SAPIENZA UNIVERSITÀ DI ROMA
Dipartimento di Matematica, Dipartimento di Metodi e Modelli Matematici per le Scienze Applicate
ISTITUTO NAZIONALE DI ALTA MATEMATICA "FRANCESCO SEVERI"

RENDICONTI DI MATEMATICA

E DELLE SUE APPLICAZIONI

Serie VII - Volume 30 - Fascicolo III-IV - 2010

Le modalità per la presentazione dei lavori e le condizioni di vendita si trovano alla web page.

Submission of papers and subscription rate information could be found on the web page.

ADDRESS:
RENDICONTI DI MATEMATICA
DIPARTIMENTO DI MATEMATICA - PIAZZALE A. MORO, 2 - 00185 ROMA, ITALIA

Fax.: +39 06 49913052
e-mail: rendmat@mat.uniroma1.it
web page: www.dmmm.uniroma1.it/˜rendiconti

Direttore Responsabile: Prof. ALESSANDRO SILVA

Autorizzazione del Tribunale di Roma del 22-11-2005 (n. 456/05)

Composizione tipografica effettuata con il programma TEX © dalla
CompoMat – Loc. Braccone – 02040 Configni (RI) – Telefax 39 0746 672240

Stampato a cura del *Centro Stampa d'Ateneo* della Sapienza Università di Roma

Rendiconti di Matematica, Serie VII
Volume 30, Roma (2010), 239–247

g-Natural metrics on unit tangent sphere bundles via a Musso-Tricerri process

MOHAMED TAHAR KADAOUI ABBASSI

Dedicated to the memory of Professor F. Tricerri

Abstract: *E. Musso and F. Tricerri had given a process of construction of Riemannian metrics on tangent bundles and unit tangent bundles, over m-dimensional Riemannian manifolds (M, g), from some special quadratic forms an $OM \times \mathbb{R}^m$ and OM, respectively, where OM is the bundle of orthonormal frames [7]. We prove in this note that every Riemannian g-natural metric on the unit tangent sphere bundle over a Riemannian manifold can be constructed by the Musso-Tricerri's process. As a corollary, we show that every Riemannian g-natural metric on the unit tangent bundle, over a two-point homogeneous space, is homogeneous.*

Let (M, g) be a Riemannian manifold and TM its tangent bundle. Considering TM as a vector bundle associated with the bundle of orthonormal frames OM, E. Musso and F. Tricerri have constructed an interesting class of *Riemannian* metrics on TM [7]. This construction is not a classification *per se*, but it is a construction process of *Riemannian* metrics on TM from symmetric, positive semi-definite tensor fields Q of type (2, 0) and rank $2m$ on $OM \times \mathbb{R}^m$, which are basic for the natural submersion $\Phi : OM \times \mathbb{R}^m \to TM$, $\Phi(v\varepsilon) = (x, \sum_i \varepsilon^e v_i)$, for $v = (x; v_1, \ldots, v_m) \in OM$ and $\varepsilon = (\varepsilon^1, \ldots, \varepsilon^m) \in \mathbb{R}^m$. Recall that Q is *basic*

Key Words and Phrases: *Unit tangent (sphere) bundle – Einstein manifold – g-natural metric*
A.M.S. Classification: 53C07, 53C25.

means that Q is $O(m)$-invariant and $Q(X,Y)=0$, if X is tangent to a fiber of Φ. The construction can be presented as follows:

Proposition 1 ([7]). *Let Q be a symmetric, positive semi-definite tensor field of type $(2,0)$ and rank $2m$ on $OM\times \mathbb{R}^m$, which is basic for the natural submersion $\Phi: OM\times\mathbb{R}^m \to TM$. Then there is a unique Riemannian metric G^Q on TM such that $\Phi^*(G^Q)=Q$. It is given by*

$$G^Q_{(x,u)}(X,Y) = Q_{(v,\varepsilon)}(X',Y'), \tag{1}$$

where (v,ε) belongs to the fiber $\Phi^{-1}(x,u), X, Y$ are elements of $(TM)_(x,u)$ and X',Y' are any tangent vectors to $OM\times\mathbb{R}^m$ at (v,ε) such that $\Phi_*(X')=X$ and $\Phi_*(Y')=Y$.

On the other hand, Musso and Tricerri proposed a similar process for constructing Riemannian metrics on the unit tangent sphere bundle T_1M from symmetric, positive semi-definite tensor fields $\tilde{Q}$ of type $(2,0)$ and rank $2m-1$ on OM, which are basic for the natural submersion $\psi_m : OM \to T_1M$, $\psi_m(v) = (x, v_m)$, for $v=(x;v_1,\dots,v_m)\in OM$. Recall that $\tilde{Q}$ is *basic* means that $\tilde{Q}$ is $O(m-1)$-invariant and $\tilde{Q}(X,Y)=0$, if X is tangent to a fiber of ψ_m. Note that ψ_m is a submersion whose fibers can be identified with the subgroup $O(m-1)$ of $O(m)$ given by the matrices of the form $\begin{pmatrix} A & 0 \\ 0 & 1\end{pmatrix}$, $A\in O(m-1)$. Then T_1M can be regarded as the quotient space $OM/O(m-1)$, and ψ_m is the natural projection. The construction can be stated as follows:

Proposition 2 ([7]). *Let $\tilde{Q}$ be a symmetric, positive semi-definite tensor field of type $(2,0)$ and rank $2m$ on OM, which is basic for the natural submersion $\psi_m : OM \to T_1M$. Then there is a unique Riemannian metric $\tilde{G}^{\tilde{Q}}$ on T_1M such that $\psi_m^*(\tilde{G}^{\tilde{Q}})=\tilde{Q}$. It is given by*

$$\tilde{G}^{\tilde{Q}}_{(x,u)}(X,Y) = \tilde{Q}_{(v)}(X',Y'), \tag{2}$$

where v belongs to the fiber $\psi_m^{-1}(x,u), X, Y$ are elements of $(T_1M)_(x,u)$ and X',Y' are any tangent vectors to OM at v such that $(\psi_m)_*(X')=X$ and $(\psi_m)_*(Y')=Y$.

The Musso-Tricerri processes described by Propositions 1 and 2, respectively, are compatible in the following sense:

Proposition 3. *If a Riemannian metric G on TM is induced from a bilinear form Q on $OM\times\mathbb{R}^m$ by the Musso-Tricerri process described in Proposition 1, i.e., $\Phi^*(G)=Q$, then the induced metric $\tilde{G}:=i^*(G)$ on T_1M, where i: $T_1M\to TM$ is the canonical injection, can be obtained from the bilinear form $\tilde{Q}:=i^*(Q)$ on OM by the Musso-Tricerri process described in Proposition 2.*

PROOF. Denote by i_m the map $OM \to OM \times \mathbb{R}^m, v \mapsto (v, 0, \dots, 0, 1)$. Then the following diagram

$$\begin{array}{ccc} OM & \overset{i_m}{\to} & OM \times \mathbb{R}^m \\ \psi_m \downarrow & & \downarrow \Phi \\ T_1M & \overset{i}{\hookrightarrow} & TM \end{array} \tag{3}$$

commutes. If we consider $\tilde{Q} := i_m^* Q$, then $\tilde{Q}$ is a symmetric, semi-positive definite, tensor field of type (0, 2) on OM. We can prove by a bit longer routine computation that is basic for ψ_m and it is of rank $2m-1$. Furthermore, we have, by virtue of (3), that $\psi_m^*(\tilde{G}) = \psi_m^*(i^*(G)) = (i \circ \psi_m)^*(G) = i_m^*(\Phi^*(G)) = i_m^*(Q) = \tilde{Q}$. ∎

Now we shall prove that every Riemannian g-natural metric on the unit tangent sphere bundle T_1M of a Riemannian manifold (M, g) can be constructed by the Musso-Tricerri's scheme, given by Proposition 2. For this, let us recall some basic definition.

Let ∇ the Levi-Civita connection of g. Then the tangent space of TM at any point $(x, u) \in TM$ split into the horizontal and vertical subspaces with respect to ∇:

$$(TM)_{(x,u)} = H_{(x,u)} \oplus V_{(x,u)}.$$

If $(x, u) \in TM$ is given then, for any vector $X \in M_x$, there exists a unique vector $X^h \in H_{(x,u)}$ such that $p_* X^h = X$, $p : TM \to M$ is the natural projection. We call X^h the *horizontal lift* of X to the point $(x, u) \in TM$. The *vertical lift* of a vector $X \in M_x$ to $(x, u) \in TM$ is a vector $X^v \in V_{(x,u)}$ such that $X^v(df) = Mf$, for all functions f on M. Here we consider 1-forms df on M as functions on TM (*i.e.*, $(df)(x, u) = uf$). Note that the map $X \to X^h$ is an isomorphism between the vector spaces M_x and $H_{(x,u)}$. Similarly, the map $X \to X^v$ is an isomorphism between the vector spaces M_x and $V_{(x,v)}$. Obviously, each tangent vector $\widetilde{Z} \in (TM)_{(x,u)}$ can be written in the form $\widetilde{Z} = X^h + Y^v$, where X, $Y \in M_x$ are uniquely determined vectors.

In an obvious way we can define horizontal and vertical lifts of vector fields on M.

If we fix an *F-metric* ξ on M, *i.e.*, a mapping $TM \oplus TM \oplus TM \to \mathbb{R}$ which is linear in the second and the third argument and smooth in the first argument, then there are three distinguished constructions of metrics on the tangent bundle TM, which are given as follows [5]:

(a) If we suppose that ξ is symmetric with respect to the last two arguments, then the *Sasaki lift* ξ^s of ξ is defined as follows:

$$\begin{cases} \xi^s_{(x,u)}(X^h, Y^h) = \xi(u; X, Y), \\ \xi^s_{(x,u)}(X^v, Y^h) = 0, \end{cases} \qquad \begin{cases} \xi^s_{(x,u)}(X^h, Y^v) = 0, \\ \xi^s_{(x,u)}(X^v, Y^v) = \xi(u; X, Y), \end{cases}$$

for all X, $Y \in M_x$. If ξ is non degenerate and positive definite with respect to the last two arguments for each fixed u, then ξ^s is a Riemannian metric on TM.

(b) The *horizontal lift* ξ^h of ξ is a pseudo-Riemannian metric on TM which is given by:

$$\begin{cases} \xi^h_{(x,u)}(X^h, Y^h) = 0, \\ \xi^h_{(x,u)}(X^v, Y^h) = \xi(u; X, Y), \end{cases} \quad \begin{cases} \xi^h_{(x,u)}(X^h, Y^v) = \xi(u; X, Y), \\ \xi^h_{(x,u)}(X^v, Y^v) = 0, \end{cases}$$

for all X, $Y \in M_x$. If ξ is positive definite with respect to the last two arguments, then ξ^s is of signature (m, m).

(c) The *vertical lift* ξ^v of ξ is a degenerate metric on TM which is given by:

$$\begin{cases} \xi^v_{(x,u)}(X^h, Y^h) = \xi(u; X, Y), \\ \xi^v_{(x,u)}(X^v, Y^h) = 0, \end{cases} \quad \begin{cases} \xi^v_{(x,u)}(X^h, Y^v) = 0, \\ \xi^v_{(x,u)}(X^v, Y^v) = 0, \end{cases}$$

for all X, $Y \in M_x$. For each fixed u, the rank of ξ^v is exactly that of ξ.

If $\xi = g$ is a Riemannian metric on M, then the three lifts of ξ just constructed coincide with the three well-known classical lifts of the metric g to TM.

Let (M, g) be non-oriented. Then it is known that all *natural* F-metrics are of the form

$$F(u; X, Y) = \alpha(\|u\|^2)g(X, Y) + \beta(\|u\|^2)g(X, u)g(Y, u),$$

where $\alpha(t)$, $\beta(t)$ are smooth functions on $[0, +\infty)$ and $\|u\| = \sqrt{g(u, u)}$ (see [4] and [2]). The three lifts above of *natural* F-metrics generate the class of g-natural metrics on TM (*cf.* [5] and [2] for the classification and the definition of g-natural metrics and [4] for the general definition of naturality).

More precisely, we have

Proposition 4. *Let (M, g) be a Riemannian manifold. Every g-natural metric G on TM is given by*

$$G = (\alpha_1 g + \beta_1 k)^s + (\alpha_2 g + \beta_2 k)^h + (\alpha_3 g + \beta_3 k)^v, \tag{4}$$

where α_i, β_i, $i = 1, 2, 3$, are smooth functions on $[0, +\infty)$, and k is the natural F-metric on M defined by

$$k(u; X, Y) = g(u, X)g(u, Y), \quad \textit{for all} \quad (u, x, Y) \in TM \oplus TM \oplus TM. \tag{5}$$

If we restrict an arbitrary g-natural metric (4) to a tangent sphere bundle $T_rM(r>0)$, then we obtain the metric $\widetilde{G}$ of the form

$$\widetilde{G} = a \cdot \widetilde{g^d} + b \cdot \widetilde{g^h} + c \cdot \widetilde{g^v} + d \cdot \widetilde{k^v}, \tag{6}$$

where $a = \alpha_1(r^2)$, $b = \alpha_2(r^2)$, $c = \alpha_3(r^2)$, $d = \beta_3(r^2)$ and $\widetilde{g^s}$, $\widetilde{g^h}$, $\widetilde{g^v}$ and $\widetilde{k^v}$ are the metrics on T_rM induced by g^s, g^h, g^v and k^v, respectively. We call such metrics on T_rM, induced by g-natural metrics, *g-natural metrics on T_rM*.

Riemannian g-natural metrics on tangent sphere bundles are characterized by

Proposition 5 ([1]). *Let $r>0$ and (M,g) be a Riemannian manifold. Then every Riemannian g-natural metric $\widetilde{G}$ on T_rm induced form a (possibly degenerate) g-natural G on TM, is of the form* (6), *where a, b, c and d are constants satisfying the inequalities $a>0$, $a(a+c)-b^2>0$ and $a+c+dr^2>0$.*

Let $\theta = (\theta^1, \dots, \theta^m)$ denote the canonical 1-form on OM, and let π denote the natural projection $OM \xrightarrow{\pi} M$. Then

$$d\pi_v(X) = \sum_i \theta^i(X) v_i, \qquad v = (x; v_1, \dots, v_m).$$

If we denote with $\omega = (\omega^i_j)$ the connection form on OM, then we find that the forms

$$\pi_1^*\theta^i,\ i=1,\dots,m;\ \pi_1^*\omega^i_j,\ 1 \le i \le j \le m;\ d\varepsilon^i,\ i=1,\dots,m,$$

where $\pi_1 : OM \times \mathbb{R}^m \to OM$ denotes the first natural projection, determine an absolute parallelism on $OM \times \mathbb{R}^m$. We consider the 1-forms $\nabla\varepsilon^i$ on $OM \times \mathbb{R}^m$ defined by

$$\nabla\varepsilon^i = d\varepsilon^i + \sum_j \varepsilon^j \pi_1^* \omega^i_j. \tag{7}$$

The first author an M. Sarih have proved the following

Proposition 6 ([2]). *Every g-natural metric on the tangent bundle TM of a Riemannian manifold (M,g) can be constructed by the Musso-Tricerri's generalized scheme, given by Proposition* 1.

More precisely, and arbitrary g-natural metric G on TM, which is of the form (4) *by Proposition* 4, *is induced by the symmetric tensor field Q of type* $(2,0)$ *on* $OM \times \mathbb{R}^m$ *given by*

$$Q = (\alpha_1+\alpha_3)(r^2)\sum_i(\pi_1^*\theta^i)^2 + (\beta_1+\beta_3)(r^2)\left(\sum_i \varepsilon^i\pi_1^*\theta^i\right)^2 +$$
$$(8) \qquad + \alpha_1(r^2)\sum_i(D\varepsilon^i)^2 + \beta_1(r^2)\left(\sum_i \varepsilon^i D\varepsilon^i\right)^2 +$$
$$+ 2\alpha_2(r^2)\sum_i \pi_i^*\theta^i D\varepsilon^i + 2\beta_2(r^2)\left(\sum_i \varepsilon^i\pi_1^*\theta^i\right)\left(\sum_i \varepsilon^i D\varepsilon^i\right),$$

where $r^2 = \sum_i(\varepsilon^i)^2$.

Note that (8) is exactly the expression (3.4) of [2] with the abuse of notation $\theta = \pi_1^*\theta$ (*cf.* [2, p. 8, line 7 from below]).

Let us mention that in the proof of this result in [2], there occurred a little misprint which did not influence the correctness of the statement.

Combining this last proposition with Proposition 3, we obtain

Theorem 1. *Every Riemannian g-natural metric on the unit tangent sphere bundle T_1M of a Riemannian manifold (M,g) can be constructed by the Musso-Tricerri's scheme, given by Proposition* 2.

More precisely, if $\widetilde{G} = a\cdot\widetilde{g^s} + b\cdot\widetilde{g^h} + c\cdot\widetilde{g^v} + d\cdot\widetilde{k^v}$, *is an arbitrary Riemannian g-natural metric on T_1M, then $\widetilde{G}$ is induced, via the Musso-Tricerri process, by the* $(0,2)$*-tensor field* $\widetilde{Q} = (a+c)\sum_{i=1}^{m-1}(\theta^i)^2 + (a+c+d)(\theta^m)^2 + a\sum_{i=1}^{m-1}(\omega_m^i)^2 + 2b\sum_{i=1}^{m-1}\theta^i\omega_m^i$ *on* OM.

Proof. By Proposition 5, every Riemannian g-natural metric on T_1M is of the form $\widetilde{G} = a\cdot\widetilde{g^s} + b\cdot\widetilde{g^h} + c\cdot\widetilde{g^v} + d\cdot\widetilde{k^v}$,where a, b, c and d are constants such that $a>0$, $a(a+c)-b^2>0$ and $a+c+d>0$. Such a metric on T_1M is obviously induced by the g-natural metric $G = a\cdot g^s + b\cdot g^h + c\cdot g^v + d\cdot k^v$ on TM. If we consider, in Proposition 6, constant functions α_i, β_i; $i=1,2,3$, such that $\alpha_1 = a$, $\alpha_2 = b$, $\alpha_3 = c$, $\beta_3 = d$ and $\beta_1 = \beta_2 = 0$, then our G is induced by the symmetric tensor filed Q of type $(2,0)$ on $OM\times\mathbb{R}^m$ given by $Q = (a+c)\sum_{i=1}^m(\pi_1^*\theta^i)^2 + d(\sum_{i=1}^m\varepsilon^i\pi_1^*\theta^i)^2 + a\sum_{i=1}^m(\nabla\varepsilon^i)^2 + 2b\sum_{i=1}^m\pi_1^*\theta^i\nabla\varepsilon^i$, where $r^2 = \sum_{i=1}^m(\varepsilon^i)^2$. From Proposition 3, G is induced, via the Musso-Tricerri process, by the bilinear form $\widetilde{Q} = (i_m) * Q$ on OM, *i.e.*, by the form

$$\widetilde{Q} = (a+c)\sum_{i=1}^{m}((\pi_1 \circ i_m)^*\theta^i)^2 + d\left(\sum_{i=1}^{m}(\varepsilon^i \circ i_m\right)(\pi_1 \circ i_m)^*\theta^i)^2 + \\ + a\sum_{i=1}^{m}((i_m)^*\nabla\varepsilon^i)^2 + 2b\sum_{i=1}((\pi_1 \circ i_m)^*\theta^i)((i_m)^*\nabla\varepsilon^i). \tag{9}$$

But, it is easy to check that $\varepsilon^i \circ i_m = \delta^i_m$, where (δ^i_j) denote the Kronecker symbols. Then

$$r^2 \circ i_m = \sum_{i=1}^{m}(\varepsilon^i \circ i_m)^2 = 1 \quad \text{and} \quad (i_m)^*(d\varepsilon^i) = 0, \tag{10}$$

and it follows from (7) that

$$(i_m)^*(\nabla\varepsilon^i) = i_m^*(\pi_1^*\omega^i_m) = \omega^i_m, \tag{11}$$

and we have also

$$i_m^*(\pi_1^*\theta^i) = \theta^i, \tag{12}$$

since $\pi_1 \circ i_m = \mathrm{Id}_{OM}$. Hence, substituting from (10)-(12) into (9) and using the fact that $\omega^m_m = 0$ by the skew-symmetry of (ω^i_j), we obtain $\widetilde{Q} = (a + c)\sum_{i=1}^{m-1}(\theta^i)^2 + (a+c+d)(\theta^m)^2 + a\sum_{i=1}^{m-1}(\omega^i_m)^2 + 2b\sum_{i=1}^{m-1}\theta^i\omega^i_m$. ∎

Remark 1. Theorem 1 is a kind of weak generalization of the Main theorem in [1], where the base manifold (M, g) was a round sphere S^m. In our weaker analogy, the base manifold (M, g) is arbitrary. (*Cf.* [1], Section 4 and the formulas (3.1), (3.2)).

Now, we prove that any Riemannian g-natural metric on the unit tangent bundle of a two-point homogeneous space is homogeneous. This will generalize a theorem proved in [7, p. 10] for the induced Sasaki metric.

Theorem 2. *Let (M, g) be a two-point homogeneous space and let $\widetilde{G}$ be a Riemannian g-natural metric on T_1M. Then $(T_1M, \widetilde{G})$ is a homogeneous Riemannian space.*

Proof. Let $I(M,g)$ denote the group of isometries of (M,g). Then there is a natural left action of $I(M,g)$ on T_1M and OM, respectively, defined by the formulas

$$L_f(x,u) = (f(x), f_*u), \tag{13}$$

$$L_f(v) = (f(x), f_*u_1, \dots, f_*u_m), \tag{14}$$

where $f \in I(M,g)$, $(x,u) \in T_1M$ and $v := (x, u_1, \dots, u_m) \in OM$.

We claim that $\widetilde{G}$ is $I(M,g)$-invariant with respect to the action (13). It is well-known that the canonical 1-form *theta* and the Levi-Civita connection form ω are $I(M,g)$-invariant, *i.e.*,

$$L_f^*(\theta^i) = \theta^i, \tag{15}$$

$$L_f^*(\omega_j^i) = \omega_j^i. \tag{16}$$

Now, $\widetilde{G}$ is induced by the (0,2)-tensor filed $\widetilde{Q}$ from Theorem 1. By using (15) and (16) we obtain that $L_f^*(\widetilde{Q}) = \widetilde{Q}$. Moreover, $\widetilde{Q} = \psi_m^*(\widetilde{G})$ holds by the proof of Proposition 3.

We deduce that $\psi_m^*(\widetilde{G}) = L_f^*(\psi_m^*(\widetilde{G})) = (\psi_m \circ L_f)^*(\widetilde{G})$. But it is straightforward, form (13) and (14), that $\psi_m \circ L_f = L_f \circ \psi_m$. It follows then that $\psi_m^*(\widetilde{g}) = (L_f \circ \psi_m)^*(\widetilde{G}) = \psi_m^*(L_f^*(\widetilde{g}))$. Since ψ_m is a submersion, then $L_f^*(\widetilde{g}) = \widetilde{G}$, for all $f \in I(M,g)$. This proves our claim.

Next, it is classical that $I(M,g)$ is transitive on T_1M if and only if (M,g) is a two-point homogeneous space (*cf.* [9], p. 289). Hence if (M,g) is a two-point homogeneous space, then $I(M,g)$ acts transitively on T_1, M, as an isometry group. Consequently, $(T_1M, \widetilde{G})$ is a homogeneous Riemannian space.

For an alternative proof of Theorem 2 see [6]. □

Acknowledgements

The author would like to thank Professor O. Kowalski for his valuable comments and suggestions on this paper.

REFERENCES

[1] K. M. T. Abbassi – O. Kowalski: *Naturality of homogeneous metrics on Stiefel manifolds $SO(m+1)/SO(m-1)$*, Diff. Geometry and Appl., **28** (2010), 131–139

[2] K. M. T. Abbassi – M. Sarih: *On natural metrics on tangent bundles of Riemannian manifolds*, Archivum Math. (Brno), **41** (2005), 71–92.

[3] K. M. T. Abbassi – M. Sarih: *On some hereditary properties of Riemannian g-natural metrics on tangent bundles of Riemannian manifolds*, Diff. Geometry and Appl., **22** (1) (2005), 19–47.

[4] I. Kolář – P. W. Michor – J. Slovák: *Natural operations in differential geometry*, Springer-Verlag, Berlin, 1993.

[5] O. Kowalski – M. Sekizawa: *Natural transformations of Riemannian metrics on manifolds to metrics on tangent bundles – a classification –*, Bull. Tokyo Gakugei Univ., **40** (4) (1988), 1–29.

[6] O. Kowalski – M. Sekizawa: *Invariance of g-natural metrics on tangent bundles, differential geometry and its applications*, in: Proc. Conf. in Honour of Leonhard Euler, Olomouc, August 2007, World Scientific Publishing Company, 2088, 171–181.

[7] E. Musso – F. Tricerri: *Riemannian metrics on tangent bundles*, Ann. Math. Pura Appl., **150** (4) (1988), 1–20.

[8] S. Sasaki: *On the differential geometry of tangent bundles of Riemannian manifolds*, Tôhoku Math. J., I: **10** (1958, 338–354; II: **14** (1962), 146–155.

[9] J. A. Wolf: *Spaces of constant curvature*, Publich or Perish, IV edition, Berkeley, 1977.

Lavoro pervenuto alla redazione il 20 dicembre 2009
ed accettato per la pubblicazione il 10 maggio 2010.
Bozze licenziate il 26 novembre 2010

Indirizzo dell'Autore:

Mohamed Tahar Kadaoui Abbassi – Département des mathématiques – faculté des sciences Dhar El Mahraz – Université Sidi Mohamed Ben Abdallah – B.P. 1796 – Fés-Atlas – Fés – Morocco
E-mail: mtk_abbassi@yahoo.fr

Rendiconti di Matematica, Serie VII
Volume 30, Roma (2010), 249–274

Trigonometric approach to convolution formulae of Bernoulli and Euler numbers

WENCHANG CHU – CHENYING WANG

Abstract: *Summation formulae involving Bernoulli and Euler numbers as well as their convolutions are systematically reviewed by applying four classically elementary trigonometric identities.*

The Bernoulli and Euler numbers are important classical numbers and have wide applications in mathematics and physics. They can be defined, respectively, through the following trigonometric generating functions (see [12, Section 3.1.4], [14, Section 7.58] and [15, Section 2.5] for example)

$$x\cot x=\sum_{n=0}^{\infty}(-1)^n\frac{(2x)^{2n}}{(2n)!}B_{2n}, \tag{1}$$

$$\sec x=\sum_{n=0}^{\infty}(-1)^n\frac{x^{2n}}{(2n)!}E_{2n}. \tag{2}$$

According to the two elementary trigonometric relations

$$\tan x=\cot x-2\cot(2x)\quad\text{and}\quad\csc x=\cot x+\tan\frac{x}{2}$$

Key Words and Phrases: *Bernoulli numbers – Euler numbers – Trigonometric sums*
A.M.S. Classification: 11B68, 05A19.

the following two power series expansions can easily be shown

$$x\tan x=\sum_{n=1}^{\infty}(-1)^n\frac{(2x)^{2n}}{(2n)!}B'_{2n}, \tag{3}$$

$$x\csc x=\sum_{n=0}^{\infty}(-1)^n\frac{x^{2n}}{(2n)!}B''_{2n}; \tag{4}$$

where B'_{2n} and B''_{2n} denote, respectively, the two variants of Bernoulli numbers $B'_{2n}:=(1-4^n)B_{2n}$ and $B''_{2n}:=(2-4^n)B_{2n}$ in order to shorten lengthy expressions.

It is well-known that the sum of n^{th} powers of the first m natural numbers can be expressed in terms of Bernoulli numbers (*cf.* [5, Section 3.9] and [8, Section 6.5]):

$$\sum_{k=1}^{m}k^n=\frac{m+1}{n+1}\sum_{i=0}^{n}(m+1)^i\binom{n+1}{i+1}B_{n-i}.$$

Similar relations have recently been found by Liu and Luo [10] for the first m odd positive integers, which motivated the authors [3] to work out four classes of arithmetic identities involving Bernoulli and Euler numbers.

Observe that the above mentioned arithmetic identities have been accomplished entirely by manipulating elementary trigonometric sums. This encourages the authors to explore thoroughly the trigonometric approach to the arithmetic sums involving Bernoulli and Euler numbers as well as their convolutions. Our investigation will be carried out by employing exclusively four basic trigonometric sum identities. In fact, the rest of the paper will be structured into four sections according to these trigonometric relations with each of them having five different reformulations, that result logically in further division of each section into five subsections. Each subsection will prove a general theorem of arithmetic convolution sum involving Bernoulli and/or Euler numbers, followed by several concrete identities.

Throughout the paper, we shall assume $\delta=0,1$ and $m,n\in\mathbb{N}_0$. In addition, the following Taylor series for sine and cosine functions

$$\sin x=\sum_{n=0}^{\infty}(-1)^n\frac{x^{2n+1}}{(2n+1)!}\quad\text{and}\quad\cos x=\sum_{n=0}^{\infty}(-1)^n\frac{x^{2n}}{(2n)!} \tag{5}$$

will frequently be appealed without explanation.

Apart from the arithmetic formulae treated in this paper, there exist vast mathematical literature dealing with different approaches and identities for Bernoulli and Euler numbers as well as polynomials. The interested reader may consult, for instance, [1], [3], [4], [6] for convolution formulae, [7], [11] for Miki-type identities and [2], [13] for Bernoulli and Euler polynomials, as well as the handbook by Hansen [9, Section 50 and Section 51].

1 – Trigonometric sum concerning $\cos(2k+\gamma)x$

According to the following well-known formula

$$2\sin\alpha\cos\beta = \sin(\alpha+\beta) - \sin(\beta-\alpha)$$

we can evaluate via telescoping method the trigonometric sum

$$2\sin x\sum_{k=1}^{m}\cos(2k+\gamma)x = \sin(2m+\gamma+1)x - \sin(\gamma+1)x. \tag{6}$$

By means of five different reformulations of this identity, this section will investigate arithmetic sums involving Bernoulli and Euler numbers as well as their convolutions.

1.1 – Firstly, it is obvious that (6) is equivalent to the equation

$$2\sum_{k=1}^{m}\cos(2k+\gamma)x = \csc x\sin(2m+\gamma+1)x - \csc x\sin(\gamma+1)x.$$

Applying (4) and (5), we get the following power series expansion

$$2\sum_{n=0}^{\infty}\sum_{k=1}^{m}(-1)^n\frac{(2k+\gamma)^{2n}}{(2n)!}x^{2n} = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}(-1)^{i+j}\frac{(2m+\gamma+1)^{2j+1}}{(2i)!(2j+1)!}B''_{2i}x^{2i+2j} + \\ -\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}(-1)^{i+j}\frac{(\gamma+1)^{2j+1}}{(2i)!(2j+1)!}B''_{2i}x^{2i+2j}.$$

Comparing the coefficients of x^{2n}, we find immediately the following identity.

THEOREM 1 ($m \geq 0$ and $n \geq 0$).

$$\sum_{k=1}^{m}(2k+\gamma)^{2n} = \frac{(2m+\gamma+1)^{2n+1}}{2(2n+1)}\sum_{i=0}^{n}\binom{2n+1}{2i}\frac{B''_{2i}}{(2m+\gamma+1)^{2i}} + \\ -\frac{(\gamma+1)^{2n+1}}{2(2n+1)}\sum_{i=0}^{n}\binom{2n+1}{2i}\frac{B''_{2i}}{(\gamma+1)^{2i}}.$$

This general theorem contains several interesting identities as special cases.

COROLLARY 2 ($m = 1$ and $\gamma = -1$ in Theorem 1: Liu and Luo [10, Equation 8]).

$$\sum_{k=0}^{n}\frac{4^n}{4^k}\binom{2n+1}{2k}B''_{2k} = 2n+1.$$

Corollary 3 ($m=1$ and $\gamma=-2$ in Theorem 1: Liu–Luo [10, Equation 5]).

$$\sum_{k=0}^{n}\binom{2n+1}{2k}B''_{2k}=0 \quad \textit{where} \quad n>0.$$

According to (2), (4) and (5), extracting the coefficients of x^{2n} across the trigonometric equation $2\sin x\csc 2x=\sec x$, we recover another similar identity.

Lemma 4 (Chu–Wang [3, Equation 19a]).

$$\sum_{i=0}^{n}4^i\binom{2n+1}{2i}B''_{2i}=(2n+1)E_{2n}.$$

In Theorem 1, letting $\gamma=-\delta-1/2$ with $\delta=0,1$ and then simplifying the resulting equation through the last identity, we get the following transformation formula.

Proposition 5 ($\delta=0,1$ and $m,n\geq 0$).

$$\frac{(1-2\delta)(n+1/2)}{4^n(2m-\delta+1/2)^{2n+1}}E_{2n}=\sum_{i=0}^{n}\binom{2n+1}{2i}\frac{B''_{2i}}{(2m-\delta+1/2)^{2i}} -\frac{2(2n+1)}{(2m-\delta+1/2)^{2n+1}}\sum_{k=1}^{m}(2k-\delta-1/2)^{2n}.$$

When $\delta=0$ and $m=1,2$, this proposition yields the following two identities

$$\sum_{k=0}^{n}\binom{2n+1}{2k}\Big(\frac{2}{5}\Big)^{2k}B''_{2k}=\frac{2n+1}{5^{2n+1}}E_{2n}+4\frac{2n+1}{5^{2n+1}}3^{2n},$$
$$\sum_{k=0}^{n}\binom{2n+1}{2k}\Big(\frac{2}{9}\Big)^{2k}B''_{2k}=\frac{2n+1}{9^{2n+1}}\Big\{E_{2n}+4\cdot 3^{2n}+4\cdot 7^{2n}\Big\}.$$

Instead for $\delta=1$ and $m=1,2$, we get similarly two other identities

$$\sum_{k=0}^{n}\binom{2n+1}{2k}\Big(\frac{2}{3}\Big)^{2k}B''_{2k}=\frac{2n+1}{3^{2n+1}}(4-E_{2n}),$$
$$\sum_{k=0}^{n}\binom{2n+1}{2k}\Big(\frac{2}{7}\Big)^{2k}B''_{2k}=\frac{2n+1}{7^{2n+1}}\Big\{4(1+5^{2n})-E_{2n}\Big\}.$$

1.2 – Secondly, the identity (6) may equivalently be restated as

$$2\csc(2m+\gamma+1)x\sum_{k=1}^{m}\cos(2k+\gamma)x$$
$$=\csc x-\csc x\sin(\gamma+1)x\csc(2m+\gamma+1)x.$$

By means of (4) and (5), extracting the coefficients of x^{2n-1} from the last equation leads us to the following transformation theorem.

THEOREM 6 ($m\geq 0$ and $n\geq 0$).

$$\frac{B''_{2n}}{2(2m+\gamma+1)^{2n-1}}-\sum_{k=1}^{m}\sum_{i=0}^{n}\binom{2n}{2i}\frac{(2k+\gamma)^{2i}}{(2m+\gamma+1)^{2i}}B''_{2n-2i}$$
$$=\frac{(\gamma+1)^{2n+1}}{2(2n+1)(2m+\gamma+1)^{2n}}\sum_{0\leq i+j\leq n}\binom{2n+1}{2i,2j}\frac{(2m+\gamma+1)^{2j}}{(\gamma+1)^{2i+2j}}B''_{2i}B''_{2j}.$$

Several interesting identities follow immediately from this theorem.

COROLLARY 7 ($m=1$ and $\gamma=-1$ in Theorem 6: Liu–Luo [10, Equation 13]).

$$\sum_{k=0}^{n}4^k\binom{2n}{2k}B''_{2k}=B''_{2n}.$$

Next letting $m=\gamma=0$ in Theorem 6 gives directly the formula

$$\sum_{0\leq i+j\leq n}\binom{2n+1}{2i,2j}B''_{2i}B''_{2j}=(2n+1)B''_{2n}.$$

In fact, combining the series rearrangement with Corollary 3 we can show the following more general result.

COROLLARY 8 ($W\neq 0$).

$$\sum_{0\leq i+j\leq n}\binom{2n+1}{2i,2j}\frac{B''_{2i}B''_{2j}}{W^{2n-2i}}=(2n+1)B''_{2n}.$$

Finally, letting $\gamma = \delta - 1/2$ in Theorem 6 and then applying Lemma 4, we find the following transformation formula.

Proposition 9 ($\delta = 0, 1$ and $m, n \geq 0$).

$$\frac{B''_{2n}}{(2m-\delta+1/2)^{2n-1}} = 2\sum_{i=0}^{n}\sum_{k=1}^{m}\binom{2n}{2i}\Big(\frac{2k-\delta-1/2}{2m-\delta+1/2}\Big)^{2i}B''_{2n-2i} - (\delta-1/2)\sum_{i=0}^{n}\binom{2n}{2i}\frac{E_{2i}B''_{2n-2i}}{(4m-2\delta+1)^{2i}}.$$

For $\delta = 0$ and $m = 0, 1, 2$, this proposition reduces to the following three identities

$$\sum_{k=0}^{n}\binom{2n}{2k}B''_{2k}E_{2n-2k} = 4^n B''_{2n},$$

$$\sum_{k=0}^{n}\binom{2n}{2k}\frac{B''_{2n-2k}}{5^{2k}}\Big\{3^{2k}+\frac{E_{2k}}{4}\Big\} = \frac{2^{2n-2}}{5^{2n-1}}B''_{2n},$$

$$\sum_{k=0}^{n}\binom{2n}{2k}\frac{B''_{2n-2k}}{9^{2k}}\Big\{3^{2k}+7^{2k}+\frac{E_{2k}}{4}\Big\} = \frac{2^{2n-2}}{9^{2n-1}}B''_{2n}.$$

Similarly, when $\delta = 1$ and $m = 1, 2$, we have two further formulae

$$\sum_{k=0}^{n}\binom{2n}{2k}\frac{B''_{2n-2k}}{3^{2k}}\Big\{1-\frac{E_{2k}}{4}\Big\} = \frac{2^{2n-2}}{3^{2n-1}}B''_{2n},$$

$$\sum_{k=0}^{n}\binom{2n}{2k}\frac{B''_{2n-2k}}{7^{2k}}\Big\{1+5^{2k}-\frac{E_{2k}}{4}\Big\} = \frac{2^{2n-2}}{7^{2n-1}}B''_{2n}.$$

1.3 – Thirdly, rewrite (6) equivalently in the following manner

$$2\csc(2m+\gamma+1)x\sum_{k=1}^{m}\cos x\cos(2k+\gamma)x = \cot x - \cot x\sin(\gamma+1)x\csc(2m+\gamma+1)x$$

and then recall the relation

$$2\cos x\cos(2k+\gamma)x = \cos(2k+\gamma+1)x + \cos(2k+\gamma-1)x. \tag{11}$$

In view of (1), (4) and (5), equating the coefficients of x^{2n-1} across the penultimate equation gives rise to the following identity.

Theorem 10 ($m \geq 0$ and $n \geq 0$).

$$\sum_{k=1}^{m}\sum_{i=0}^{n}\binom{2n}{2i}\frac{(2m+\gamma+1)^{2i-1}}{(\gamma+1)^{2n+1}}B''_{2i}\left\{\begin{array}{c}(2k+\gamma+1)^{2n-2i}\\+(2k+\gamma-1)^{2n-2i}\end{array}\right\}$$
$$=\frac{4^n B_{2n}}{(\gamma+1)^{2n+1}}-\sum_{0\leq i+j\leq n}\binom{2n+1}{2i,2j}\frac{4^i(2m+\gamma+1)^{2j-1}}{(2n+1)(\gamma+1)^{2i+2j}}B_{2i}B''_{2j}.$$

As special cases of this theorem, three identities are displayed below. First letting $m=\gamma=0$ in Theorem 10, we have

$$\sum_{0\leq i+j\leq n}4^i\binom{2n+1}{2i,2j}B_{2i}B''_{2j}=(2n+1)4^n B_{2n}.$$

However, considering Corollary 3, we can show the following more general result.

Corollary 11 ($W\neq 0$).

$$\sum_{0\leq i+j\leq n}4^i\binom{2n+1}{2i,2j}\frac{B_{2i}B''_{2j}}{W^{2n-2i}}=(2n+1)4^n B_{2n}.$$

Corollary 12 ($m=1$ and $\gamma=-1$ in Theorem 10: Chu-Wang [3, Equation 9a]).

$$\sum_{k=0}^{n}\binom{2n}{2k}B''_{2k}=4^n B_{2n}.$$

Corollary 13 ($m=1$ and $\gamma=-3$ in Theorem 10: $n>0$).

$$\sum_{k=0}^{n}\binom{2n+1}{2k}B_{2k}=n+\frac{1}{2}.$$

1.4 – Fourthly, the identity (6) may equivalently be expressed as

$$2\cot(2m+\gamma+1)x\sum_{k=1}^{m}\cos(2k+\gamma)x=$$
$$=\csc x\cos(2m+\gamma+1)x-\csc x\sin(\gamma+1)x\cot(2m+\gamma+1)x.$$

On account of (1), (4) and (5), equating the coefficients of x^{2n-1} across the last equation results in the following transformation theorem.

THEOREM 14 ($m \geq 0$ and $n \geq 0$).

$$\begin{aligned}
&\sum_{i=0}^{n}\binom{2n}{2i}(2m+\gamma+1)^{2n-2i}B''_{2i}\\
&=2\sum_{k=1}^{m}\sum_{i=0}^{n}4^i\binom{2n}{2i}(2m+\gamma+1)^{2i-1}(2k+\gamma)^{2n-2i}B_{2i}\\
&\quad+\sum_{0\leq i+j\leq n}4^j\binom{2n+1}{2i,2j}(2m+\gamma+1)^{2j-1}\frac{(\gamma+1)^{2n+1-2i-2j}}{2n+1}B''_{2i}B_{2j}.
\end{aligned}$$

When $m = 1$ and $\gamma = -1$, Theorem 14 reduces to the following equality

$$\sum_{k=0}^{n}\frac{4^n}{4^k}\binom{2n}{2k}B''_{2k}=\sum_{\ell=0}^{n}16^\ell\binom{2n}{2\ell}B_{2\ell}.$$

By extracting the coefficients of x^{2n-1} across the following equation

$$\csc x\cos 2x = 2\cot 2x\cos x = \csc x - 2\sin x$$

we derive the two identities together.

COROLLARY 15 ($m = 1$ and $\gamma = -1$ in Theorem 14: CHU-WANG [3, Equation 10a]).

$$\sum_{k=0}^{n}\frac{4^n}{4^k}\binom{2n}{2k}B''_{2k}=\sum_{\ell=0}^{n}16^\ell\binom{2n}{2\ell}B_{2\ell}=B''_{2n}+4n.$$

1.5 – Finally, reformulate (6) equivalently as the following equality

$$\begin{aligned}
&2\cot(2m+\gamma+1)x\sum_{k=1}^{m}\cos x\cos(2k+\gamma)x\\
&=\cot x\cos(2m+\gamma+1)x-\cot x\sin(\gamma+1)x\cot(2m+\gamma+1)x.
\end{aligned}$$

With the help of (1), (5) and (11), extracting the coefficients of x^{2n-1} across this equation, we get the following identity.

Theorem 16 ($m \geq 0$ and $n \geq 0$).

$$\sum_{k=0}^{n} 4^k \binom{2n}{2k} (2m+\gamma+1)^{2n-2k} B_{2k}$$
$$= \sum_{k=1}^{m} \sum_{i=0}^{n} 4^i \binom{2n}{2i} (2m+\gamma+1)^{2i-1} B_{2i} \left\{ \begin{array}{c} (2k+\gamma+1)^{2n-2i} \\ +(2k+\gamma-1)^{2n-2i} \end{array} \right\}$$
$$+ \sum_{0 \leq i+j \leq n} 4^{i+j} \binom{2n+1}{2i, 2j} \frac{B_{2i} B_{2j}}{2n+1} (2m+\gamma+1)^{2j-1} (\gamma+1)^{2n-2i-2j+1}.$$

First letting $m = \gamma = 0$ in Theorem 16 results in the following relation

$$\sum_{0 \leq i+j \leq n} 4^{i+j} \binom{2n+1}{2i, 2j} \frac{B_{2i} B_{2j}}{2n+1} = \sum_{k=0}^{n} 4^k \binom{2n}{2k} B_{2k}$$

which can also be verified by applying the following lemma.

Lamma 17.

$$\sum_{k=0}^{n} 4^k \binom{2n+1}{2k} B_{2k} = 2n+1.$$

This identity follows easily by equating the coefficients of x^{2n} across the trigonometric equation $\sin x \cot x = \cos x$. Instead, by extracting the coefficients of x^{2n-1} across the equalities

$$\cot^2 x \sin x = \cos x \cot x = \csc x - \sin x$$

we get the following closed formulae.

Corollary 18 ($m = \gamma = 0$ in Theorem 16).

$$\sum_{0 \leq i+j \leq n} 4^{i+j} \binom{2n+1}{2i, 2j} \frac{B_{2i} B_{2j}}{2n+1} = \sum_{k=0}^{n} 4^k \binom{2n}{2k} B_{2k} = B''_{2n} + 2n.$$

Finally we examine the case of Theorem 16 with $m = 1$ and $\gamma = -1$

$$2 \sum_{k=0}^{n} \binom{2n}{2k} B_{2k} = 4^n B_{2n} + \sum_{i=0}^{n} 4^i \binom{2n}{2i} B_{2i}.$$

Recalling Corollary 18, we find another similar closed formula.

Corollary 19 ($m = 1$ and $\gamma = -1$ in Theorem 16).

$$\sum_{k=0}^{n} \binom{2n}{2k} B_{2k} = n + B_{2n}.$$

2 – Trigonometric sum concerning $\sin(2k+\gamma)x$

By means of the trigonometric relation

$$2\sin\alpha\sin\beta = \cos(\alpha-\beta) - \cos(\alpha+\beta)$$

it is not hard to compute the finite sum

$$2\sin x \sum_{k=1}^{m} \sin(2k+\gamma)x = \cos(\gamma+1)x - \cos(2m+\gamma+1)x. \tag{12}$$

According to five different reformulations of this identity, this section will investigate summation formulae involving Bernoulli and Euler numbers.

2.1 – Firstly, it is obvious that (12) is equivalent to the equation

$$2\sum_{k=1}^{m} \sin(2k+\gamma)x = \csc x\cos(\gamma+1)x - \csc x\cos(2m+\gamma+1)x.$$

Applying (4) and (5), we have the power series expansion

$$\begin{aligned} 2\sum_{n=0}^{\infty}\sum_{k=1}^{m}(-1)^n\frac{(2k+\gamma)^{2n+1}}{(2n+1)!}x^{2n+1} = {} & \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}(-1)^{i+j}\frac{(\gamma+1)^{2j}}{(2i)!(2j)!}B''_{2i}x^{2i+2j-1} \\ & - \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}(-1)^{i+j}\frac{(2m+\gamma+1)^{2j}}{(2i)!(2j)!}B''_{2i}x^{2i+2j-1}. \end{aligned}$$

Extracting the coefficients of x^{2n-1} from both sides of the last equation and then simplifying the result, we derive the following formula.

Theorem 20 ($m \geq 0$ and $n \geq 1$).

$$\sum_{k=1}^{m}(2k+\gamma)^{2n-1} = \sum_{i=0}^{n}\frac{B''_{2i}}{4n}\binom{2n}{2i}\frac{(2m+\gamma+1)^{2n}}{(2m+\gamma+1)^{2i}} - \sum_{i=0}^{n}\frac{B''_{2i}}{4n}\binom{2n}{2i}\frac{(\gamma+1)^{2n}}{(\gamma+1)^{2i}}.$$

According to Corollary 7, letting $\gamma = -\delta - 1/2$ in this theorem yields the formula.

Proposition 21 ($\delta = 0, 1$ and $m,\, n \geq 0$).

$$\sum_{i=0}^{n}\binom{2n}{2i}\frac{B''_{2i}}{(2m-\delta+1/2)^{2i}} = 4n\sum_{k=1}^{m}\frac{(2k-\delta-1/2)^{2n-1}}{(2m-\delta+1/2)^{2n}} + \frac{B''_{2n}}{(4m-2\delta+1)^{2n}}.$$

For $\delta = 0$ and $m = 1, 2$, the last theorem reduces to the following two identities

$$\sum_{k=0}^{n} \binom{2n}{2k} \left(\frac{2}{5}\right)^{2k} B''_{2k} = \frac{1}{5^{2n}} \{8n \cdot 3^{2n-1} + B''_{2n}\},$$

$$\sum_{k=0}^{n} \binom{2n}{2k} \left(\frac{2}{9}\right)^{2k} B''_{2k} = \frac{1}{9^{2n}} \{8n(3^{2n-1} + 7^{2n-1}) + B''_{2n}\}.$$

Similarly, when $\delta = 1$ and $m = 1, 2$, we get from Theorem 20 two further interesting identities

$$\sum_{k=0}^{n} \binom{2n}{2k} \left(\frac{2}{3}\right)^{2k} B''_{2k} = \frac{1}{3^{2n}} \Big\{8n + B''_{2n}\Big\},$$

$$\sum_{k=0}^{n} \binom{2n}{2k} \left(\frac{2}{7}\right)^{2k} B''_{2k} = \frac{1}{7^{2n}} \Big\{8n(1 + 5^{2n-1}) + B''_{2n}\Big\}.$$

2.2 – Secondly, the identity (12) may equivalently be restated as

$$\begin{aligned} &2\sec(2m+\gamma+1)x \sum_{k=1}^{m} \sin(2k+\gamma)x = \\ &= \csc x \cos(\gamma+1)x \sec(2m+\gamma+1)x - \csc x. \end{aligned}$$

By means of (2), (4) and (5), extracting the coefficients of x^{2n-1} across this equation yields the following identity.

Theorem 22 ($m \geq 0$ and $n \geq 1$).

$$\begin{aligned} &\frac{B''_{2n}}{4n(2m+\gamma+1)^{2n}} - \sum_{k=1}^{m} \sum_{i=1}^{n} \binom{2n-1}{2i-1} \frac{(2k+\gamma)^{2i-1}}{(2m+\gamma+1)^{2i}} E_{2n-2i} \\ &= \frac{(\gamma+1)^{2n}}{4n(2m+\gamma+1)^{2n}} \sum_{0 \leq i+j \leq n} \binom{2n}{2i, 2j} \frac{(2m+\gamma+1)^{2j}}{(\gamma+1)^{2i+2j}} B''_{2i} E_{2j}. \end{aligned}$$

Letting $m = \gamma = 0$, Theorem 22 gives directly the formula

$$\sum_{0 \leq i+j \leq n} \binom{2n}{2i, 2j} B''_{2i} E_{2j} = B''_{2n}.$$

In fact, applying Corollary 45, we can show the following more general result.

Corollary 23 ($W \neq 0$).

$$\sum_{0 \leq i+j \leq n} \binom{2n}{2i, 2j} \frac{B''_{2i} E_{2j}}{W^{2n-2i}} = B''_{2n}.$$

Taking $\gamma = \delta - 1/2$ in Theorem 22 and then appealing to Corollary 7, we derive the following transformation.

PROPOSITION 24 ($\delta = 0$, and $m \geq 0$, $n \geq 1$).

$$\sum_{k=1}^{m}\sum_{i=1}^{n}\binom{2n-1}{2i-1}\frac{(2k-\delta-1/2)^{2i-1}}{(2m-\delta+1/2)^{2i}}E_{2n-2i}$$
$$=\frac{B''_{2n}}{4n(2m-\delta+1/2)^{2n}}-\frac{1}{4n}\sum_{i=0}^{n}\binom{2n}{2i}\frac{B''_{2i}E_{2n-2i}}{4^i(2m-\delta+1/2)^{2i}}.$$

2.3– Thirdly, rewrite (12) equivalently in the following manner

$$2\sec(2m+\gamma+1)x\sum_{k=1}^{m}\cos x\sin(2k+\gamma)x$$
$$=\cot x\cos(\gamma+1)x\sec(2m+\gamma+1)x-\cot x$$

and then recall the trigonometric relation

$$2\cos x\sin(2k+\gamma)x=\sin(2k+\gamma+1)x+\sin(2k+\gamma-1)x. \tag{15}$$

In view of (1), (2) and (5), extracting the coefficients of x^{2n+1} from the penultimate equation and then simplifying the result, we derive the following arithmetic formula.

THEOREM 25 ($m \geq 0$ and $n \geq 0$).

$$\sum_{0\leq i+j\leq n+1}4^i\binom{2n+2}{2i,2j}(2m+\gamma+1)^{2j}\frac{(\gamma+1)^{2n+2-2i-2j}}{2n+2}B_{2i}E_{2j}$$
$$=4^{n+1}\frac{B_{2n+2}}{2n+2}-\sum_{k=1}^{m}\sum_{i=0}^{n}\binom{2n+1}{2i}(2m+\gamma+1)^{2i}E_{2i}\left\{\begin{array}{c}(2k+\gamma+1)^{2n+1-2i}\\+(2k+\gamma-1)^{2n+1-2i}\end{array}\right\}.$$

Two examples of this theorem are given below as applications.

Taking $m = \gamma = 0$ in Theorem 25, we have directly the formula

$$\sum_{0\leq i+j\leq n}4^i\binom{2n}{2i,2j}B_{2i}E_{2j}=4^nB_{2n}.$$

In fact, by means of Corollary 45, we can show the following more general result.

COROLLARY 26 ($W \neq 0$).

$$\sum_{0\leq i+j\leq n}4^i\binom{2n}{2i,2j}\frac{B_{2i}E_{2j}}{W^{2n-2i}}=4^nB_{2n}.$$

Letting $m = 1$ and $\gamma = -1$ in Theorem 10, we have the following transformation

$$\sum_{i+j=n+1} \binom{2+2n}{2i, 2j} B_{2i} E_{2j} = B_{2n+2} - (n+1) \sum_{i=0}^{n} \binom{2n+1}{2i} E_{2i}.$$

Evaluating the last sum by Corollary 34 and then replacing n by $n-1$, we get the following convolution formula between Bernoulli and Euler numbers.

Corollary 27.

$$\sum_{k=0}^{n} \binom{2n}{2k} B_{2k} E_{2n-2k} = B_{2n} \Big\{ 1 + 2^{2n-1} - 2^{4n-1} \Big\}.$$

This can also be verified by equating the coefficients x^{2n-1} across the following trigonometric equation

$$\sec x \cot \tfrac{x}{2} = \tan \tfrac{x}{2} + \cot \tfrac{x}{2}.$$

2.4 – Fourthly, the identity (12) may equivalently be expressed as

$$\begin{aligned} &2 \tan(2m+\gamma+1)x \sum_{k=1}^{m} \sin(2k+\gamma)x \\ &= -\csc x \sin(2m+\gamma+1)x + \csc x \cos(\gamma+1)x \tan(2m+\gamma+1)x. \end{aligned}$$

On account of (3), (4) and (5), we can equate the coefficients of x^{2n} across the last equation and obtain the following arithmetic formula.

Theorem 28 ($m \geq 0$ and $n \geq 0$).

$$\begin{aligned} &2 \sum_{k=1}^{m} \sum_{i=0}^{n} 4^i \binom{2n+1}{2i} (2m+\gamma+1)^{2i-1} (2k+\gamma)^{2n+1-2i} B'_{2i} \\ &= - \sum_{0 \leq i+j \leq n+1} 4^i \binom{2n+2}{2i, 2j} (2m+\gamma+1)^{2i-1} \frac{(\gamma+1)^{2n+2-2i-2j}}{2n+2} B'_{2i} B''_{2j} \\ &\quad - \sum_{i=0}^{n} \binom{2n+1}{2i} (2m+\gamma+1)^{2n+1-2i} B''_{2i}. \end{aligned}$$

Letting $m = \gamma = 0$ in this theorem and then keeping in mind of Corollary 3, we find the following strange-looking identity.

Corollary 29 ($n > 1$).

$$\sum_{0\le i+j\le n} 4^i \binom{2n}{2i, 2j} B'_{2i} B''_{2j} = 0.$$

2.5 – Finally, reformulate (6) equivalently as the following equality

$$\begin{aligned} &2\tan(2m+\gamma+1)x \sum_{k=1}^{m} \cos x \sin(2k+\gamma)x = \\ &= \cot x \cos(\gamma+1)x \tan(2m+\gamma+1)x - \cot x \sin(2m+\gamma+1)x. \end{aligned}$$

With the help of the trigonometric relation (15), we can extract, according to (1), (3) and (5), the coefficients of x^{2n} across the last equation and establish the following formula.

Theorem 30 ($m \ge 0$ and $n \ge 0$).

$$\begin{aligned} &\sum_{k=1}^{m}\sum_{i=0}^{n} 4^i \binom{2n+1}{2i} (2m+\gamma+1)^{2i-1} B'_{2i} \left\{ \begin{array}{c} (2k+\gamma+1)^{2n+1-2i} \\ +(2k+\gamma-1)^{2n+1-2i} \end{array} \right\} \\ &= -\sum_{i=0}^{n} 4^i \binom{2n+1}{2i} (2m+\gamma+1)^{2n-2i+1} B_{2i} \\ &\quad - \sum_{0\le i+j\le n+1} 4^{i+j} \binom{2n+2}{2i, 2j} \frac{B_{2i} B'_{2j}}{2n+2} (2m+\gamma+1)^{2j-1} (\gamma+1)^{2n+2-2i-2j}. \end{aligned}$$

When $m = \gamma = 0$, the last expression yields the following transformation

$$\sum_{0\le i+j\le n+1} 4^{i+j} \binom{2n+2}{2i, 2j} B_{2i} B'_{2j} = -(2n+2) \sum_{i=0}^{n} 4^i \binom{2n+1}{2i} B_{2i}.$$

Evaluating the last sum by Lemma 17 and then replacing n by $n-1$, we get the following convolution formula for Bernoulli numbers.

Corollary 31.

$$\sum_{0\le i+j\le n} 4^{i+j} \binom{2n}{2i, 2j} B_{2i} B'_{2j} = 2n(1-2n).$$

This can also be proved by equating the coefficients x^{2n-2} across the following trigonometric equation

$$\cos x \tan x \cot x = \cos x.$$

3 – Alternating sum concerning $\sin(2k+\gamma)x$

Recalling the trigonometric formula

$$2\sin\alpha\cos\beta = \sin(\alpha+\beta) + \sin(\alpha-\beta)$$

we have the finite trigonometric sum

$$2\cos x \sum_{k=1}^{m} (-1)^k \sin(2k+\gamma)x = (-1)^m \sin(2m+\gamma+1)x - \sin(\gamma+1)x. \tag{16}$$

By means of five different reformulations of this identity, this section will investigate convolution formulae involving Bernoulli and Euler numbers.

3.1 – Firstly, it is obvious that (16) is equivalent to the equation

$$2\sum_{k=1}^{m} (-1)^k \sin(2k+\gamma)x = (-1)^m \sec x \sin(2m+\gamma+1)x - \sec x \sin(\gamma+1)x.$$

According to (2) and (5), equating the coefficients of x^{2n+1} across the last equation, we find the following formula.

Theorem 32 ($m \geq 0$ and $n \geq 0$).

$$2\sum_{k=1}^{m} (-1)^k (2k+\gamma)^{2n+1} = (-1)^m \sum_{i=0}^{n} \binom{2n+1}{2i+1} (2m+\gamma+1)^{2i+1} E_{2n-2i} - \sum_{i=0}^{n} \binom{2n+1}{2i+1} (\gamma+1)^{2i+1} E_{2n-2i}.$$

Two known identities can be recovered directly from this theorem.

Corollary 33 ($m = 1$ and $\gamma = -1$ in Theorem 32: Chu-Wang [3, Equation 17a]).

$$\sum_{i=0}^{n} \frac{4^n}{4^i} \binom{2n+1}{2i} E_{2i} = 1.$$

Comparing the case $\gamma=-\delta$ of this theorem with the identity due to Chu and Wang [3, Theorem 7], we recover another formula.

Corollary 34 (Hansen [9, Equation 51.1.2] and Chu-Wang [3, Equation 16a: $n>0$]).

$$\sum_{i=0}^{n}\binom{2n-1}{2i}E_{2i}=-4^n\frac{B'_{2n}}{2n}.$$

Letting $m=1,2$ in this theorem, we get respectively the following two identities

$$\sum_{k=0}^{n}\binom{2n+1}{2k+1}\Big\{(\gamma+1)^{2k+1}+(\gamma+3)^{2k+1}\Big\}E_{2n-2k}=2(\gamma+2)^{2n+1},$$

$$\sum_{k=0}^{n}\binom{2n+1}{2k+1}\Big\{(\gamma+5)^{2k+1}-(\gamma+1)^{2k+1}\Big\}E_{2n-2k}=2\Big\{(\gamma+4)^{2n+1}-(\gamma+2)^{2n+1}\Big\}.$$

3.2 – Secondly, the identity (16) may equivalently be restated as

$$\begin{aligned}&2\csc(2m+\gamma+1)x\sum_{k=1}^{m}(-1)^k\sin(2k+\gamma)x\\&=(-1)^m\sec x-\sec x\sin(\gamma+1)x\csc(2m+\gamma+1)x.\end{aligned}$$

By means of (2), (4) and (5), equating the coefficients of x^{2n} across this equation and then simplifying the result, we get the following identity.

Theorem 35 ($m\geq 0$ and $n\geq 0$).

$$\begin{aligned}&\frac{(-1)^m(2n+1)}{(2m+\gamma+1)^{2n-1}}E_{2n}-2\sum_{k=1}^{m}\sum_{i=0}^{n}(-1)^k\binom{2n+1}{2i+1}\frac{(2k+\gamma)^{2i+1}}{(2m+\gamma+1)^{2i}}B''_{2n-2i}\\&=\frac{(\gamma+1)^{2n+1}}{(2m+\gamma+1)^{2n}}\sum_{0\leq i+j\leq n}\binom{2n+1}{2i,2j}\frac{(2m+\gamma+1)^{2i}}{(\gamma+1)^{2i+2j}}B''_{2i}E_{2j}.\end{aligned}$$

Letting $m=\gamma=0$ in Theorem 35 gives directly the formula

$$\sum_{0\leq i+j\leq n}\binom{2n+1}{2i,2j}B''_{2i}E_{2j}=(2n+1)E_{2n}.$$

In fact, applying Corollary 3, we can prove the following more general result.

Corollary 36 ($W \neq 0$).

$$\sum_{0\le i+j\le n} \binom{2n+1}{2i,2j} \frac{B''_{2i}E_{2j}}{W^{2n-2j}} = (2n+1)E_{2n}.$$

3.3 – Thirdly, rewrite (16) equivalently in the following manner

$$\begin{aligned} &2\csc(2m+\gamma+1)x \sum_{k=1}^{m}(-1)^k \sin x \sin(2k+\gamma)x \\ &= (-1)^m \tan x - \tan x \sin(\gamma+1)x \csc(2m+\gamma+1)x \end{aligned}$$

and recall the trigonometric relation

$$2\sin x \sin(2k+\gamma)x = \cos(2k+\gamma-1)x - \cos(2k+\gamma+1)x.$$

In view of (3), (4) and (5), extracting the coefficient of x^{2n-1} across the penultimate equation, we get the identity.

Theorem 37 ($m \ge 0$ and $n \ge 0$).

$$\begin{aligned} &\sum_{k=1}^{m}\sum_{i=0}^{n}(-1)^k \binom{2n}{2i} \frac{(2m+\gamma+1)^{2i-1}}{(\gamma+1)^{2n+1}} B''_{2i} \left\{ \begin{matrix} (2k+\gamma-1)^{2n-2i} \\ -(2k+\gamma+1)^{2n-2i} \end{matrix} \right\} \\ &= (-1)^m \frac{4^n B'_{2n}}{(\gamma+1)^{2n+1}} - \sum_{0\le i+j\le n} \binom{2n+1}{2i,2j} \frac{4^i(2m+\gamma+1)^{2j-1}}{(2n+1)(\gamma+1)^{2i+2j}} B'_{2i}B''_{2j}. \end{aligned}$$

When $m = \gamma = 0$, Theorem 37 yields the following identity

$$\sum_{0\le i+j\le n} 4^i \binom{2n+1}{2i,2j} B'_{2i}B''_{2j} = (2n+1)4^n B'_{2n}.$$

This identity can also be shown by equating the coefficients x^{2n-1} across the following trigonometric equation

$$\tan x \csc x \sin x = \tan x.$$

Furthermore, we can verify through Corollary 3, the following more general result.

corollary 38 ($W \neq 0$).

$$\sum_{0\le i+j\le n} 4^i \binom{2n+1}{2i,2j} \frac{B'_{2i}B''_{2j}}{W^{2n-2i}} = (2n+1)4^n B'_{2n}.$$

Corollary 39 ($m = 1$ and $m = -3$ in Theorem 37: $n > 0$).

$$\sum_{i=0}^{n} \binom{2n+1}{2i} B'_{2i} = -n - 1/2.$$

We remark that this identity is also the linear combination of Corollary 13 and Lemma 17.

3.4 – Fourthly, the identity (16) may equivalently be expressed as

$$\begin{aligned} &2\cot(2m+\gamma+1)x \sum_{k=1}^{m} (-1)^k \sin(2k+\gamma)x = \\ &= (-1)^m \sec x \cos(2m+\gamma+1)x - \sec x \sin(\gamma+1)x \cot(2m+\gamma+1)x. \end{aligned}$$

On account of (1), (2) and (5), equating the coefficients of x^{2n} across this equation and then simplifying the result, we get the following identity.

Theorem 40 ($m \geq 0$ and $n \geq 0$).

$$\begin{aligned} &\sum_{k=1}^{m} \sum_{i=0}^{n} (-1)^k \binom{2n+1}{2i} 2^{2i+1} (2m+\gamma+1)^{2i-1} (2k+\gamma)^{2n+1-2i} B_{2i} \\ &= (2n+1) \sum_{i=0}^{n} (-1)^m \binom{2n}{2i} (2m+\gamma+1)^{2n-2i} E_{2i} \\ &\quad - \sum_{0 \leq i+j \leq n} 4^j \binom{2n+1}{2i, 2j} (2m+\gamma+1)^{2j-1} (\gamma+1)^{2n+1-2i-2j} E_{2i} B_{2j}. \end{aligned}$$

When $m = \gamma = 0$, it yields the following expression

$$\sum_{0 \leq i+j \leq n} 4^j \binom{2n+1}{2i, 2j} E_{2i} B_{2j} = (2n+1) \sum_{i=0}^{n} \binom{2n}{2i} E_{2i}.$$

According to Corollary 45, this gives rise to the following formula.

Corollary 41 ($n > 0$).

$$\sum_{0 \leq i+j \leq n} 4^i \binom{2n+1}{2i, 2j} B_{2i} E_{2j} = 0.$$

3.5 – Finally, reformulate (16) equivalently as the following equality

$$\begin{aligned}&2\cot(2m+\gamma+1)x\sum_{k=1}^{m}(-1)^k\sin x\sin(2k+\gamma)x\\&=(-1)^m\tan x\cos(2m+\gamma+1)x+\tan x\sin(\gamma+1)x\cot(2m+\gamma+1)x.\end{aligned}$$

With the help of (1), (3) and (5), extracting the coefficient of x^{2n-1} across the last equation leads us to the following identity.

Theorem 42 ($m \geq 0$ and $n \geq 0$).

$$\begin{aligned}&\sum_{k=1}^{m}\sum_{i=0}^{n}(-1)^k\binom{2n}{2i}\frac{4^i(2m+\gamma+1)^{2i-1}}{(\gamma+1)^{2n+1}}B_{2i}\left\{\begin{array}{c}(2k+\gamma-1)^{2n-2i}\\-(2k+\gamma+1)^{2n-2i}\end{array}\right\}\\&=\sum_{i=0}^{n}(-1)^m\binom{2n}{2i}\frac{4^i(2m+\gamma+1)^{2n-2i}}{(\gamma+1)^{2n+1}}B'_{2i}\\&\quad-\sum_{0\leq i+j\leq n}\binom{2n+1}{2i,2j}\frac{4^{i+j}(2m+\gamma+1)^{2j-1}}{(2n+1)(\gamma+1)^{2i+2j}}B'_{2i}B_{2j}.\end{aligned}$$

When $m=\gamma=0$, it yields the following transformation

$$\sum_{0\leq i+j\leq n}4^{i+j}\binom{2n+1}{2i,2j}\frac{B'_{2i}B_{2j}}{2n+1}=\sum_{i=0}^{n}4^i\binom{2n}{2i}B'_{2i}.$$

By extracting the coefficients of x^{2n-1} across the expansion of the trigonometric relation

$$\tan x\cot x\sin x=\cos x\tan x=\sin x,$$

we can show further the following two closed formulae.

Corollary 43.

$$\sum_{0\leq i+j\leq n}4^{i+j}\binom{2n+1}{2i,2j}\frac{B'_{2i}B_{2j}}{2n+1}=\sum_{i=0}^{n}4^i\binom{2n}{2i}B'_{2i}=-2n.$$

4 – Alternating sum concerning $\cos(2k+\gamma)x$

In view of the following trigonometric relation

$$2\cos\alpha\cos\beta = \cos(\alpha+\beta) + \cos(\alpha-\beta)$$

it is almost trivial to derive that

$$2\cos x \sum_{k=1}^{m}(-1)^k \cos(2k+\gamma)x = (-1)^m \cos(2m+\gamma+1)x - \cos(\gamma+1)x. \tag{18}$$

According to five different reformulations of this identity, this section will investigate convolution identities involving Bernoulli and Euler numbers.

4.1 – Firstly, it is obvious that (18) is equivalent to the equation

$$2\sum_{k=1}^{m}(-1)^k \cos(2k+\gamma)x = (-1)^m \sec x\cos(2m+\gamma+1)x - \sec x\cos(\gamma+1)x.$$

According to (2) and (5), we have the following power series expansions

$$\begin{aligned}2\sum_{k=1}^{m}\sum_{n=0}^{\infty}(-1)^{n+k}\frac{(2k+\gamma)^{2n}}{(2n)!}x^{2n} + \sum_{i,j\geq 0}(-1)^{i+j}\frac{(\gamma+1)^{2j}}{(2i)!(2j)!}E_{2i}x^{2i+2j}\\ = \sum_{i,j\geq 0}(-1)^{m+i+j}\frac{(2m+\gamma+1)^{2j}}{(2i)!(2j)!}E_{2i}x^{2i+2j}.\end{aligned}$$

Equating the coefficients of x^{2n} across this equation, we find the transformation.

Theorem 44 ($m \geq 0$ and $n \geq 0$).

$$\begin{aligned}2\sum_{k=1}^{m}(-1)^k(2k+\gamma)^{2n} = (-1)^m\sum_{i=0}^{n}\binom{2n}{2i}(2m+\gamma+1)^{2n-2i}E_{2i}\\ -\sum_{i=0}^{n}\binom{2n}{2i}(\gamma+1)^{2n-2i}E_{2i}.\end{aligned}$$

Comparing the case $\gamma = -\delta$ of this theorem with the identity due to Chu and Wang [3, Theorem 10], we recover the following well-known identity.

Corollary 45 (Stromberg [14, Section 7.58]: $n > 0$).

$$\sum_{i=0}^{n} \binom{2n}{2i} E_{2i} = 0.$$

For $m = 1$ and $\gamma = -3$, the last theorem recovers another interesting identity.

Corollary 46 ($m = 1$ and $\gamma = -3$ in Theorem 44: Chu-Wang [3, Equation 23a]).

$$\sum_{i=0}^{n} \frac{4^n}{4^i} \binom{2n}{2i} E_{2i} = 2 - E_{2n}.$$

4.2 – Secondly, the identity (18) may equivalently be restated as

$$\begin{aligned} &2 \sec(2m+\gamma+1)x \sum_{k=1}^{m} (-1)^k \cos(2k+\gamma)x \\ &= (-1)^m \sec x - \sec x \sec(2m+\gamma+1)x \cos(\gamma+1)x. \end{aligned}$$

By means of (2), (4) and (5), equating the coefficients of x^{2n} across this equation yields the following identity.

Theorem 47 ($m \geq 1$ and $n \geq 0$).

$$\begin{aligned} &\frac{(-1)^m E_{2n}}{(2m+\gamma+1)^{2n}} - 2 \sum_{k=1}^{m} \sum_{i=0}^{n} (-1)^k \binom{2n}{2i} \frac{(2k+\gamma)^{2i}}{(2m+\gamma+1)^{2i}} E_{2n-2i} \\ &= \frac{(\gamma+1)^{2n}}{(2m+\gamma+1)^{2n}} \sum_{0 \leq i+j \leq n} \binom{2n}{2i, 2j} \frac{(2m+\gamma+1)^{2i}}{(\gamma+1)^{2i+2j}} E_{2i} E_{2j}. \end{aligned}$$

When $m = \gamma = 0$, it reduces to the following identity

$$\sum_{0 \leq i+j \leq n} \binom{2n}{2i, 2j} E_{2i} E_{2j} = E_{2n}.$$

By exchanging the summation order and then applying Corollary 45, we can show the following more general result.

Corollary 48 ($W \neq 0$).

$$\sum_{0 \leq i+j \leq n} \binom{2n}{2i, 2j} \frac{E_{2i} E_{2j}}{W^{2n-2i}} = E_{2n}.$$

4.3 – Thirdly, rewrite (18) equivalently in the following manner

$$\begin{aligned}&2\sec(2m+\gamma+1)x\sum_{k=1}^{m}(-1)^k\sin x\cos(2k+\gamma)x\\&=(-1)^m\tan x-\tan x\sec(2m+\gamma+1)x\cos(\gamma+1)x\end{aligned}$$

and recall to the trigonometric relation

$$2\sin x\cos(2k+\gamma)x=\sin(2k+\gamma+1)x-\sin(2k+\gamma-1)x. \tag{19}$$

In view of (2), (3) and (5), extracting the coefficients of x^{2n+1} across the penultimate equation results in the following general transformation.

Theorem 49 ($m\geq 0$ and $n\geq 0$).

$$\begin{aligned}&\sum_{k=1}^{m}\sum_{i=0}^{n}(-1)^k\binom{2n+1}{2i}\frac{(2m+\gamma+1)^{2i}}{(\gamma+1)^{2n+2}}E_{2i}\begin{Bmatrix}(2k+\gamma+1)^{2n-2i+1}\\-(2k+\gamma-1)^{2n-2i+1}\end{Bmatrix}\\&=\frac{(-1)^{m+1}2^{2n+1}}{(n+1)(\gamma+1)^{2n+2}}B'_{2n+2}+\sum_{0\leq i+j\leq n+1}\binom{2n+2}{2i,2j}\frac{4^i(2m+\gamma+1)^{2j}}{(2n+2)(\gamma+1)^{2i+2j}}B'_{2i}E_{2j}.\end{aligned}$$

Three identities can be derived from this theorem as consequences.

Firstly, when $m=\gamma=0$, the theorem yields the following identity

$$\sum_{0\leq i+j\leq n}4^i\binom{2n}{2i,2j}B'_{2i}E_{2j}=4^nB'_{2n}.$$

Applying Corollary 45, this can be generalized to the following general formula.

Corollary 50 ($W\neq 0$).

$$\sum_{0\leq i+j\leq n}4^i\binom{2n}{2i,2j}\frac{B'_{2i}E_{2j}}{W^{2n-2i}}=4^nB'_{2n}.$$

Secondly, taking $m=1$ and $\gamma=-1$, we have the transformation expression

$$\sum_{i=0}^{n+1}\binom{2n+2}{2i}B'_{2i}E_{2n+2-2i}=-B'_{2n+2}-(n+1)\sum_{i=0}^{n}\binom{2n+1}{2i}E_{2i}.$$

Evaluating the last sum by Corollary 34 and then replacing n by $n-1$, we get the following interesting convolution formula.

Corollary 51.

$$\sum_{i=0}^{n}\binom{2n}{2i}B'_{2i}E_{2n-2i}=(2^{2n-1}-1)B'_{2n}.$$

This can also be verified by equating the coefficients of x^{2n-1} across the following trigonometric equation

$$\tan x \sec x = \tan x - \tan \frac{x}{2}.$$

Finally, letting $m = 1$ and $\gamma = -3$, we find another closed formula, which is , in fact, also a linear combination of Corollary 18 and Corollary 19.

Corollary 52 ($m = 1$ and $\gamma = -3$ in Theorem 49).

$$\sum_{i=0}^{n} \binom{2n}{2i} B'_{2i} = -n - B'_{2n}.$$

4.4 – Fourthly, the identity (18) may equivalently be expressed as

$$\begin{aligned} &2\tan(2m+\gamma+1)x \sum_{k=1}^{m} (-1)^k \cos(2k+\gamma)x \\ &= (-1)^m \sec x \sin(2m+\gamma+1)x - \sec x \tan(2m+\gamma+1)x \cos(\gamma+1)x. \end{aligned}$$

On account of (19), we can extract, via (2), (3) and (5), the coefficients of x^{2n-1} across this equation. Simplifying the result gives the following identity.

Theorem 53 ($m \geq 0$ and $n \geq 0$).

$$\begin{aligned} &2\sum_{k=1}^{m} \sum_{i=0}^{n+1} (-1)^k \binom{2n+2}{2i} 4^i (2m+\gamma+1)^{2i-1} (2k+\gamma)^{2n+2-2i} B'_{2i} \\ &= (2n+2) \sum_{i=0}^{n} (-1)^{m+1} \binom{2n+1}{2i} (2m+\gamma+1)^{2n+1-2i} E_{2i} \\ &\quad - \sum_{0 \leq i+j \leq n+1} 4^i \binom{2n+2}{2i, 2j} (2m+\gamma+1)^{2i-1} (\gamma+1)^{2n+2-2i-2j} B'_{2i} E_{2j}. \end{aligned}$$

When $m = -\gamma = 1$, this theorem reduced a simplified transformation.

Corollary 54 ($m = 1$ and $\gamma = -1$ in Theorem 53).

$$\sum_{i=0}^{n} 4^{2i} \binom{2n}{2i} B'_{2i} E_{2n-2i} = 8n + 2 \sum_{i=0}^{n} 4^{2i} \binom{2n}{2i} B'_{2i}.$$

4.5 – Finally, reformulate (18) equivalently as the following equality

$$\begin{aligned}&2\tan(2m+\gamma+1)x\sum_{k=1}^{m}(-1)^k\sin x\cos(2k+\gamma)x\\&=(-1)^m\tan x\sin(2m+\gamma+1)x+\tan x\tan(2m+\gamma+1)x\cos(\gamma+1)x.\end{aligned}$$

Similarly with the help of (3) and (5), extracting the coefficient of x^{2n} across the last equation, we establish the following identity.

THEOREM 55 ($m \geq 0$ and $n \geq 0$).

$$\begin{aligned}&\sum_{k=1}^{m}\sum_{i=0}^{n}(-1)^k\binom{2n+1}{2i}4^i\frac{(2m+\gamma+1)^{2i-1}}{(\gamma+1)^{2n+2}}B'_{2i}\left\{\begin{matrix}(2k+\gamma+1)^{2n-2i+1}\\-(2k+\gamma-1)^{2n-2i+1}\end{matrix}\right\}\\&=\sum_{i=0}^{n}(-1)^m\binom{2n+1}{2i}4^i\frac{(2m+\gamma+1)^{2n-2i+1}}{(\gamma+1)^{2n+2}}B'_{2i}\\&\quad+\sum_{0\leq i+j\leq n+1}4^{i+j}\binom{2n+2}{2i,2j}\frac{(2m+\gamma+1)^{2j-1}}{(2n+2)(\gamma+1)^{2i+2j}}B'_{2i}B'_{2j}.\end{aligned}$$

When $m=\gamma=0$, it yields the following strange identity

$$\sum_{0\leq i+j\leq n+1}4^{i+j}\binom{2n+2}{2i,2j}B'_{2i}B'_{2j}=-(2n+2)\sum_{i=0}^{n}4^i\binom{2n+1}{2i}B'_{2i}.$$

By extracting the coefficients of x^{2n} across the expansion of the trigonometric relation

$$\tan^2 x\cos x=\tan x\sin x=\sec x-\cos x,$$

we have the following two convolution formulae.

COROLLARY 56.

$$\sum_{i=0}^{n}4^i\binom{2n+1}{2i}B'_{2i}=(2n+1)(E_{2n}-1).$$

COROLLARY 57 ($n>0$).

$$\sum_{0\leq i+j\leq n}4^{i+j}\binom{2n}{2i,2j}B'_{2i}B'_{2j}=2n(2n-1)(1-E_{2n-2}).$$

Taking $m = -\gamma = 1$ in Theorem 55 and then replacing n by $n-1$, we have the transformation formula

$$\sum_{i=0}^{n} \frac{4^n}{4^i} \binom{2n}{2i} B'_{2i} B'_{2n-2i} = n \sum_{i=0}^{n} \binom{2n-1}{2i} (2-4^i) B'_{2i}.$$

Evaluating the last sum by Corollary 39 and Corollary 56, we derive further the following convolution identity.

Corollary 58 ($m = 1$ and $m = -1$ in Theorem 55: $n > 1$).

$$\sum_{i=0}^{n} \frac{4^n}{4^i} \binom{2n}{2i} B'_{2i} B'_{2n-2i} = n(1-2n) E_{2n-2}.$$

REFERENCES

[1] T. Agoh – K. Dilcher: *Convolution identities and lacunary recurrences for Bernoulli numbers*, J. Number Theory, **124** (2007), 105–122.

[2] G. S. Cheon: *A note on the Bernoulli and Euler polynomials*, Appl. Math. Lett., **16** (2003), 365–368.

[3] W. Chu – C. Y. Wang: *Arithmetic identities involving Bernoulli and Euler numbers*, Resultate der Mathematik, **55** (2009), 65–77.

[4] W. Chu – C. Y. Wang: *Convolution formulae for Bernoulli numbers*, Integral Transforms and Special Functions **21** (2010), 437–457.

[5] L. Comtet: *Advanced Combinatorics*, Dordrecht-Holland, The Netherlands,1974.

[6] K. Dilcher: *Sums of products of Bernoulli numbers*, J. Number Theory, **60** (1996), 23–41.

[7] I. M. Gessel: *On Miki's identity for Bernoulli numbers*, J. Number Theory, **110** (2005), 75–82.

[8] R. L. Graham – D. E. Knuth – O. Patashnik: *Concrete Mathematics* (second edition, Addison-Wesley Publ. Company, Reading, MA, 1994.

[9] E. R. Hansen: *A Table of Seriefs and Products*, Prentice-Hall Inc., Englewood Cliffs, New Jersey, 1975.

[10] G. D. Liu – H. Luo: *Some identities involving Bernoulli numbers*, Fibonacci Quart., **43** (2005), 208–212.

[11] H. Miki: *A relation between Bernoulli numbers*, J. Number Theory, **10** (1978), 297–302.

[12] K. H. Rosen: *Handbook of Discrete and Combinatorial Mathematics*, CRC Press, Boca Raton, 2000.

[13] H. M. Srivastava – A. Pinter: *Remarks on some relationships between the Bernoulli and Euler polynomials*, Appl. Math. Lett., **17** (2004), 375–380.

[14] K. R. Stromberg: *Introduction to Classical Real Analysis*, Wadsworth INC. Belmont, California, 1981.

[15] H. S. Wilf: *Generatingfunctionology* (second edition), Academic Press, Inc., Boston, MA, 1994.

Lavoro pervenuto alla redazione il 18 aprile 2010
ed accettato per la pubblicazione il 1 luglio 2010.
Bozze licenziate il 26 novembre 2010

Indirizzo degli Autori:

Wenchang Chu – Dipartimento di Matematica – Università del Salento – Lecce – Arnesano P. O. Box 193 – 73100 Lecce, Italia
E-mail: chu.wenchang@unile.it

Chenying Wang – College of Mathematics and Physics – Nanjing University of Information Science and Technology – Nanjing 210044, P. R. China
E-mail: wang.chenying@163.com

Rendiconti di Matematica, Serie VII
Volume 30, Roma (2010), 275–283

Some landmarks in the history of the tangential Cauchy Riemann equations

R. MICHAEL RANGE

ABSTRACT: *We discuss the origins of the tangential Cauchy Riemann equation beginning with W. Wirtinger in 1926, and trace the largely unknown early developments until the emergence of the* $\overline{\partial}_b-$ *Neumann complex in the 1960s.*

Vienna is a most appropriate venue for a program centered on the $\overline{\partial}-$ Neumann Problem. Not only did the calculus of the differential operators $\partial/\partial z_j$ and $\partial/\partial\overline{z}_j$ originate in the work of Wilhelm Wirtinger, Professor at the University of Vienna, but to my knowledge Wirtinger also was the first person to have thought of what today we call the *tangential Cauchy Riemann equations* and the corresponding notion of *(tangential) Cauchy-Riemann* ($= CR$) *functions*. Since much of the modern literature seems to be unaware of this work and of other early work on "tangential analytic functions", it may be useful to trace the path from these origins to the modern theory of the tangential $\overline{\partial}-$Neumann Complex as developed by J. J. Kohn and H. Rossi in the 1960s.

KEY WORDS AND PHRASES: *Tangential Cauchy-Riemann functions – CR Extension – Global CR Extension theorem – Local CR embedding problem*
A.M.S. CLASSIFICATION: 3503, 32V25, 0102, 35N15, 5F35.
This paper is based on a lecture given during the program "The $\overline{\partial}-$Neumann Problem: Analysis, Geometry, and Potential Theory." held at the Erwin Schrödinger International Institute for Mathematical Physics in Vienna in Fall 2009. The author gratefully acknowledges the support of the ESI during his stay in Vienna.

1 – The Beginning

Wilhelm Wirtinger (1865 - 1945) was born in Ybbs on the Danube and studied mathematics at the Universität Wien. He earned his doctorate in 1887 with Emil Weyr and Gustav Ritter von Escherich, working on triple evolutions in the plane. For the next three years he expanded his mathematical horizons in Berlin and Göttingen, where he was strongly influenced by F. Klein. In 1890 he earned the Habilitation in Vienna, and after a few years as assistant he was appointed to a chair at the University of Innsbruck in 1895. He returned to Vienna in 1905 to assume a chair at his alma mater, where he stayed until his retirement in 1935. Wirtinger was productive across a broad spectrum of mathematics and mathematical physics, ranging from complex analysis and number theory to relativity theory and capillary waves. He was well recognized internationally as one of the leading mathematicians of his days. Among his nine doctoral students are W. Blaschke (1908, Wien) and L. Vietoris (1920, Wien). Other well known mathematicians such as Schreier, Gödel, Radon, and Tausky-Todd studied with him.

Most relevant for the present discussion is Wirtinger's 1926 paper *Zur formalen Theorie der Funktionen von mehr komplexen Veränderlichen* [Wir]. Starting with a (smooth) function $F(x_1, ..., x_{2n})$ of the real variables x_β, $\beta = 1, ..., 2n$, Wirtinger introduces the complex functions $z_j = x_{2j-1} + ix_{2j}$ and their conjugates $\overline{z_j}$, $j = 1, ..., n$ and thinks of F as a function of the z_j and $\overline{z_j}$ via $x_{2j-1} = \frac{1}{2}(z_j + \overline{z_j})$ and $x_{2j} = \frac{1}{2i}(z_j - \overline{z_j})$. Formal application of the chain rule leads to

$$\frac{\partial}{\partial z_j} = \frac{1}{2}\left(\frac{\partial}{\partial x_{2j-1}} + \frac{1}{i}\frac{\partial}{\partial x_{2j}}\right) \text{ and } \frac{\partial}{\partial \overline{z_j}} = \frac{1}{2}\left(\frac{\partial}{\partial x_{2j-1}} - \frac{1}{i}\frac{\partial}{\partial x_{2j}}\right).$$

F is then an *analytic* function of $z_1, ..., z_n$ precisely when F satisfies the Cauchy-Riemann equations

$$\frac{\partial F}{\partial \overline{z_j}} = 0, \text{ or, equivalently, } \frac{\partial \overline{F}}{\partial z_j} = 0,\ j = 1, ..., n.$$

As a first elegant application of this point of view, Wirtinger notes that if W is the real part of an analytic function, or more generally, a linear combination $aF + b\overline{G}$, where F and G are analytic in $z_1, ..., z_n$, then obviously

$$\frac{\partial^2 W}{\partial z_j \partial \overline{z_k}} = 0 \text{ for all } j, k = 1, ..., n.$$

Conversely, if W satisfies these equations, W must be such a linear combination, at least locally. In fact, the $1-$form $\omega_1 = \sum \frac{\partial W}{\partial z_j} dz_j$ has analytic coefficients and is clearly closed, hence it is (locally) the differential dF of a function $F = \int \omega_1$

which is analytic in $z_1, ..., z_n$. Similarly, $\omega_2 = \sum \frac{\partial W}{\partial \overline{z_j}} d\overline{z_j}$ is the differential $d\overline{G}$ of a function $\overline{G} = \int \omega_2$ which depends analytically on $\overline{z_1}, ..., \overline{z_n}$, i.e., G depends analytically on $z_1, ..., z_n$. Since $dW = \omega_1 + \omega_2 = d(F + \overline{G})$, W and $F + \overline{G}$ differ by a constant. In particular, if W is *real* valued, then $\omega_2 = \overline{\omega_1}$, and hence $G = F$, so W is the real part of an analytic function.

Influenced by Riemann's point of view, who considered functions $f(x, y)$ on a $2-$ dimensional manifold which are analytic in $z = x + iy$, in the sense that their differential df is just a multiple of dz, Wirtinger generalizes this idea to the setting of an $m-$ dimensional manifold M_m, with (real) coordinates $t = (t_1, ..., t_m)$. Given a positive integer n, with $n < m \leq 2n$, he introduces $2n$ real functions $x_\gamma(t),\ y_\gamma(t),\ 1 \leq \gamma \leq n$ on M_m, and the corresponding complex valued functions $z_\gamma = x_\gamma + iy_\gamma$, subject to the nondegeneracy condition

$$rank \begin{bmatrix} \frac{\partial z_\gamma}{\partial t_\lambda} & \frac{\partial \overline{z_\gamma}}{\partial t_\lambda} \end{bmatrix}_{m \times 2n} = m,$$

so that the points on M_m are uniquely determined by the values of the z_j and $\overline{z_j}$. Furthermore, the functions $z_1(t), ..., z_n(t)$ are assumed to be independent, i.e., $dz_1 \wedge .... \wedge dz_n \neq 0$. Wirtinger then introduces the concept of a *complex valued function* $\Phi(t)$ *on* M_m *which depends on* $z_1, ..., z_n$ (" ... *eine Funktion* $\Phi(t)$ [welche] *als Funktion der* z_γ ... *dargestellt werden kann*" [Wir, p. 364]) by the condition that

$$rank \begin{bmatrix} \frac{\partial \Phi}{t_\lambda} \\ \frac{\partial z_\gamma}{\partial t_\lambda} \end{bmatrix} < n + 1.$$

In the language of differential forms, this means that $d\Phi \wedge dz_1 \wedge .... \wedge dz_n \equiv 0$ on M_m, or $d\Phi = \sum a_\gamma(t)\, dz_\gamma$. Equivalently, the partial derivatives $\frac{\partial \Phi}{\partial t_\lambda}$, $\lambda = 1, ..., m$, must satisfy a system of $m - n > 0$ linear equations, i.e., there exist linear differential operators $X_k = \sum X_k^\lambda \frac{\partial}{\partial t_\lambda}$, $k = 1, ..., m - n$, with complex valued coefficients X_k^λ on M_m, such that Φ is "analytic in $z_1, ..., z_n$" if and only if

$$X_k(\Phi) = \sum_{\lambda=1}^{m} X_k^\lambda \frac{\partial \Phi}{\partial t_\lambda} = 0,\ k = 1, ..., m - n.$$

This system so generated has two basic properties:

a) The only *real* valued solutions are constants.
b) *span* $\{X_1, ..., X_{m-n}\}$ is closed under Lie brackets.

Conversely, starting with such a system which satisfies a) and b), Wirtinger notes that if there exist n independent solutions $z_1, ..., z_n$, then all solutions of $X_k(\Phi) = 0,\ k = 1, ..., m - n$, on M_m can be thought of as analytic functions of these independent solutions. The key problem thus involves proving the existence of such independent solutions. Motivated by the classical case $m = 2$, $n = 1$, which Riemann studied by means of extremal properties, i.e, via the Dirichlet problem, Wirtinger first attempts to introduce appropriate variational problems and integral invariants in the higher dimensional case $n > 1$, $m = 2n$. In modern language, Wirtinger was considering an integrable almost complex structure on M_{2n}, and he was trying to extend Riemann's methods to prove that the given data defined a complex manifold M_{2n}. But he quickly realized that *"bis zu bestimmten Existenzsätzen noch ein weiter Weg ist.*[(1)] ([Wir], p. 372). He surely was right: it took over 30 years until the problem was eventually solved by A. Newlander and L. Nirenberg ([NeNi]).

Wirtinger then turned to the case $m < 2n$ and set up some explicit computations in the case $m = 3$, $n = 2$, thereby attempting to outline a strategy to solve what eventually became known as the *(local) embedding problem for abstract CR -structures.* He seemed prescient, as he stated that such investigations, *if they can be carried out at all*, would be much more difficult and complicated (op. cit, p. 375). In fact, moving ahead half a century, L. Nirenberg showed in 1974 that there is in general no solution in this particular dimension, even assuming a definite Levi form [Nir]. On the other hand, in a remarkable tour de force, M. Kuranishi [Kur] proved in 1982 that the answer is positive in the hypersurface case $m = 2n - 1$ with definite Levi form, provided $n \geq 5$. Subsequently, T. Akahori was able to extend Kuranishi's work to the case $n = 4$ [Aka]. The case $n = 3$ (i.e. $m = 5$) remains open to this date. Wirtinger's intuition thus was remarkably accurate. Realizing these difficulties with continuing along the path initiated by Riemann, Wirtinger ended his paper with the statement *"Vielleicht hätte Riemann auch Ideen zur Überwindung dieser Schwierigkeiten gehabt."*[(2)]

2 – Early *CR* Extension Results

As noted above, there was no progress for a long time regarding the deep question about *existence* of solutions to the system of partial differential equations introduced by Wirtinger. However, Wirtinger's idea of "analytic functions of the complex variables $z_1, ..., z_n$" on a real manifold led to other important developments. Remarkable results were obtained just a few years after Wirtinger's paper in the concrete setting in which the real manifold M_{2n-1} is a submanifold of $\mathbb{C}^n$, where the complex coordinates $z_1, ..., z_n$ trivially provide n independent solutions of Wirtinger's system. Here a most natural question is to examine the relationship between functions analytic on M_{2n-1} in Wirtinger's sense, and the functions analytic in $z_1, ..., z_n$ in the ambient space in the classical sense.

(1) *The path to specific existence theorems is still long.*

(2) *Perhaps Riemann would also have had ideas to overcome these difficulties.*

Clearly restrictions to M_{2n-1} of such classical analytic functions, as well as suitable boundary values of such functions defined on only one side of M_{2n-1} , are solutions of the corresponding Wirtinger system. The obvious question then is whether all solutions are, essentially, of this type.

Making reference to Wirtinger's 1926 paper, Francesco Severi gave an affirmative answer in 1931 in the real analytic category [Sev].

THEOREM (SEVERI 1931). *If $M_{2n-1} \subset \mathbb{C}^n$ is real analytic, and if f is a real analytic function on M_{2n-1} which satisfies the Wirtinger condition $df \wedge dz_1 \wedge \ldots \wedge dz_n = 0$ in a neighborhood of a point $P \in M_{2n-1}$, then there exists a function $F \in \mathcal{O}(U)$ on an open neighborhood U of P in $\mathbb{C}^n$, such that $F|_{M_{2n-1} \cap U} = f$.*

The proof, which is essentially trivial in case $n = 1$, involves an elegant application of Severi's method to pass from real to complex variables in appropriate power series. Severi proved the theorem in case $n = 2$, but his proof works in higher dimensions as well with the obvious modifications. Via the identity theorem, the result is easily globalized. By applying the classical Hartogs extension theorem, Severi thus obtains the following generalization of the Hartogs theorem to the case of Wirtinger's tangential analytic functions.

GLOBAL *CR* EXTENSION THEOREM, REAL ANALYTIC CASE .*If $n > 1$ and the bounded region $D \subset \mathbb{C}^n$ has connected real analytic boundary bD, then any real analytic function f which satisfies $df \wedge dz_1 \wedge \ldots \wedge dz_n = 0$ on bD has a holomorphic extension to D.*

The *local* extension theory in the *differentiable* case is considerably more complicated. Apparently unaware of Severi's work, in 1936 Helmuth Kneser studied the problem on M_3 in $\mathbb{C}^2$ and produced examples to show that differentiable functions satisfying the Wirtinger condition are not necessarily the boundary values of classical holomorphic functions [Kne]. In fact, Kneser considered a generalization of Wirtinger's differential condition on $M_3 = M$ to a Morera type condition (A) for *continuous* functions f, as follows. *A continuous function f satisfies condition (A) on the 3-dimensional manifold M if $\int_{bG} f\, dz_1 \wedge dz_2 = 0$ for every subregion $G \subset M$ with C^1 boundary bG.* Kneser showed that for f of class C^1, condition (A) is equivalent to Wirtinger's differential condition. More significantly, in analogy to the E. E. Levi extension phenomenon for holomorphic functions, Kneser proved a deep *local* one-sided extension result for such continuous *CR* functions near a strictly Levi pseudoconvex boundary point, as follows.

THEOREM (KNESER 1936). *Assume that $P \in bD$ and that D is strictly Levi pseudoconvex at P.*[(3)] *Then there exist neighborhoods $V \subset\subset U$ of P, such*

(3) This implies in particular that bD is of class C^2 near P.

that every continuous function f on $U \cap bD$ which satisfies condition (A) can be extended continuously to a function holomorphic on $V \cap D$.

In the proof, Kneser first showed that the geometric hypothesis implied that after a local holomorphic change of coordinates one could assume that the boundary was strictly Euclidean convex.(4) In this geometrically simple setting Kneser then produced the holomorphic extension via an explicit integral formula which was a suitably adapted variant of the Cauchy integral formula for polydiscs. Condition (A) is the critical ingredient that makes the proof work.

Incidentally, just as Severi had done in the real analytic case, Kneser also proved the corresponding *global* version.

Global CR Extension Theorem, Strictly Pseudoconvex Case. *If the bounded region $D \subset \mathbb{C}^2$ has connected strictly pseudoconvex boundary bD then any continuous (weakly) CR function f on bD has a holomorphic extension to D.*

To my knowledge, Kneser's result is the the first global CR extension theorem in the differentiable category, albeit under some restrictive geometric conditions.

Unfortunately, the phenomenal progress in global complex function theory in higher dimensions achieved by K. Oka and H. Cartan beginning in the mid 1930s, as well as the political climate in Germany and the disruptions of the second world war, relegated the investigations begun by Wirtinger, Severi, and Kneser to the sidelines, to the extent that for all practical purposes they were forgotten and did not get proper recognition for a long time.

3 – Results of Lewy and Fichera in the 1950s

In the early 1950s there was renewed interest in fundamental investigations in the theory of partial differential equations. One major result of this period was the proof of existence of fundamental solutions for every linear partial differential operator with *constant* coefficients, obtained independently by L. Ehrenpreis and B. Malgrange. Furthermore, the multivariable classical Cauchy-Riemann equations presented a central example of an *overdetermined* system which required new methods for its study. Lastly, the system of linear partial differential equations introduced by Wirtinger in 1926 provided natural important classes of examples which were not covered by the Ehrenpreis - Malgrange theory. Although it is not clear how much Wirtinger's ideas were known in those days,

(4) This seems to be the earliest explicit occurence of what has become a well known standard tool.

times were certainly ripe for studying such more general equations. In particular, Hans Lewy, who had earned his doctorate in Göttingen with R. Courant in 1925 and had moved to the University of California in Berkeley in 1935 after he was forced to emigrate from Germany, began to investigate a linear differential equation with non-constant coefficients which is equivalent to Wirtinger's equation for "analytic" functions on submanifolds in the case of a 3-dimensional submanifold in $\mathbb{C}^2$ [Lew 1]. While Lewy did not mention Wirtinger's name in his paper, he made explicit reference to Severi's 1931 extension theorem in the real analytic case, and thus it is most likely that he also knew Wirtinger's work, which is prominently cited in Severi's paper. Unaware of Hellmuth Kneser's 1936 work, Lewy proved Kneser's extension theorem for continously differentiable CR functions in the strictly pseudoconvex case. Shortly thereafter, Lewy used an explicit example of the equation studied earlier to produce the first example - and at that time quite unexpected - of a smooth first order complex linear partial differential equation in 3 real variables without any solutions [Lew 2]. Lewy's results generated much interest and became widely known.

These developments probably contributed to overshadowing the remarkable extension result for CR functions obtained in Italy by G. Fichera [Fic] around the same time. Motivated by Severi, Fichera showed in 1957 that Severi's *"global CR extension theorem"* (i.e., the CR version of Hartogs' Theorem) remained true *without* assuming real analyticity and without any geometric restrictions. Fichera's proof, based on the solution of the Dirichlet problem, required the given data to be of class $C^{1+\varepsilon}$. His work subsequently inspired E. Martinelli to modify his 1942 integral formula proof of the classical Hartogs Theorem to produce a simple proof of the Severi-Fichera global extension result in the C^1 category [Mar].(5) However, these results about the *global* CR extension problem remained virtually unrecognized outside of Italy for a long time.

4 – The modern theory

The global CR extension theorem came to the forefront in 1965, when J.J. Kohn and H. Rossi, inspired by Lewy's local extension theorem, introduced tangential (p, q) forms and the $\overline{\partial}_b-$complex on smooth boundaries of domains in complex manifolds [KoRo]. By using Kohn's then new deep regularity results for the $\overline{\partial}-$Neumann problem, Kohn and Rossi proved the holomorphic extension of C^∞ global CR functions from the connected boundary of domains in Stein manifolds, assuming that the Levi form has at least one positive eigenvalue at each point on the boundary. They also proved corresponding extension results for $\overline{\partial}_b-$ closed forms in higher degree. Their work marks the beginning of the *modern* theory of tangential CR functions and forms, either incorporated in the

(5) More recently, this author used the Bochner-Martinelli kernel to give a simple proof of the *local* Kneser-Lewy extension theorem for continuous (weak) CR functions [Ran 2].

$\overline{\partial}_b-$complex and in the theory of the $\overline{\partial}_b-$Neumann problem, or as the principal object of study in numerous settings. The reader may consult the recent monographs by M. S. Baouendi, P. Ebenfelt, and L. Rothschild [BER] and So-Chin Chen and Mei-Chi Shaw [ChSh] for an overview of many developments since then.

Unfortunately, Kohn and Rossi were apparently unaware of the earlier work on the global CR extension theorem by Severi, Fichera, and Martinelli. Furthermore, a remark in the introduction of their 1965 paper connected the global CR extension theorem to S. Bochner's 1943 proof of the classical Hartogs extension theorem. Shortly thereafter this linkage led L. Hörmander to crediting Bochner with the proof of the global CR extension theorem in his well known 1966 monograph. This erroneous attribution became widely accepted since then, even though there is no evidence in the published record that Bochner stated and proved such a theorem, nor that he had even been thinking about tangential CR functions. The historical record was eventually corrected beginning in 1999, when this author learned of the long forgotten 1936 paper of H. Kneser. The interested reader should consult [Ran 1, 3] for more details.

REFERENCES

[Aka] T. Akahori: *A New Approach to the Local Embedding Theorem of CR−Structures for $n \geq 4$*, Mem. of the Amer. Math. Soc. **366**, Providence, R. I., 1987.

[BER] M. S. Baouendi, P. Ebenfelt, and L. P Rothschild: *Real Submanifolds in Complex Space and their Mappings*, Princeton Math. Ser. **47**, Princeton U. Press, Princeton, N. J. 1999.

[ChSh] S. C. Chen and M. C. Shaw: *Partial Differential Equations in Several Complex Variables.*, Amer. Math. Soc. and Int. Press, Providence, R. I., 2001.

[Fic] G. Fichera: *Caratterizazione della traccia, sulla frontiera di un campo, di una funzione analitica di più variabili complesse*, Rend. Acc. Naz. Lincei VII, **23** (1957), 706–715.

[Kne] H. Kneser: *Die Randwerte einer analytischen Funktion zweier Veränderlichen*, Monatsh. für Math. u. Phys. **43** (1936), 364–380.

[Kur] M. Kuranishi: *Strongly pseudoconvex CR structures over small balls*, Ann. Math. **115** and **116** (1982), 451-500, 1-64, and 249–330.

[Lew 1] H. Lewy: *On the local character of the solutions of an atypical linear differential equation in three variables and a related theorem for regular functions of two complex variables*, Ann. Math. **64** (1956), 514–522.

[Lew 2] H. Lewy: *An example of a smooth linear partial differential equation without solutions*, Ann. Math. **66** (1957), 155–158.

[Mar] E. Martinelli: *Sulla determinazione di una funzione analitica di più variabili complesse in un campo, assegnatane la traccia sulla frontiera*, Ann. Mat. Pura e Appl. **55** (1961), 191–202.

[NeNi] A. Newlander and L. Nirenberg: *Complex coordinates in almost complex manifolds*, Ann. Math. **64** (1957), 391–404.

[Nir] L. Nirenberg: *On a question of Hans Lewy*, Russian Math. Surv. **29** (1974), 251–262.

[KoRo] J. J. Kohn and H. Rossi: *On the extension of holomorphic functions from the boundary of a complex manifold*, Ann. Math. **81** (1965), 451–472.

[Ran 1] R. M. Range: *Extension Phenomena in Multidimensional Complex Analysis: Correction of the Historical Record*, Math. Intell. **24** (2002), 4–12.

[Ran 2] R. M. Range: *On the decomposition of holomorphic functions by integrals and the local CR extension theorem*, Adv. Studies in Pure Math. **42** (2004), 269–273.

[Ran 3] R. M. Range: *Kneser's paper on the boundary values of analytic functions in two variables*, In: Hellmuth Kneser, Gesammelte Abhandlungen, G. Betsch and K. Hofmann, Eds., 872–876, De Gruyter, Berlin 2005.

[Sev] F. Severi: *Risoluzione generale del problema di Dirichlet per le funzioni biarmoniche*, Rend. Reale Accad. Lincei **23** (1931), 795–804.

[Wir] W. Wirtinger: *Zur formalen Theorie der Funktionen von mehr komplexen Veränderlichen*, Math. Ann. **97** (1926), 357–375.

Lavoro pervenuto alla redazione il 6 luglio 2010
ed accettato per la pubblicazione il 19 luglio 2010.
Bozze licenziate il 26 novembre 2010

Indirizzo dell'Autore:

R. Michael Range– Department of Mathematics and Statistics – State University of New York at Albany – Albany, New York 12222, USA
E-mail: range@math.albany.edu

Rendiconti di Matematica, Serie VII
Volume 30, Roma (2010), 285–327

Dualities in convex algebraic geometry

PHILIPP ROSTALSKI – BERND STURMFELS

Abstract: *Convex algebraic geometry concerns the interplay between optimization theory and real algebraic geometry. Its objects of study include convex semialgebraic sets that arise in semidefinite programming and from sums of squares. This article compares three notions of duality that are relevant in these contexts: duality of convex bodies, duality of projective varieties, and the Karush-Kuhn-Tucker conditions derived from Lagrange duality. We show that the optimal value of a polynomial program is an algebraic function whose minimal polynomial is expressed by the hypersurface projectively dual to the constraint set. We give an exposition of recent results on the boundary structure of the convex hull of a compact variety, we contrast this to Lasserre's representation as a spectrahedral shadow, and we explore the geometric underpinnings of semidefinite programming duality.*

1 – Introduction

Dualities are ubiquitous in mathematics and its applications. This article compares several notions of duality that are relevant for the interplay between convexity, optimization, and algebraic geometry. It is primarily expository, and is intended for a diverse audience, ranging from graduate students in mathematics to practitioners of optimization who are based in engineering.

Duality for vector spaces lies at the heart of linear algebra and functional analysis. Duality in convex geometry is an involution on the set of convex bodies: for instance, it maps the cube to the octahedron and vice versa (Figure 1). Duality in optimization, known as *Lagrange duality*, plays a key role in designing

Key Words and Phrases: *Optimization – Duality – Semidefinite programming – Spectrahedron – Convexity – Real algebraic geometry*
A.M.S. Classification: 90C22, 14P05, 52A05.

efficient algorithms for the solution of various optimization problems. In projective geometry, points are dual to hyperplanes, and this leads to a natural notion of *projective duality* for algebraic varieties.

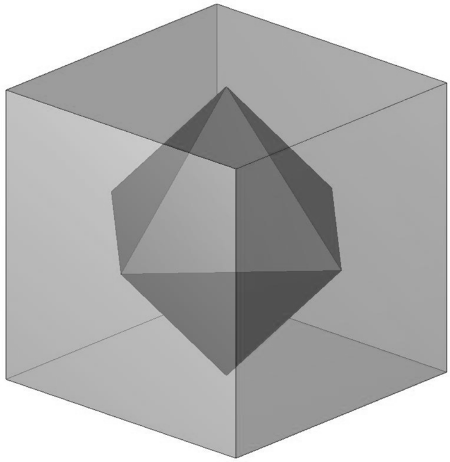

Fig. 1: The cube is dual to the octahedron.

Our aim here is to explore these dualities and their interconnections in the context of polynomial optimization and semidefinite programming. Towards the end of the Introduction, we shall discuss the context and organization of this paper. At this point, however, we jump right in and present a concrete three-dimensional example that illustrates our perspective on these topics.

1.1 – How to Dualize a Pillow

We consider the following symmetric matrix with three indeterminate entries:

$$Q(x,y,z) = \begin{pmatrix} 1 & x & 0 & x \\ x & 1 & y & 0 \\ 0 & y & 1 & z \\ x & 0 & z & 1 \end{pmatrix}. \tag{1.1}$$

This symmetric 4×4-matrix specifies a 3-dimensional compact convex body

$$P = \left\{ (x,y,z) \in \mathbb{R}^3 \mid Q(x,y,z) \succeq 0 \right\}. \tag{1.2}$$

The notation "$\succeq 0$" means that the matrix is *positive semidefinite*, i.e., all four eigenvalues are non-negative real numbers. Such a *linear matrix inequality* always defines a closed convex set which is referred to as a *spectrahedron*.

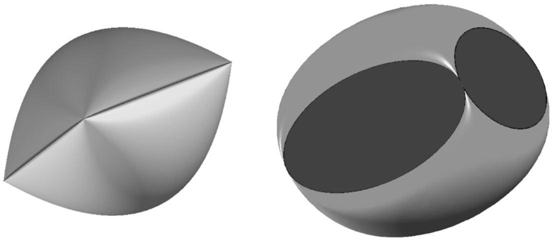

Fig. 2: A 3-dimensional spectrahedron P and its dual convex body P^{Δ}.

Our spectrahedron P looks like a pillow. It is shown on the left in Figure 2. The *algebraic boundary* of P is the surface specified by the determinant

$$\det(Q(x,y,z)) \quad = \quad x^2(y-z)^2 - 2x^2 - y^2 - z^2 + 1 \quad = \quad 0.$$

The interior of P represents all matrices $Q(x,y,z)$ whose four eigenvalues are positive. At all smooth points on the boundary of P, precisely one eigenvalue vanishes, and the rank of the matrix $Q(x,y,z)$ drops from 4 to 3. However, the rank drops further to 2 at the four singular points

$$(x,y,z) \;=\; \frac{1}{\sqrt{2}}(1,1,-1),\ \frac{1}{\sqrt{2}}(-1,-1,1),\ \frac{1}{\sqrt{2}}(1,-1,1),\ \frac{1}{\sqrt{2}}(-1,1,-1).$$

We find these from a *Gröbner basis* of the ideal of 3×3-minors of $Q(x,y,z)$:

$$\big\{\, 2x^2-1,\, 2z^2-1,\, y+z \,\big\}.$$

The linear polynomial $y+z$ in this Gröbner basis defines the symmetry plane of the pillow P. The four corners form a square in that plane. Its edges are also edges of P. All other faces of P are exposed points. These come in two families, called *protrusions*, one above the plane $y+z=0$ and one below it.

Like all convex bodies, our pillow P has an associated *dual convex body*

$$P^{\Delta} = \big\{\, (a,b,c) \in \mathbb{R}^3 \mid ax+by+cz \le 1 \ \text{ for all } (x,y,z) \in P \,\big\}, \qquad (1.3)$$

consisting of all linear forms that evaluate to at most one on P. Our notation P^{Δ} is chosen to be consistent with that in Ziegler's text book [29, §2.3].

The dual pillow P^{Δ} is shown on the right in Figure 2. Note the association of faces under duality. The pillow P has four 1-dimensional faces, four singular 0-dimensional faces, and two smooth families of 0-dimensional faces. The corresponding dual faces of P^{Δ} have dimensions 0, 2 and 0 respectively.

Semidefinite programming is the computational problem of minimizing a linear function over a spectrahedron. For our pillow P, this takes the form

$$p^*(a,b,c) = \underset{(x,y,z)\in\mathbb{R}^3}{\text{Maximize}} \;\; ax+by+cz$$
$$\text{subject to} \quad Q(x,y,z) \succeq 0. \tag{1.4}$$

We regard this as a parametric optimization problem: we are interested in the optimal value and optimal solution of (1.4) as a function of $(a,b,c) \in \mathbb{R}^3$. This function can be expressed in terms of the dual body P^Δ as follows:

$$p^*(a,b,c) = \underset{\lambda\in\mathbb{R}}{\text{Minimize}} \lambda$$
$$\text{subject to} \quad \frac{1}{\lambda}\cdot(a,b,c) \in P^\Delta. \tag{1.5}$$

We distinguish this formulation from the duality in semidefinite programming. The *dual* to (1.4) is the following program with 7 decision variables:

$$d^*(a,b,c) \;=\; \underset{u\in\mathbb{R}^7}{\text{Minimize}}\, u_1+u_4+u_6+u_7$$
$$\text{subject to} \quad \begin{pmatrix} 2u_1 & 2u_2 & u_3 & -2u_2-a \\ 2u_2 & 2u_4 & -b & 2u_5 \\ 2u_3 & -b & 2u_6 & -c \\ -2u_2-a & 2u_5 & -c & 2u_7 \end{pmatrix} \succeq 0. \tag{1.6}$$

Since (1.4) and (1.6) are both strictly feasible, strong duality holds [4, §5.2.3], i.e. the two programs attain the same optimal value: $p^*(a,b,c) = d^*(a,b,c)$. Hence, problem (1.6) can be derived from (1.5), as we shall see in Section 5.

We write $M(u;a,b,c)$ for the 4×4-matrix in (1.6). The following equations and inequalities, known as the *Karush-Kuhn-Tucker conditions* (KKT), are necessary and sufficient for any pair of optimal solutions:

$$\begin{aligned} Q(x,y,z)\cdot M(u;a,b,c) \;&=\; 0, \qquad \text{(complementary slackness)} \\ Q(x,y,z) \;&\succeq\; 0, \\ M(u;a,b,c) \;&\succeq\; 0. \end{aligned}$$

We relax the inequality constraints and consider the system of equations

$$\lambda = ax+by+cz \quad \text{and} \quad Q(x,y,z)\cdot M(u;a,b,c) \;=\; 0.$$

This is a system of 11 equations. Using computer algebra, we eliminate the 10 unknowns $x,y,z,u_1,\ldots,u_7$. The result is a polynomial in a,b,c and λ. Its factors, shown in (1.7)-(1.8), express the optimal value λ^* in terms of a,b,c.

At the optimal solution, the product of the two 4×4-matrices $Q(x,y,z)$ and $M(u;a,b,c)$ is zero, and their respective ranks are either $(3,1)$ or $(2,2)$. In the former case the optimal value λ^* is one of the two solutions of

$$(1.7)\quad (b^2+2bc+c^2)\cdot\lambda^2 - a^2b^2 - a^2c^2 - b^4 - 2b^2c^2 - 2bc^3 - c^4 - 2b^3c \;=\; 0.$$

In the latter case it comes from the four corners of the pillow, and it satisfies

$$(1.8)\qquad \begin{array}{rl} & (2\lambda^2 - a^2 + 2ab - b^2 + 2bc - c^2 - 2ac) \\ \cdot & (2\lambda^2 - a^2 - 2ab - b^2 + 2bc - c^2 + 2ac) \quad = \quad 0. \end{array}$$

These two equations describe the algebraic boundary of the dual body P^Δ. Namely, after setting $\lambda = 1$, the irreducible polynomial in (1.7) describes the quartic surface that makes up the curved part of the boundary of P^Δ, as seen in Figure 2. In addition, there are four planes spanned by flat 2-dimensional faces of P^Δ. The product of the four corresponding affine-linear forms equals (1.8). Indeed, each of the two quadrics in (1.8) factors into two linear factors. These two characterize the planes spanned by opposite 2-faces of P^Δ.

The two equations (1.7) and (1.8) also offer a first glimpse at the concept of *projective duality* in algebraic geometry. Namely, consider the surface in projective space $\mathbb{P}^3$ defined by $\det(Q(x,y,z)) = 0$ after replacing the ones along the diagonals by a homogenization variable. Then (1.7) is its dual surface in the dual projective space $(\mathbb{P}^3)^*$. The surface (1.8) in $(\mathbb{P}^3)^*$ is dual to the 0-dimensional variety in $\mathbb{P}^3$ cut out by the 3×3-minors of $Q(x,y,z)$.

The optimal value function of the optimization problem (1.4) is given by the algebraic surfaces dual to the boundary of P and its singular locus. We have seen two different ways of dualizing (1.4): the dual optimization problem (1.6), and the optimization problem (1.5) on P^Δ. These two formulations are related as follows. If we regard (1.6) as specifying a 10-dimensional spectrahedron, then the dual pillow P^Δ is a projection of that spectrahedron:

$$P^\Delta \;=\; \big\{(a,b,c)\in\mathbb{R}^3 \,|\, \exists u \in \mathbb{R}^7 \;:\; M(u;a,b,c)\succeq 0 \text{ and } u_1+u_4+u_6+u_7 \;=\; 1\big\}.$$

Linear projections of spectrahedra are called *spectrahedral shadows*. These objects play a prominent role in the interplay between semidefinite programming and convex algebraic geometry. The dual body to a spectrahedron is generally not a spectrahedron, but it is always a spectrahedral shadow.

1.2 – Context and Outline

Duality is a central concept in convexity and convex optimization, and numerous authors have written about their connections and their interplay with other notions of duality and polarity. Relevent references include Barvinok's text book [1, §4] and the survey by Luenberger [19]. The latter focuses on dualities used in engineering, such as duality of vector spaces, polytopes, graphs, and control systems. The objective of this article is to revisit the theme of duality in the context of *convex algebraic geometry*. This emerging field aims to exploit algebraic structure in convex optimization problems, specifically in semidefinite programming and polynomial optimization. In algebraic geometry, there is a natural notion of projective duality, which associates to every algebraic variety a dual variety. One of our goals is to explore the meaning of projective duality for optimization theory.

Our presentation is organized as follows. In Section 2 we cover preliminaries needed for the rest of the paper. Here the various dualities are carefully defined and their basic properties are illustrated by means of examples. In Section 3 we derive the result that the optimal value function of a polynomial program is represented by the defining equation of the hypersurface projectively dual to the manifold describing the boundary of all feasible solutions. This highlights the important fact that the duality best known to algebraic geometers arises very naturally in convex optimization. Section 4 concerns the convex hull of a compact algebraic variety in $\mathbb{R}^n$. We discuss recent work of Ranestad and Sturmfels [24, 25] on the hypersurfaces in the boundary of such a convex body, and we present several new examples and applications.

In Section 5 we focus on semidefinite programming (SDP), and we offer a concise geometric introduction to SDP duality. This leads to the concept of algebraic degree of SDP [8, 22], or, geometrically, to projective duality for varieties defined by rank constraints on symmetric matrices of linear forms.

A *spectrahedral shadow* is the image of a spectrahedron under a linear projection. Its dual body is a linear section of the dual body to the spectrahedron. In Section 6 we examine this situation in the context of sums-of-squares programming, and we discuss linear families of non-negative polynomials.

2 – Ingredients

In this section we review the mathematical preliminaries needed for the rest of the paper, we give precise definitions, and we fix more of the notation. We begin with the notion of duality for vector spaces and cones therein, then move on to convex bodies, polytopes, Lagrange duality in optimization, the KKT conditions, projective duality in algebraic geometry, and discriminants.

2.1 – Vector Spaces and Cones

We fix an ordered field K. The primary example is the field of real numbers, $K = \mathbb{R}$, but it makes much sense to also allow other fields, such as the rational numbers $K = \mathbb{Q}$ or the real Puiseux series $K = \mathbb{R}\{\{\epsilon\}\}$. For a finite dimensional K-vector space V, the dual vector space is the set $V^* = \mathrm{Hom}(V, K)$ of all linear forms on V. Let V and W be vector spaces and $\varphi : V \to W$ a linear map. The *dual map* $\varphi^* : W^* \to V^*$ is the linear map defined by $\varphi^*(w) = w \circ \varphi \in V^*$ for every $w \in W^*$. If we fix bases of V and W then φ is represented by a matrix A. The dual map φ^* is represented, relative to the dual bases for W^* and V^*, by the *transpose* A^t of the matrix A.

A subset $C \subset V$ is a *cone* if it is closed under multiplication with positive scalars. A cone C need not be convex, but its *dual cone*

$$C^* = \{\, l \in V^* \mid \forall x \in C \,:\, l(x) \geq 0 \,\} \tag{2.1}$$

is always closed and convex in V^*. If C is a *convex cone* then the second dual $(C^*)^*$ is the closure of C. Thus, if C is a closed convex cone in V then

$$(C^*)^* \;=\; C. \tag{2.2}$$

This important relationship is referred to as *biduality*.

Every linear subspace $L \subset V$ is also a cone. Its dual cone is the orthogonal complement of the subspace:

$$L^* \;=\; L^\perp \;=\; \{\, l \in V^* \mid \forall x \in L \,:\, l(x) = 0 \,\}.$$

The dual map to the inclusion $L \subset V$ is the projection $\pi_L : V^* \to V^*/L^\perp$. Given any cone $C \subset V$, the intersection $C \cap L$ is a cone in L. Its dual cone $(C \cap L)^*$ is the projection of the cone C into $V^*/L^\perp$. More precisely,

$$(C \cap L)^* \;=\; C^* + L^\perp \quad \text{in } V^*.$$

Now, it makes sense to consider this convex set modulo $L^\perp$. We thus obtain

$$(C \cap L)^* \;=\; \overline{\pi_L(C^*)} \quad \text{in } V^*/L^\perp. \tag{2.3}$$

This identity shows that projection and intersections are dual operations.

A subset $F \subseteq C$ of a convex set C is a *face* if F is itself convex and contains any line segment $L \subset C$ whose relative interior intersects F. We say that F is an *exposed face* if there exists a linear functional l that attains its minimum over C precisely at F. Clearly, an exposed face is a face, but the converse does not hold. For instance, the edges of the red triangle in Figure 6 are non-exposed faces of the 3-dimensional convex body shown there.

An exposed face F of a cone C determines a face of the dual cone C^* via

$$F^\diamond \;=\; \{\, l \in C^* \mid l \text{ attains its minimum over } C \text{ at } F \,\}.$$

The dimensions of the faces F of C and $F^\diamond$ of C^* satisfy the inequality

$$\dim(F) + \dim(F^\diamond) \;\leq\; \dim(V). \tag{2.4}$$

If C is a polyhedral cone then C^* is also polyhedral. In that case, the number of faces F and $F^\diamond$ is finite and equality holds in (2.4). On the other hand, most cones considered in this article are not polyhedral, they have infinitely many faces, and the inequality in (2.4) is usually strict. For instance, the *second order cone* $C = \{\,(x,y,z) \in \mathbb{R}^3 : \sqrt{x^2+y^2} \leq z\}$ is self-dual, each proper face F of C is 1-dimensional, and the formula (2.4) says $1 + 1 \leq 3$.

2.2 – Convex Bodies and their Algebraic Boundary

A *convex body* in V is a full-dimensional convex set that is closed and bounded. If C is a cone and $z \in \text{int}(C^*)$ then $C \cap \{z = 1\}$ is a convex body in the hyperplane $\{ z = 1 \}$ of V. In this manner, every pointed d-dimensional cone gives rise to a $(d-1)$-dimensional convex body, and vice versa. These transformations, known as *homogenization* and *dehomogenization*, respect faces and algebraic boundaries. They allow us to go back and fourth between convex bodies and cones in the next higher dimension. For instance, the 3-dimensional body P in (1.2) corresponds to the cone in $\mathbb{R}^4$ we get by multiplying the constants 1 on the diagonal in (1.1) with a new variable.

Let P be a full-dimensional convex body in V and assume that $0 \in \text{int}(P)$. Dehomogenizing the definition for cones, we obtain the *dual convex body*

$$P^{\Delta} \;=\; \{ \ell \in V^* \mid \forall x \in P \,:\, \ell(x) \leq 1 \} . \tag{2.5}$$

This is derived from (2.1) using the identification $l(x) = z - \ell(x)$ for $z = 1$.

Just as in the case of convex cones, if P is closed then biduality holds:

$$(P^{\Delta})^{\Delta} = P.$$

The definition (2.5) makes sense for arbitrary subsets P of V. That is, P need not be convex or closed. A standard fact from convex analysis [26, Cor. 12.1.1 and §14] says that the double dual is the closure of the convex hull with the origin:

$$(P^{\Delta})^{\Delta} = \overline{\text{conv}(P \cup 0)}.$$

All convex bodies discussed in this article are *semialgebraic*, that is, they can be described by polynomial inequalities. We note that if P is semialgebraic then its dual body P^{Δ} is also semialgebraic. This is a consequence of Tarski's theorem on quantifier elimination in real algebraic geometry [2, 3].

The *algebraic boundary* of a semialgebraic convex body P, denoted $\partial_a P$, is the smallest algebraic variety that contains the boundary ∂P. In geometric language, $\partial_a P$ is the *Zariski closure* of ∂P. It is identified with the squarefree polynomial f_P that vanishes on ∂P. Namely, $\partial_a P = V(f_P)$ is the zero set of the polynomial f_P. Note that f_P is unique up to a multiplicative constant. Thus $\partial_a P$ is an algebraic hypersurface which contains the boundary ∂P.

A *polytope* is the convex hull of a finite subset of V. If P is a polytope then so is its dual P^{Δ} [29]. The boundary of P consists of finitely many *facets* F. These are the faces $F = v^{\diamond}$ dual to the *vertices* v of P^{Δ}. The algebraic boundary $\partial_a P$ is the arrangement of hyperplanes spanned by the facets of P. Its defining polynomial f_P is the product of the linear polynomials $v - 1$.

Example 2.1. A polytope known to everyone is the three-dimensional cube

$$P \;=\; \mathrm{conv}\{(\pm 1, \pm 1, \pm 1)\} \;=\; \{-1 \leq x, y, z \leq 1\}.$$

Figure 1 illustrates the familiar fact that its dual polytope is the octahedron

$$P^{\Delta} \;=\; \{-1 \leq a \pm b \pm c \leq 1\} \;=\; \mathrm{conv}\{\pm e_1, \pm e_2, \pm e_3\}.$$

Here e_i denotes the ith unit vector. The eight vertices of P correspond to the facets of P^{Δ}, and the six facets of P correspond to the vertices of P^{Δ}. The algebraic boundary of the cube is described by a degree 6 polynomial

$$\partial_a P \;=\; V\left((x^2-1)(y^2-1)(z^2-1)\right).$$

The algebraic boundary of the octahedron is given by a degree 8 polynomial

$$\partial_a P^{\Delta} \;=\; V\left(\prod (1 - a \pm y \pm c) \prod (a \pm b \pm c + 1)\right).$$

Note that P and P^{Δ} are the unit balls for the norms L_∞ and L_1 on $\mathbb{R}^3$. □

Recall that the L_p-*norm* on $\mathbb{R}^n$ is defined by $\|x\|_p \;=\; (\sum_{i=1}^n |x_i|^p)^{1/p}$ for $x \in \mathbb{R}^n$. The *dual norm* to the L_p-norm is the L_q-norm for $\frac{1}{p} + \frac{1}{q} = 1$, that is,

$$\|y\|_q \;=\; \sup\{\langle y, x\rangle \,|\, x \in \mathbb{R}^n,\; \|x\|_p \leq 1\}.$$

Geometrically, the unit balls for these norms are dual as convex bodies.

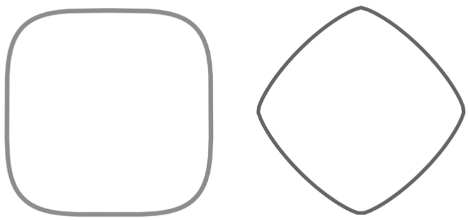

Fig. 3: The unit balls for the L_4 norm and the $L_{4/3}$ norm are dual. The curve on the left has degree 4, while its dual curve on the right has degree 12.

Example 2.2. Consider the case $n = 2$ and $p = 4$. Here the unit ball equals

$$P \;=\; \{\,(x, y) \in \mathbb{R}^2 \,:\, x^4 + y^4 \leq 1\,\}.$$

This planar convex set is shown in Figure 3. In this example, since the curve is convex, the ordinary boundary coincides with the algebraic boundary, $\partial_a P = \partial P$, and is represented by the defining quartic polynomial $x^4 + y^4 - 1$.

The dual body is the unit ball for the $L_{4/3}$-norm on $\mathbb{R}^2$:

$$P^{\Delta} \;=\; \{(a,b) \in \mathbb{R}^2 \,:\, |a|^{4/3} + |b|^{4/3} \leq 1\}\,.$$

The algebraic boundary of P^{Δ} is an irreducible algebraic curve of degree 12,

$$\partial_a P^{\Delta} \;=\; V\big(a^{12}+3a^8b^4+3a^4b^8+b^{12}-3a^8+21a^4b^4-3b^8+3a^4+3b^4-1\big), \tag{2.6}$$

which again coincides precisely with the (geometric) boundary ∂P^{Δ}. This dual polynomial is easily produced by the following one-line program in the computer algebra system `Macaulay2` due to Grayson and Stillman [9]:

```
R = QQ[x,y,u,v]; eliminate(x,y,ideal(x^4+y^4-1,x^3-u,y^3-v))
```

In Subsection 2.4 we shall introduce the algebraic framework for performing such duality computations, not just for curves, but for arbitrary varieties. □

2.3 – Lagrange Duality in Optimization

We now come to a standard concept of duality in optimization theory. Let us consider the following general nonlinear polynomial optimization problem:

$$\begin{aligned} &\underset{x}{\text{Minimize}}\ f(x) \\ &\text{subject to}\quad g_i(x) \leq 0, \quad i = 1, \dots, m, \\ &\phantom{\text{subject to}}\quad h_j(x) = 0, \quad j = 1, \dots, p. \end{aligned} \tag{2.7}$$

Here the $g_1, \dots, g_m, h_1, \dots, h_p$ and f are polynomials in $\mathbb{R}[x_1, \dots, x_n]$. The *Lagrangian* associated to the optimization problem (2.7) is the function

$$\begin{array}{rcl} L \,:\, \mathbb{R}^n \times \mathbb{R}^m_+ \times \mathbb{R}^p & \rightarrow & \mathbb{R}^n \\ (x, \lambda, \mu) & \mapsto & f(x) + \sum_{i=1}^m \lambda_i g_i(x) + \sum_{j=1}^p \mu_j h_j(x) \end{array}$$

The scalars $\lambda_i \in \mathbb{R}_+$ and $\mu_j \in \mathbb{R}$ are the *Lagrange multipliers* for the constraints $g_i(x) \leq 0$ and $h_j(x) = 0$. The Lagrange function $L(x, \lambda, \mu)$ can be interpreted as an augmented cost function with penalty terms for the constraints. For more information on the above formulation see [4, §5.1].

One can show that problem (2.7) is equivalent to finding

$$u^* \;=\; \underset{x \in \mathbb{R}^n}{\text{Minimize}}\ \underset{\mu \in \mathbb{R}^p \text{ and } \lambda \geq 0}{\text{Maximize}}\quad L(x, \lambda, \mu).$$

The key observation here is that any positive evaluation of one of the polynomials $g_i(x)$, or any non-zero evaluation of one of the polynomials $h_j(x)$, would render the inner optimization problem unbounded.

The *dual optimization problem* to (2.7) is obtained by exchanging the order of the two nested optimization subproblems in the above formulation:

$$v^* \;=\; \underset{\mu\in\mathbb{R}^p \text{ and } \lambda\geq 0}{\text{Maximize}} \; \underbrace{\underset{x\in\mathbb{R}^n}{\text{Minimize}} \quad L(x,\lambda,\mu)}_{\phi(\lambda,\mu)}\,.$$

The function $\phi(\lambda,\mu)$ is known as the *Lagrange dual function* to our problem. This function is always concave, so the dual is always a convex optimization problem. It follows from the definition of the dual function that $\phi(\lambda,\mu) \leq u^*$ for all λ,μ. Hence the optimal values satisfy the inequality

$$v^* \;\leq\; u^*.$$

If equality happens, $v^* = u^*$, then we say that *strong duality* holds. A necessary condition for strong duality is $\lambda_i^* g_i(x^*) = 0$ for all $i = 1,\ldots,m$, where x^*, λ^* denote the primal and dual optimizer. We see this by inspecting the Lagrangian and taking into account the fact that $h_j(x) = 0$ for all feasible x.

Collecting all inequality and equality constraints in the primal and dual optimization problems yields the following optimality conditions:

Theorem 2.3. (Karush-Kuhn-Tucker (KKT) conditions) *Let (x^*,λ^*,μ^*) be primal and dual optimal solutions with $u^* = v^*$ (strong duality). Then*

$$\nabla_x f\Big|_{x^*} + \sum_{i=1}^{m} \lambda_i^* \cdot \nabla_x g_i\Big|_{x^*} + \sum_{j=1}^{p} \mu_j^* \cdot \nabla_x h_j\Big|_{x^*} \;=\; 0,$$

$$\begin{aligned} g_i(x^*) \;&\leq\; 0 \quad \text{for } i = 1,\ldots,m,\\ \lambda_i^* \;&\geq\; 0 \quad \text{for } i = 1,\ldots,m,\\ h_j(x^*) \;&=\; 0 \quad \text{for } j = 1,\ldots,p,\\ \textit{Complementary slackness:}\quad \lambda_i^* \cdot g_i(x^*) \;&=\; 0 \quad \text{for } i = 1,\ldots,m. \end{aligned} \tag{2.8}$$

For a derivation of this theorem see [4, §5.5.2]. Several comments on the KKT conditions are in order. First, we note that complementary slackness amounts to a case distinction between active ($g_i = 0$) and inactive inequalities ($g_i < 0$). For any index i with $g_i(x^*) \neq 0$ we need $\lambda_i = 0$, so the corresponding inequality does not play a role in the gradient condition. On the other hand, if $g_i(x^*) = 0$, then this can be treated as an equality constraint.

From an algebraic point of view, it is natural to relax the inequalities and to focus on the *KKT equations*. These are the polynomial equations in (2.8):

$$h_1(x) \;=\; \cdots \;=\; h_p(x) \;=\; \lambda_1 g_1(x) = \;\cdots\; = \lambda_m g_m(x) \;=\; 0. \tag{2.9}$$

If we wish to solve our optimization problem exactly then we must compute the algebraic variety in $\mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p$ that is defined by these equations.

In what follows we explore Lagrange duality and the KKT conditions in two special cases, namely in optimizing a linear function over an algebraic variety (Section 3) and in semidefinite programming (Section 5).

2.4 – Projective varieties and their duality

In algebraic geometry, it is customary to work over an algebraically closed field, such as the complex numbers $\mathbb{C}$. All our varieties will be defined over a subfield K of the real numbers $\mathbb{R}$, and their points have coordinates in $\mathbb{C}$. It is also customary to work in projective space $\mathbb{P}^n$ rather than affine space $\mathbb{C}^n$, i.e., we work with equivalence classes $x \sim \lambda x$ for all $\lambda \in \mathbb{C}\backslash\{0\}$, $x \in \mathbb{C}^{n+1}\backslash\{0\}$. Points $(x_0 : x_1 : \cdots : x_n)$ in projective space $\mathbb{P}^n$ are lines through the origin in $\mathbb{C}^{n+1}$, and the usual affine coordinates are obtained by *dehomogenization* with respect to x_0 (i.e. setting $x_0 = 1$). All points with $x_0 = 0$ are then considered as points at infinity. We refer to [6, Chapter 8] for an elementary introduction to projective algebraic geometry.

Let $I = \langle h_1, \ldots, h_p\rangle$ be a homogeneous ideal in the polynomial ring $K[x_0, x_1, \ldots, x_n]$. We write $X = V(I)$ for its variety in the projective space $\mathbb{P}^n$ over $\mathbb{C}$. The *singular locus* $\mathrm{Sing}(X)$ is a proper subvariety of X. It is defined inside X by the vanishing of the $c \times c$-minors of the $m\times(n+1)$-Jacobian matrix $J(X) = \big(\partial h_i/\partial x_j\big)$, where $c = \mathrm{codim}(X)$. See [6, §9.6] for background on singularities and dimension. While the matrix $J(X)$ depends on our choice of ideal generators h_i, the singular locus of X is independent of that choice. Points in $\mathrm{Sing}(X)$ are called *singular points* of X. We write $X_{\mathrm{reg}} = X\backslash\mathrm{Sing}(X)$ for the set of *regular points* in X. We say that the projective variety X is smooth if $\mathrm{Sing}(X) = \emptyset$, or equivalently, if $X = X_{\mathrm{reg}}$.

The dual projective space $(\mathbb{P}^n)^*$ parametrizes hyperplanes in $\mathbb{P}^n$. A point $(u_0 : \cdots : u_n) \in (\mathbb{P}^n)^*$ represents the hyperplane $\big\{x \in \mathbb{P}^n \mid \sum_{i=0}^n u_i x_i = 0\big\}$. We say that u is *tangent* to X at a regular point $x \in X_{\mathrm{reg}}$ if x lies in that hyperplane and its representing vector $(u_1, \ldots, u_n)$ lies in the row space of the Jacobian matrix $J(X)$ at the point x.

We define the *conormal variety* $\mathrm{CN}(X)$ of X to be the closure of the set

$$\big\{(x, u) \in \mathbb{P}^n \times (\mathbb{P}^n)^* \mid x \in X_{\mathrm{reg}} \text{ and } u \text{ is tangent to } X \text{ at } x\,\big\}.$$

The projection of $\mathrm{CN}(X)$ onto the second factor is denoted X^* and is called the dual variety. More precisely, the *dual variety* X^* is the closure of the set

$$\big\{\, u \in (\mathbb{P}^n)^* \mid \text{the hyperplane } u \text{ is tangent to } X \text{ at some regular point}\,\big\}.$$

Proposition 2.4.. *The conormal variety* $\mathrm{CN}(X)$ *has dimension* $n - 1$.

Proof. We may assume that X is irreducible. Let $c = \mathrm{codim}(X)$. There are $n{-}c$ degrees of freedom in picking a point x in X_{reg}. Once the regular point x is fixed, the possible tangent vectors u to X at x form a linear space of dimension $c{-}1$. Hence the dimension of $\mathrm{CN}(X)$ is $(n{-}c) + (c{-}1) = n{-}1$. ∎

Since the dual variety X^* is a linear projection of the conormal variety $\mathrm{CN}(X)$, Proposition 2.4 implies that the dimension of X^* is at most is $n-1$. We typically expect X^* to have dimension $n-1$, i.e. regardless of the dimension of X, the dual variety X^* is typically a hypersurface in $(\mathbb{P}^n)^*$.

Example 2.5. [Example 2.2 cont.] Fix coordinates $(x{:}y{:}z)$ on $\mathbb{P}^2$ and consider the ideal $I = \langle x^4 + y^4 - z^4 \rangle$. Then $X = V(I)$ is the projectivization of the quartic curve in Example 2.2. The dual curve X^* is the projectivization of the curve $\partial_a P^\Delta$ in (2.6). Hence, X^* is a curve of degree 12 in $(\mathbb{P}^2)^*$. □

To compute the dual X^* of a given variety X, we can utilize Gröbner bases [6, 9] as follows. We augment the ideal I with the bilinear polynomial $\sum_{i=0}^n u_i x_i$ and all the $(c+1)\times(c+1)$-minors of the matrix obtained from $\mathrm{Jac}(X)$ by adding the extra row u. Let J' denote the resulting ideal in $K[x_0, \ldots, x_n, u_0, \ldots, u_n]$. In order to remove the singular locus of X from the variety of J', we replace J' with the *saturation ideal*

$$J \;:=\; \big(J' \,:\, \langle\, c \times c\text{-minors of } \mathrm{Jac}(X)\,\rangle^\infty\big).$$

See [6, Exercise 8 in §4.4] for the definition of saturation of ideals.

The ideal J is bi-homogeneous in x and u respectively. Its zero set in $\mathbb{P}^n \times (\mathbb{P}^n)^*$ is the conormal variety $\mathrm{CN}(X)$. The ideal of the dual variety X^* is finally obtained by eliminating the variables $x_0, \ldots, x_n$ from J:

$$\text{ideal of the dual variety } X^* \;=\; J \cap K[u_0, u_1, \ldots, u_n]. \tag{2.10}$$

As was remarked earlier, the expected dimension of X^* is $n-1$, so the elimination ideal (2.10) is expected to be principal. We seek to compute its generator. We shall see many examples of such dual hypersurfaces later on.

Theorem 2.6. (Biduality, [7, Theorem 1.1]) *Every irreducible projective variety $X \subset \mathbb{P}^n$ satisfies*

$$(X^*)^* \;=\; X.$$

Proof idea The main step in proving this important theorem is that the conormal variety is self-dual, in the sense that $\mathrm{CN}(X) = \mathrm{CN}(X^*)$. In this identity the roles of $x \in \mathbb{P}^n$ and $u \in (\mathbb{P}^n)^*$ are swapped. It implies $(X^*)^* = X$. A proof for the self-duality of the conormal variety is found in [7, §I.1.3]. □

Example 2.7. Suppose that $X \subset \mathbb{P}^n$ is a general smooth hypersurface of degree d. Then X^* is a hypersurface of degree $d(d-1)^{n-1}$ in $(\mathbb{P}^n)^*$. A concrete instance for $d = 4$ and $n = 2$ was seen in Examples 2.2 and 2.5. □

Example 2.8. Let X be the variety of symmetric $m \times m$ matrices of rank at most r. Then X^* is the variety of symmetric $m \times m$ matrices of rank at most $m - r$ [7, §I.1.4]. Here the conormal variety $\mathrm{CN}(X)$ consists of pairs of symmetric matrices A and B such that $A \cdot B = 0$. This conormal variety will be important for our discussion of duality in semidefinite programming. □

An important class of examples, arising from toric geometry, is featured in the book by Gel'fand, Kapranov and Zelevinsky [7]. A *projective toric variety* X_A in $\mathbb{P}^n$ is specified by an integer matrix $A = (a_0, a_1, \ldots, a_n)$ of format $d \times (n+1)$ and rank d whose row space contains the vector $(1, 1, \ldots, 1)$. We define X_A as the closure in $\mathbb{P}^n$ of the set $\{(t^{a_0} : t^{a_1} : \cdots : t^{a_n}) \,|\, t \in (\mathbb{C}\backslash\{0\})^d\}$.

The dual variety X_A^* is called the *A-discriminant*. It is usually a hypersurface, in which case we identify the *A-discriminant* with the irreducible polynomial Δ_A that vanishes on X_A^*. The A-discriminant is indeed a discriminant in the sense that its vanishing characterizes Laurent polynomials

$$p(t) \;=\; \sum_{j=0}^{n} c_j \cdot t_1^{a_{1j}} t_2^{a_{2j}} \ldots t_d^{a_{dj}}$$

with the property that the hypersurface $\{p(t) = 0\}$ has a singular point in $(\mathbb{C}\backslash\{0\})^d$. In other words, we can define (and compute) the A-discriminant as

$$\Delta_A \;=\; \overline{\left\{c \in (\mathbb{P}^n)^* \mid \exists\, t \in (\mathbb{C}\backslash\{0\})^d \text{ with } p(t) = \frac{\partial p}{\partial t_1} = \cdots = \frac{\partial p}{\partial t_d} = 0\right\}}.$$

Example 2.9. Let $d = 2$, $n = 4$, and fix the matrix

$$A = \begin{pmatrix} 4 & 3 & 2 & 1 & 0 \\ 0 & 1 & 2 & 3 & 4 \end{pmatrix}$$

The associated toric variety is the rational normal curve

$$\begin{aligned} X_A \;&=\; \{(t_1^4 : t_1^3 t_2 : t_1^2 t_2^2 : t_1 t_2^3 : t_2^4) \in \mathbb{P}^4 \mid (t_1 : t_2) \in \mathbb{P}^1\} \\ &=\; V(x_0x_2 - x_1^2, x_0x_3 - x_1x_2, x_0x_4 - x_2^2, x_1x_3 - x_2^2, x_1x_4 - x_2x_3, x_2x_4 - x_3^2). \end{aligned}$$

A hyperplane $\{\sum_{j=0}^{4} c_j x_j = 0\}$ is tangent to X_A if and only if the binary form

$$p(t_1, t_2) \quad = \quad c_0 t_2^4 + c_1 t_1 t_2^3 + c_2 t_1^2 t_2^2 + c_3 t_1^3 t_2 + c_4 t_1^4$$

has a linear factor of multiplicity ≥ 2. This is controlled by the A-discriminant

$$\Delta_A \;=\; \frac{1}{c_4} \cdot \det \begin{pmatrix} c_0 & c_1 & c_2 & c_3 & c_4 & 0 & 0 \\ 0 & c_0 & c_1 & c_2 & c_3 & c_4 & 0 \\ 0 & 0 & c_0 & c_1 & c_2 & c_3 & c_4 \\ c_1 & 2c_2 & 3c_3 & 4c_4 & 0 & 0 & 0 \\ 0 & c_1 & 2c_2 & 3c_3 & 4c_4 & 0 & 0 \\ 0 & 0 & c_1 & 2c_2 & 3c_3 & 4c_4 & 0 \\ 0 & 0 & 0 & c_1 & 2c_2 & 3c_3 & 4c_4 \end{pmatrix}, \tag{2.11}$$

given here in form of the determinant of a Sylvester matrix. The sextic hypersurface $X_A^* = V(\Delta_A)$ is the dual variety of the curve X_A. □

3 – The Optimal Value Function

In this section we examine the optimization problem (2.7) under the hypotheses that the cost function $f(x)$ is linear and that there are no inequality constraints $g_i(x)$. The purposes of these restrictions is to simplify the presentation and focus on the key ideas. Our analysis can be extended to the general problem (2.7) and we discuss this briefly at the end of this section.

We consider the problem of optimizing a linear cost function over a compact real algebraic variety X in $\mathbb{R}^n$:

$$\begin{aligned} c_0^* \;&=\; \underset{x}{\text{minimize}} \quad \langle c, x\rangle \\ \text{subject to} \quad & x \in X = \{v \in \mathbb{R}^n \mid h_1(v) = \cdots = h_p(v) = 0\}\,. \end{aligned} \tag{3.1}$$

Here $h_1, h_2, \ldots, h_p$ are fixed polynomials in n unknowns $x_1, \ldots, x_n$. The expression $\langle c, x\rangle = c_1x_1 + \cdots + c_nx_n$ is a linear form whose coefficients $c_1, \ldots, c_n$ are unspecified parameters. Our aim is to compute the *optimal value function* c_0^*. Thus, we regard the optimal value c_0^* as a function $\mathbb{R}^n \to \mathbb{R}$ of the parameters $c_1, \ldots, c_n$, and we seek to determine this function.

The hypothesis that X be compact has been included to ensure that the optimal value function c_0^* is well-defined on all of $\mathbb{R}^n$. Again, also this hypothesis can be relaxed. We assume compactness here just for convenience.

Our problem is equivalent to that of describing the dual convex body P^Δ of the convex hull $P = \text{conv}(X)$, assuming that the latter contains the origin in its interior. A small instance of this was seen in (1.5). Since our convex hull P is a semi-algebraic set, Tarski's theorem on quantifier elimination in real algebraic geometry [2, 3] ensures that the dual body P^Δ is also semialgebraic. This implies that the optimal value function c_0^* is an algebraic function, i.e., there exists a polynomial $\Phi(c_0, c_1, \ldots, c_n)$ in $n + 1$ variables such that

$$\Phi(c_0^*, c_1, \ldots, c_n) \quad = \quad 0. \tag{3.2}$$

Our aim is to compute such a polynomial Φ of least possible degree. The input consists of the polynomials $h_1, \ldots, h_p$ that cut out the variety X. The degree of Φ in the unknown c_0 is called the *algebraic degree* of the optimization problem (2.7). This number is an intrinsic algebraic complexity measure for the problem of optimizing a linear function over X. For instance, if $c_1, \ldots, c_n$ are rational numbers then the algebraic degree indicates the degree of the field extension K over $\mathbb{Q}$ that contains the coordinates of the optimal solution.

We illustrate our discussion by computing the optimal value function and its algebraic degree for the trigonometric space curve featured in [24, §1].

Example 3.1. Let X be the curve in $\mathbb{R}^3$ with parametric representation

$$(x_1, x_2, x_3) \quad = \quad \big(\cos(\theta), \sin(2\theta), \cos(3\theta)\big).$$

In terms of equations, our curve can be written as $X = V(h_1, h_2)$, where

$$h_1 \;=\; x_1^2 - x_2^2 - x_1x_3 \quad \text{and} \quad h_2 \;=\; x_3 - 4x_1^3 + 3x_1.$$

The optimal value function for maximizing $c_1x_1+c_2x_2+c_3x_3$ over X is given by

$$\begin{aligned}
\Phi \;=\;\; & (11664c_3^4)\cdot c_0^6 \;+\; (864c_1^3c_3^3 + 1512c_1^2c_2^2c_3^2 - 19440c_1^2c_3^4 \\
& +576c_1c_2^4c_3 - 1296c_1c_2^2c_3^3 + 64c_2^6 - 25272c_2^2c_3^4 - 34992c_3^6)\cdot c_0^4 \\
& + (16c_1^6c_3^2 + 8c_1^5c_2^2c_3 - 1152c_1^5c_3^3 - 1920c_1^4c_2^2c_3^2 + 8208c_1^4c_3^4 - 724c_1^3c_2^4c_3 + 144c_1^3c_2^2c_3^3 \\
& +c_1^4c_2^4 - 17280c_1^3c_3^5 - 80c_1^2c_2^6 - 2802c_1^2c_2^4c_3^2 - 3456c_1^2c_2^2c_3^4 + 3888c_1^2c_3^6 - 1120c_1c_2^6c_3 \\
& +540c_1c_2^4c_3^3 + 55080c_1c_2^2c_3^5 - 128c_2^8 - 208c_2^6c_3^2+15417c_2^4c_3^4+15552c_2^2c_3^6+34992c_3^8)\cdot c_0^2 \\
& + (-16c_1^8c_3^2 - 8c_1^7c_2^2c_3 + 256c_1^7c_3^3 - c_1^6c_2^4 + 328c_1^6c_2^2c_3^2 - 1600c_1^6c_3^4 + 114c_1^5c_2^4c_3 \\
& -2856c_1^5c_2^2c_3^3 + 4608c_1^5c_3^5 + 12c_1^4c_2^6 - 1959c_1^4c_2^4c_3^2 + 9192c_1^4c_2^2c_3^4 - 4320c_1^4c_3^6 \\
& -528c_1^3c_2^6c_3 + 7644c_1^3c_2^4c_3^3 - 7704c_1^3c_2^2c_3^5 - 6912c_1^3c_3^7 - 48c_1^2c_2^8 + 3592c_1^2c_2^6c_3^2 \\
& -4863c_1^2c_2^4c_3^4 - 13608c_1^2c_2^2c_3^6 + 15552c_1^2c_3^8 + 800c_1c_2^8c_3 - 400c_1c_2^6c_3^3 - 10350c_1c_2^4c_3^5 \\
& +16200c_1c_2^2c_3^7 + 64c_2^10 + 80c_2^8c_3^2 - 1460c_2^6c_3^4 + 135c_2^4c_3^6 + 9720c_2^2c_3^8 - 11664c_3^{10}).
\end{aligned}$$

The optimal value function c_0^* is the algebraic function of c_1, c_2, c_3 obtained by solving $\Phi = 0$ for the unknown c_0. Since c_0 has degree 6 in Φ, we see that the algebraic degree of this optimization problem is 6. Note that we can write c_0^* in terms of radicals in c_1, c_2, c_3 because there are no odd powers of c_0 in Φ, which ensures that the Galois group of c_0^* over $\mathbb{Q}(c_1, c_2, c_3)$ is solvable. □

We now come to the main result in this section. It will explain what the polynomial Φ means and how it was computed in the previous example. For the sake of simplicity, we shall first assume that the given variety X is smooth, i.e. $X = X_{\text{reg}}$, where the set X_{reg} denotes all regular points on X.

THEOREM 3.2. *Let $X^* \subset (\mathbb{P}^n)^*$ be the dual variety to the projective closure of X. If X is irreducible, smooth and compact in $\mathbb{R}^n$ then X^* is an irreducible hypersurface, and its defining polynomial equals $\Phi(-c_0, c_1, \ldots, c_n)$ where Φ represents the optimal value function as in* (3.2) *of the optimization problem* (3.1). *In particular, the algebraic degree of* (3.1) *is the degree in c_0 of the irreducible polynomial that vanishes on the dual hypersurface X^*.*

Here the change of sign in the coordinate c_0 is needed because the equation $c_0 = c_1x_1 + \cdots + c_nx_n$ for the objective function value in $\mathbb{R}^n$ becomes the homogenized equation $(-c_0)x_0 + c_1x_1 + \cdots + c_nx_n = 0$ when we pass to $\mathbb{P}^n$.

PROOF. Since X is compact, for every cost vector c there exists an optimal solution x^*. Our assumption that X is smooth ensures that x^* is a regular point of X, and c lies in the span of the gradient vectors $\nabla_x h_i\big|_{x^*}$ for $i = 1, \ldots, p$. In other words, the KTT conditions are necessary at the point x^*:

$$\begin{aligned}
c \;&=\; \sum_{i=1}^{p} \lambda_i^* \cdot \nabla_x h_i\big|_{x^*}, \\
h_i(x^*) \;&=\; 0 \qquad \text{for} \;\; i = 1, 2, \ldots, p.
\end{aligned}$$

The scalars $\lambda_1^*, \ldots, \lambda_p^*$ express c as a vector in the orthogonal complement of the tangent space of X at x^*. In other words, the affine hyperplane $\{x \in \mathbb{R}^n : \langle c, x\rangle = c_0^*\}$ contains the tangent space of X at x^*. This means that the pair $\big(x^*, (-c_0^* : c_1 : \cdots : c_n)\big)$ lies in the conormal variety $\mathrm{CN}(X) \subset \mathbb{P}^n \times (\mathbb{P}^n)^*$ of the projective closure of X. By projection onto the second factor, we see that $(-c_0^* : c_1 : \cdots : c_n)$ lies in the dual variety X^*.

Our argument shows that the boundary of the dual body P^Δ is a subset of X^*. Since that boundary is a semialgebraic set of dimension $n-1$, we conclude that X^* is a hypersurface. If we write its defining equation as $\Phi(-c_0, c_1, \ldots, c_n) = 0$, then the polynomial Φ satisfies (3.2), and the statement about the algebraic degree follows as well. □

The KKT condition for the optimization problem (3.1) involves three sets of variables, two of which are dual variables, to be carefully distinguished:

1. Primal variables $x_1, \ldots, x_n$ to describe the set X of feasible solutions.
2. (Lagrange) dual variables $\lambda_1, \ldots, \lambda_p$ to parametrize the linear space of all hyperplanes that are tangent to X at a fixed point x^*.
3. (Projective) dual variables $c_0, c_1, \ldots, c_n$ for the space of all hyperplanes. These are coordinates for the dual variety X^* and the dual body P^Δ.

We can compute the equation Φ that defines the dual hypersurface X^* by eliminating the first two groups of variables $x = (x_1, \ldots, x_n)$ and $\lambda = (\lambda_1, \ldots, \lambda_p)$ from the following system of polynomial equations:

$$c_0 \;=\; \langle c, x\rangle \;\text{ and }\; h_1(x) = \cdots = h_p(x) = 0 \;\text{ and }\; c \;=\; \lambda_1 \nabla_x h_1 + \cdots + \lambda_p \nabla_x h_p.$$

Example 3.3. (Example 2.2 cont.) We consider (3.1) with $n = 2$, $p = 1$ and $h_1 = x_1^4 + x_2^4 - 1$. The KKT equations for maximizing the function

$$c_0 \;=\; c_1 x_1 + c_2 x_2 \tag{3.3}$$

over the "TV screen" curve $X = V(h_1)$ are

$$c_1 \;=\; \lambda_1 \cdot 4x_1^3, \quad c_2 \;=\; \lambda_1 \cdot 4x_2^3, \quad x_1^4 + x_2^4 = 1. \tag{3.4}$$

We eliminate the three unknowns x_1, x_2, λ_1 from the system of four polynomial equations in (3.3) and (3.4). The result is the polynomial $\Phi(-c_0, c_1, c_2)$ of degree 12 which expresses the optimal value c_0^* as an algebraic function of c_1 and c_2. We note that $\Phi(1, a, b)$ is precisely the polynomial in (2.6). □

It is natural to ask what happens with Theorem 3.2 when X fails to be smooth or compact, or if there are additional inequality constraints. Let us first consider the case when X is no longer smooth, but still compact. Now, X_{reg} is a proper (open, dense) subset of X. The optimal value function c_0^* for the problem

(3.1) is still perfectly well-defined on all of $\mathbb{R}^n$, and it is still an algebraic function of $c_1, \dots, c_n$. However, the polynomial Φ that represents c_0^* may now have more factors than just the equation of the dual variety X^*.

EXAMPLE 3.4. Let $n = 2$ and $p = 1$ as in Example 3.3, but now we consider a singular quartic. The *bicuspid curve*, shown in Figure 4, has the equation

$$h_1 \;=\; (x_1^2 - 1)(x_1 - 1)^2 + (x_2^2 - 1)^2.$$

The algebraic degree of optimizing a linear function $c_1x_1 + c_2x_2$ over $X = V(h_1)$ equals 8.

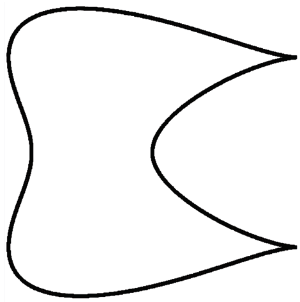

Fig. 4: The bicuspid curve in Example 3.4.

The optimal value function $c_0^* = c_0^*(c_1, c_2)$ is represented by

$$\Phi \;=\; \big(c_0 - c_1 + c_2\big)\cdot\big(c_0 - c_1 - c_2\big)\cdot\big(16c_0^6 - 48(c_1^2 + c_2^2)c_0^4 + (24c_1^2c_2^2 + 21c_2^4 + 64c_1^4)c_0^2 + (54c_1c_2^4 + 32c_1^5)c_0 + 8c_1^4c_2^2 - 3c_1^2c_2^4 + 11c_2^6\big).$$

The first two linear factors correspond to the singular points of the bicuspid curve X , and the larger factor of degree six represents the dual curve X^*. □

This example shows that, when X has singularities, it does not suffice to just dualize the variety X but we must also dualize the singular locus of X. This process is recursive, and we must also consider the singular locus of the singular locus etc. We believe that, in order to characterize the value function Φ, it always suffices to dualize all irreducible varieties occurring in a *Whitney stratification* of X but this has not been worked out yet. In our view, this topic requires more research, both on the theoretical side and on the computational side. The following result is valid for any variety X in $\mathbb{R}^n$.

COROLLARY 3.5. *If the dual variety of X is a hypersurface then its defining polynomial contributes a factor to the value function of the problem (3.1).*

This result can be extended to an arbitrary optimization problem of the form (2.7). We obtain a similar characterization of the optimal value c_0^* as a semi-algebraic function of $c_1, c_2, \ldots, c_n$ by eliminating all primal variables $x_1, \ldots, x_n$ and all dual (optimization) variables x, λ, μ from the KKT equations. Again, the optimal value function is represented by a unique square-free polynomial $\Phi(c_0, c_1, \ldots, c_n)$, and each factor of this polynomial is the dual hypersurface Y^* of some variety Y that is obtained from X by setting $g_i(x) = 0$ for some of the inequality constraints, by recursively passing to singular loci. In Section 5 we shall explore this for semidefinite programming.

We close this section with a simple example involving A-discriminants.

Example 3.6. Consider the calculus exercise of minimizing a polynomial

$$q(t) \quad = \quad c_1 t + c_2 t^2 + c_3 t^3 + c_4 t^4$$

of degree four over the real line $\mathbb{R}$. Equivalently, we wish to minimize

$$c_0 \;=\; c_1 x_1 + c_2 x_2 + c_3 x_3 + c_4 x_4$$

over the *rational normal curve* $X_A \cap \{x_0 = 1\} \;=\; V(x_1^2 - x_2, x_1^3 - x_3, x_1^4 - x_4)$, seen in Example 2.9. The optimal value function c_0^* is given by the equation $\Delta_A(-c_0, c_1, c_2, c_3, c_4) = 0$, where Δ_A is the discriminant in (2.11). Hence the algebraic degree of this optimization problem is equal to three. □

4 – An Algebraic View of Convex Hulls

The problem of optimizing arbitrary linear functions over a given subset of $\mathbb{R}^n$, discussed in the previous section, leads naturally to the geometric question of how to represent the convex hull of that subset. In this section we explore this question from an algebraic perspective. To be precise, we shall study the algebraic boundary $\partial_a P$ of the convex hull $P = \operatorname{conv}(X)$ of a compact real algebraic variety X in $\mathbb{R}^n$. Biduality of projective varieties (Theorem 2.6) will play an important role in understanding the structure of $\partial_a P$. The results to be presented are drawn from [24, 25]. In Section 6 we shall discuss the alternative representation of P as a spectrahedral shadow.

We begin with the seemingly easy example of a plane quartic curve.

Example 4.1. We consider the following smooth compact plane curve

$$(4.1)\quad X \;=\; \big\{(x,y) \in \mathbb{R}^2 \mid 144x^4 + 144y^4 - 225(x^2+y^2) + 350x^2y^2 + 81 \;=\; 0\,\big\}.$$

This curve is known as the *Trott curve*. It was first constructed by Michael Trott in [28], and is illustrated above in Figure 5. A classical result of algebraic geometry states that a general quartic curve in the complex projective plane $\mathbb{P}^2$

has 28 bitangent lines, and the Trott curve X is an instance where all 28 lines are real and have a coordinatization in terms of radicals over $\mathbb{Q}$. Four of the 28 bitangents form edges of $\operatorname{conv}(X)$. These special bitangents are

$$\{(x,y) \in \mathbb{R}^2 \mid \pm x \pm y = \gamma\}, \quad \text{where } \gamma = \frac{\sqrt{48050 + 434\sqrt{9889}}}{248} = 1.2177...$$

The boundary of $\operatorname{conv}(X)$ alternates between these four edges and pieces of the curve X. The eight transition points have the floating point coordinates

$$(\pm 0.37655..., \pm 0.84122...) \ , \ \ (\pm 0.84122..., \pm 0.37655...).$$

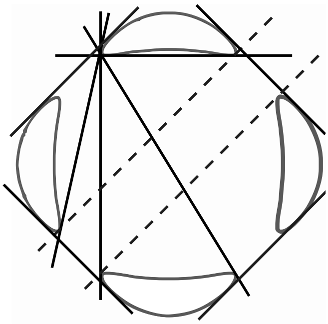

Fig. 5: A quartic curve in the plane can have up to 28 real bitangents.

These coordinates lie in the field $\mathbb{Q}(\gamma)$ and we invite the reader to write them in the form $q_1 + q_2\gamma$ where $q_i \in \mathbb{Q}$. The $\mathbb{Q}$-Zariski closure of the 4 edge lines of $\operatorname{conv}(X)$ is a curve Y of degree 8. Its equation has two irreducible factors:

$$(992x^4 - 3968x^3y + 5952x^2y^2 - 3968xy^3 + 992y^4 - 1550x^2 + 3100xy - 1550y^2 + 117)$$
$$(992x^4 + 3968x^3y + 5952x^2y^2 + 3968xy^3 + 992y^4 - 1550x^2 - 3100xy - 1550y^2 + 117)$$

Each reduces over $\mathbb{R}$ to four parallel lines, two of which contribute to the boundary. The point of this example is to stress the role of the (arithmetic of) bitangents in any exact description of the convex hull of a plane curve. □

We now present a general formula for the algebraic boundary of the convex hull of a compact variety X in $\mathbb{R}^n$. The key observation is that the algebraic boundary of $P = \operatorname{conv}(X)$ will consist of different types of components, resulting from planes that are simultaneously tangent at k different points of X, for various values of the integer k. For the Trott curve X in Example 4.1, the relevant integers were $k = 1$ and $k = 2$, and we demonstrated that the algebraic boundary of its convex hull P is a reducible curve of degree 12:

$$\partial_a(P) = X \cup Y. \tag{4.2}$$

In the following definitions we regard X as a complex projective variety in $\mathbb{P}^n$.

Let $X^{[k]}$ be the variety in the dual projective space $(\mathbb{P}^n)^*$ which is the closure of the set of all hyperplanes that are tangent to X at k regular points which span a $(k-1)$-plane in $\mathbb{P}^n$. This definition makes sense for $k = 1, 2, \dots, n$. Note that $X^{[1]}$ coincides with the dual variety X^*, and $X^{[2]}$ parametrizes all hyperplanes that are tangent to X at two distinct points. Typically, $X^{[2]}$ is an irreducible component of the singular locus of $X^* = X^{[1]}$. We have the following nested chain of projective varieties in the dual space:

$$X^{[n]} \subseteq X^{[n-1]} \subseteq \cdots \subseteq X^{[2]} \subseteq X^{[1]} \subseteq (\mathbb{P}^n)^*.$$

We now dualize each of the varieties in this chain. The resulting varieties $(X^{[k]})^*$ live in the primal projective space $\mathbb{P}^n$. For $k = 1$ we return to our original variety, i.e., we have $(X^{[1]})^* = X$ by biduality (Theorem 2.6). In the following result we assume that X is smooth as a complex variety in $\mathbb{P}^n$, and we require one technical hypothesis concerning tangency of hyperplanes.

Theorem 4.2. [25, Theorem 1.1] *Let X be a smooth and compact real algebraic variety that affinely spans $\mathbb{R}^n$, and such that only finitely many hyperplanes are tangent to X at infinitely many points. The algebraic boundary $\partial_a P$ of its convex hull, $P = conv(X)$, can be computed by biduality as follows:*

$$\partial_a P \subseteq \bigcup_{k=1}^{n} (X^{[k]})^*. \tag{4.3}$$

Since $\partial_a P$ is pure of codimension one, in the union we only need indices k having property that $(X^{[k]})^*$ is a hypersurface in $\mathbb{P}^n$. As argued in [25], this leads to the following lower bound on the relevant values to be considered:

$$k \geq \left\lceil \frac{n}{\dim(X) + 1} \right\rceil. \tag{4.4}$$

The formula (4.3) computes the algebraic boundary $\partial_a P$ in the following sense. For each relevant k we check whether $(X^{[k]})^*$ is a hypersurface, and, if yes, we determine its irreducible components (over the field K of interest). For each component we then check, usually by means of numerical computations, whether it meets the boundary ∂P in a regular point. The irreducible hypersurfaces which survive this test are precisely the components of $\partial_a X$.

Example 4.3. When X is a plane curve in $\mathbb{R}^2$ then (4.3) says that

$$\partial_a P \subseteq X \cup (X^{[2]})^*. \tag{4.5}$$

Here $X^{[2]}$ is the set of points in $(\mathbb{P}^2)^*$ that are dual to the bitangent lines of X, and $(X^{[2]})^*$ is the union of those lines in $\mathbb{P}^2$. If we work over $K = \mathbb{Q}$ and the curve X is general enough then we expect equality to hold in (4.5). For special curves the inclusion can be strict. This happens for the Trott curve (4.1) since Y is a proper subset of $(X^{[2]})^*$. Namely, Y consists of two of the six $\mathbb{Q}$-components of $(X^{[2]})^*$. However, a small perturbation of the coefficients in (4.1) leads to a curve X with equality in (4.5), as the relevant Galois group acts transitively on the 28 points in $X^{[2]}$ for general quartics X. Now, the algebraic boundary over $\mathbb{Q}$ is a reducible curve of degree $32 = 28 + 4$. ∎

If we are given the variety X in terms of equations or in parametric form, then we can compute equations for $X^{[k]}$ by an elimination process similar to our computation of the dual variety X^*. However, expressing the tangency condition at k different points requires a larger number of additional variables (which need to be eliminated afterwards) and thus the computations are quite involved. The subsequent step of dualizing $X^{[k]}$ to get the right hand side of (4.3) is even more forbidding. The resulting hypersurfaces $(X^{[k]})^*$ tend to have high degree and their defining polynomials are very large when $n \geq 3$.

The article [24] offers a detailed study of the case when X is a space curve in $\mathbb{R}^3$. Here the lower bound (4.4) tells us that $\partial_a X \subseteq (X^{[2]})^* \cup (X^{[3]})^*$. The surface $(X^{[2]})^*$ is the *edge surface* of the curve X, and $(X^{[3]})^*$ is the union of all *tritangent planes* of X. The following example illustrates these objects.

Example 4.4. We consider the trigonometric curve X in $\mathbb{R}^3$ which has the parametrization $x = \cos(\theta)$, $y = \cos(2\theta)$, $z = \sin(3\theta)$. This is an algebraic curve of degree six. Its implicit representation equals $X = V(h_1, h_2)$ where

$$h_1 \;=\; 2x^2 - y - 1 \;\text{ and }\; h_2 \;=\; 4y^3 + 2z^2 - 3y - 1.$$

The edge surface $(X^{[2]})^*$ has three irreducible components. Two of the components are the quadric $V(h_1)$ and the cubic $V(h_2)$. The third and most interesting component of $(X^{[2]})^*$ is the surface of degree 16 with equation $h_3 =$

$$\begin{aligned}
&- 419904x^{14}y^2 + 664848x^{12}y^4 - 419904x^{10}y^6 + 132192x^8y^8 - 20736x^6y^{10} + 1296x^4y^{12}\\
&- 46656x^{14}z^2 + 373248x^{12}y^2z^2 - 69984x^{10}y^4z^2 - 22464x^8y^6z^2 + 4320x^6y^8z^2 + 31104x^{12}z^4\\
&+ 5184x^{10}y^2z^4 + 4752x^8y^4z^4 + 1728x^{10}z^6 + 699840x^{14}y - 46656x^{12}y^3 - 902016x^{10}y^5\\
&+ 694656x^8y^7 - 209088x^6y^9 - 1150848x^{10}y^3z^2 + 279936x^8y^5z^2 + 17280x^6y^7z^2 - 4032x^4y^9z^2\\
&- 98496x^{10}yz^4 + 27072x^4y^{11} - 1152x^2y^{13} - 419904x^{12}yz^2 - 25920x^8y^3z^4 - 4608x^6y^5z^4\\
&- 1728x^8yz^6 - 291600x^{14} - 169128x^{12}y^2 - 256608x^{10}y^4 + 956880x^8y^6 - 618192x^6y^8\\
&+ 148824x^4y^{10} - 13120x^2y^{12} + 256y^{14} + 392688x^{12}z^2 + 671976x^{10}y^2z^2 + 1454976x^8y^4z^2\\
&- 292608x^6y^6z^2 - 4272x^4y^8z^2 + 1016x^2y^{10}z^2 - 116208x^{10}z^4 + 135432x^8y^2z^4 + 18144x^6y^4z^4\\
&+ 1264x^4y^6z^4 - 5616x^8z^6 + 504x^6y^2z^6 - 1108080x^{12}y + 925344x^{10}y^3 + 215136x^8y^5\\
&- 672192x^6y^7 + 331920x^4y^9 - 54240x^2y^{11} + 2304y^{13} + 273456x^{10}yz^2 + 282528x^8y^3z^2\\
&- 1185408x^6y^5z^2 + 149376x^4y^7z^2 - 368x^2y^9z^2 - 32y^{11}z^2 + 273456x^8yz^4 - 67104x^6y^3z^4
\end{aligned}$$

$$
\begin{aligned}
&- 4704x^4y^5z^4 - 64x^2y^7z^4 + 4752x^6yz^6 - 32x^4y^3z^6 + 747225x^{12} + 636660x^{10}y^2\\
&- 908010x^8y^4 - 65340x^6y^6 + 291465x^4y^8 - 101712x^2y^{10} + 8256y^{12} - 818100x^{10}z^2\\
&- 1405836x^8y^2z^2 - 905634x^6y^4z^2 + 583824x^4y^6z^2 - 39318x^2y^8z^2 + 368y^{10}z^2 + 193806x^8z^4\\
&- 282996x^6y^2z^4 + 15450x^4y^4z^4 + 716x^2y^6z^4 + y^8z^4 + 6876x^6z^6 - 1140x^4y^2z^6 + 2x^2y^4z^6\\
&+ x^4z^8 + 507384x^{10}y - 809568x^8y^3 + 569592x^6y^5 - 27216x^4y^7 - 71648x^2y^9 + 13952y^{11}\\
&+ 555768x^8yz^2 + 869040x^6y^3z^2 + 688512x^4y^5z^2 - 154128x^2y^7z^2 + 4416y^9z^2 - 343224x^6yz^4\\
&+ 127360x^4y^3z^4 - 1656x^2y^5z^4 - 64y^7z^4 - 4536x^4yz^6 + 48x^2y^3z^6 - 775170x^{10} - 191808x^8y^2\\
&+ 599022x^6y^4 - 245700x^4y^6 + 31608x^2y^8 + 7872y^{10} + 765072x^8z^2 + 589788x^6y^2z^2\\
&- 66066x^4y^4z^2 - 234252x^2y^6z^2 + 16632y^8z^2 - 173196x^6z^4 + 248928x^4y^2z^4 - 26158x^2y^4z^4\\
&- 32y^6z^4 - 3904x^4z^6 + 804x^2y^2z^6 + 2y^4z^6 - 2x^2z^8 + 5832x^8y + 98280x^6y^3 - 219456x^4y^5\\
&+ 72072x^2y^7 - 8064y^9 - 724032x^6yz^2 - 515760x^4y^3z^2 - 99672x^2y^5z^2 + 29976y^7z^2\\
&+ 225048x^4yz^4 - 76216x^2y^3z^4 + 1912y^5z^4 + 1696x^2yz^6 - 32y^3z^6 + 411345x^8 - 66096x^6y^2\\
&- 62532x^4y^4 + 29388x^2y^6 - 11856y^8 - 365346x^6z^2 + 19812x^4y^2z^2 + 104922x^2y^4z^2 + 24636y^6z^2\\
&+ 85090x^4z^4 - 104580x^2y^2z^4 + 8282y^4z^4 + 1014x^2z^6 - 144y^2z^6 + z^8 - 39744x^6y + 61992x^4y^3\\
&+ 2304x^2y^5 + 576y^7 + 305328x^4yz^2 + 86640x^2y^3z^2 + 960y^5z^2 - 73480x^2yz^4 + 16024y^3z^4\\
&- 200yz^6 - 114966x^6 + 24120x^4y^2 - 5958x^2y^4 + 6192y^6 + 85494x^4z^2 - 39696x^2y^2z^2\\
&- 11970y^4z^2 - 21610x^2z^4 + 16780y^2z^4 - 94z^6 - 3672x^4y - 11024x^2y^3 + 272y^5\\
&- 46904x^2yz^2 - 4632y^3z^2 + 9368yz^4 + 15246x^4 - 84x^2y^2 - 1908y^4 - 6892x^2z^2\\
&+ 2204y^2z^2 + 2215z^4 + 3216x^2y + 168y^3 + 904yz^2 - 664x^2 + 292y^2 - 282z^2 - 96y + 9.
\end{aligned}
$$

The boundary of $P = \mathrm{conv}(X)$ contains patches from all three surfaces $V(h_1)$, $V(h_2)$ and $V(h_3)$. There are also two triangles, with vertices at $(\sqrt{3}/2, 1/2, \pm 1)$, $(\sqrt{3}/2, 1/2, \pm 1)$ and $(0, -1, \pm 1)$. They span two of the tritangent planes of X, namely, $z = 1$ and $z = -1$. The union of all tritangent planes equals $(X^{[3]})^*$. Only one triangle is visible in Figure 6. It is colored in white. The curved black patch adjacent to one of the edges of the triangle is given by the cubic h_3, while the other two edges of the triangle lie in the degree 16 surface $V(h_3)$. The curve X has two singular points at $(x, y, z) = (\pm 1/2, -1/2, 0)$. Around these two singular points, the boundary is given by four alternating patches from the quadric $V(h_1)$ highlighted in dark grey and the degree 16 surface $V(h_3)$ in light gray. We conclude that the edge surface $(X^{[2]})^* = V(h_1h_2h_3)$ is reducible of degree $21 = 2 + 3 + 16$, and the algebraic boundary $\partial_a(P)$ is a reducible surface of degree $23 = 2 + 21$. □

In our next example we examine the convex hull of space curves of degree four that are obtained as the intersection of two quadratic surfaces in $\mathbb{R}^3$.

Example 4.5. Let $X = V(h_1, h_2)$ be the intersection of two quadratic surfaces in 3-space. We assume that X has no singularities in $\mathbb{P}^3$. Then X is a curve of genus one. According to recent work of Scheiderer [27], the convex body $P = \mathrm{conv}(X)$ can be represented exactly using Lasserre relaxations, a

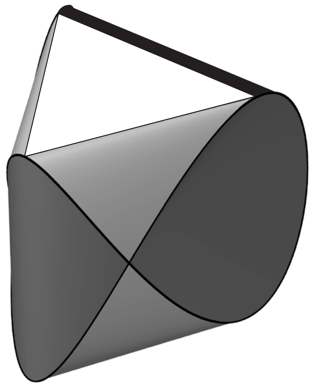

Fig. 6: The convex hull of the curve $(\cos(\theta), \cos(2\theta), \sin(3\theta))$ in $\mathbb{R}^3$.

topic we shall return to when discussing spectrahedral shadows in Section 6. If we are willing to work over $\mathbb{R}$ then P is in fact a spectrahedron, as shown in [24, Example 2.3]. We here derive that representation for a concrete example.

Lazard *et al.* [18, §8.2] examine the curve X cut out by the two quadrics

$$h_1 = x^2 + y^2 + z^2 - 1 \quad \text{and} \quad h_2 = 19x^2 + 22y^2 + 21z^2 - 20.$$

Figure 7 shows the two components of X on the unit sphere $V(h_1)$.

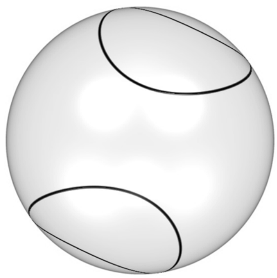

Fig. 7: The curve on the unit sphere discussed in Example 4.5 and 4.6.

The dual variety X^* is a surface of degree 8 in $(\mathbb{P}^3)^*$. The singular locus of X^* contains the curve $X^{[2]}$ which is the union of four quadratic curves. The duals of these four plane curves are the singular quadratic surfaces defined by

$$h_3 = x^2 - 2y^2 - z^2,\ h_4 = 2x^2 - y^2 - 1,\ h_5 = 3y^2 + 2z^2 - 1,\ h_6 = 3x^2 + z^2 - 2.$$

The edge surface of X is the union of these four quadrics:

$$(X^{[2]})^* \;=\; V(h_3) \cup V(h_4) \cup V(h_5) \cup V(h_6).$$

The algebraic boundary of P consists of the last two among these quadrics:

$$\partial_a P \;=\; V(h_5) \cup V(h_6).$$

These two quadrics are convex. From this we derive a representation of P as a spectrahedron by applying Schur complements to the quadrics h_5 and h_6:

$$P \quad = \quad \left\{ (x,y,z) \in \mathbb{R}^3 \;\middle|\; \begin{pmatrix} 1{+}\sqrt{3}y & \sqrt{2}z & 0 & 0 \\ \sqrt{2}z & 1{-}\sqrt{3}y & 0 & 0 \\ 0 & 0 & \sqrt{2}{-}x & \sqrt{3}x \\ 0 & 0 & \sqrt{3}x & \sqrt{2}{+}x \end{pmatrix} \succeq 0 \right\}.$$

5 – Spectrahedra and Semidefinite Programming

Spectrahedra and semidefinite programming (SDP) have already surfaced a few times in our discussion. In this section we take a systematic look at these topics and their dualities. We write $\mathcal{S}^n$ for the space of real symmetric $n{\times}n$-matices and $\mathcal{S}^n_+$ for the cone of positive semidefinite matrices in $\mathcal{S}^n \simeq \mathbb{R}^{\binom{n+1}{2}}$. This cone is self-dual with respect to the inner product $\langle U, V\rangle = \mathrm{trace}(U \cdot V)$.

A *spectrahedron* is the intersection of the cone $\mathcal{S}^n_+$ with an affine subspace

$$\mathcal{K} \;=\; C + \underbrace{\mathrm{Span}(A_1, A_2, \ldots, A_m)}_{\mathcal{W}}.$$

Here $\mathcal{W}$ is a linear subspace of dimension m in $\mathcal{S}^n$, and the spectrahedron is

$$P = \big\{\, x \in \mathbb{R}^m \mid C - \sum_{i=1}^{m} x_i A_i \succeq 0 \big\} \simeq\; \mathcal{K} \cap \mathcal{S}^n_+. \tag{5.1}$$

We shall assume that C is positive definite or, equivalently, that $0 \in \mathrm{int}(P)$. The dual body to our spectrahedron is written in the coordinates on $\mathbb{R}^m$ as

$$P^{\Delta} \;=\; \{\, y \in \mathbb{R}^m \mid \forall\, x \in P \text{ with } \langle y, x\rangle \le 1 \}\,.$$

We can express P^{Δ} as a projection of the $\binom{n+1}{2}$-dimensional spectrahedron

$$Q = \{\, U \in \mathcal{S}^n_+ \mid \langle U, C\rangle \le 1 \,\}. \tag{5.2}$$

Namely, consider the linear map dual to the inclusion of the linear subspace $\mathcal{W} = \mathrm{Span}(A_1, A_2, \ldots, A_m)$ in the $\binom{n+1}{2}$-dimensional real vector space $\mathcal{S}^n$:

$$\begin{aligned} \pi_{\mathcal{W}} : \mathcal{S}^n \;&\rightarrow\; \mathcal{S}^n/\mathcal{W}^{\perp} \;\simeq\; \mathbb{R}^m \\ U \;&\mapsto\; \big(\langle U, A_1\rangle, \langle U, A_2\rangle, \ldots, \langle U, A_m\rangle\big). \end{aligned}$$

REMARK 5.1. The convex body P^{Δ} dual to the spectrahedron P is affinely isomorphic to the closure of the image of the spectrahedron (5.2) under the linear map $\pi_{\mathcal{W}}$, i.e. $P^{\Delta} \simeq \overline{\pi_{\mathcal{W}}(Q)}$. This result is due to Ramana and Goldman [23]. In summary, while the dual to a spectrahedron is generally not a spectrahedron, it is always a *spectrahedral shadow*. See also Theorem 6.1.

EXAMPLE 5.2. The *elliptope* $\mathcal{E}_n$ is the spectrahedron consisting of all *correlation matrices* of size n, see [15]. These are the positive semidefinite symmetric $n{\times}n$-matrices whose diagonal entries are 1. We consider the case $n = 3$:

$$\mathcal{E}_3 \;=\; \left\{ (x,y,z) \in \mathbb{R}^3 \;\middle|\; \begin{pmatrix} 1 & x & y \\ x & 1 & z \\ y & z & 1 \end{pmatrix} \succeq 0 \right\}. \tag{5.3}$$

This spectrahedron of dimension $m = 3$ is shown on the left in Figure 8. The algebraic boundary of $\mathcal{E}_3$ is the cubic surface X defined by the vanishing of the 3×3-determinant in (5.3). That surface has four isolated singular points

$$X_{\mathrm{sing}} = \{(1,1,1), (1,-1,-1), (-1,1,-1), (-1,-1,1)\}.$$

The six edges of the tetrahedron $\mathrm{conv}(X_{\mathrm{sing}})$ are edges of the elliptope $\mathcal{E}_3$. The dual body, shown on the right of Figure 8, is the spectrahedral shadow

$$\mathcal{E}_3^{\Delta} \;=\; \left\{ (a,b,c) \in \mathbb{R}^3 \;\middle|\; \exists\, u,v \in \mathbb{R} \,:\, \begin{pmatrix} u & a & b \\ a & v & c \\ b & c & 2{-}u{-}v \end{pmatrix} \succeq 0 \right\}. \tag{5.4}$$

The algebraic boundary of $\mathcal{E}_3^{\Delta}$ can be computed by the following method. We form the ideal generated by the determinant in (5.4) and its derivatives with respect to u and v, and we eliminate u, v. This results in the polynomial

$$(a^2b^2 + b^2c^2 + a^2c^2 - 2abc)(a{+}b{+}c{+}1)(a{-}b{-}c{+}1)(a{-}b{+}c{-}1)(a{+}b{-}c{-}1).$$

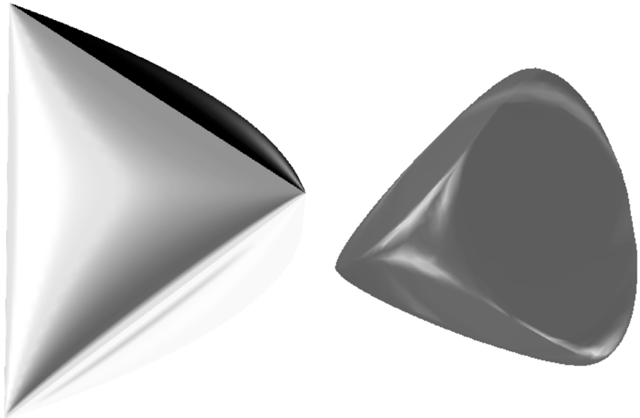

Fig. 8: The elliptope $P = \mathcal{E}_3$ and its dual convex body P^Δ.

The first factor is the equation of the *Steiner quartic* surface X^* which is dual to *Cayley cubic* surface $X = \partial_a \mathcal{E}_3$. The four linear factors represent the arrangement $(X_{\text{sing}})^*$ of the four planes dual to the four singular points.

Thus the algebraic boundary of the dual body $\mathcal{E}_3^\Delta$ is the reducible surface

$$\partial_a \mathcal{E}_3^\Delta \;=\; X^* \cup (X_{\text{sing}})^* \quad \subset \quad (\mathbb{P}^3)^*. \tag{5.5}$$

We note that $\mathcal{E}_3^\Delta$ is not a spectrahedron as it fails to be a *basic semi-algebraic set*, i.e. no polynomial ϕ satisfies $\mathcal{E}_3^\Delta = \{(a,b,c) \in \mathbb{R}^3 \,:\, \phi(a,b,c) \geq 0\}$. This is seen from the fact that the Steiner surface intersects the interior of $\mathcal{E}_3^\Delta$. □

Semidefinite programming (SDP) is the branch of convex optimization that is concerned with maximizing a linear function b over a spectrahedron:

$$\begin{aligned} p^* \;&=\; \underset{x}{\text{Maximize}}\ \langle b, x\rangle \\ &\text{subject to}\ \ x \in P. \end{aligned} \tag{5.6}$$

Here P is as in (5.1). As the semidefiniteness of a matrix is equivalent to the simultaneous non-negativity of its principal minors, SDP is an instance of the polynomial optimization problem (2.7). Lagrange duality theory applies here by [4, §5]. We shall derive the optimization problem dual to (5.6) from

$$d^* \;=\; \underset{\lambda}{\text{Minimize}}\ \lambda \ \text{ subject to }\ \frac{1}{\lambda} b \,\in\, P^\Delta\,. \tag{5.7}$$

Since we assumed $0 \in \text{int}(P)$, strong duality holds and we have $p^* = d^*$.

The fact that P^{Δ} is a spectrahedral shadow implies that the dual optimization problem is again a semidefinite optimization problem. In light of Remark 5.1, the condition $\frac{1}{\lambda}b \in P^{\Delta}$ can be expressed as follows:

$$\exists\, U \; U \succeq 0\,,\; \langle C, U\rangle \leq 1 \text{ and } b_i = \lambda\langle A_i, U\rangle \text{ for } i = 1, 2, \dots, m.$$

Since the optimal value of (5.7) is attained at the boundary of P^{Δ}, we can here replace the condition $\langle C, U\rangle \leq 1$ with $\langle C, U\rangle = 1$. This is in fact what was done to obtain (5.4). If we now set $Y = \lambda U$, then (5.7) translates into

$$\begin{aligned} d^* \;=\; & \underset{Y \in \mathcal{S}^n_+}{\text{Minimize}} \quad \langle C, Y\rangle \\ & \text{subject to} \quad \langle A_i, Y\rangle \;=\; b_i \text{ for } i = 1, \dots, m \\ & \text{and } Y \succeq 0. \end{aligned} \tag{5.8}$$

We recall that $\mathcal{W} = \text{Span}(A_1, A_2, \dots, A_m)$ and we fix any matrix $B \in \mathcal{S}^n$ with $\langle A_i, B\rangle = b_i$ for $i = 1, \dots, m$. Then (5.8) can be written as follows:

$$d^* \;=\; \underset{Y \in \mathcal{S}^n_+}{\text{Minimize}}\, \langle C, Y\rangle \text{ subject to } Y \in (B + \mathcal{W}^{\perp}) \cap \mathcal{S}^n_+ \tag{5.9}$$

The following reformulation of (5.6) highlights the symmetry between the primal and dual formulations of our semidefinite programming problem:

$$p^* \;=\; \underset{X \in \mathcal{S}^n_+}{\text{Maximize}}\, \langle B, C - X\rangle \text{ subject to } X \in (C + \mathcal{W}) \cap \mathcal{S}^n_+ \tag{5.10}$$

Then the following variant of the Karush-Kuhn-Tucker Theorem holds:

THEOREM 5.3. [4, §5.9.2] *If both the primal problem* (5.10) *and its dual* (5.9) *are strictly feasible, then the KKT conditions take the following form:*

$$\begin{aligned} X &\in (C + \mathcal{W}) \cap \mathcal{S}^n_+ \\ Y &\in (B + \mathcal{W}^{\perp}) \cap \mathcal{S}^n_+ \\ X \cdot Y &= 0 \quad \textit{(complementary slackness).} \end{aligned}$$

These conditions characterize all the pairs (X, Y) *of optimal solutions.*

This theorem can be related to the general optimality conditions (2.8) by regarding the entries of $Y \in \mathcal{S}^n$ as the (Lagrangian) dual variables to the positive semidefinite constraint $X = C - \sum_{i=1}^m x_i A_i \succeq 0$. The three conditions are both necessary and sufficient since semidefinite programming is a convex problem and every local optimum is also a global optimal solution.

In order to study algebraic and geometric properties of SDP, we will relax the conic inequalities $X, Y \in \mathcal{S}^n_+$ and focus only on the *KKT equations*

$$X \in C + \mathcal{W}\,,\ Y \in B + \mathcal{W}^\perp \text{ and } X \cdot Y = 0. \tag{5.11}$$

Given B, C and $\mathcal{W}$, our problem is to solve the polynomial equations (5.11). The theorem ensures that, among its solutions (X, Y), there is precisely one pair of positive semidefinite matrices. That pair is the one desired in SDP.

Example 5.4. Consider the problem of minimizing a linear function $Y \mapsto \langle C, Y\rangle$ over the set of all correlation matrices Y, that is, over the elliptope $\mathcal{E}_n$ of Example 5.2. Here $m = n$, B is the identity matrix, $\mathcal{W}$ is the space of all diagonal matrices, and $\mathcal{W}^\perp$ consists of matrices with zero diagonal. The dual problem is to maximize the trace of $C - X$ over all matrices $X \in \mathcal{S}^n_+$ such that $C - X$ is diagonal. Equivalently, we seek to find the minimum trace t^* of any positive semidefinite matrix that agrees with C in its off-diagonal entries.

For $n = 4$, the KKT equations (5.11) can be written in the form

$$X \cdot Y \ = \ \begin{pmatrix} x_1 & c_{12} & c_{13} & c_{14} \\ c_{12} & x_2 & c_{23} & c_{24} \\ c_{13} & c_{23} & x_3 & c_{34} \\ c_{14} & c_{24} & c_{34} & x_4 \end{pmatrix} \cdot \begin{pmatrix} 1 & y_{12} & y_{13} & y_{14} \\ y_{12} & 1 & y_{23} & y_{24} \\ y_{13} & y_{23} & 1 & y_{34} \\ y_{14} & y_{24} & y_{34} & 1 \end{pmatrix} \ = \ 0. \tag{5.12}$$

This is a system of 10 quadratic equations in 10 unknowns. For general values of the 6 parameters c_{ij}, these equations have 14 solutions. Eight of these solutions have $\mathrm{rank}(X) = 3$ and $\mathrm{rank}(Y) = 1$ and they are defined over $\mathbb{Q}(c_{ij})$. The other six solutions form an irreducible variety over $\mathbb{Q}(c_{ij})$ and they satisfy $\mathrm{rank}(X) = \mathrm{rank}(Y) = 2$. This case distinction reflects the boundary structure of the dual body to the six-dimensional elliptope $\mathcal{E}_4$:

$$\partial_a \mathcal{E}_4^\Delta \quad = \quad \{\mathrm{rank}(Y) \le 2\}^* \,\cup\, \{\mathrm{rank}(Y) = 1\}^*. \tag{5.13}$$

Indeed, the boundary of $\mathcal{E}_4$ is the quartic hypersurface $\{\mathrm{rank}(Y) \le 3\}$, its singular locus is the degree 10 threefold $\{\mathrm{rank}(Y) \le 2\}$, and, finally, the singular locus of that threefold consists of eight matrices of rank one:

$$\{\mathrm{rank}(Y) = 1\} \ = \ \big\{ (u_1, u_2, u_3, u_4)^T \cdot (u_1, u_2, u_3, u_4) \ : \ u_i \in \{-1, +1\} \big\}.$$

The last two strata are dual to the hypersurfaces in (5.13). The second component in (5.13) consists of eight hyperplanes, while the first component is irreducible of degree 18. The corresponding projective hypersurface is defined by an irreducible homogeneous polynomial of degree 18 in seven unknowns $c_{12}, c_{13}, c_{14}, c_{23}, c_{24}, c_{34}, t^*$. That polynomial has degree 6 in the special unknown t^*. Hence, the algebraic degree of our SDP is 6 when $\text{rank}(Y) = 2$. ∎

In algebraic geometry, it is natural to regard the matrix pairs (X, Y) as points in the product of projective spaces $\mathbb{P}(\mathcal{S}^n) \times \mathbb{P}(\mathcal{S}^n)^*$. This has the advantage that solutions of (5.11) are invariant under scaling, i.e. whenever (X, Y) is a solution, then so is $(\lambda X, \mu Y)$ for any nonzero $\lambda, \mu \in \mathbb{R}$. In that setting, there are no worries about complications due to solutions at infinity.

For the algebraic formulation we assume that, without loss of generality,

$$b_1 = 1, b_2 = 0, b_3 = 0, \ldots, b_m = 0.$$

This means that $\langle A_1, X\rangle = 1$ plays the role of the homogenizing variable. Our SDP instance is specified by two linear subspaces of symmetric matrices:

$$\mathcal{L} = \text{Span}(A_2, A_3, \ldots, A_m) \ \subset\ \mathcal{U} = \text{Span}(C, A_1, A_2, \ldots, A_m) \ \subset\ \mathcal{S}^n.$$

Note that we have the following identifications:

$$\mathbb{R}C + \mathcal{W} \ = \mathcal{U} \quad \text{and} \quad \mathbb{R}B + \mathcal{W}^\perp \ = \ \mathbb{R}B + (\mathcal{L}^\perp \cap A_1^\perp) \ = \ \mathcal{L}^\perp.$$

With the linear spaces $\mathcal{L} \subset \mathcal{U}$, we write the *homogeneous KKT equations* as

$$X \in \mathcal{U},\ Y \in \mathcal{L}^\perp \text{ and } X \cdot Y = 0. \tag{5.14}$$

Here is an *abstract definition of semidefinite programming* that might appeal to some of our readers: Given any flag of linear subspaces $\mathcal{L} \subset \mathcal{U} \subset \mathcal{S}^n$ with $\dim(\mathcal{U}/\mathcal{L}) = 2$, locate the unique semidefinite point in the variety (5.14). For instance, in Example 5.4 the space $\mathcal{L}$ consists of traceless diagonal matrices and $\mathcal{U}/\mathcal{L}$ is spanned by the unit matrix B and one off-diagonal matrix C. We seek to solve the matrix equation $X \cdot Y = 0$ where the diagonal entries of X are constant and the off-diagonal entries of Y are proportional to C.

The formulation (5.14) suggests that we study the variety $\{XY = 0\}$ for pairs of symmetric matrices X and Y. In [22, Eqn. (3.9)] it was shown that this variety has the following decomposition into irreducible components:

$$\{XY = 0\} \ = \ \bigcup_{r=1}^{n-1} \{XY = 0\}^r \quad \subset \ \mathbb{P}(\mathcal{S}^n) \times \mathbb{P}(\mathcal{S}^n)^*.$$

Here $\{XY = 0\}^r$ denotes the subvariety consisting of pairs (X, Y) where $\text{rank}(X) \leq r$ and $\text{rank}(Y) \leq n - r$. This is irreducible because, by Example 2.8, it is the conormal variety of the variety of symmetric matrices of rank $\leq r$. The KKT equations describe sections of these conormal varieties:

$$\{XY = 0\}^r \cap \big(\mathbb{P}(\mathcal{U}) \times \mathbb{P}(\mathcal{L}^\perp)\big). \tag{5.15}$$

All solutions of a semidefinite optimization problem (and thus also the boundary of a spectrahedron and its dual) can be characterized by rank conditions. The main result in [22] describes the case when the section in (5.15) is generic:

THEOREM 5.5. [22, Theorem 7] *For generic subspaces* $\mathcal{L} \subset \mathcal{U} \subset \mathcal{S}^n$ *with* $\dim(\mathcal{L}) = m-1$ *and* $\dim(\mathcal{U}) = m+1$*, the variety (5.15) is empty unless*

$$\binom{n-r+1}{2} \leq m \quad \textit{and} \quad \binom{r+1}{2} \leq \binom{n+1}{2} - m. \tag{5.16}$$

In that case, the variety (5.15) is reduced, nonempty and zero-dimensional and at each point the rank of X *and* Y *is* $n-r$ *and* r *respectively (strict complementarity). The cardinality of this variety depends only on* m, n *and* r.

The generic choice of subspaces $\mathcal{L} \subset \mathcal{U}$ corresponds to the assumption that our matrices $A_1, A_2, \ldots, A_m, B, C$ lie in a certain dense open subset in the space of all SDP instances. The inequalities (5.16) are known as *Pataki's inequalities.* If m and n are fixed then they give a lower bound and an upper bound for the possible ranks r of the optimal matrix of a generic SDP instance. The variety (5.15) represents all complex solutions of the KTT equations for such a generic SDP instance. Its cardinality, denoted $\delta(m,n,r)$, is known as the *algebraic degree of semidefinite programming.*

COROLLARY 5.6. *Consider the variety of symmetric* $n \times n$*-matrices of rank* $\leq r$ *that lie in the generic* m*-dimensional linear subspace* $\mathbb{P}(\mathcal{U})$ *of* $\mathbb{P}(\mathcal{S}^n)$*. Its dual variety is a hypersurface if and only if Pataki's inequalities (5.16) hold, and the degree of that hypersurface is* $\delta(m,n,r)$*, the algebraic degree of SDP.*

PROOF. The genericity of $\mathcal{U}$ ensures that $\{XY=0\}^r \cap (\mathbb{P}(\mathcal{U}) \times \mathbb{P}(\mathcal{U})^*)$ is the conormal variety of the given variety. We obtain its dual by projection onto the second factor $\mathbb{P}(\mathcal{U})^* = \mathbb{P}(\mathcal{S}^n/\mathcal{U}^\perp)$. The degree of the dual hypersurface is found by intersecting with a generic line. The line we take is $\mathbb{P}(\mathcal{L}^\perp/\mathcal{U}^\perp)$. That intersection corresponds to the second factor $\mathbb{P}(\mathcal{L}^\perp)$ in (5.15). ∎

We note that the symmetry in the equations (5.14) implies the duality

$$\delta\big(m,n,r\big) \quad = \quad \delta\Big(\binom{n+1}{2} - m, n, n-r\Big),$$

first shown in [22, Proposition 9]. See also [22, Table 2]. Bothmer and Ranestad [8] derived an explicit combinatorial formula for the algebraic degree of SDP. Their result implies that $\delta(m,n,r)$ is a polynomial of degree m in n when $n-r$ is fixed. For example, in addition to [22, Theorem 11], we have

$$\delta(6,n,n-2) \quad = \quad \frac{1}{72}\big(11n^6 - 81n^5 + 185n^4 - 75n^3 - 196n^2 + 156n\big).$$

The algebraic degree of SDP represents a universal upper bound on the intrinsic algebraic complexity of optimizing a linear function over any m-dimensional spectrahedron of $n \times n$-matrices. The algebraic degree can be much smaller for families of instances involving special matrices A_i, B or C.

EXAMPLE 5.7. Fix $n = 4$ and $m = 6 = \dim(\mathcal{E}_4)$. Pataki's inequalities (5.16) state that the rank of the optimal matrix is $r = 1$ or $r = 2$, and this was indeed observed in Example 5.4. For $r = 2$ we had found the algebraic degree six when solving (5.12). However, here B is the identity matrix and A_1, A_2, A_3, A_4 are diagonal. When these are replaced by generic symmetric matrices, then the algebraic degree jumps from six to $\delta(6, 4, 2) = 30$. □

We now state a result that elucidates the decompositions in (5.5) and (5.13).

THEOREM 5.8. *If the matrices $A_1, \ldots, A_m$ and C in the definition (5.1) of the spectrahedron P are sufficiently generic, then the algebraic boundary of the dual body P^Δ is the following union of dual hypersurfaces:*

$$\partial_a P^\Delta \quad \subseteq \quad \bigcup_{r \text{ as in } (5.16)} \{X \in \mathcal{L} \mid \operatorname{rank}(X) \leq r\}^* \tag{5.17}$$

PROOF. Let $\mathcal{Y}$ be any irreducible component of $\partial_a P^\Delta \subset (\mathbb{P}^m)^*$. Then $\mathcal{Y} \cap \partial P^\Delta$ is a semi-algebraic subset of codimension 1 in P^Δ. We consider a general point in that set. The corresponding hyperplane H in the primal $\mathbb{R}^m$ supports the spectrahedron P at a unique point Z. Then $r = \operatorname{rank}(Z)$ satisfies Pataki's inequalities, by Theorem 5.5. Moreover, the genericity in our choices of $A_1, \ldots, A_m, C, H$ ensure that Z is a regular point in $\{X \in \mathcal{L} \,|\, \operatorname{rank}(X) \leq r\}$. Bertini's Theorem ensures that this determinantal variety is irreducible and that its singular locus consists only of matrices of rank $< r$. This implies that $\{X \in \mathcal{L} \,|\, \operatorname{rank}(X) \leq r\}$ is the Zariski closure of $\{X \in P \,|\, \operatorname{rank}(X) = r\}$, and hence also of a neighborhood of Z in that rank stratum. Likewise, $\mathcal{Y}$ is the Zariski closure in $(\mathbb{P}^m)^*$ of $\mathcal{Y} \cap \partial P^\Delta$. An open dense subset of points in $\mathcal{Y} \cap \partial P^\Delta$ corresponds to hyperplanes that support P at a rank r matrix. We conclude $\mathcal{Y}^* = \{X \in \mathcal{L} \,|\, \operatorname{rank}(X) \leq r\}$. Biduality completes the proof. □

Theorem 5.8 is similar to Theorem 4.2 in that it characterizes the algebraic boundary in terms of dual hypersurfaces. Just like in Section 4, we can apply this result to compute $\partial_a P^\Delta$. For each rank r in the Pataki range (5.16), we need to check whether the corresponding dual hypersurface meets the boundary of P^Δ. The indices r which survive this test determine $\partial_a P^\Delta$.

When the data that specify the spectrahedron P are not generic but special then the computation of $\partial_a P^\Delta$ is more subtle and we know of no formula as simple as (5.17). This issue certainly deserves further research.

We close this section with an interesting 3-dimensional example.

Example 5.9. The *cyclohexatope* is a spectrahedron with $m = 3$ and $n = 5$ that arises in the study of chemical conformations [10]. Consider the following *Schönberg matrix* for the pairwise distances $\sqrt{D_{ij}}$ among six carbon atoms:

$$\begin{pmatrix} 2D_{12} & D_{12}{+}D_{13}{-}D_{23} & D_{12}{+}D_{14}{-}D_{24} & D_{12}{+}D_{15}{-}D_{25} & D_{12}{+}D_{16}{-}D_{26} \\ D_{12}{+}D_{13}{-}D_{23} & 2D_{13} & D_{13}{+}D_{14}{-}D_{34} & D_{13}{+}D_{15}{-}D_{35} & D_{13}{+}D_{16}{-}D_{36} \\ D_{12}{+}D_{14}{-}D_{24} & D_{13}{+}D_{14}{-}D_{34} & 2D_{14} & D_{14}{+}D_{15}{-}D_{45} & D_{14}{+}D_{16}{-}D_{46} \\ D_{12}{+}D_{15}{-}D_{25} & D_{13}{+}D_{15}{-}D_{35} & D_{14}{+}D_{15}{-}D_{45} & 2D_{15} & D_{15}{+}D_{56}{-}D_{56} \\ D_{12}{+}D_{16}{-}D_{26} & D_{13}{+}D_{16}{-}D_{36} & D_{14}{+}D_{16}{-}D_{46} & D_{15}{+}D_{56}{-}D_{56} & 2D_{16} \end{pmatrix}$$

The D_{ij} are the squared distances among six points in $\mathbb{R}^3$ if and only if this matrix is positive-semidefinite of rank ≤ 3. The points represent the carbon atoms in *cyclohexane* C_6H_{12} if and only if $D_{i,i+1} = 1$ and $D_{i,i+2} = 8/3$ for all indices i, understood cyclically. The three diagonal distances are unknowns, so, for cyclohexane conformations, the above Schönberg matrix equals

$$\mathbf{C}_6(x,y,z) \quad = \quad \begin{pmatrix} 2 & 8/3 & x-5/3 & 11/3-y & -2/3 \\ 8/3 & 2 & 5/3+x & 8/3 & 11/3-z \\ x-5/3 & 5/3+x & 16/3 & x+5/3 & x-5/3 \\ 11/3-y & 8/3 & x+5/3 & 2y & 8/3 \\ -2/3 & 11/3-z & x-5/3 & 8/3 & 16/3 \end{pmatrix}.$$

The *cyclohexatope* Cyc_6 is the spectrahedron in $\mathbb{R}^3$ defined by $\mathbf{C}_6(x,y,z) \succeq 0$. Its algebraic boundary decomposes as $\partial_a \mathrm{Cyc}_6 = V(f) \cup V(g)$, where

$$\begin{aligned} f &= 27xyz - 75x - 75y - 75z - 250 \qquad \text{and} \\ g &= 3xy + 3xz + 3yz - 22x - 22y - 22z + 121. \end{aligned}$$

The conformation space of cyclohexane is the real algebraic variety

$$\big\{(x,y,z) \in \mathrm{Cyc}_6 \mid \mathrm{rank}(\mathbf{C}_6(x,y,z)) \leq 3\,\big\} \quad = \quad V(f,g) \;\cup\; V(g)_{\mathrm{sing}}.$$

The first component is the closed curve of all *chair conformations*. The second component is the *boat conformation* point $(x,y,z) = \big(\frac{11}{3}, \frac{11}{3}, \frac{11}{3}\big)$. These are well-known to chemists [10]. Remarkably, the cyclohexatope coincides with the convex hull of these two components. This spectrahedron is another example of a convex hull of a space curve, now with an isolated point. Semidefinite programming over the cyclohexatope means computing the conformation which minimizes a linear function in the squared distances D_{ij}. □

6 – Spectrahedral Shadows

A *spectrahedral shadow* is the image of a spectrahedron under a linear map. The class of spectrahedral shadows is much larger than the class of spectrahedra. In fact, it has even been conjectured that every convex basic semialgebraic set in $\mathbb{R}^n$ is a spectrahedral shadow [13]. Our point of departure is the result, known to optimization experts, that the convex body dual to a spectrahedral shadow is again a spectrahedral shadow [11, Proposition 3.3]

THEOREM 6.1. *The class of spectrahedral shadows is closed under duality.*

Construction A spectrahedral shadow can be written in the form

$$P \;=\; \left\{ x \in \mathbb{R}^m \,\middle|\, \exists\, y \in \mathbb{R}^p \text{ with } C + \sum_{i=1}^{m} x_i A_i + \sum_{j=1}^{p} y_j B_j \succeq 0 \right\}.$$

An expression for the dual body P^Δ is obtained by the following variant of the construction in Remark 5.1. We consider the same linear map as before:

$$\pi : \mathcal{S}^n_+ \to \mathbb{R}^m,\; U \mapsto (\langle A_1, U\rangle, \ldots, \langle A_m, U\rangle).$$

We apply this linear map π to the spectrahedron

$$Q \;=\; \left\{ U \in \mathcal{S}^n_+ \mid \langle C, U\rangle \le 1 \text{ and } \langle B_1, U\rangle = \cdots = \langle B_p\,, U\rangle = 0 \right\}.$$

The closure of the spectrahedral shadow $\pi(Q)$ equals the dual convex body P^Δ. This closure is itself a spectrahedral shadow, by [11, Corollary 3.4]. □

We now consider the following problem: Given a real variety $X \subset \mathbb{R}^n$, find a representation of its convex hull $\mathrm{conv}(X)$ as a spectrahedral shadow. A systematic approach to computing such representations was introduced by Lasserre [16], and further developed by Gouveia *et al.* [12]. It is based on the relaxation of non-negative polynomial functions on X as sums of squares in the coordinate ring $\mathbb{R}[X]$. This approach is known as *moment relaxation*, in light of the duality between positive polynomials and moments of measures.

We shall begin by exploring these ideas for homogeneous polynomials of even degree $2d$ that are non-negative on $\mathbb{R}^n$. These form a cone in a real vector space of dimension $N = \binom{d+n-1}{d}$. Inside that cone lies the smaller *SOS cone* of polynomials p that are sums of squares of polynomials of degree d:

$$p \;=\; q_1^2 + q_2^2 + \cdots + q_N^2. \tag{6.1}$$

By Hilbert's Theorem [20], this inclusion of convex cones is strict unless $(n, 2d)$ equals $(1, 2d)$ or $(n, 2)$ or $(3, 4)$. The SOS cone is easily seen to be a spectrahedral shadow. Indeed, consider an unknown symmetric matrix $Q \in \mathcal{S}^N$ and write $p = v^T Q v$ where v is the vector of all N monomials of degree d. The matrix Q is positive semidefinite if it has a Cholesky factorization $Q = C^T C$. The resulting identity $p = (Cv)^T(Cv)$ can be rewritten as (6.1). Hence the SOS cone is the image of $\mathcal{S}^N_+$ under the linear map $Q \mapsto v^T Q v$.

Recent work of Nie [21] studies the boundaries of our two cones via computations with the discriminants we encountered at the end of Section 2.4.

Proposition 6.2. (Theorem 4.1 in [21]) *The algebraic boundary of the cone of homogeneous polynomials p of degree $2d$ that are non-negative on $\mathbb{R}^n$ is given by the discriminant of a polynomial p with unknown coefficients. This discriminant is the irreducible hypersurface dual to the Veronese embedding*

$$\mathbb{P}^{n-1} \hookrightarrow \mathbb{P}^{N-1},\ (x_1 : \cdots : x_n) \mapsto (x_1^d : x_1^{d-1}x_2 : \cdots : x_n^d)$$

The degree of this discriminant is $n(2d-1)^{n-1}$.

Proof. The discriminant of p vanishes if and only if there exists $x \in \mathbb{P}^{n-1}$ with $p(x) = 0$ and $\nabla p\big|_x = 0$. If p is in the boundary of the cone of positive polynomials then such a real point x exists. For the degree formula see [7]. □

Results similar to Proposition 6.2 hold when we restrict to polynomials p that lie in linear subspaces. This is why the *A-discriminants* Δ_A from Section 2.4 are relevant. We show this for a 2-dimensional family of polynomials.

Example 6.3. Consider the two-dimensional family of ternary quartics

$$f_{a,b}(x,y,z) \quad = \quad x^4 + y^4 + ax^3z + ay^2z^2 + by^3z + bx^2z^2 + (a+b)z^4.$$

Here a and b are parameters. Such a polynomial is non-negative on $\mathbb{R}^3$ if and only if it is a sum of squares, by Hilbert's Theorem. This condition defines a closed convex region $\mathcal{C}$ in the (a,b)-plane $\mathbb{R}^2$. It is non-empty because $(0,0) \in \mathcal{C}$. Its boundary $\partial_a(\mathcal{C})$ is derived from the *A-discriminant* Δ_A, where

$$A \quad = \quad \begin{pmatrix} 4 & 0 & 3 & 0 & 0 & 2 & 0 \\ 0 & 4 & 0 & 2 & 3 & 0 & 0 \\ 0 & 0 & 1 & 2 & 1 & 2 & 4 \end{pmatrix}.$$

This A-discriminant is an irreducible homogeneous polynomial of degree 24 in the seven coefficients. What we are interested in here is the *specialized discriminant* which is obtained from Δ_A by substituting the vector of coefficients $(1,1,a,a,b,b,a+b)$ corresponding to our polynomial $f_{a,b}$. The specialized discriminant is an inhomogeneous polynomial of degree 24 in the two unknowns a and b, and it is no longer irreducible. A computation reveals that it is the product of four irreducible factors whose degrees are 1, 5, 5 and 13.

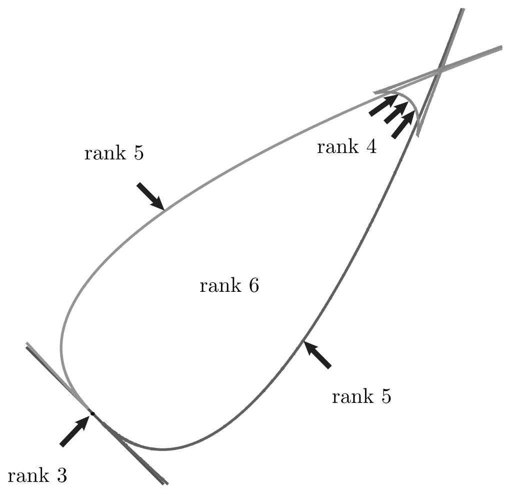

Fig. 9: The discriminant in Example 6.3 defines a curve in the (a, b)-plane. The *spectrahedral shadow* $\mathcal{C}$ is the set of points where the ternary quartic $f_{a,b}$ is SOS. The ranks of the corresponding SOS matrices Q are indicated.

The linear factor equals $a + b$. The two factors of degree 5 are

$$256a^2-27a^5+512ab+144a^3b-27a^4b+256b^2-128ab^2+144a^2b^2-128b^3-4a^2b^3+16b^4,$$
$$256a^2-128a^3+16a^4+512ab-128a^2b+256b^2+144a^2b^2-4a^3b^2+144ab^3-27ab^4-27b^5.$$

Finally, the factor of degree 13 in the specialized discriminant equals

$$\begin{gathered}
2916a^{11}b^2 + 19683a^9b^4 + 19683a^8b^5 + 2916a^7b^6 + 2916a^6b^7 + 19683a^5b^8 \\
+19683a^4b^9 + 2916a^2b^{11} - 11664a^{12} - 104976a^{10}b^2 - 136080a^9b^3 - 27216a^8b^4 \\
-225504a^7b^5 - 419904a^6b^6 - 225504a^5b^7 - 27216a^4b^8 - 136080a^3b^9 \\
-104976a^2b^{10} - 11664b^{12} + 93312a^{11} + 217728a^{10}b + 76032a^9b^2 \\
+1133568a^8b^3 + 1976832a^7b^4 + 891648a^6b^5 + 891648a^5b^6 + 1976832a^4b^7 \\
+1133568a^3b^8 + 76032a^2b^9 + 217728ab^{10} + 93312b^{11} - 241920a^{10} \\
-1368576a^9b - 2674944a^8b^2 - 1511424a^7b^3 - 4729600a^6b^4 - 9369088a^5b^5 \\
-4729600a^4b^6 - 1511424a^3b^7 - 2674944a^2b^8 - 1368576ab^9 - 241920b^{10} \\
+663552a^9 + 2949120a^8b + 10539008a^7b^2 + 17727488a^6b^3 + 9981952a^5b^4 \\
+9981952a^4b^5 + 17727488a^3b^6 + 10539008a^2b^7 + 2949120ab^8 + 663552b^9 \\
-2719744a^8 - 8847360a^7b - 14974976a^6b^2 - 36503552a^5b^3 - 56360960a^4b^4 \\
-36503552a^3b^5 - 14974976a^2b^6 - 8847360ab^7 - 2719744b^8 + 4587520a^7 \\
+25821184a^6b + 52035584a^5b^2 + 50724864a^4b^3+50724864a^3b^4+52035584a^2b^5 \\
+25821184ab^6 + 4587520b^7 - 6291456a^6 - 31457280a^5b - 94371840a^4b^2 \\
-138412032a^3b^3 - 94371840a^2b^4 - 31457280ab^5 - 6291456b^6 + 16777216a^5 \\
+50331648a^4b + 67108864a^3b^2 + 67108864a^2b^3 + 50331648ab^4 + 16777216b^5 \\
-16777216a^4 - 67108864a^3b - 100663296a^2b^2 - 67108864ab^3 - 16777216b^4.
\end{gathered}$$

The relevant pieces of these four curves in the (a, b)-plane are depicted in Figure 9. The line $a + b = 0$ is seen in the lower left, the degree 13 curve is the swallowtail in the upper right, and the two quintic curves form the upper-left and lower-right boundary of the enclosed convex region $\mathcal{C}$.

For each $(a,b) \in \mathcal{C}$, the ternary quartic $f_{a,b}$ has an SOS representation

$$f_{a,b}(x,y,z) \;=\; (x^2, xy, y^2, xz, yz, z^2) \cdot Q \cdot (x^2, xy, y^2, xz, yz, z^2)^T,$$

where Q is a positive semidefinite $6{\times}6$-matrix. This identity gives 15 independent linear constraints which, together with $Q \succeq 0$, define an 8-dimensional spectrahedron in the $(21+2)$-dimensional space of parameters (Q,a,b). The projection of this spectrahedron onto the (a,b)-plane is our convex region $\mathcal{C}$. This proves that $\mathcal{C}$ is a spectrahedral shadow. If (a,b) lies in the interior of $\mathcal{C}$ then the fiber of the projection is a 6-dimensional spectrahedron. If (a,b) lies in the boundary $\partial\mathcal{C}$ then the fiber consists of a single point. The ranks of these unique matrices are indicated in Figure 9. Notice that $\partial\mathcal{C}$ has three singular points, at which the rank drops from 5 to 4 and 3 respectively. ∎

We now shift towards a functional analytic point of view. The degree d is no longer fixed, and we consider all polynomials, not just homogeneous ones. Polynomials that are non-negative on $\mathbb{R}^n$ form a convex cone $\mathcal{C}$ in the infinite-dimensional real vector space $\mathbb{R}[x_1,\ldots,x_n]$. Its dual cone $\mathcal{C}^*$ is the set of all linear functionals $\mathbb{R}[x_1,\ldots,x_n] \to \mathbb{R}$ that are non-negative on $\mathcal{C}$. We consider functionals that evaluate to 1 on the constant polynomial 1. These are represented by the moments of probability measures μ on $\mathbb{R}^n$:

$$y_\alpha \;=\; \int_{\mathbb{R}^n} x^\alpha d\mu \qquad \text{for } \; \alpha \in \mathbb{N}^n.$$

Points in $\mathcal{C}^*$ are moment sequences $(y_\alpha) \in \mathbb{R}^{\mathbb{N}^n}$ of Borel measures μ on $\mathbb{R}^n$.

This setup allows for an elegant and fruitful interpretation of Lagrange duality for polynomial optimization problems (2.7). To keep the exposition and notation simple, we restrict ourselves to the unconstrained problem

$$\underset{x\in\mathbb{R}^n}{\text{minimize}} \qquad f(x) = \sum_\alpha f_\alpha x^\alpha$$

Here we assume that f is bounded from below, say $f \geq \epsilon$, and $\deg(f) = 2d$. Our problem is equivalent to finding the best possible lower bound:

$$\begin{aligned} &\underset{t\in\mathbb{R}}{\text{maximize}} \quad t \\ &\text{subject to} \quad f(x) - t \geq 0 \;\text{ for all }\; x \in \mathbb{R}^n. \end{aligned} \tag{6.2}$$

The Lagrange dual of the problem (6.2) reads

$$\underset{\mu\in\mathcal{P}}{\text{minimize}} \quad \int_{\mathbb{R}^n} f(x) d\mu,$$

where $\mathcal{P}$ is the convex set of all Borel probability measures on $\mathbb{R}^n$. We can rewrite this now as an infinite-dimensional linear optimization problem:

$$\begin{aligned} &\underset{y\in\mathcal{Y}}{\text{minimize}} \quad \sum_\alpha f_\alpha y_\alpha \\ (6.3)\qquad &\text{where} \quad \mathcal{Y} := \left\{ y \in \mathbb{R}^{\mathbb{N}^n} \,\middle|\, y_\alpha = \int_{\mathbb{R}^n} x^\alpha d\mu \text{ with } \mu \in \mathcal{P} \right\}. \end{aligned}$$

The two dual problems (6.2) and (6.3) are as difficult to solve as our original optimization problem. There is a natural relaxation which is easier, and we can express this either on the primal side or on the dual side. In (6.2) we replace the constraint that $f(x) - t$ be non-negative on $\mathbb{R}^n$ with the easier constraint that $f(x)-t$ be a sum of squares. We relax the dual (6.3) by enlarging the convex set $\mathcal{Y}$ to the *infinite-dimensional spectrahedron* consisting of all positive semidefinite moment matrices

$$M(y) \,=\, (\, y_{\alpha+\beta}\,)_{\alpha,\beta\in\mathbb{N}^n} \succeq 0.$$

These two relaxations are again related by Lagrange duality, but now they represent a dual pair of semidefinite programs. Of course, when we solve such an SDP in practise, we always restrict to a finite submatrix of $M(y)$, usually that indexed by all monomials x^α, x^β of some bounded degree $\leq d$. The question of when such a relaxtion is exact and, if not, how large the gap can be, is an active area of research in convex algebraic geometry [12, 17, 27].

We now turn our attention to a variant of the above procedure which approximates the convex hull of a variety by a nested family of spectrahedral shadows. Let I be an ideal in $\mathbb{R}[x_1, \ldots, x_n]$ and $V_\mathbb{R}(I)$ its variety in $\mathbb{R}^n$. Consider the set of affine-linear polynomials that are non-negative on $V_\mathbb{R}(I)$:

$$\mathrm{NN}(I) \quad = \quad \{\, f \in \mathbb{R}[x_1, \ldots, x_n]_{\leq 1} \mid f(x) \geq 0 \text{ for all } x \in V_\mathbb{R}(I)\}.$$

In light of the biduality theorem for convex sets (cf. Section 2.2), we can characterize the (closure of) the convex hull of our variety as follows:

$$\overline{\mathrm{conv}(V_\mathbb{R}(I))} \quad = \quad \{x \in \mathbb{R}^n \mid f(x) \geq 0 \text{ for all } f \in \mathrm{NN}(I)\}.$$

The geometry behind this formula is shown in Figure 10.

Following Gouveia *et al.* [12], we now replace the hard constraint that $f(x)$ be non-negative on $V_\mathbb{R}(I)$ with the (hopefully easier) constraint that $f(x)$ be a sum of squares in the coordinate ring $\mathbb{R}[x_1, \ldots, x_n]/I$. Introducing a parameter d that indicates the degree of the polynomials allowed in that SOS representation, we consider the following set of affine-linear polynomials:

$$\mathrm{SOS}_d(I) \;=\; \{\, f \mid f - q_1^2 - \cdots - q_r^2 \,\in\, I \;\text{ for some}\, q_i \in \mathbb{R}[x_1, \ldots, x_n]_{\leq d}\,\}.$$

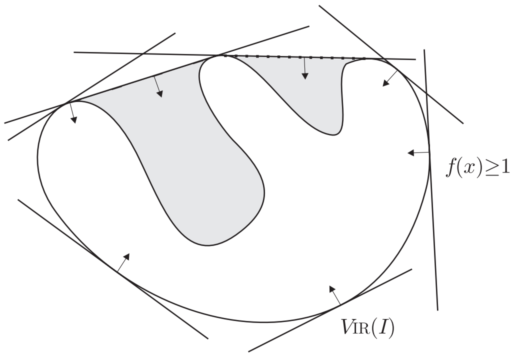

Fig. 10: Convex hull as intersection of half spaces.

The following chain of inclusions holds:

$$\mathrm{SOS}_1(I) \subseteq \mathrm{SOS}_2(I) \subseteq \mathrm{SOS}_3(I) \subseteq \cdots \subseteq \mathrm{NN}(I). \tag{6.4}$$

We now dualize the situation by considering the subsets of $\mathbb{R}^n$ where the various f are non-negative. The *d-th theta body* of the ideal I is the set

$$\mathrm{TH}_d(I) \;=\; \big\{ x \in \mathbb{R}^n \,|\, f(x) \geq 0 \text{ for all } f \in \mathrm{SOS}_d(I) \big\}.$$

The following reverse chain of inclusions holds among subsets in $\mathbb{R}^n$:

$$\mathrm{TH}_1(I) \supseteq \mathrm{TH}_2(I) \supseteq \mathrm{TH}_3(I) \supseteq \cdots \supseteq \overline{\mathrm{conv}(V_{\mathbb{R}}(I))}. \tag{6.5}$$

This chain of outer approximations can fail to converge in general, but there are various convergence results when the geometry is nice. For instance, if the real variety $V_{\mathbb{R}}(I)$ is compact then *Schmüdgen's Positivstellensatz* [27, §3] ensures asymptotic convergence. When $V_{\mathbb{R}}(I)$ is a finite set, so that $\mathrm{conv}(V_{\mathbb{R}}(I))$ is a polytope, then we have finite convergence, that is, $\exists\, d : \mathrm{TH}_d(I) = \mathrm{conv}(V_{\mathbb{R}}(I))$. This was shown in [14]. For more information on theta bodies see [12]. The main point we wish to record here is the following:

Theorem 6.4. ([12, 17]) *Each theta body* $\mathrm{TH}_d(I)$ *is a spectrahedral shadow.*

Proof. We may assume, without loss of generality, that the origin 0 lies in the interior of $\mathrm{conv}(V_{\mathbb{R}}(I))$. Then $\mathrm{SOS}_d(I)$ is the cone over the convex set dual to $\mathrm{TH}_d(I)$. Since the class of spectrahedral shadows is closed under duality, and under intersecting with affine hyperplanes, it suffices to show that $\mathrm{SOS}_d(I)$ is a spectrahedral shadow. But this follows from the formula $f - q_1^2 - \cdots - q_r^2 \in I$, by an argument similar to that given after (6.1). ∎

In this article we have seen two rather different representations of the convex hull of a real variety, namely, the characterization of the algebraic boundary in Section 4, and the representation as a theta body suggested above. The relationship between these two is not yet well understood. A specific question is how to best compute the algebraic boundary of a spectrahedral shadow. This leads to problems in elimination theory that seem to be particularly challenging for current computer algebra systems.

We conclude by revisiting one of the examples we had seen in Section 4.

EXAMPLE 6.5. (EXAMPLE 4.5 CONT.) We revisit the curve $X = V(h_1, h_2)$ with

$$h_1 = x^2 + y^2 + z^2 - 1,$$
$$h_2 = 19x^2 + 21y^2 + 22z^2 - 20.$$

Scheiderer [27] proved that finite convergence holds in (6.5) whenever I defines a curve of genus 1, such as X. We will show that $d = 1$ suffices in our example, i.e. we will show that $\mathrm{TH}_1(I) = \mathrm{conv}(X)$ for the ideal $I = \langle h_1, h_2 \rangle$.

We are interested in affine-linear forms f that admit a representation

$$f \;=\; 1 + ux + vy + wz \;=\; \mu_1 h_1 + \mu_2 h_2 + \sum_i q_i^2. \tag{6.6}$$

Here μ_1 and μ_2 are real parameters. Moreover, we want f to lie in $\mathrm{SOS}_1(I)$, so we require $\deg q_i = 1$ for all i. The sum of squares can be written as

$$\sum_i q_i^2 \;=\; (1, x, y, z) \cdot Q \cdot (1, x, y, z)^T, \qquad \text{where } Q \in \mathcal{S}_+^4.$$

After matching coefficients in (6.6), we obtain the spectrahedral shadow

$$\mathrm{SOS}_1(I) = \Big\{ (u, v, w) \in \mathbb{R}^3 \,\Big|\, \exists \mu_1, \mu_2 : \begin{pmatrix} 1 + \mu_1 + 20\mu_2 & u & v & w \\ u & -\mu_1 - 19\mu_2 & 0 & 0 \\ v & 0 & -\mu_1 - 21\mu_2 & 0 \\ w & 0 & 0 & -\mu_1 - 22\mu_2 \end{pmatrix} \succeq 0 \Big\}.$$

Dual to this is the theta body $\mathrm{TH}_1(I) = \mathrm{SOS}_1(I)^{\Delta}$. It has the representation

$$\mathrm{TH}_1(I) = \Big\{ (x, y, z) \in \mathbb{R}^3 \Big|\, \exists u_1, u_2, u_3, u_4 : \begin{pmatrix} 1 & x & y & z \\ x & \frac{2}{3} - \frac{1}{3}u_4 & u_1 & u_2 \\ y & u_1 & \frac{1}{3} - \frac{2}{3}u_4 & u_3 \\ z & u_2 & u_3 & u_4 \end{pmatrix} \succeq 0 \Big\}.$$

We consider the ideal generated by this 4×4-determinant and its derivatives with respect to u_1, u_2, u_3, u_4, we saturate by the ideal of 3×3-minors, and then we eliminate u_1, u_2, u_3, u_4. The result is the principal ideal $\langle h_4 h_5 h_6 \rangle$, with h_i as in Example 4.5. This computation reveals that the algebraic boundary of $\mathrm{conv}(X)$ consists of quadrics, and we can conclude that $\mathrm{TH}_1(I) = \mathrm{conv}(X)$.

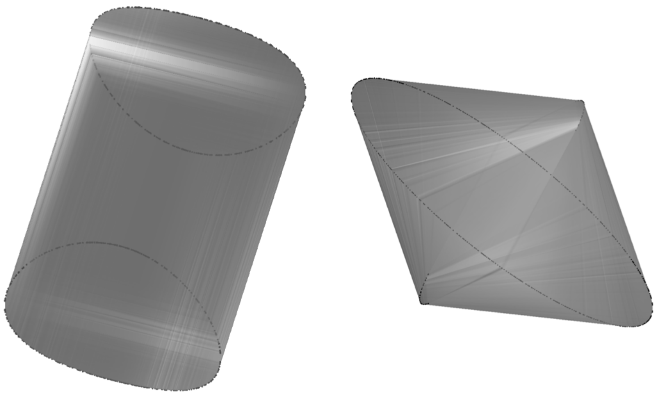

Fig. 11: Convex hull of the curve in Figure 7 and its dual convex body.

Pictures of our convex body and its dual are shown in Figure 11. Diagrams such as these can be drawn fairly easily for any spectrahedral shadow in $\mathbb{R}^3$. To be precise, the matrix representation of $\mathrm{TH}_1(I)$ and $\mathrm{SOS}_1(I)^\Delta$ given above can be used to rapidly sample the boundaries of these convex bodies, by maximizing many linear functions via semidefinite programming. □

Acknowledgements

This article grew out of three lectures given by Bernd Sturmfels on March 22-24, 2010, at the spring school on Linear Matrix Inequalities and Polynomial Optimization (LMIPO) at UC San Diego. We thank Jiawang Nie and Bill Helton for organizing that event. We are grateful to João Gouveia and Raman Sanyal and Rekha Thomas for their comments on a draft of this paper. Philipp Rostalski was supported by the Alexander-von-Humboldt Foundation through a Feodor Lynen postdoctoral fellowship. Bernd Sturmfels was partially supported by the U.S. National Science Foundation through the grants DMS-0456960 and DMS-0757207.

REFERENCES

[1] A. I. Barvinok: *A Course in Convexity*, Graduate Studies in Mathematics, **54**, American Mathematical Society, Providence, 2002.

[2] S. Basu – R. Pollack – M-F. Roy: *Algorithms in Real Algebraic Geometry*, Springer, Berlin, 2006.

[3] J. Bochnak – M. Coste – M.-F. Roy: *Géométrie Algébrique Réelle*, Ergebnisse der Mathematik und ihrer Grenzgebiete, **12**, Springer, Berlin, 1987.

[4] S. Boyd – L. Vandenberghe: *Convex Optimization*, Cambridge University Press, 2004.

[5] S. Boyd – L. Vandenberghe: *Semidefinite programming*, SIAM Review, **38** (1996), 49–95.

[6] D. Cox – J. Little – D. O'Shea: *Ideals, Varieties and Algorithms*, Undergraduate Texts in Math., 3rd edition, Springer, New York, 2007.

[7] I. Gelfand – M. Kapranov – A. Zelevinsky: *Discriminants, Resultants and Multidimensional Determinants*, Birkhäuser, Boston, 1994.

[8] H-C. Graf von Bothmer – K. Ranestad: *A general formula for the algebraic degree in semidefinite programming*, Bull. Lond. Math. Soc. **41** (2009), 193–197.

[9] D. Grayson – M. Stillman: *Macaulay2, a software system for research in algebraic geometry*, Available at `http://www.math.uiuc.edu/Macaulay2/`.

[10] N. Go – H.A. Scheraga: *Ring closure in chain molecules with C_n,I, and S_{2n} symmetry*, Macromolecules **6(2)** (1973), 273–281.

[11] J. Gouveia – T. Netzer: *Positive polynomials and projections of spectrahedra*, `arXiv:0911.2750`.

[12] J. Gouveia – P. Parrilo – R. Thomas: *Theta bodies for polynomial ideals*, SIAM J. Optim. **20** (2010), 2097–2118.

[13] J.W. Helton – J. Nie: *Semidefinite representation of convex sets*, Math. Program., **122** (2010), Ser. A, 21–64.

[14] M. Laurent – J.B. Lassarre – P. Rostalski: *Semidefinite characterization and computation of zero-dimensional real radical ideals*, Found. Comp. Math. **8** (2008), 607–647.

[15] M. Laurent – S. Poljak: *On the facial structure of the set of correlation matrices*, SIAM J. Matrix Anal. Appl., **17** (1996), 530–547.

[16] J.B. Lasserre: *Global optimization with polynomials and the problem of moments*, SIAM J. Optimization, **11** (2001), 796–817.

[17] J.B. Lasserre: *Moments, Positive Polynomials and their Applications*, Imperial College Press, London, 2010.

[18] S. Lazard – L.M. Peñaranda – S. Petitjean: *Intersecting quadrics: an efficient and exact implementation*, Comput. Geometry **35** (2006), 74–99.

[19] D. G. Luenberger: *A double look at duality*, IEEE Trans. Automat. Control, **73** (1992), 1474–1482.

[20] M. Marshall: *Positive Polynomials and Sums of Squares*, American Math. Soc, 2008.

[21] J. Nie: *Discriminants and nonnegative polynomials*, `arXiv:1002.2230`.

[22] J. Nie – K. Ranestad – B. Sturmfels: *The algebraic degree of semidefinite programming*, Mathematical Programming, **122** (2010), 379–405.

[23] M. Ramana – A.J. Goldman: *Some geometric results in semidefinite programming*, J. Global Optimization, **7** (1995), 33–50.

[24] K. Ranestad – B. Sturmfels: *On the convex hull of a space curve*, to appear in Advances in Geometry, `arXiv:0912.2986`.

[25] K. Ranestad – B. Sturmfels: *The convex hull of a variety*, `arXiv:1004.3018`.

[26] R.T. Rockefeller: *Convex Analysis*, Princeton University Press, 1970.

[27] C. Scheiderer: *Convex hulls of curves of genus one*, `arXiv:1003.4605`.

[28] M. Trott: *Applying GroebnerBasis to three problems in geometry*, Mathematica in Education and Research, **6** (1997), 15–28.

[29] G. Ziegler: *Lectures on Polytopes*, Graduate Texts in Mathematics, Springer, New York, 1995.

Lavoro pervenuto alla redazione il 10 giugno 2010
ed accettato per la pubblicazione il 25 giugno 2010.
Bozze licenziate il 26 novembre 2010

Indirizzo degli Autori:

Philipp Rostalski– Department of Mathematics – University of California – Berkeley, CA 94720, USA
E-mail: philipp@math.berkeley.edu

Bernd Sturmfels– Department of Mathematics – University of California – Berkeley, CA 94720, USA
E-mail: bernd@math.berkeley.edu

Rendiconti di Matematica, Serie VII
Volume 30, Roma (2010), 329–343

Regularity results for planar quasilinear equations

GABRIELLA ZECCA

Abstract: *We study the Dirichlet problem for the quasilinear elliptic equation*

$$-\text{div } A(x, \nabla v) = f$$

in a planar domain Ω*, when* f *belongs to the Zygmund space* $L(\log L)^{\frac{1}{2}}(\log\log L)^{\frac{1}{2}}(\Omega)$*. We prove that the gradient of the variational solution* $v \in W_0^{1,2}(\Omega)$ *belongs to the Zygmund space* $L^2 \log\log L(\Omega)$*.*

1 – Introduction

Let $\Omega \subset \mathbb{R}^2$ be a bounded open set with C^1-boundary. We consider the following Dirichlet problem

$$\begin{cases} -\text{div } A(x, \nabla v) = f & \text{in } \Omega \\ v \in W_0^{1,2}(\Omega), \end{cases} \tag{1.1}$$

where $A : \Omega \times \mathbb{R}^2 \to \mathbb{R}^2$ is a mapping such that:

(1,2) $\quad x \to A(x, \xi)$ is measurable for any $\xi \in \mathbb{R}^2$;

(1.3) $\quad \xi \to A(x, \xi)$ is continuous for almost every $x \in \Omega$.

Key Words and Phrases: *Elliptic equations – Zygmund spaces – Gradient regularity*
A.M.S. Classification: 35B65, 46E30.

Moreover we assume that there exists $K \geqslant 1$ such that for almost every $x \in \Omega$ we have

$$|A(x,\xi) - A(x,\eta)| \leqslant K|\xi - \eta| \qquad \text{(Lipschitz continuity)} \tag{1.4}$$

$$|\xi - \eta|^2 \leqslant K\langle A(x,\xi) - A(x,\eta), \xi - \eta\rangle \qquad \text{(strong monotonicity)} \tag{1.5}$$

$$A(x,0) = 0 \tag{1.6}$$

for any vectors ξ and η in $\mathbb{R}^2$ (see [18]).

In [9] an existence and uniqueness theorem for the Dirichlet problem for the equation div $A(x,\nabla v) = f$ is proved where $f \in L^1(\Omega)$ and the solution v belongs to the so called *grand Sobolev space* $W_0^{1,2)}(\Omega)$ i.e. the space of function $v \in W_0^{1,1}(\Omega)$ whose gradient $|\nabla v|$ satisfies

$$\sup_{1<s<2} \left[(2-s)\int_\Omega |\nabla v|^s dx\right]^{\frac{1}{s}} = \|v\|_{W_0^{1,2)}} < \infty.$$

Note that the space of such functions $W_0^{1,2)}(\Omega)$ is slightly larger than $W_0^{1,2}(\Omega)$ and this is the appropriate space when the right-hand side f is assumed to be only L^1−integrable (see [9], [11] for more details).

In this paper we study cases where the solution v is the variational $W_0^{1,2}(\Omega)$-solution, under the assumption

$$f \in L(\log L)^{\frac{1}{2}}(\log\log L)^{\frac{1}{2}}(\Omega) \subset L(\log L)^{\frac{1}{2}}(\Omega). \tag{1.7}$$

Let us observe that by the Sobolev-Trudinger imbedding in the plane

$$W_0^{1,2}(\Omega) \hookrightarrow EXP_2(\Omega), \tag{1.8}$$

hypothesis (1.7) guarantees that f belongs to the dual space of $W_0^{1,2}(\Omega)$ and then, at least in the linear case $A(x,\xi) = \mathcal{A}(x)\xi$ the Lax-Milgram Theorem ensure that there exists a unique solution $v \in W_0^{1,2}(\Omega)$.

The case where f belongs to the Zygmund space

$$f \in L(\log L)^\delta(\Omega) \subset L^1(\Omega), \qquad \text{for } \frac{1}{2} \leqslant \delta \leqslant 1 \tag{1.9}$$

is treated in [3] (see also [2], [21] for the case $\delta = 1$) where e.g. the authors prove that under the assumption (1.9), there is a unique solution $v \in W_0^{1,2}(\Omega)$ to the Dirichlet problem (1.1) with $\nabla v \in L^2(\log L)^{2\delta-1}$ and

$$\|\nabla v\|_{L^2(\log L)^{2\delta-1}(\Omega)} \leqslant c(K)\, \|f\|_{L(\log L)^\delta(\Omega)}\,, \tag{1.10}$$

where $c(K) > 0$ depends only on K.

We prove the following

THEOREM 1.1. *Let $A = A(x, \xi)$ satisfy conditions (1.2)-(1.6) and let $f \in L(\log L)^{\frac{1}{2}}(\log\log L)^{\frac{1}{2}}(\Omega)$. Then, there exists an unique $v \in W_0^{1,2}(\Omega)$ solution to*

$$\begin{cases} -\text{div}\, A(x, \nabla v) = f & \textit{in } \Omega \\ v \in W_0^{1,2}(\Omega), \end{cases} \tag{1.11}$$

such that $\nabla v \in L^2(\log\log L)(\Omega)$ and

$$\|\nabla v\|_{L^2(\log\log L)(\Omega)} \leqslant C(K)\,\|f\|_{L(\log L)^{\frac{1}{2}}(\log\log L)^{\frac{1}{2}}(\Omega)}\,.$$

Note that by imbedding theorems for Orlicz-Sobolev spaces, (see [5]) we obtain in particular that the solution v in Theorem 1.1 belongs to the Orlicz space $L^{\Lambda}(\Omega)$ generated by the Young function $\Lambda(t) = \exp\left\{t^2\log(e+t)\right\} - 1$.

It is worth to point out that under the assumptions of Theorem 1.1 we cannot expect the boundedness of the solution u. In fact in [2] is proved that $f \in L\log L(\Omega)$ is a sufficient condition for the boundedness (and continuity) of the solution u and in [3] there are examples where $f \in L\log^{\delta} L(\Omega)$, $\delta \in [\frac{1}{2}, 1[$, and the solution u is not bounded.

In Section 5 we prove that also approaching $L\log L(\Omega)$ in the scale of spaces $L\log L(\log\log L)^{\alpha}$, $L\log L(\log\log\ \log L)^{\alpha}(\Omega)$, $L\log L(\log\log\log\ldots\log L)^{\alpha}(\Omega)$, $\alpha < 0$, we cannot obtain the boundedness of the solution.

The case $n \geqslant 3$ is extensively treated for the n-harmonic equations in the recent papers [14] and [12].

2 – Young's functions and Orlicz spaces

Let $\Phi : [0, +\infty) \to [0, +\infty)$ be a Young's function, i.e. a convex function of type $\Phi(t) = \int_0^t \varphi(s)ds$, $t > 0$, where $\varphi : [0, \infty[\to \mathbb{R}$ is nondecreasing, right-continuous and such that

$$\varphi(s) > 0 \qquad \forall s > 0, \qquad \varphi(0) = 0, \qquad \lim_{s\to\infty} \varphi(s) = +\infty. \tag{2.1}$$

The Young's function $\tilde{\Phi}(t)$, complementary to $\Phi(t)$, is defined by $\tilde{\Phi}(t) = \sup\{st - \Phi(s) : s > 0\}$ and it is easy to see that $\tilde{\tilde{\Phi}} = \Phi$.

In the sequel we will deal with a particular class of Young functions Φ verifying a suitable sub-homogeneity property at infinity called Δ_2-*condition*. Namely, Definition 1. A young function Φ satisfies the Δ_2-condition (we will write $\Phi \in \Delta_2$) if there exists a constant $l > 0$ such that

$$\Phi(\lambda t) \leqslant \lambda^l \Phi(t), \quad \forall \lambda \geqslant 1, \quad \forall t \geqslant t_0, \tag{2.2}$$

where $t_0 \geqslant 0$ is a suitable large constant.

Let Ω be an open and bounded set in $\mathbb{R}^n$, $n \geqslant 1$. The Orlicz class $\Lambda^\Phi(\Omega)$ is the set of all measurable functions $u : \Omega \to \mathbb{R}$ satisfying

$$\int_\Omega \Phi(|u(x)|)dx < \infty$$

The Orlicz Space $L^\Phi = L^\Phi(\Omega)$ is the linear hull of $\Lambda^\Phi(\Omega)$ and the equality $L^\Phi(\Omega) \equiv \Lambda^\Phi(\Omega)$ holds if and only if $\Phi \in \Delta_2$.

Define the functional $\|u\|_{L^\Phi(\Omega)} : L^\Phi(\Omega) \to [0, +\infty[$ by

$$\|u\|_{L^\Phi(\Omega)} = \inf\left\{K > 0 : \int_\Omega \Phi\left(\frac{|u(x)|}{K}\right) dx \leqslant 1\right\}. \tag{2.3}$$

It is a norm, called the *Luxemburg norm*, and $L^\Phi(\Omega)$ is a Banach space when endowed with it. When no confusion arise we will simply write $\|u\|_{L^\Phi}$ or $\|u\|_\Phi$ instead of $\|u\|_{L^\Phi(\Omega)}$.

We recall that:

i) If $\Phi(t) = t^p$ and $1 \leqslant p < \infty$ then $L^\Phi(\Omega) = L^p(\Omega)$, the classical Lebesgue space and $\|\cdot\|_{L^\Phi(\Omega)} = \|\cdot\|_{L^p}$.

ii) If $\Phi(t) = t^p(\log(e+t))^q$ where either $p > 1$ and $-\infty < q < \infty$ or $p = 1$ and $q \geqslant 0$, then the Orlicz space $L^\Phi(\Omega)$ is the Zygmund space $L^p(\log L)^q(\Omega)$, and the norm (2.3) is equivalent to the quantity (see [16])

$$[v]_{L^p(\log L)^q(\Omega)} = \left[\fint_\Omega |v|^p \log^q\left(e + \frac{|v|}{\left(\fint_\Omega |v|^p dx\right)^{\frac{1}{p}}}\right) dx \right]^{\frac{1}{p}} \tag{2.4}$$

where, for all Lebesgue measurable set E with positive measure, we denote by $\fint_E f dx$ the mean value of f taken over the set E, i.e. $\fint_E f dx = f_E = \frac{1}{|E|}\int_E f dx$, where $|E|$ denotes the Lebesgue measure of E.

iii) If $\Phi(t) = e^{t^a} - 1$, $a > 0$, then the Orlicz space $L^\Phi(\Omega)$ reproduces the space of exponentially integrable functions $EXP(\Omega)$ when $a = 1$ and $EXP_a(\Omega)$ otherwise.

iv) If $\Phi(t) = t^p(\log\log(e^e + t))^q$ where either $p > 1$ and $-\infty < q < \infty$ or $p = 1$ and $q \geqslant 0$, then the Orlicz space $L^\Phi(\Omega)$ is the space $L^p(\log\log L)^q(\Omega)$.

The closure of $C_0^\infty(\Omega)$ in $L^\Phi(\Omega)$ is denoted by $E^\Phi(\Omega)$ and the inclusions

$$E^\Phi(\Omega) \subseteq \Lambda^\Phi(\Omega) \subseteq L^\Phi(\Omega) \tag{2.5}$$

are trivial with equality holding if and only if $\Phi \in \Delta_2$.

The *Orlicz-Sobolev* space $W^{1,\Phi}(\Omega)$ is defined as

$$W^{1,\Phi}(\Omega) = \left\{u \in W^{1,1}(\Omega) \cap L^\Phi(\Omega) : |Du| \in L^\Phi(\Omega)\right\},$$

and, equipped with the norm

$$\|u\|_{W^{1,\Phi}} = \|u\|_\Phi + \|Du\|_\Phi$$

it is a Banach space.

By $W_0^{1,\Phi}(\Omega)$ we denote the subspace of $W^{1,\Phi}(\Omega)$ of those functions whose continuation by 0 outside Ω belongs to $W^{1,\Phi}(\mathbb{R}^n)$. Properties of Orlicz-Sobolev spaces are presented in [7], [20].

The Orlicz space $L^\Phi(\Omega)$ is isometrically isomorphic to the dual space of $E^{\tilde{\Phi}}(\Omega)$ (see [17], [20]) and $[L^\Phi(\Omega)]' \simeq L^{\tilde{\Phi}}(\Omega)$ if and only if $\Phi \in \Delta_2$. In particular the space $L^\Phi(\Omega)$ is reflexive if and only if both Φ and $\tilde{\Phi}$ belong to class Δ_2.

Here below we recall the explicit expression of the dual spaces of some Orlicz space (see [4] and [8]) which will be useful in the sequel

i) for any $1 < p < \infty$ and $-\infty < q < \infty$ it is

$$(L^p(\log L)^q(\Omega))' \cong \frac{L^{p'}}{(\log L)^{\frac{q}{p-1}}}(\Omega)$$

where p' is the conjugate exponent of p, i.e. $\frac{1}{p} + \frac{1}{p'} = 1$

ii) for any $1 < p < \infty$ and $-\infty < q < \infty$ it is

$$(L^p(\log\log L)^q(\Omega))' \cong \frac{L^{p'}}{(\log\log L)^{\frac{q}{p-1}}}(\Omega)$$

iii) for $p = 1$ and $q > 0$ it is

$$(L(\log L)^q(\Omega))' \cong EXP_{\frac{1}{q}}(\Omega) \tag{2.6}$$

The following partial ordering relation between functions is involved in imbedding theorems between Orlicz spaces associated with different Young functions.

Definition 2. The function Ψ is said to dominate the function Φ globally (respectively near infinity) if there exists $c > 0$ such that

$$\Phi(t) \leqslant \Psi(ct) \tag{2.7}$$

for any $t \geqslant 0$ (respectively for any t greater than some positive number).

The functions Φ and Ψ are called equivalent globally (respectively near infinity) if each dominates the other globally (respectively near infinity).

Lemma 2.1. *Let $\Theta(t) = \exp\left\{\frac{t^2}{\log(e+t)}\right\} - 1$. Then the conjugate Young function $\tilde{\Theta}(t)$ of Θ is equivalent, near infinity, to the function*

$$\Psi(t) = t \log^{\frac{1}{2}}(e+t)(\log\log(e+t))^{\frac{1}{2}}.$$

Proof. Let us start the proof by observing that the derivative function of Θ

$$\theta(t) = \Theta'(t) = \exp\left\{\frac{t^2}{\log t}\right\} \frac{2t\log t - t}{\log^2 t}$$

is equivalent near infinity to Θ. In fact, for any t sufficiently large we have

$$\theta(t) \cong \exp\left\{\frac{t^2}{\log t}\right\} \frac{2t}{\log t}$$

and

$$\exp\left\{\frac{t^2}{\log t}\right\} \leqslant \exp\left\{\frac{t^2}{\log t}\right\} \frac{2t}{\log t} \leqslant \exp\left\{\frac{(ct)^2}{\log ct}\right\},$$

for some constant $c > 1$. On the other hand it is not hard to see that the inverse function θ^{-1} of θ is equivalent near infinity to the function

$$\psi(s) = \frac{1}{\sqrt{2}} \log^{\frac{1}{2}} s (\log\log s)^{\frac{1}{2}}.$$

Hence, near infinity we have

$$\tilde{\Theta}(y) = \int_0^y \theta^{-1}(s)ds \cong y \log^{\frac{1}{2}} y (\log\log y)^{\frac{1}{2}}$$

as we claimed. □

Theorem 2.1. *The continuous imbedding $L^\Psi(\Omega) \to L^\Phi(\Omega)$ holds if and only if either Ψ dominates Φ globally or $|\Omega| < \infty$ and Ψ dominates Φ near infinity.*

In particular, for any Young function $\Psi = \Psi(t)$ which is dominated (near infinity) by the Young function

$$\Theta(t) = \exp\left\{\frac{t^2}{\log(e+t)}\right\} - 1,$$

by Theorem 2.1 we have

$$EXP_2(\Omega) \to L^{\Theta}(\Omega) \to L^{\Psi}(\Omega). \tag{2.8}$$

Moreover, for any $0 < \varepsilon < p < \infty$ and $-\infty < a < b < \infty$ the following imbedding are obvious

$$L^{p+\varepsilon}(\Omega) \to L^p(\log L)^b(\Omega) \to L^p(\log L)^a(\Omega) \to L^{p-\varepsilon}(\Omega)$$

$$L^p(\log L)^{\varepsilon}(\Omega) \to L^p(\Omega) \to L^p(\log L)^{-\varepsilon}(\Omega).$$

The following Sobolev-Trudinger type embedding holds

$$W_0\frac{L^2}{(\log L)^a}(\Omega) \hookrightarrow EXP_{\frac{2}{1+a}}(\Omega) \qquad \text{for } a < 1, \tag{2.9}$$

(see [22], [10], [5]), where we denote by $W_0\frac{L^2}{(\log L)^a}(\Omega)$ the space $W_0^{1,\Phi}(\Omega)$ where $\Phi(t) = t^2 \log^{-a}(e+t)$. It is worth to point out that in case $a = 0$ imbedding (1.8) follows.

We will finish this section by recalling the following result (see [5], Example 2 pag. 43)

Lemma 2.2. *Let $\Omega \subset \mathbb{R}^2$ be an open bounded set with C^1-boundary. If we consider Young functions $\Phi(t)$ which are equivalent to $t^p(\log\log(e+t))^q$ near infinity, where either $p > 1$ and $q \in \mathbb{R}$ or $p = 1$ and $q \geqslant 0$, then*

$$W^{1,\Phi}(\Omega) \to C_b(\Omega)$$

if $p > 2$ and

$$W^{1,\Phi}(\Omega) \to L^{\Phi_2}(\Omega) \tag{2.10}$$

otherwise, where Φ_2 is equivalent near infinity to

$$\begin{cases} t^{\frac{2p}{2-p}}(\log\log(t))^{\frac{2q}{2-q}} & \text{if } 1 \leqslant p < 2 \\ e^{t^2(\log(t))^q} & \text{if } p = 2 \end{cases}$$

(Here $C_b(\Omega)$ denotes the space of continuous bounded functions on Ω).

For more details and proofs of results about Young function and Orlicz spaces we refer the reader to [1], [5], [6], [17], [20], [23].

3 – Preliminaries

The results we are going to obtain in this section are true in all dimensions. Hence, here we assume $A = A(x,\xi)$ to be defined on $\Omega \times \mathbb{R}^n$, where conditions (1.2)–(1.6) hold for $x \in \Omega \subset \mathbb{R}^n$ and $\xi, \eta \in \mathbb{R}^n$. Let us recall the following regularity result for the solution to quasilinear elliptic problem with the right-hand side in divergence form (see Theorem 3.2 of [3]).

THEOREM 3.1. *Let $\Omega \subset \mathbb{R}^n$ be a bounded open set with C^1-boundary and let $A = A(x,\xi)$ be as before. Then for $\psi_1, \psi_2 \in \frac{L^2}{(\log L)^a}(\Omega;\mathbb{R}^n)$ with $0 \leqslant a \leqslant 1$, each of the two problems*

$$\begin{cases} \operatorname{div} A(x,\nabla\varphi_1) = \operatorname{div} \psi_1 & \text{in } \Omega \\ \varphi_1 \in W_0^{1,1}(\Omega) \end{cases}$$

$$\begin{cases} \operatorname{div} A(x,\nabla\varphi_2) = \operatorname{div} \psi_2 & \text{in } \Omega \\ \varphi_2 \in W_0^{1,1}(\Omega) \end{cases}$$

has a unique solution and

$$\|\nabla\varphi_1 - \nabla\varphi_2\|_{\frac{L^2}{(\log L)^a}(\Omega)} \leqslant c(K)\|\psi_1 - \psi_2\|_{\frac{L^2}{(\log L)^a}(\Omega)} \tag{3.1}$$

where $c(K) > 0$ depends only on K.

We prove the following

3.2. *Let $A = A(x,\xi)$ satisfy hypotheses (1.2)–(1.6). Then for $\psi_1, \psi_2 \in \frac{L^2}{\log\log L}(\Omega)$ each of the two problems*

$$\begin{cases} \operatorname{div} A(x,\nabla\varphi_1) = \operatorname{div} \psi_1 & \text{in } \Omega \\ \varphi_1 \in W_0^{1,1}(\Omega) \end{cases}$$
$$\begin{cases} \operatorname{div} A(x,\nabla\varphi_2) = \operatorname{div} \psi_2 & \text{in } \Omega \\ \varphi_2 \in W_0^{1,1}(\Omega) \end{cases} \tag{3.2}$$

has a unique solution and

$$\|\nabla\varphi_1 - \nabla\varphi_2\|_{\frac{L^2}{\log\log L}(\Omega)} \leqslant c(K)\|\psi_1 - \psi_2\|_{\frac{L^2}{\log\log L}(\Omega)} \tag{3.3}$$

Proof. For $i = 1, 2$ let $\psi_i \in \frac{L^2}{\log\log L}(\Omega)$. Then obviously ψ_i belong to $\frac{L^2}{(\log L)^a}(\Omega)$, $0 < a \leqslant 1$. Hence, by Theorem 3.1, there exists a unique solution φ_i to the Dirichlet Problem (3.2) and the estimate

$$\|\nabla\varphi_1 - \nabla\varphi_2\|_{\frac{L^2}{(\log L)^a}(\Omega)} \leq c(K)\|\psi_1 - \psi_2\|_{\frac{L^2}{(\log L)^a}(\Omega)} \tag{3.4}$$

holds uniformly with respect to $a \in]0, 1]$.

Now we claim that the following inequality holds true:

$$\begin{aligned}\left(1 - \frac{1}{e}\right)\int_\Omega \frac{k(x)^2}{\log\log(k(x) + e^e)}dx \leqslant \int_0^1 da \int_\Omega \frac{k(x)^2}{\log^a(k(x) + e^e)}dx \leqslant \\ \leqslant \int_\Omega \frac{k(x)^2}{\log\log(k(x) + e^e)}dx.\end{aligned} \tag{3.5}$$

Indeed by

$$\int_0^1 \frac{1}{\log^a(e^e + k(x))}da = \left[1 - \frac{1}{\log(k(x) + e^e)}\right]\frac{1}{\log\log(k(x) + e^e)},$$

we have

$$\left(1 - \frac{1}{e}\right)\frac{1}{\log\log(k(x) + e^e)} \leqslant \int_0^1 \frac{1}{\log^a(e^e + k(x))}da \leqslant \frac{1}{\log\log(k(x) + e^e)}$$

so that Inequality (3.5) follows.

Integrating both sides of (3.4) with respect to $0 \leqslant a \leqslant 1$ and using suitably (2.4) and (3.5) with $k(x) = |\nabla\varphi_1 - \nabla\varphi_2|$ and $k(x) = |\psi_1 - \psi_2|$ the thesis follows. $\square$

4 – The main result

In this Section we will give the proof of Theorem 1.1. Here and below we assume

$$\Phi(t) = t\log^{\frac{1}{2}}(e + t)(\log\log(e + t))^{\frac{1}{2}}.$$

Proof of Theorem 1.1. We start the proof by using the linearization procedure contained in [15] (see also [3]) which we report for the convenience of the reader. So, let $v \in W_0^{1,2)}(\Omega)$ be the solution to quasilinear problem

$$\begin{cases} -\text{div}\, A(x, \nabla v) = f & \text{in } \Omega \\ v = 0 & \text{on } \partial\Omega, \end{cases} \tag{4.1}$$

which exists and is unique because $f \in L^\Phi(\Omega) \subset L^1(\Omega)$ (see [9], [15]). We will determine a symmetric measurable matrix valued function $\mathcal{A} = \mathcal{A}(x)$ such that v satisfies the linear problem

$$\begin{cases} -\text{div}\, \mathcal{A}(x)\nabla v = f & \text{in } \Omega \\ v = 0 & \text{on } \partial\Omega, \end{cases} \tag{4.2}$$

and $\mathcal{A}$ verifying

$$\frac{|\xi|^2}{C(K)} \leqslant \langle \mathcal{A}(x)\xi, \xi\rangle \leqslant C(K)|\xi|^2, \tag{4.3}$$

for any $\xi \in \mathbb{R}^2$, a.e.$x \in \Omega$, and where $C(K)$ is a constant depending only upon K.

Setting

$$B = A(x, \nabla v(x)), \qquad E = \nabla v(x). \tag{4.4}$$

one obtain, by assumptions (1.4)-(1.6)

$$|B| \leqslant K|E|, \qquad |E|^2 \leqslant K|\langle B, E\rangle|. \tag{4.5}$$

Moreover if we set,

$$\lambda = \frac{\langle B, E\rangle}{|E|^2}, \qquad \Lambda = \frac{|B|}{|E|} \qquad (|E| > 0)$$

by (4.5) we have

$$\frac{1}{K} \leqslant \lambda \leqslant \Lambda \leqslant K \qquad \text{and} \qquad \frac{|B|^2 + |E|^2}{\langle B, E\rangle} = \frac{1 + \Lambda^2}{\lambda}. \tag{4.6}$$

Define $H \geqslant 1$ by solving the equation

$$H + \frac{1}{H} = \frac{1 + \Lambda^2}{\lambda}$$

that is,

$$H = \frac{1}{2}\left[\frac{1 + \Lambda^2}{\lambda} + \sqrt{\left(\frac{1 + \Lambda^2}{\lambda} - 4\right)^2}\,\right].$$

Then, consider the 2×2 matrix defined by

$$\mathcal{A} = HI_d + \left(\frac{1}{H} - H\right)\frac{B - HE}{|B - HE|} \otimes \frac{B - HE}{|B - HE|},$$

where for $z = (x, t)$, we have used the shorthand notation

$$z \otimes z = \begin{pmatrix} x^2 & xt \\ xt & t^2 \end{pmatrix}$$

and $I_d = (\delta_{ij})$ is identical matrix. It holds (see [15])

$$\mathcal{A}E = B \tag{4.7}$$

and

$$\frac{|\xi|^2}{H} \leqslant \langle \mathcal{A}(x)\xi, \xi \rangle \leqslant H|\xi|^2, \qquad \forall \xi \in \mathbb{R}^2. \tag{4.8}$$

By (4.4) and (4.8), we have

$$\mathcal{A}(x)\nabla v(x) = B$$

which implies (4.2). Finally, by (4.8) and observing that it holds

$$H(x) \leqslant C(K),$$

(4.3) follows, with

$$C(K) = \frac{1}{2}\left[(K + K^3) + \sqrt{(K + K^3)^2 - 4}\right].$$

Now, let

$$L\cdot = -\mathrm{div}\ \mathcal{A}(x)\nabla \cdot .$$

Since $f \in L^\Phi(\Omega)$ then v is the variational solution in $W_0^{1,2}(\Omega)$ to the equation $Lv = f$. Hence we have

$$\int_\Omega \langle \mathcal{A}(x)\nabla v, \nabla\varphi \rangle dx = \int_\Omega \varphi f dx$$

for any $\varphi \in W_0^{1,2}(\Omega)$.

Now, let us fix $\psi \in \mathcal{C}^1(\bar{\Omega}; \mathbb{R}^2)$ with

$$\|\psi\|_{\frac{L^2}{(\log\log L)}(\Omega)} \leqslant 1 \tag{4.9}$$

and let φ be the (unique) solution to the Dirichlet problem

$$\begin{cases} L\varphi = \mathrm{div}\ \psi & \text{in } \Omega \\ \varphi = 0 & \text{on } \partial\Omega. \end{cases}$$

given by Theorem 3.2. Note that φ verifies

$$\|\nabla\varphi\|_{\frac{L^2}{(\log\log L)}(\Omega;\mathbb{R}^2)} \leqslant c(K)\|\psi\|_{\frac{L^2}{(\log\log L)}(\Omega;\mathbb{R}^2)} \leqslant c(K). \tag{4.10}$$

We have

$$|\langle\nabla v,\psi\rangle| = \left|\int_\Omega \langle\mathcal{A}(x)\nabla v,\nabla\varphi\rangle\,dx\right| = \left|\int_\Omega \varphi f\,dx\right|. \tag{4.11}$$

On the other hand, using Lemma 2.2 with $p = 2$ and $q = -1$, the Orlicz-Sobolev imbedding

$$W_0^{1,}\frac{L^2}{\log\log L}(\Omega) \to L^\Theta(\Omega) \qquad \text{where} \qquad \Theta(t) = \exp\frac{t^2}{\log(e+t)} - 1$$

holds. Moreover, by Lemma 2.1 the conjugate Young function $\tilde\Theta$ of Θ is equivalent (near infinity) to the Young function Φ and then

$$L^{\tilde\Theta}(\Omega) = L^\Phi(\Omega). \tag{4.12}$$

Thus, for any $\psi \in \mathcal{C}^1(\bar\Omega,\mathbb{R}^2)$ verifying (4.9), by (4.11) and using Hölder inequality between associated Orlicz spaces (see for example [1]), we obtain

$$|\langle\nabla v,\psi\rangle| \leqslant c\|\varphi\|_{L^\Theta(\Omega)}\|f\|_{L^\Phi(\Omega)} \tag{4.13}$$

Taking the supremum under conditions $\psi \in \mathcal{C}^1(\bar\Omega;\mathbb{R}^2)$ and $\|\psi\|_{\frac{L^2}{(\log\log L)}(\Omega)} \leqslant 1$, the estimates (4.10) and (4.13) give

$$\sup\left\{|\langle\nabla v,\psi\rangle| : \psi \in \mathcal{C}^1(\bar\Omega;\mathbb{R}^2) \text{ and } \|\psi\|_{\frac{L^2}{(\log\log L)}(\Omega)} \leqslant 1\right\} \leqslant c(K,|\Omega|)\|f\|_{L^\Phi(\Omega)}$$

and the thesis follows. In fact it is now sufficient to observe that

$$\|\nabla v\|_{L^2(\log\log L)(\Omega)} = \sup_{\|\psi\|_{\frac{L^2}{(\log\log L)}(\Omega)}\leqslant 1} |\langle\nabla v,\psi\rangle|.$$

and that by (2.5) the space $C^1(\bar\Omega)$ is dense in $\frac{L^2}{\log\log L}(\Omega)$. ∎

Remark 4.1 It is evident that the thesis of Theorem 1.1 remains invaried whenever $f \in L^\Psi(\Omega)$, Ψ any Young function verifying

$$\Psi(t) \geqslant t\log^{\frac12}(e+t)(\log\log(e+t))^{\frac12}$$

for any $t > 0$ sufficiently large.

5 – On the boundedness of the solution

In this section we show with an example that we cannot expect the boundedness of the solution under the assumptions of Theorem 1.1 (see also [3], [14]).

Example 1. Let

$$u(x) = \log\log\log\frac{1}{|x|}$$

and let $\Omega = \{x \in \mathbb{R}^2 : |x| < e^{-e}\}$. Then, the unbounded function u verifies $|\nabla u| \in L^2 \log\log L(\Omega)$ and solves the Dirichlet problem

$$\begin{cases} -\Delta u = f & \text{in } \Omega \\ u \in W_0^{1,2}(\Omega), \end{cases} \tag{5.1}$$

where

$$f := \frac{1}{|x|^2 \log^2 \frac{1}{|x|} \log\log \frac{1}{|x|}} \left(1 + \frac{1}{\log\log\frac{1}{|x|}}\right) \in L(\log L)(\log\log L)^\alpha(\Omega), \quad \forall \alpha < 0.$$

Proof. We have

$$\nabla u(x) = \frac{-x}{|x|^2 \log\frac{1}{|x|}\log\log\frac{1}{|x|}}, \qquad \forall\, x \neq 0,$$

so that

$$|\Delta u(x)| = |\mathrm{div}\ \nabla u(x)| = \frac{1}{|x|^2 \left(\log\frac{1}{|x|}\right)^2 \log\log\frac{1}{|x|}} \left(1 + \frac{1}{\log\log\frac{1}{|x|}}\right).$$

Hence, by $|f| = |\Delta u|$ we have, for any $\alpha < 0$,

$$\begin{aligned} \int_\Omega |f| \log(|f|)\,(\log\log|f|)^\alpha\, dx \leqslant \\ &\leqslant c\int_\Omega \frac{1}{|x|^2 \log\frac{1}{|x|}\left(\log\log\frac{1}{|x|}\right)^{1-\alpha}} dx = \\ &= c\int_0^{e^{-e}} \frac{1}{\rho\log\frac{1}{\rho}(\log\log\frac{1}{\rho})^{1-\alpha}} d\rho = \\ &= \frac{c}{-\alpha}\left[\left(\log\log\frac{1}{\rho}\right)^\alpha\right]_0^{e^{-e}} < \infty, \end{aligned}$$

so that f belongs to $L\log L(\log\log L)^\alpha(\Omega)$ for any $\alpha < 0$. Note that for $\alpha = 0$ first integral in last inequality is infinite. ∎

In a similar way we have the following

Example 2. Let

$$u(x) = \log\log\log\log \frac{1}{|x|}$$

and let $\Omega = \left\{x \in \mathbb{R}^2 : |x| < e^{-e^e}\right\}$. Then, the unbounded function u verifies $\nabla u \in L^2 \log\log L(\Omega)$ and solves the Dirichlet problem

$$\begin{cases} -\Delta u = f & \text{in } \Omega \\ v \in W_0^{1,2}(\Omega), \end{cases} \tag{5.2}$$

where

$$f := \frac{1}{|x|^2 \log^2 \frac{1}{|x|} \log\log \frac{1}{|x|} \log\log\log \frac{1}{|x|}} \left(1 + \frac{1}{\log\log \frac{1}{|x|}} + \frac{1}{\log\log \frac{1}{|x|} \log\log\log \frac{1}{|x|}}\right)$$

and holds

$$f \in L(\log L)(\log\log\log L)^\alpha(\Omega), \quad \forall \alpha < 0.$$

By continuing in the same way, we can conclude that if by one hand $f \in L\log L$ is a sufficient condition to obtain the boundedness of the solution u (see [2]) by the other hand slightly weaker condition $f \in L\log L(\log\log\log ... \log L)^\alpha(\Omega)$, $\alpha < 0$, is insufficient.

Acknowledgements

The author wishes to express her thanks to Prof. A. Cianchi and to Prof. A. Verde for the helpful suggestions during the preparation of the paper.

REFERENCES

[1] R. A. Adams: *Sobolev Spaces*, Academic Press, New York, 1975.

[2] A. Alberico, V. Ferone: *Regularity properties of solutions of elliptic equations in* $\mathbb{R}^2$ *in limit cases*, Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur. Rend. Lincei (9) Suppl., **6** (1995), 237–250.

[3] A. Alberico, T. Alberico, C. Sbordone: *Regularity results for planar quasi-linear equations with right hand side in* $L(\log L)^\delta$, (2010) to appear.

[4] C. Bennett, K. Rudnick: *On Lorentz-Zygmund spaces*, Dissert. Math. **175** (1980), 1–67.

[5] A. Cianchi: *A Sharp Embedding Theorem for Orlicz-Sobolev Spaces*, Indiana University Mathematics J., **45** (1996), 39-65.

[6] A. Cianchi: *Continuity Properties of Functions from Orlicz-Sobolev Spaces and Embedding Theorems*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. **23** (1996), 575–608.

[7] T. K. Donaldson, N. S. Trudinger: *Orlicz-Sobolev spaces and imbedding theorems*, J. Funct. Anal., **8** (1971), 52–75.

[8] D. E. Edmunds, H. Triebel: *Function spaces, entropy numbers, differential operators*, Cambridge University Press, Cambridge, 1996.

[9] A. Fiorenza, C. Sbordone: *Existence and uniqueness results for solutions of nonlinear equations with right hand side in L^1*, Studia Mathematica **127** (1998), 223–231.

[10] N. Fusco, P. L. Lions, C.Sbordone: *Sobolev imbedding theorems in borderline cases*, Proc. Amer. Math. Soc. **124** (1996), 561–565.

[11] L. Greco: *A remark on the equality det Df= Det Df*, Diff. Int. Eq. **6** (1993), 1089–1100.

[12] N. Ioku: *Brezis-Merle type inequality for a weak solution to the N-Laplace equation in Lorentz-Zygmund spaces*, Diff. Int. Eq., n. 5-6 **22** (2009), 495–518.

[13] T. Iwaniec, G. Martin: *Geometric function theory and nonlinear analysis*, Oxford Math. Monographs (2001).

[14] T. Iwaniec, J. Onninen: *Continuity Estimates for n-harmonic Equations*, Indiana Univ. Math. J., **56** (2007), 805–824.

[15] T. Iwaniec, C. Sbordone: *Quasiharmonic Fields*, Ann. Inst. Poincaré Anal. Non Linéaire 18, 5 (2001), 519-572.

[16] T. Iwaniec, A. Verde: *On the operator $\mathcal{L}(f) = f \log |f|$,*, J. Funct. Anal., **169** (1999), 391–420.

[17] M. A. Krasnosel'skiĭ, Ya. B. Rutickiĭ: *Convex Functions and Orlicz Spaces.*, Noordhoff, Groningen, 1961.

[18] J. Leray, J. L. Lions: *Quelques résultats de Višik sur les problèmes elliptiques non linéaires par les méthodes de Minty-Browder*, Bull. Soc. Math. France **93** (1965), 97–107.

[19] A. Passarelli di Napoli, C. Sbordone: *Elliptic equations with right-hand side in $L(\log L)^\alpha$*, Rend. Accad. Sci. Fis. Mat. Napoli (4), **62** (1995), 301–314.

[20] M. Rao, Z. D. Ren: *Theory of Orlicz spaces*, Monographs and Textbooks in Pure and Applied Mathematics 146 Marcel Dekker, Inc., New York, 1991.

[21] G. Stampacchia: *Some limit cases of L^p-estimates for solutions of second order elliptic equations*, Comm. Pure Appl. Math. **16** (1963), 505–510.

[22] N. S. Trudinger: *On imbeddings into Orlicz spaces and applications*, J. Math. Mech. **17** (1967), 473–483.

[23] A. Verde, G. Zecca: *On the higher integrability for certain nonlinear problems*, Differential and Integral Equations, **21**(2008), 247–263

Lavoro pervenuto alla redazione il 20 luglio 2010
ed accettato per la pubblicazione il 28 ottobre 2010.
Bozze licenziate il 26 novembre 2010

Indirizzo dell'Autore:

Gabriella Zecca – Dipartimento di Matematica e Applicazioni "R. Caccioppoli" – Università degli Studi di Napoli "Federico II" – Via Cintia - Complesso Universitario Monte S. Angelo – 80126 Napoli – Italy
E-mail: g.zecca@unina.it

INDICE DEL VOLUME 30