geltenden Recht hierfür erforderlichen Voraussetzungen nicht vorliegen. (Gesetz zur Wiederherstellung des Berufsbeamten-
tums)

For the restoration of a national professional civil service and in order to simplify administration, civil servants can, in accordance with the following provisions, be dismissed even if the required legal conditions are not fulfilled. (My translation)

§ 6

Zur Vereinfachung der Verwaltung können Beamte in den Ruhestand versetzt werden, auch wenn sie noch nicht dienstunfähig sind. [...] (Gesetz zur Wiederherstellung des Berufsbeamten-
tums)

In order to simplify administration, employees can be transferred to the state of retirement, even if they are not yet unsuitable for service. [...] (My translation)

12.3. Hilbert

David Hilbert (1862–1943) was the founder of the theory of Hilbert spaces and known for having presented in Paris on 1900 August 08 ten of a total of twenty-three problems that were to keep mathematicians busy for a long time. He was a professor at Göttingen and is reported to have summed up the situation there soon after Hitler’s Machtübernahme thus:

Sitting next to the Nazis’ newly appointed minister of education at a banquet, he was asked, “And how is mathematics in Göttingen now that it has been freed of the Jewish influence?”

“Mathematics in Göttingen?” Hilbert replied. “There is really none any more.” (Reid 1970:205)

Reinhard Siegmund-Schultze (personal communication 2014 August 19) remarks, however, that there is no documentation showing that Hilbert actually pronounced these words in the presence of the minister of education.

12.4. Einstein, Weyl, Fenchel, and many others

Albert Einstein (1879–1955) resigned his position at the Preußische Akademie der Wissenschaften in a letter dated 1933 March 28, before he was ousted (Pais 1982:450). He accepted an offer (one of many) from the Institute for Advanced Study in Princeton, NJ. Most dismissed scientists could of course not find something similar.

Hermann Weyl (1885–1955) succeeded Hilbert in Göttingen in 1930. He was not Jewish, but his wife Friederike Bertha Helene Weyl (1893–1948) was. The law said nothing about non-Aryan descent of spouses, but, as noted above, it opened

In spite of Einstein’s own decision to leave, a caricature exhibited at the Einstein Museum in Bern (a part of the Musée d’histoire de Berne, Bernisches Historisches Museum) shows Einstein being kicked down a staircase by a boot, mentioning that he mistakenly thought he was a Prussian: “Der Hausknecht der Deutschen Gesellschaft in Brüssel wurde beauftragt, einen dort herumlungernden Asiaten von der Wahnvorstellung, er sei ein Preuße, zu heilen.” ‘The house servant of the German embassy in Brussels got the commission to free an Asian tramp from his misconceived idea that he was a Prussian.’
wider possibilities based on “Vereinfachung der Verwaltung.” Weyl did not want to leave Germany and therefore declined an early offer to join the Institute for Advanced Study, but later changed his mind and accepted a second offer in 1933. Another victim of this “Vereinfachung” was Michael A. Sadowsky (1902–1967), who got his doctorate at the Technische Hochschule Berlin in 1927 and later lost his teaching permit there: “In 1933 Michael Sadowsky’s teaching permit was revoked because his wife was of Jewish descent. However, the legal base for such a decision was established only in 1937.” (Brüning et al. 1998:6; see also Siegmund-Schultze 2009:37–38).

Among the mathematicians forced to retire or who were dismissed were Otto Blumenthal (1876–1944), Edmund Landau (1877–1938), Felix Bernstein (1878–1956), Emmy Noether (1882–1935), Max Born (1882–1970), Richard Courant (1888–1972), and Stefan Bergman (1895–1977).

Of all these scientists, I came to know one personally: Werner Fenchel.

12.5. Zentralblatt and Mathematical Reviews

A reviewing journal called Zentralblatt für Mathematik und ihre Grenzgebiete was founded in 1931 by Otto Neugebauer (1899–1990). It provided reviews and abstracts for articles in pure and applied mathematics. In 1938, the publisher, Springer-Verlag, asked for written assurances that no Jews would act as reviewers (Jackson 1997:330). A number of reviewers in England and the United States resigned in protest (Lehmer 1988:265). Neugebauer left the Zentralblatt in 1938 when the name of a Jewish editor (Levi-Civita) was, without further notice, eliminated from the title page. (Mehrtens 1987:217)

The Zentralblatt no longer registered all mathematical publications. Something had to be done. A new reviewing journal, named Mathematical Reviews, was started in the United States. It was founded in 1940 by the same Otto Neugebauer, now at Brown University. (Nowadays most users rely on its online version MathSciNet.)

12.6. The end of a glorious era

It is clear that the glorious era for mathematics in Germany ended well before the beginning of the war—and independently of the outcome of the war as well as of the rise of the United States.

12.7. Hermann Minkowski

In another epoch, people moved to Germany rather than out of it: Hermann Minkowski, who was born in the Kovno Governorate in the Russian Empire (now Kaunas, Lithuania), moved as a young boy with his family to Königsberg in 1872

10“Legal base” . . . of a very special kind.
11Harald Bohr, who worked with Landau in Göttingen, wrote: “Landau made a strong impression on everyone who came into contact with him. His baroque sense of humour and his exceptional vitality, characterized equally his scientific research and his teaching. His thinking was amazingly quick and sure and his standards for precision and exactitude of exposition were absolute and inexorable.” (Bohr 1952:xxii).
12Tullio Levi Civita (1873–1941).
to avoid persecution. He became a brilliant mathematician, an early pioneer of convexity theory, and later taught in Bonn, Königsberg, and Zürich (where Einstein studied for him in the Eidgenössische Polytechnische Schule), and finally in Göttingen.

Acknowledgments

Tom Fenchel, son of Käte Fenchel, née Sperling, and Werner Fenchel, allowed me to publish a translation of his father’s letter. In addition, he kindly sent me three photos of his father and allowed them to be published here. And it was Tinne Hoff Kjeldsen who brought me into contact with Tom!

Reinhard Siegmund-Schultze provided important historical information, and Seidon Alsaody, Sharon Rider, and Arild Stubhaug made valuable comments to earlier versions of the manuscript.

Michael D. Gordin encouraged me to write this paper separately from my earlier paper on language choice in doctoral theses presented at Uppsala University and in the journal *Acta Mathematica*, and sent me important comments on both manuscripts.

Staffan Rodhe sent me a link to the seminal article by le Gendre [Legendre] (1789). Jean-Baptiste Hiriart-Urruty sent encouraging comments, drew my attention to the nice result of Shiri Artstein-Avidan and Vitali Milman (2009), and sent me a copy of a letter from Jean-Jacques Moreau. Julinsz Brzezinski sent me comments and corrections.

For all this invaluable help I am most grateful.

References


Hörmander, Lars. 1963. *Konveksa och subharmoniska funktioner. Seminarier, Stockholms universitet ht 63* [Convex and subharmonic functions. Seminars, Stockholm University, Fall Term 1963].


Иоффе, А. Д.; Тихомиров, В. М. Двойственность выпуклых функций и экстремальные задачи. *Успехи математических наук*. т. XXIII, вып. 6 (144), 51–116.


Jacoby, Ruth. 2015. Så skapades den makabra bilden av nazisternas ”möblerläger” [“Thus was created the macabre image of the Nazis’ “Model Camp”]. *Dagens Nyheter* 2015 May 10.


Kiselman, Christer O. 2017. Language choice in scientific writing: The case of mathematics at Uppsala University and a Nordic journal. *Noram* **61**(02), ??–??.


Uppsala University, Department of Information Technology

Paper address: P. O. Box 337, SE-751 05 Uppsala

Amber addresses: kiselman@it.uu.se, christer@kiselman.eu

URL: www.cb.uu.se/~kiselman