

The fundamental theorem of calculus for Lipschitz functions*

SANJO ZLOBEC[†]

Abstract. *Every smooth function in several variables with a Lipschitz derivative, when considered on a compact convex set, is the difference of a convex function and a convex quadratic function. We use this result to decompose anti - derivatives of continuous Lipschitz functions and augment the fundamental theorem of calculus. The augmentation makes it possible to convexify and monotonize ordinary differential equations and obtain possibly new results for integrals of scalar functions and for line integrals. The result is also used in linear algebra where new bounds for the determinant and the spectral radius of symmetric matrices are obtained.*

Key words: *Lipschitz function, decomposition of functions, convexification, fundamental theorem of calculus, definite integral, ODE, Newton's second law, Green's theorem, determinant*

AMS subject classifications: 26A24, 26A36, 26B12, 26B25

1. Introduction

This paper is based on a decomposition result from [21]. The result says that every smooth, i.e., continuously differentiable, function f of several variables with Lipschitz derivative, when considered on a compact convex set, is the difference of a convex function and a convex quadratic function; see Section 2. The main results of the paper are contained in Section 3 where we use the decomposition to augment the fundamental theorem of calculus for continuous Lipschitz functions. The classic theorem is “unquestionably the most important theorem in calculus . . . ” [15, p. 287]. It has two parts. When taken together they say that differentiation and integration are inverse processes: each undoes that the other does. It is stated in the literature for continuous functions. In this paper we look at the theorem under the assumption that the continuous function is Lipschitz. This yields additional information on anti - derivatives and integrals which are incorporated here into the

*Research partly supported by NSERC of Canada. Presented in part in plenary talks at IX International Symposium on Parametric Programming and Related Topics, Cienfuegos, Cuba, March 2007 and at the 4th Croatian Mathematical Congress, Osijek, Croatia, June 2008.

[†]McGill University, Department of Mathematics and Statistics, Burnside Hall, 805 Sherbrooke Street West, Montreal, Quebec, Canada, e-mail: zlobec@math.mcgill.ca

classic theorem. The augmentation shows that, in particular, ordinary differential equations (ODEs) can be reduced to equivalent ODEs with only convex or monotone solutions. The fact that one knows a priori that the solution is convex or monotone is helpful for numerical methods, e.g. [4]. There are first - order ODEs that can not be “convexified”, but the reduction works for the second - (and higher -) order ODEs. This is important because ODEs that follow from Newton’s Second Law of mechanics, e.g., [8, 14], have twice differentiable solutions and hence they can be both convexified and monotonized. The reduction is illustrated on the movement of a pendulum. Since the augmentation is given in terms of convex functions, it may yield endlessly many results on Lipschitz functions and their anti - derivatives. One only has to apply general properties of convex functions to the convex terms in the augmentation. In Section 5, the line integral formulation of the fundamental theorem of calculus is recalled from [2]. Its augmentation yields possibly new properties of line integrals, in particular, for Green’s theorem.

The decomposition result from [21] can be used in two directions: One is to “convexify” smooth problems, e.g. [21, 23, 24, 25], see also [11]. We follow this direction to augment the fundamental theorem of calculus in Sections 2 - 5. Along the other direction one can extend properties of convex functions to non - convexity. This was demonstrated, e.g., for Jensen’s inequalities in [20, 22]. Here, in Section 6, a well - known inequality for convex functions of symmetric matrices from [12, 24] is extended beyond convexity. The extension yields new bounds for the determinant of symmetric matrices. In particular, a lower bound of Hadamard’s determinantal inequality and a lower bound of the spectral radius of symmetric matrices are obtained.

Instead of decomposing functions into differences of convex functions and convex quadratics one can essentially do the same for concave functions and concave quadratics. The assumption, that we are working with Lipschitz functions, is not too restrictive for the applied mathematician because, loosely speaking, “almost all” meaningful (having a real life interpretation) continuous functions on compact sets *are* Lipschitz.

2. The basic decomposition result

In this section we recall the decomposition result from [21] and prove it for scalar functions. First, a continuous scalar function $f : \mathbb{R} \rightarrow \mathbb{R}$ is said to be Lipschitz, or that it has the Lipschitz property, on an interval $I = [a, b]$ if

$$|f(s) - f(t)| \leq L|s - t|$$

for every s and t in I . The number L is a Lipschitz constant. Similarly, a continuous vector function $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is said to be Lipschitz on a set S in \mathbb{R}^n if

$$\|F(x) - F(y)\| \leq L\|x - y\|$$

for every x and y in S for some constant L . The vector norm is chosen to be Euclidean. Continuously differentiable functions are called “smooth functions”. We will study these on compact convex sets K in \mathbb{R}^n . Their derivatives may or may

not have the Lipschitz property. The derivatives do have the Lipschitz property on K if

$$\|\nabla^T f(x) - \nabla^T f(y)\| \leq \tilde{L}\|x - y\| \tag{2.1}$$

for every x and y in K for some constant \tilde{L} . Note that $L \neq \tilde{L}$; $\nabla^T f(x)$ denotes the transposed of the (row) gradient $\nabla f(x)$. A continuous $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is convex on a convex set K if

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y) \text{ for every } 0 \leq \lambda \leq 1 \text{ and } x \in K, y \in K.$$

This paper is based on the following decomposition result from [21].

Theorem 2.1. (Basic Decomposition of Smooth Functions in Several Variables with a Lipschitz Derivative) *Consider a smooth function in several variables $f : \mathbb{R}^n \rightarrow \mathbb{R}$ on a compact convex set K contained in its open domain. If the derivative of f is Lipschitz on K , then there exists a number α^* such that for every $\alpha \leq \alpha^*$*

$$f(x) = C(\alpha, x) - q(\alpha, x)$$

where $C(\alpha, x)$ is a convex function and $q(x) = -\frac{1}{2}\alpha x^T x$ is a convex quadratic function on K .

Numbers α and α^* in *Theorem 2.1*, which “make” $C(\alpha, \cdot)$ convex, are called convexifiers of f on K . If α' is a convexifier, so is every $\alpha < \alpha'$. Function $C(\alpha, x)$ is called a convexification of f on K . A continuous function f is convexifiable (or “weakly convex”, e.g., [17]) on K if it has a convexification for some α . If $\alpha \leq \alpha^*$ is replaced by $\alpha < \alpha^*$, then the phrase: “ $C(\alpha, x)$ is a convex function and $q(x) = -\frac{1}{2}\alpha x^T x$ is a convex quadratic” can be replaced by “ $C(\alpha, x)$ is a strictly convex function and $q(x) = -\frac{1}{2}\alpha x^T x$ is a strictly convex quadratic”. Note that the claim of the theorem can be rephrased: “If the derivative of f is Lipschitz on K , then f becomes convex after adding to it a particular convex quadratic function q ”, e.g. [10, 19].

For two important classes of smooth functions one can determine convexifiers explicitly. These classes are:

- i) Continuously differentiable functions with Lipschitz derivatives. For these functions, (2.1) holds and one can choose $\alpha^* = -\tilde{L}$, e.g., [21].
- ii) Twice continuously differentiable functions. If f is such function, then one can choose

$$\alpha^* = \min_{x \in K} \min_{\lambda} \lambda(\nabla^2 f(x))$$

where $\lambda = \lambda(\nabla^2 f(x))$ is an eigenvalue of the Hessian $\nabla^2 f(x)$ of f at a given x , e.g. [19].

Illustration 2.2. The graphs of $f(x_1, x_2) = \sin x_1 \cos x_2$ and of its convexification $C(-4, x_1, x_2) = f(x_1, x_2) + 2(x_1^2 + x_2^2)$ on

$$K = \{x = (x_i) \in \mathbb{R}^2 : -\frac{\pi}{2} \leq x_1, x_2 \leq \frac{\pi}{2}\}$$

are depicted in *Figure 1*. Function f is neither convex nor concave on K .

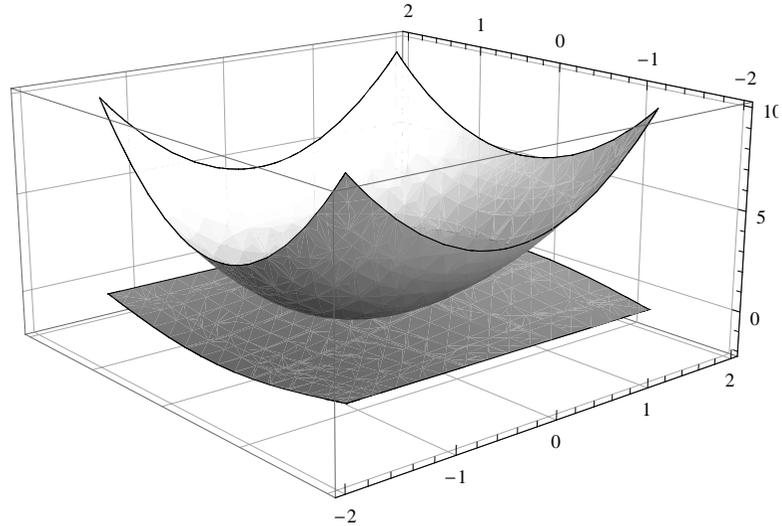


Figure 1. Function f and its convexification C

Let us prove a scalar version of *Theorem 2.1*. For scalar variables, we use the notation $s, t \in \mathbb{R}$ instead of $x, y \in \mathbb{R}^n$.

Corollary 2.3. (Decomposition of Smooth Scalar Functions with a Lipschitz Derivative) *Consider a smooth scalar function $y : \mathbb{R} \rightarrow \mathbb{R}$ on a compact interval I of its open domain. If the derivative of y has the Lipschitz property on I , then there exists a number α^* such that for every $\alpha \leq \alpha^*$*

$$y(s) = C(\alpha, s) - q(\alpha, s) \tag{2.2}$$

where $C(\alpha, s)$ is a convex function and $q(s) = -\frac{1}{2}\alpha s^2$ is a convex quadratic function on I . In particular, one can specify $\alpha^* = -\tilde{L}$, where \tilde{L} is a Lipschitz constant of the derivative of y on I .

Proof. We know that for s and t , $s \neq t$ in I

$$|[y'(s) - y'(t)](s - t)| \leq \tilde{L}(s - t)^2$$

since the derivative y' is assumed to be Lipschitz. Using the absolute value property, this implies

$$-\tilde{L} \leq \frac{[y'(s) - y'(t)](s - t)}{(s - t)^2} \leq \tilde{L}.$$

Hence

$$\alpha(s - t)^2 < [y'(s) - y'(t)](s - t) \text{ for every } \alpha < -\tilde{L}.$$

This yields

$$\{[y'(s) - y'(t)] - \alpha(s - t)\}(s - t) > 0. \tag{2.3}$$

Consider $\varphi(t) = y(t) - \frac{1}{2}\alpha t^2$. The derivative of $\varphi(t)$ is $\varphi'(t) = y'(t) - \alpha t$. Since (2.3) can be written

$$[\varphi'(s) - \varphi'(t)](s - t) > 0,$$

we conclude that φ is strictly convex, hence convex, on I , e.g. [12]. □

Theorem 2.1 and (2.2) provide a link between smooth functions with Lipschitz derivatives and convex functions. In particular, loosely speaking, they say that every smooth function with a Lipschitz derivative is a convex quadratic “away” from the set of convex functions for all sufficiently small α 's. This proximity to convexity is used in, e.g. [24] to formulate canonical forms of mathematical programs. In the illustration below, we depict this proximity for the sine function.

Illustration 2.4. Consider $f(s) = \sin s$. We have

$$f(s) = [f(s) + q(\alpha, s)] - q(\alpha, s) = C(\alpha, s) - q(\alpha, s)$$

where $C(\alpha, s) = f(s) - \frac{1}{2}\alpha s^2$ and $q(\alpha, s) = -\frac{1}{2}\alpha s^2$ are convex for every $\alpha \leq -1$.

Note that lower the parameter α , the “more convex” $C(\alpha, s)$.

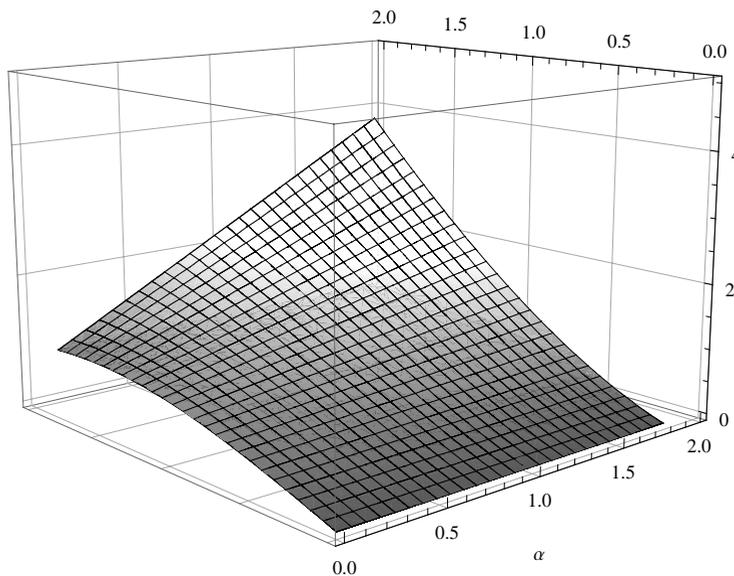


Figure 2. *Function $f(s) = \sin s$ and its convexifications $C(\alpha, s)$*

A characterization of continuous convexifiable functions is given in [23].

3. Augmentation of the fundamental theorem of calculus

The fundamental theorem of calculus has two parts. The first part states that for a continuous scalar function $f : \mathbb{R} \rightarrow \mathbb{R}$ on an interval $[a, b]$ the function

$$y(s) = \int_a^s f(\xi) d\xi, \quad a \leq s \leq b \tag{3.1}$$

is continuous on $[a, b]$ and differentiable on (a, b) and its derivative is $y'(s) = f(s)$. (Thus y is an anti - derivative of f .) We wish to know what can be said about y if f is both continuous *and* Lipschitz. An answer follows from *Corollary 2.3*. However, we need a compact interval I in (a, b) . On this I , y is the difference of a convex function and a convex quadratic function. The theorem below recalls the classic version in (i) and it gives its augmentation in (ii).

Theorem 3.1. (The fundamental theorem of calculus: First part)

- (i) Consider a scalar function $f : \mathbb{R} \rightarrow \mathbb{R}$ that is continuous on an interval $[a, b]$. The function $y(s)$ defined by (3.1) is continuous on $[a, b]$ and differentiable on (a, b) and its derivative is $y'(s) = f(s)$.
- (ii) Assume that f is continuous and Lipschitz on $[a, b]$ with a Lipschitz constant L . Then for every compact subinterval I in (a, b) there is a number α^* such that $y(s)$ can be decomposed as

$$y(s) = C(\alpha, s) - q(\alpha, s) \tag{2.2}$$

for every $\alpha \leq \alpha^*$, where $C(\alpha, s)$ is a convex function and $q(\alpha, s) = -\frac{1}{2}\alpha s^2$ is a convex quadratic on I . In particular, one can specify $\alpha^* = -L$. Moreover, $f(s) = M(\alpha, s) - \alpha s$ for every $\alpha \leq \alpha^*$, where $M(\alpha, s)$ is a monotone increasing function.

Proof. For the proof of (i) see, e.g. [15, p. 280]. In order to prove (ii) we use (i). We note that for a continuous *and* Lipschitz f , y is smooth and it has a Lipschitz derivative. Therefore, by *Corollary 2.3*, y is convexifiable. Furthermore, we know that $\alpha^* = -L$ is a convexifier of y , by, e.g. [21]. (Observe that $y'(s) = f(s)$!) From the properties of convex functions it follows that also every $\alpha \leq \alpha^*$ is a convexifier. Finally, the claim on monotonicity follows from the property of differentiable convex functions, e.g. [12]. \square

Illustration 3.2. Consider $f(s) = \cos s$.

- (i) An anti - derivative of f is

$$y(s) = \int_0^s \cos \xi d\xi = \sin s.$$

- (ii) (Augmentation:) Since f is continuous and Lipschitz with, e.g., $L = 1$ on any interval, one can specify $\alpha^* = -1$. The anti - derivative y can be decomposed $y(s) = C(\alpha, s) - q(\alpha, s)$, where $C(\alpha, s) = \sin s + q(\alpha, s)$ and $q(\alpha, s) = -\frac{1}{2}\alpha s^2$ are convex for every $\alpha \leq -1$. Moreover, $f(s) = M(\alpha, s) - \alpha s$ where $M(\alpha, s) = C'(\alpha, s) = \cos s - \alpha s$ is monotone increasing for every $\alpha \leq -1$.

Warning. The decomposition (2.2) is not trivial because it requires convexity of the two terms. It does not work for, e.g., y given in *Illustration 4.3* below.

The second part of the fundamental theorem of calculus says that every continuous f on $[a, b]$ satisfies

$$\int_a^b f(s) ds = y(b) - y(a) \tag{3.2}$$

where y is any anti - derivative of f , that is, $y' = f$. The result follows from the first part of *Theorem 3.1*, e.g., [15]. Using the fact that a continuous and Lipschitz f has a decomposable anti - derivative $y(s) = C(\alpha, s) - \frac{1}{2}\alpha s^2$ on compact intervals we have the following result.

Theorem 3.3. (The fundamental theorem of calculus: Second part)

- (i) Consider a scalar function $f : \mathbb{R} \rightarrow \mathbb{R}$ that is continuous on a compact interval $I = [a, b]$. The equation (3.2) holds for an arbitrary anti - derivative y of f .
- (ii) Assume that f is continuous and Lipschitz on an interval J with a Lipschitz constant L . Then for every compact subinterval $I = [a, b]$ in the interior of J there is a number α^* such that

$$\int_a^b f(s)ds = C(\alpha, b) - C(\alpha, a) + \frac{1}{2}\alpha(b^2 - a^2) \tag{3.3}$$

for every $\alpha \leq \alpha^*$, where $C(\alpha, s)$ is a convexification of an arbitrary anti - derivative $y(s)$ on I . In particular, one can specify $\alpha^* = -L$.

4. Applications of the augmentation

There are many applications of the augmentation. A straightforward one is to ordinary differential equations (ODE). If a solution y of an ODE on a compact interval is a smooth function with a Lipschitz derivative then y is convexifiable. Hence, using the transformation $y(s) = C(\alpha, s) - q(\alpha, s)$, every ODE with such a solution can be reduced to an ODE with a convex solution. We denote these transformed ODEs by C - ODE (standing for “convexified ODE”).

Illustration 4.1. The movement of a pendulum of length $l = g$ under small vibrations and no resisting forces is described in time t essentially by

$$\Theta''(t) = -\Theta(t), \quad \Theta'(0) = 0, \quad \Theta(0) = \Theta_0$$

e.g., [8, 14]. The solution $\Theta(t)$ is of the form $\Theta(t) = C(\alpha, t) + \frac{1}{2}\alpha t^2$, by *Theorem 3.1* (ii), and it is convex for every α sufficiently small. Using this substitution, the corresponding C - ODE is

$$C'''(\alpha, t) = -C(\alpha, t) - \frac{1}{2}\alpha t^2 - \alpha, \quad C'(\alpha, 0) = 0, C(\alpha, 0) = \Theta_0.$$

Its solution is convex and it can be found in the “region of convexification”, i.e., in a region with “sufficiently small” α . It is $C(\alpha, t) = \Theta_0 \cos t - \frac{1}{2}\alpha t^2$. After back - substitution, one finds the solution of the original problem: $\Theta(t) = \Theta_0 \cos t$. Graphs of the solutions $C(\alpha, t)$ of the above C - ODE are depicted in *Figure 3* for different values of α . For sufficiently negative $\alpha \leq -L$ the solutions are convex and for sufficiently positive $\alpha \geq L$ they are concave. The transformation from ODE to C - ODE does not require that α be numerically specified.

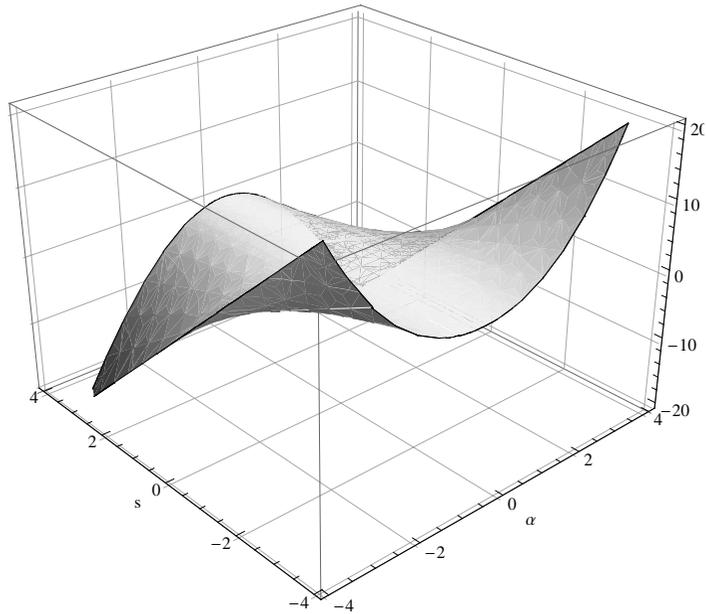


Figure 3. Solutions of C - ODE for the pendulum problem

The trajectories of objects that are described by Newton’s Second Law in theoretical mechanics are solutions of the second order ODE. These trajectories are smooth and have a Lipschitz derivative. Therefore they are convexifiable and hence one concludes that ODE of theoretical mechanics can be transformed to C - ODE.

Theorem 3.1 (ii) says that for a Lipschitz f :

$$f(s) = M(\alpha, s) - \alpha s \tag{4.1}$$

for every $\alpha \leq \alpha^*$, where $M(\alpha, s)$ is an increasing monotone function in s . In particular, one can specify $\alpha^* = -L$. Function $M(\alpha, s) = f(s) + \alpha s$ for $\alpha \leq \alpha^*$ is called a monotonization of f on I . Since differentiable functions are both continuous and Lipschitz, one can use (4.1) to “monotonize” ODE of any order. The general method of reducing an ODE to an ODE by monotonization is denoted by M - ODE.

Illustration 4.2. Using M - ODE and the substitution $\Theta(s) = M(\alpha, s) - \alpha s$, the pendulum problem from *Illustration 4.1* becomes

$$M''(\alpha, s) = -M(\alpha, s) + \alpha s, \quad M'(\alpha, 0) = \alpha, M(\alpha, 0) = \Theta_0.$$

Its monotone increasing solution is $M(\alpha, s) = \Theta_0 \cos s + \alpha s$. After back substitution, the solution to the original ODE is $\Theta(t) = \Theta_0 \cos t$.

There exist ODEs of the *first* order for which the solutions can not be decomposed as in (2.2). For such equations C - ODE does not work.

Illustration 4.3. The solution of $y'(s) = f(s)$, where

$$f(s) = \begin{cases} 2s \sin \frac{1}{s} - \cos \frac{1}{s}, & \text{if } s \neq 0, \\ 0, & \text{if } s = 0 \end{cases}$$

is

$$y(s) = \begin{cases} s^2 \sin \frac{1}{s}, & \text{if } s \neq 0, \\ 0, & \text{if } s = 0. \end{cases}$$

The function f is not continuous at $s = 0$, hence y is not smooth. In fact, y is *not* the difference of a convex function and a convex quadratic. Indeed, if y were decomposable as in (2.2), then we would have $y(s) = C(\alpha, s) + \frac{1}{2}\alpha s^2$ and hence $f(s) = C'(\alpha, s) + \alpha s$. Since the derivative of a strictly convex function is monotone increasing, f would have to be the sum of a monotone increasing function and a linear function, which is not true. The graphs of these particular f and y are depicted in [18].

Illustration 4.4. The initial - value problem $y'(s) = f(s)$, $y(0) = 0$, where $f(s)$ is given in *Illustration 4.3*, can *not* be solved by C - ODE. However, it can be solved by M - ODE.

Illustration 4.5. A continuous monotone function is differentiable almost everywhere on a compact interval I , by Lebesgue's theorem, e.g. [9, p. 321]. Hence, using *Theorem 3.1 (ii)*, we conclude that every continuous Lipschitz function is differentiable almost everywhere. This result, known as Rademacher's theorem, extends to functions in several variables.

The augmentations of the fundamental theorem of calculus in *Theorems 3.1* and *3.3* relate f and its anti - derivative y to convexity. Since convex functions are well studied, these augmentations provide wealth of information on f and y . One can simply pick any result on general convex functions from, e.g. [12, 16, 20, 22] etc. and apply it to C . For example, we know that, with a fixed α , a differentiable C is convex on I if, and only if,

$$C'(\alpha, s)(t - s) \leq C(\alpha, t) - C(\alpha, s) \quad \text{for every } s \text{ and } t \text{ in } I.$$

What does this property imply here? Since

$$C(\alpha, s) = y(s) - \frac{1}{2}\alpha s^2, \quad C(\alpha, t) = y(t) - \frac{1}{2}\alpha t^2 \quad \text{and} \quad C'(\alpha, s) = f(s) - \alpha s$$

we have the following situation:

$$\begin{aligned} y(t) - y(s) &= C(\alpha, t) - C(\alpha, s) + \frac{1}{2}\alpha(t^2 - s^2) \\ &\geq C'(\alpha, s)(t - s) + \frac{1}{2}\alpha(t^2 - s^2), \quad \text{by convexity} \\ &= (t - s)f(s) + \frac{1}{2}\alpha(t - s)^2, \quad \text{after substitution and rearrangement.} \end{aligned}$$

Specifying, e.g. $\alpha = -L$, we have

$$y(t) - y(s) \geq (t - s)f(s) - \frac{L}{2}(t - s)^2.$$

Similarly, using concavity of $y(s) + \frac{1}{2}Ls^2$, e.g., [19, 21], we find that $y(t) - y(s) \leq (t - s)f(s) + \frac{L}{2}(t - s)^2$. Thus the following bound is obtained:

Corollary 4.6. *Consider a scalar function $f : \mathbb{R} \rightarrow \mathbb{R}$ that is continuous and Lipschitz on $[a, b]$ with a constant L . Then an arbitrary anti - derivative y of f satisfies*

$$|y(t) - y(s) - (t - s)f(s)| \leq \frac{L}{2}(t - s)^2 \quad (4.2)$$

for every s and t on a compact subinterval I of (a, b) .

After interchanging s and t and a rearrangement, (4.2) yields stricter bounds:

Corollary 4.7. *Consider a scalar function $f : \mathbb{R} \rightarrow \mathbb{R}$ that is continuous and Lipschitz on $[a, b]$ with a constant L . Then every anti - derivative y of f satisfies*

$$\begin{aligned} -\frac{L}{2}(s - t)^2 + \max\{(t - s)f(s), (t - s)f(t)\} &\leq y(t) - y(s) \leq \frac{L}{2}(s - t)^2 \\ &+ \min\{(t - s)f(s), (t - s)f(t)\} \end{aligned}$$

for every s and t on a compact subinterval I of (a, b) .

Using Theorem 3.3 and Corollary 4.6 we obtain the following estimate of a definite integral.

Corollary 4.8. *Assume that f is continuous and Lipschitz on an interval J with a Lipschitz constant L . Then for each compact interval $I = [a, b]$ in the interior of J*

$$\left| \int_a^b f(s)ds - (b - a)f(a) \right| \leq \frac{L}{2}(b - a)^2. \quad (4.3)$$

The inequality (4.3) gives a relationship between the definite integral of f on $I = [a, b]$, the rectangle $(b - a)f(a)$, and a triangle $\frac{L}{2}(b - a)^2$ with a basis $(b - a)$ and the height $L(b - a)$. If f is a linear function, then the inequality is replaced by an equation. If f is smooth, then one can specify $L = \max_{s \in I} |f'(s)|$. In order to obtain "good" estimates, one should choose L (or another convexifier) relative to given I .

Illustration 4.9. Consider the function $y(s)$ from Illustration 4.3. This function has a complicated anti - derivative, e.g. [13, p. 68]. Here (4.3), regardless of $b > 0$, gives

$$\left| \int_0^b f(s)ds \right| \leq \frac{b^2}{2}.$$

Let us apply Artin's characterization of convex functions, e.g., [12], to a convexification $C(-L, s)$. Following the approach used in [22] we find a connection between a Lipschitz constant of f and its arbitrary anti - derivative y . The connection is given in terms of the absolute value of a determinant!

Corollary 4.10. *If y is an anti - derivative of a Lipschitz function f on I with a constant L , then*

$$\left| \det \begin{pmatrix} 1 & 1 & 1 \\ s & t & \xi \\ y(s) & y(t) & y(\xi) \end{pmatrix} \right| \leq \frac{L}{2}(s-t)(t-\xi)(\xi-s)$$

for every $s < t < \xi$ on I .

One can use *Corollary 4.10* in situation where $y(s)$ and $y(t)$ are known to estimate $y(\xi)$. The above also yields all kinds of inequalities. Take, e.g. $f(s) = \cos s$. Then $y(s) = \sin s$ and $L = 1$ regardless of I and the corollary yields various inequalities for the sine function.

5. Augmentation of the fundamental theorem of calculus for line integrals

In this section we will use *Theorem 2.1* to augment the fundamental theorem of calculus for functions in several variables. Only in this section, we will use the notation and terminology from [2]. In particular, the bold symbols, such as \mathbf{X} , represent “vectors” from the origin to the point $x = (x_i) \in \mathbb{R}^n$. They appear in [2] with an arrow sign; the symbol “ \bullet ” stands for the dot (inner) product. Using the vector notation, the value of $f : \mathbb{R}^n \rightarrow \mathbb{R}$ at \mathbf{X} is denoted by $f(\mathbf{X})$ instead of $f(x_1, \dots, x_n)$ or $f(x)$. Function f is then said to be a scalar - valued function of a vector variable or, briefly, a *scalar field*. Similarly, the “value” of $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ at \mathbf{X} is denoted $\mathbf{F}(\mathbf{X})$ and F is then called a *vector field* and denoted \mathbf{F} . We consider only fields that are continuous and bounded on open connected sets S . The results given below are meaningful in n - space. However, following [2], we are interested only in the case when $m = 2$ or $m = 3$.

Apostol in [2, Section 5.10] gives a method for constructing a potential function of a vector field that is different from the one given in, e.g. [1, Section 4.19]. He uses a step polygon to connect points in S . Since S is connected, each point \mathbf{X} in S can be reached from any fixed point \mathbf{A} in S by such a curve. He says: “When a potential is constructed by line integration the calculations are often simplified by connecting \mathbf{A} and \mathbf{X} by a step polygon” [2, p. 235]. We wish to integrate along a curve C from a fixed point \mathbf{A} in S to an arbitrary point \mathbf{X} . Denote by Φ the scalar field

$$\Phi(\mathbf{X}) = \int_{\mathbf{A}}^{\mathbf{X}} \mathbf{F} \bullet d\mathbf{r} \tag{5.1}$$

where $\mathbf{r} : \mathbb{R} \rightarrow \mathbb{R}^3$ describes C . For the definition of $\Phi(\mathbf{X})$ to be unambiguous, we need to know that the integral depends on \mathbf{X} and not on the particular path used to join \mathbf{A} to \mathbf{X} . Therefore it is natural to require the line integral to be independent of the path in S . (See [2, Section 5.8] for details.) Under these conditions, an extension of the first part of the fundamental theorem of calculus to line integrals takes the following form:

Theorem 5.1. (First part of the fundamental theorem for line integrals; [2, p. 235]) *Let \mathbf{F} be a vector field that is continuous on an open connected set S . Let \mathbf{A} be a fixed point of S . Assume that for every \mathbf{X} in S the line integral of \mathbf{F} from \mathbf{A} to \mathbf{X} has the same value for every simple step polygon connecting \mathbf{A} to \mathbf{X} . Define a scalar field Φ on S by the equation (5.1), where the line integral is taken along any simple step polygon in S joining \mathbf{A} to \mathbf{X} . Then Φ is differentiable on S and its gradient, $\nabla\Phi$, is equal to \mathbf{F} ; that is*

$$\nabla\Phi(\mathbf{X}) = \mathbf{F}(\mathbf{X}) \quad \text{for every } \mathbf{X} \text{ in } S. \tag{5.2}$$

In order to apply *Theorem 2.1* we need to place a compact convex set K in S :

Corollary 5.2. (Augmentation of the first part of the fundamental theorem for line integrals.) *In addition to the assumptions given in Theorem 5.1, consider a compact convex set K in S . Assume that \mathbf{F} has the Lipschitz property on K with a constant L . Take a fixed point \mathbf{A} in K and consider $\Phi(\mathbf{X})$ with \mathbf{X} in K . Then the following two statements hold:*

(i) *For every $\alpha \leq -L$, $\Phi(\mathbf{X})$ is of the form*

$$\Phi(\mathbf{X}) = C(\alpha, \mathbf{X}) - q(\alpha, \mathbf{X}) \tag{5.3}$$

where $C(\alpha, \mathbf{X})$ is a scalar convex field and $q(\alpha, \mathbf{X}) = -\frac{1}{2}\alpha\mathbf{X} \bullet \mathbf{X}$ is a scalar convex quadratic field on K .

(ii) *The vector field $\mathbf{F}(\mathbf{X})$ is the difference of a component - wise monotone increasing vector field $\mathbf{M}(\alpha, \mathbf{X}) = \nabla C(\alpha, \mathbf{X})$ and a linear field, i.e., $\mathbf{F}(\mathbf{X}) = \mathbf{M}(\alpha, \mathbf{X}) - (-\alpha\mathbf{X})$ for every $\alpha \leq -L$ on K .*

Proof. (i) The assumption on \mathbf{F} and the equation (5.2) imply that Φ is smooth and that its (Fréchet) derivative is Lipschitz. The decomposition (5.3) now follows from *Theorem 2.1* and the fact that $\alpha^* = -L$ is a convexifier of Φ on K , e.g. [21]. (ii) We have $\Phi(\mathbf{X}) = C(\alpha, \mathbf{X}) + \frac{1}{2}\alpha\mathbf{X} \bullet \mathbf{X}$, by (5.3). The monotonicity claim follows after differentiation using (5.2). □

We recall that a continuous function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is said to be component - wise monotone on a convex set K if it is monotone along every line segment that is parallel with each coordinate axis emanating from any $x \in K$, e.g. [24]. Note that $\mathbf{M}(\alpha, \mathbf{X}) = \nabla C(\alpha, \mathbf{X})$ is component - wise monotone increasing, being the derivative of a convex function. This follows from, e.g. [12].

Illustration 5.3. Consider $\mathbf{F}(\mathbf{X}) = (\cos x_1 \cos x_2)\mathbf{i} + (-\sin x_1 \sin x_2)\mathbf{j}$. Its potential function is $\Phi(\mathbf{X}) = \sin x_1 \cos x_2$. Clearly, $\mathbf{F} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ is continuous and Lipschitz on every compact convex set and any $L \geq 2$ can be chosen to be a Lipschitz constant. Therefore $\Phi(\mathbf{X})$ can be decomposed as $\Phi(\mathbf{X}) = C(\alpha, \mathbf{X}) - q(\alpha, \mathbf{X})$ where $C(\alpha, \mathbf{X}) = \sin x_1 \cos x_2 - \frac{1}{2}\alpha(x_1^2 + x_2^2)$ and $q(\alpha, \mathbf{X}) = -\frac{1}{2}\alpha(x_1^2 + x_2^2)$ are convex functions for every $\alpha \leq -L$. The graphs of $\Phi(\mathbf{X})$ and $C(\alpha, \mathbf{X})$ were depicted in *Figure 1*.

A monotonization of $\mathbf{F}(\mathbf{X})$ is $\mathbf{M}(\alpha, \mathbf{X}) = (\cos x_1 \cos x_2 - \alpha x_1)\mathbf{i} + (-\sin x_1 \sin x_2 - \alpha x_2)\mathbf{j}$ for any $\alpha \leq -L$. Graphs of the first components F_1 of $\mathbf{F}(\mathbf{X})$ and of its monotonization $M_1(\alpha, \mathbf{X}) = \cos x_1 \cos x_2 - \alpha x_1$ are depicted in *Figure 4*. $M_1(\cdot, x_2)$

is monotone increasing along x_1 for every fixed x_2 . Similarly the second component $M2(\alpha, \mathbf{X}) = -\sin x_1 \sin x_2 - \alpha x_2$ is monotone increasing along x_2 for every fixed x_1 , for every α sufficiently small.

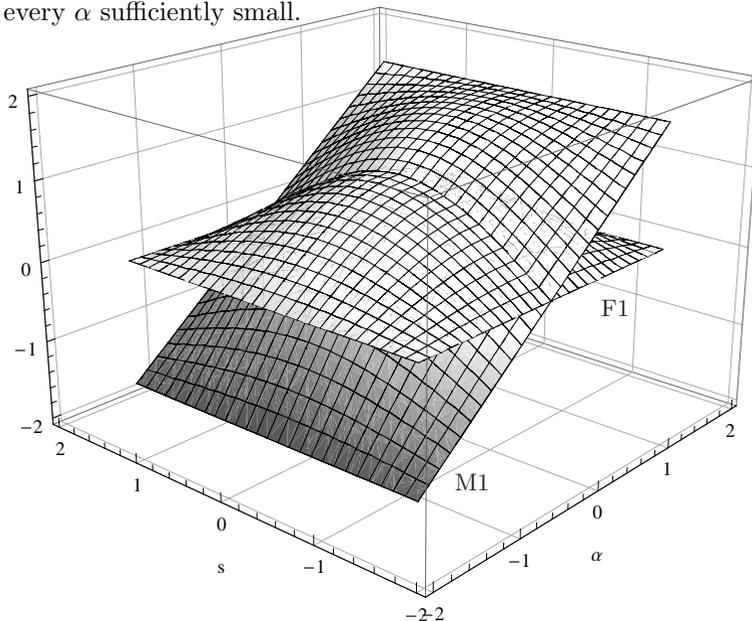


Figure 4. First components of $\mathbf{F}(\mathbf{X})$ and its monotonization

Let us now look at the second part of the fundamental theorem of calculus for line integrals from, e.g. [2] and its augmentation.

Theorem 5.4. (Second part of the fundamental theorem for line integrals; [2, p.230].) *Let f be a scalar field that is continuously differentiable on an open set S , and let ∇f denote the gradient of f . Let \mathbf{A} and \mathbf{X} be two points of S that can be connected by a piecewise smooth curve lying in S . If C is such a curve, described by a vector-valued function \mathbf{r} defined on an interval $[a, b]$, where $\mathbf{r}(a) = \mathbf{A}$ and $\mathbf{r}(b) = \mathbf{X}$, then we have*

$$\int_C \nabla f \bullet d\mathbf{r} = f(\mathbf{X}) - f(\mathbf{A}).$$

Corollary 5.5. (Augmentation of the second part of the fundamental theorem for line integrals.) *In addition to the assumptions given in Theorem 5.4, let us also assume that the derivative of the continuously differentiable f has the Lipschitz property on a compact convex subset K of S with a constant \tilde{L} . Let \mathbf{A} and \mathbf{X} be two points of K that can be connected by a piecewise smooth curve C lying in K . Then*

$$\int_C \nabla f \bullet d\mathbf{r} = C(\alpha, \mathbf{X}) - C(\alpha, \mathbf{A}) + \frac{1}{2}\alpha(\mathbf{X} \bullet \mathbf{X} - \mathbf{A} \bullet \mathbf{A}) \tag{5.4}$$

for every $\alpha \leq -\tilde{L}$, where $C(\alpha, \mathbf{X})$ is a scalar convex field on K .

Proof. We know that the line integral of ∇f along a path joining two points \mathbf{A} and \mathbf{B} in K may be expressed in terms of the values $f(\mathbf{A})$ and $f(\mathbf{B})$. Alternatively, using a decomposition of f , one can express it as in (5.4). \square

The above yields estimates that are analogues to those described in Section 4. A sample result follows.

Corollary 5.6. *Under the assumptions of Corollary 5.5, for every $\alpha \leq -\tilde{L}$*

$$\int_C \nabla f \bullet dr \geq \nabla f(\mathbf{A}) \bullet (\mathbf{X} - \mathbf{A}) + \frac{1}{2}\alpha(\mathbf{X} - \mathbf{A}) \bullet (\mathbf{X} - \mathbf{A}).$$

Proof. We have $C(\alpha, \mathbf{X}) - C(\alpha, \mathbf{A}) \geq \nabla C(\alpha, \mathbf{A}) \bullet (\mathbf{X} - \mathbf{A})$ by convexity of the field C . Also $\nabla C(\alpha, \mathbf{A}) = \nabla f(\mathbf{A}) - \alpha \mathbf{A}$ by the construction of C . We use these in (5.4). \square

A two - dimensional analog of the second part of the fundamental theorem of calculus that expresses a double integral over a region R as a line integral taken along a closed curve forming the boundary of R , is referred to as Green's theorem. The integration symbol \oint below indicates that the curve is to be traversed in the counterclockwise direction. We recall a version of Green's theorem and then give its augmentation.

Theorem 5.7. (Green's theorem for plane regions; [2, p. 244].) *Let P and Q be scalar fields that are continuously differentiable on an open set S in the xy - plane. Let C be a piecewise smooth Jordan curve, and let R denote the union of C and its interior. Assume R lies in the set S . Then we have the identity*

$$\int \int_R \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dx dy = \oint P dx + Q dy,$$

where the line integral is taken around C in the counterclockwise direction.

If the union R is a compact convex set and if the derivatives of P and Q have the Lipschitz property on R , then P and Q are convexifiable, i.e., $C_P(\alpha_P, x, y) = P(x, y) - \frac{1}{2}\alpha_P(x^2 + y^2)$ and $C_Q(\alpha_Q, x, y) = Q(x, y) - \frac{1}{2}\alpha_Q(x^2 + y^2)$ are convex on R with some convexifiers α_P and α_Q . After substituting C_Q for Q and C_P for P in Green's identity we have the following.

Corollary 5.8. (Convexification of Green's theorem for plane regions.)

Let P and Q be scalar fields that are continuously differentiable on an open set S in the xy - plane. Let C be a piecewise smooth Jordan curve, and let R denote the union of C and its interior. Assume R is a compact convex set that lies in the set S . Also assume that the Fréchet derivatives of P and Q have the Lipschitz property on R . If $C_P(\alpha_P, x, y)$ and $C_Q(\alpha_Q, x, y)$ are convexifications of P and Q on R then

$$\int \int_R \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dx dy + \int \int_R (\alpha_P y - \alpha_Q x) dx dy = \oint C_P(\alpha_P, x, y) dx + C_Q(\alpha_Q, x, y) dy$$

where the line integral is taken around C in the counterclockwise direction.

The essential difference between *Theorem 5.7* and *Corollary 5.8* is that the latter assumes convexity of R and the line integration is traversed only for *convex* functions

C_P and C_Q . This may be easier to implement than line integration for non - convex functions.

Illustration 5.9. (Adopted from [1, p. 369].) We wish to evaluate

$$I = \oint [x - y^3 + 3a(x^2 + y^2)]dx + (x^3 + y^3)dy$$

where C is the positively oriented boundary of the quarter disc $R: 0 \leq x^2 + y^2 \leq a^2, x \geq 0, y \geq 0$. Take $P(x, y) = x - y^3$ and $Q(x, y) = x^3 + y^3$. We find, from the Hessian of P , (recall the method from [19]) that a convexifier of P on R is $\alpha_P = -6a (a > 0)$. Hence $C_P(\alpha_P, x, y) = x - y^3 + 3a(x^2 + y^2)$. Since Q is convex on R , one can set $\alpha_Q = 0$ and $C_Q(\alpha_Q, x, y) = x^3 + y^3$. Hence, by *Corollary 5.8*:

$$\begin{aligned} I &= \int \int_R \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dx dy + \int \int_R (\alpha_P y - \alpha_Q x) dx dy \\ &= 3 \int \int_R (x^2 + y^2) dx dy - 6a \int \int_R y dx dy \\ &= \frac{3}{8} \pi a^4 + 2a^4 = 3.178a^4. \end{aligned}$$

One may apply convexifications and monotonicizations to double and triple integrals of certain “derivatives” of vector fields over two and three dimensional structures as given in, e.g. [1, 2]. This could be useful in modeling problems in, e.g. fluid dynamics and electrostatics.

6. Going beyond convexity in linear algebra

Theorem 2.1 can be used in two directions. On one hand, problems with smooth solutions y having Lipschitz derivatives can be “convexified” after a substitution for y , recall (2.2). These transformed problems have convex solutions in $C(\alpha, s)$. On the other hand, one can extend properties of general convex functions f to smooth, generally non - convex, functions. These are obtained after properties of convex functions are applied to the convexifications C . An illustration of this approach was given in [20] for Jensen’s inequality. In this section we will follow the main ideas from [24] where the details can be found; however some results will be sharpened. Our aim is to extend an important inequality for convex functions of symmetric matrices to non - convex functions. To this end we will work with a product function and use its convexification.

We begin with an arbitrary real symmetric matrix A . Given a fixed vector consisting of its eigenvalues $\lambda = (\lambda_i)$ in \mathbb{R}^n , denote $K_\lambda = \{x \in \mathbb{R}^n : x = M\lambda, M \in \Omega_n\}$, where Ω_n is the set of doubly stochastic $n \times n$ matrices, i.e., matrices with nonnegative entries in which the sum of elements in each row and in each column is 1, $\langle u, v \rangle = u^T v$ is the Euclidean inner product of u and v from \mathbb{R}^n .

Theorem 6.1. (Basic inequality for convex functions of symmetric matrices, [12, Theorem C, p. 202].) *Let A be an $n \times n, n \geq 2$, real symmetric matrix and let $\lambda = (\lambda_1, \dots, \lambda_n)^T$ be a vector consisting of the eigenvalues of*

A arranged in some order. Let $f : U \rightarrow R$ be a convex function on a set U containing the set K_λ . Then for any orthonormal set $\{v^1, \dots, v^n\}$ in \mathbb{R}^n there exists a permutation matrix P_λ such that $f(\langle v^1, Av^1 \rangle, \dots, \langle v^n, Av^n \rangle) \leq f(P_\lambda \lambda)$.

After applying the above theorem to a convexification $C(x, \alpha) = f(x) - \frac{1}{2}\alpha x^T x$ of f , the following non - convex analog of *Theorem 6.1* is obtained. Here $\|A\|$ denotes the Frobenius (Euclidean) norm of A .

Theorem 6.2. (Basic inequality for convexifiable functions of symmetric matrices, [24].) *Let A be an $n \times n, n \leq 2$, real symmetric matrix and let $\lambda = (\lambda_1, \dots, \lambda_2)^T$ be a vector consisting of the eigenvalues of A arranged in some order. Let $f : U \rightarrow R$ be a convexifiable function on a convex set U containing the set K_λ with a convexifier α and a concavifier β . Then for any orthonormal set $\{v^1, \dots, v^n\}$ in \mathbb{R}^n , there exists a permutation matrix P_λ such that*

$$f(\langle v^1, Av^1 \rangle, \dots, \langle v^n, Av^n \rangle) \leq f(P_\lambda \lambda) + \frac{\alpha}{2} \left\{ \sum_{i=1}^n \langle v^i, Av^i \rangle^2 - \|A\|^2 \right\} \quad \text{and}$$

$$f(\langle v^1, Av^1 \rangle, \dots, \langle v^n, Av^n \rangle) \geq f(P_\lambda \lambda) + \frac{\beta}{2} \left\{ \sum_{i=1}^n \langle v^i, Av^i \rangle^2 - \|A\|^2 \right\}.$$

After specifying f to be the product function $f(x) = x_1 x_2 \cdots x_n$, recalling that $f(P_\lambda \lambda) = \det A$, and estimating its convexifier α and concavifier β , we have :

Corollary 6.3. (Bounds on Determinants of Symmetric Matrices, [24]) *Let $A = (a_{ij})$ be an $n \times n$ real symmetric matrix and let ρ be its spectral radius. Then*

$$\prod_{i=1}^n a_{ii} - (n - 1)\rho^{n-2} \sum_{i,j=1; i < j}^n a_{ij}^2 \leq \det A \leq \prod_{i=1}^n a_{ii} + (n - 1)\rho^{n-2} \sum_{i,j=1; i < j}^n a_{ij}^2.$$

This result is interesting for several reasons: (i) If the left - hand side is positive, then we have a sufficient condition for non - singularity of $A!$ (In order to simplify the condition one can replace ρ by $\|A\|$ in the inequality.) (ii) If A is a real arbitrary square matrix, then we recall Hadamard’s inequality:

$$(\det A)^2 \leq \prod_{j=1}^n \left(\sum_{i=1}^n a_{ij}^2 \right).$$

This is considered by many to be “the most famous determinantal inequality” (e.g. Bellman’s book: Introduction to Matrix Analysis, McGraw Hill, Second Edition, 1970, p. 130). What is a left - side bound of the inequality? *Corollary 6.3* gives the answer:

$$\prod_{j=1}^n \left(\sum_{i=1}^n a_{ij}^2 \right) - (n - 1)[\rho(A^T A)]^{n-2} \sum_{i,j=1; i < j}^n \left(\sum_{k=1}^n (a_{ki} a_{kj})^2 \right) \leq (\det A)^2.$$

(iii) *Corollary 6.3* also gives a lower bound on the spectral radius ρ of a symmetric matrix:

$$|\det A - \prod_{i=1}^n a_{ii}| \leq (n-1)\rho^{n-2} \sum_{i,j=1; i < j}^n a_{ij}^2.$$

Illustration 6.4. Consider

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

Using (i) we find that A is non-singular if $0 < 4\varepsilon(1 + \varepsilon) < 1$. Using (iii) we find that the spectral radius of A satisfies $\rho \geq \frac{1}{2}$ regardless of ε .

It may be interesting to find non-convex analogs of other properties of convex functions, e.g. those from [3, 16] and numerous other texts on convexity. The idea is to apply these properties to convexifications of non-convex functions. The most difficult step in the process appears to be an estimation of convexifiers.

Let us emphasize what we have *not* done in this paper. We have *not* attempted to generalize convexity in any way. We have *not* mentioned any kind of “generalized convexity”, such as pseudo convexity or quasi convexity, Karlin’s method of generalized convexity or his total positivity, “convexifications” based on optimality conditions, etc. For the sake of simplicity we have used only the classic notion of a convex function and we have worked in a finite-dimensional setting.

References

- [1] R. A. ADAMS, *Calculus of Several Variables*, Addison-Wesley, 1987.
- [2] T. M. APOSTOL, *Calculus*, vol. II: *Calculus of Several Variables with Applications to Probability and Vector Analysis*, Blaisdell, New-York, 1965.
- [3] A. BEN-TAL, A. NEMIROVSKI, *Lectures on Modern Convex Optimization*, MPS-SIAM Series in Optimization, 2001.
- [4] W. CHENEY, D. KINCAID, *Numerical Mathematics and Computing*, Brooks/Cole, Monterey, California, 1980.
- [5] D. M. DAVIDESCU, *Convexification of Mathematical Programming Models*, M. Sc. thesis, McGill University, 2005.
- [6] C. A. FLOUDAS, *Deterministic Global Optimization*, Kluwer, Dordrecht, 2000.
- [7] C. A. FLOUDAS, *Advances in deterministic global optimization*, Plenary talk, International Conference on Computational Management, Gainesville, Florida, April 2005.

- [8] D. C. GIANCOLI, *Physics: Principles with Applications*, Prentice - Hall, New Jersey, 1985.
- [9] A. N. KOLMOGOROV, S. V. FOMIN, *Introductory Real Analysis*, Prentice - Hall, New Jersey, 1970. (Revised English edition translated and edited by R. A. Silverman.)
- [10] W. B. LIU, C. A. FLOUDAS, *A remark on the GOP algorithm for global optimization*, J. Global Optim. **3**(1993), 519-521.
- [11] P. M. PARDALOS, *Enumerative techniques in global optimization*, OR Spectrum **10**(1988), 29-35.
- [12] A. W. ROBERTS, D. E. VARBERG, *Convex Functions*, Academic Press, New York, 1973.
- [13] M. L. SMOLIANSKII, *Table of Indefinite Integrals*, Nauka, Moscow, 1963. (In Russian)
- [14] M. R. SPIEGEL, *Theory and Problems of Theoretical Mechanics*, Schaum Publ. Co., New York, 1967.
- [15] J. STEWART, *Single Variable Calculus*, Brooks/Cole, Pacific Grove, California, 1987.
- [16] VAN J. TIEL, *Convex Analysis*, Wiley, Chichester, 1984.
- [17] J. P. VIAL, *Strong and weak convexity of sets and functions*, Math. Oper. Res. **8**(1983), 231-259.
- [18] S. ZLOBEC, *Stable Parametric Programming*, Research monograph, Kluwer Academic, Series: Applied Optimization v. 56, Dodrecht, 2002.
- [19] S. ZLOBEC, *Estimating convexifiers in continuous optimization*, Math. Commun. **8**(2003), 129-137.
- [20] S. ZLOBEC, *Jensen's inequality for nonconvex functions*, Math. Commun. **9**(2004), 119-124.
- [21] S. ZLOBEC, *On the Liu-Floudas convexification of Lipschitz functions*, J. Global Optim. **32**(2005), 401-407.
- [22] S. ZLOBEC, *Convexifiable functions in integral calculus*, Glasnik Matematički **40**(60) (2005), 241-247; production erratum Ibid **41**(61) (2006), 187-188.
- [23] S. ZLOBEC, *Characterization of convexifiable functions*, Optimization **55** (2006), 251-262.
- [24] S. ZLOBEC, *On two simple decompositions of Lipschitz functions*, Optimization **57**(2008), 249-261.
- [25] S. ZLOBEC, *Characterization of a convexifiable function*, Encyclopaedia of Optimization: Second edition, Springer, 2009, 551-555.