Rostocker Mathematisches Kolloquium

Heft 57

Universität Rostock

Fachbereich Mathematik

2003

Herausgeber: Fachbereich Mathematik
der Universität Rostock

Wissenschaftlicher Beirat: H.-D. Gronau
F. Liese
P. Tákač
G. Wildenhain

Schriftleitung: W. Plischke

Herstellung der Druckvorlage: S. Dittmer
H. Schubert

Zitat–Kurztitel: Rostock. Math. Kolloq. **57** (2003)

ISSN 0138-3248

© Universität Rostock, Fachbereich Mathematik, D - 18051 Rostock

BEZUGSMÖGLICHKEITEN: Universität Rostock
Universitätsbibliothek, Schriftentausch
18051 Rostock
Tel.: +49-381-498 22 81
Fax: +49-381-498 22 68
e-mail: maria.schumacher@ub.uni-rostock.de

Universität Rostock
Fachbereich Mathematik
18051 Rostock
Tel.: +49-381-498 6551
Fax: +49-381-498 6553
e-mail: romako@mathematik.uni-rostock.de

DRUCK: Universitätsdruckerei Rostock

Rostock. Math. Kolloq. **57**, 3–51 (2003)

Subject Classification (AMS)
62J02,62F12

Fiedrich Liese, Igor Vajda

On $\sqrt{n}$–Consistency and Asymptotic Normality of Consistent Estimators in Models with Independent Observations[1]

ABSTRACT. The paper presents relatively simple verifiable conditions for $\sqrt{n}$-consistency and asymptotic normality of M-estimators of vector parameters in a wide class of statistical models. The conditions are established for the M-estimators with absolutely continuous ρ-function of locally bounded variation, and for the class of models including e.g. the linear and the nonlinear regression, the generalized linear models and the proportional hazards models as special cases. The conditions are verified on L_1 and L_2 estimators embedded into a continuum of their alternative versions, as well as on one new class of M-estimators of parameters of exponential families which are shown to be robust in the sense of bounded gross-error sensitivity. Comparisons with known conditions for special models indicate that the present general conditions are not too restrictive in special situations and that sometimes they are even weaker than the previously published special conditions.

1 Introduction and basic concepts

We consider a general parametric statistical model with independent observations. In other words, for every $n \in \mathbb{N}$ we consider a random sample $\mathbf{Y}_n = (Y_1, \ldots, Y_n)'$ of independent real valued observations,

$$\mathbf{Y}_n \sim G(y_1, \ldots, y_n) = \prod_{i=1}^{n} G(y_i|i, \theta_0), \tag{1.1}$$

where θ_0 is a true value of a parameter $\theta = (\theta_1, \ldots, \theta_m)' \in \Theta$ for open $\Theta \subset \mathbb{R}^m$ and

$$\mathcal{G}_1 = \{G(y|1, \theta) : \theta \in \Theta\}, \ldots, \mathcal{G}_n = \{G(y|n, \theta) : \theta \in \Theta\} \tag{1.2}$$

are given families of distribution functions (briefly *distributions*) possibly depending on the sample size n. This means that we admit the triangular observation schemes $(Y_1, \ldots, Y_n) = (Y_1^{(n)}, \ldots, Y_n^{(n)})$. Important particular versions of this model are discussed in Section 2.

[1]Supported by the grant A 1075101.

We study a general M-estimator of the unknown true parameter θ_0 in the above considered model. This estimator is defined as a sequence of Θ-valued measurable functions $\hat{\theta}_n = \hat{\theta}_n(\mathbf{Y}_n)$ minimizing on Θ the random functions

$$M_n(\theta) = \frac{1}{n}\sum_{i=1}^{n} \rho(Y_i - \varphi_i(\theta)), \tag{1.3}$$

where $\rho : \mathbb{R} \mapsto \mathbb{R}$ is a given function called *criterion function* and $\varphi_1 : \Theta \mapsto \mathbb{R}, \dots, \varphi_n : \Theta \mapsto \mathbb{R}$ are given functions called *locators*. The locators may depend on the sample size n, i. e. we admit triangular schemes of locators

$$(\varphi_1, \dots, \varphi_n) = (\varphi_1^{(n)}, \dots, \varphi_n^{(n)}). \tag{1.4}$$

Since the M-estimator under consideration is defined by the criterion function and locators, we use the symbols

$$\hat{\theta}_n \sim \langle \rho; \varphi_1, \dots, \varphi_n \rangle \quad \text{or briefly} \quad \hat{\theta}_n \sim \langle \rho; \varphi_i \rangle. \tag{1.5}$$

We are interested in the asymptotic properties of M-estimators $\hat{\theta}_n \sim \langle \rho; \varphi_i \rangle$ when the sample size n tends to infinity. Therefore, unless otherwise explicitly stated, all asymptotic relations, formulas and properties are automatically considered for $n \to \infty$.

Our attention is restricted to the M-estimators $\hat{\theta}_n \sim \langle \rho; \varphi_i \rangle$ with criterion functions ρ absolutely continuous on bounded intervals of $\mathbb{R}$ (briefly, absolutely continuous on $\mathbb{R}$). This means that there exists a measurable function $\psi : \mathbb{R} \mapsto \mathbb{R}$ satisfying the condition

$$\psi(y) = \frac{d\rho(y)}{dy} \quad \text{a.e.} \tag{1.6}$$

with respect to the Lebesgue measure on $\mathbb{R}$ and absolutely integrable on bounded intervals. We shall consider a right-continuous extension of ψ on $\mathbb{R}$ which is (up to a constant $\rho(0)$ playing no role in the definition of M-estimator $\hat{\theta}_n$ (cf. (1.3)) one-one related to ρ and satisfies for all a, $b \in \mathbb{R}$ the relation

$$\rho(b) - \rho(a) = \int_{(a,b]} \psi(y)dy \tag{1.7}$$

(the so-called fundamental theorem of calculus for Lebesgue integrals, cf. Theorem 18.16 in Hewitt and Stromberg [9]). Here, and in the sequel,

$$\int_{(a,b]} = -\int_{(b,a]} \quad \text{if } b < a. \tag{1.8}$$

The right-continuous function $\psi : \mathbb{R} \mapsto \mathbb{R}$ characterizes a *sensitivity* of the M-estimator $\hat{\theta}_n \sim \langle \rho; \varphi_i \rangle$ to small deviations of observations $Y_1, \dots, Y_n$ (an appropriately normed version of ψ is an *influence function* of the M-estimator, see Huber [12] or Hampel et al [8]). Due to the one-one relation between the criterion function ρ and the sensitivity function ψ mentioned above, we can replace the representation of M-estimators (1.5) by

$$\hat{\theta}_n \sim \langle \psi; \varphi_1, \ldots, \varphi_n \rangle \quad \text{or, briefly} \quad \hat{\theta}_n \sim \langle \psi; \varphi_i \rangle. \tag{1.9}$$

Our theory is restricted to the estimators $\hat{\theta}_n \sim \langle \psi; \varphi_i \rangle$ with sensitivity ψ of a locally bounded variation. This means that ψ is a difference of two nondecreasing functions ψ^+ and ψ^- which are assumed to be continuous from the right. This theory presents conditions for $\sqrt{n}-$ consistency and asymptotic normality of $\hat{\theta}_n$ in terms of the sum $\psi^{\pm} = \psi^+ + \psi^-$.

As is indicated by the title of the paper, our main results are restricted to the M-estimators $\hat{\theta}_n \sim \langle \psi; \varphi_i \rangle$ which are consistent in the standard sense

$$\hat{\theta}_n \xrightarrow{P} \theta_0 \ . \tag{1.10}$$

We present conditions on the sensitivity function ψ, locators φ_i and the model (1.1) under which $\hat{\theta}_n$ is $\sqrt{n}-$consistent in the sense

$$\lim_{y\to\infty} \lim_{n\to\infty} \mathsf{P}\left(\sqrt{n}\left\|\hat{\theta}_n - \theta_0\right\| > y\right) = 0 \tag{1.11}$$

and asymptotically normal in the sense

$$\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{\mathcal{L}} N(0, V) \tag{1.12}$$

and under which the variance-covariance $m \times m$ matrix V can be explicitly evaluated.

These main results are presented in the next Section 2. The conditions on the sensitivity function ψ, locators φ_i and the model (1.1) are formulated as regularity conditions (R1) – (R4+). Important particular versions of the general model (1.1) and sufficient conditions for (R1) – (R4+) are in Section 3.

The consistency (1.10) in reasonably general classes of M-estimators (1.9) and models (1.1) is a difficult problem. Sufficient conditions have been established e. g. in Yohai and Maronna [31], Zhao and Chen [32], Hjort and Pollard [10], Liese and Vajda [18]-[21], Zhao [33], Arcones [1]-[2] and some other references therein. Presentation of such conditions would increase the complexity and size of the paper above bearable bounds. Therefore we refer in this respect to the mentioned literature and restrict ourselves to the verification of consistency only in special cases illustrating applicability of the main result of Section 2.

In Sections 4 and 5 we illustrate the applicability of the general results of Sections 2 and 3 to special classes of M-estimators (1.9) and models (1.1). Particular attention is payed to the class of M-estimators with the criterion functions

$$\rho(y) = \rho_\beta(y) = \beta\, y\, I_{[0,\infty)}(y) - (1-\beta)\, y\, I_{(-\infty,0)}(y), \quad 0 < \beta < 1, \tag{1.13}$$

introduced by Koenker and Basset [16] and later used by many authors (e. g. Portnoy [24], Koul and Saleh [17], Jurečková and Sen [14], Hallin and Jurečková [7]).

In Section 6 are proofs of main results of Section 2. The proofs employ some general results and techniques of van der Vaart and Wellner [30], in particular their Theorems 3.2.2 and 3.2.5. The proofs use also the methods developed in [21].

The present paper differs from[21] in a considerably simpler formulations and proofs of results, and in application of these results to different special models (1.1) and/or estimators (1.9). It also differs from the classical literature studying the consistency (1.10) and the asymptotic normality (1.12) of the estimators $\hat{\theta}_n$ defined as solutions of the equations

$$\sum_{i=1}^{n} \psi(Y_i - \varphi_i(\theta))\, \nabla\varphi_i(\theta) = 0 \tag{1.14}$$

on Θ when the locators $\varphi_i(\theta)$ are differentiable on Θ with gradients $\nabla\varphi_i(\theta)$ (see the monographs of Serfling [27], [12], Singer and Sen [28], [14], and references therein). Obviously, our M-estimators $\hat{\theta}_n \sim \langle\psi;\varphi_i\rangle$ coincide with solutions of (1.14) only in special cases, e. g. if the sensitivity ψ is monotone on $\mathbb{R}$ (i. e. the criterion function ρ is convex) and the locators $\varphi_i(\theta)$ are linear in θ. This takes place e. g. if $\theta \in \Theta = \mathbb{R}$ is the location parameter, $\varphi_i(\theta) = \theta$ and

$$\rho(y) = \begin{cases} y^2 & \text{for} \quad |y| \le k, \\ 2k|y| - k^2 & \text{for} \quad |y| > k, \end{cases}$$

which is the situation studied by Huber [11]. The results about asymptotic normality of solutions $\hat{\theta}_n$ of (1.14), based on the ideas and techniques of [11, 12], are thus disjoint with our results except the relatively rare situations when solutions of (1.14) minimize the function $M_n(\theta)$ of (1.3). Such situations are trivial from the point of view of our theory which primarily intends to bring results about M-estimators $\hat{\theta}_n \sim \langle\rho;\varphi_i\rangle$ where either $\rho(y)$ is not convex in $y \in \mathbb{R}$ or $\varphi_i(\theta)$ are not linear in $\theta \in \Theta$, i. e. about situations not covered by the classical Huber-type theories.

2 Main results

In this section we consider an arbitrary model (1.1) and an arbitrary M-estimator $\hat{\theta}_n \sim \langle\psi;\varphi_i\rangle$ (equivalently, $\widehat{\theta}_n \sim \langle\rho;\varphi_i\rangle$, see (1.5) and (1.9)) with the variation of ψ *locally bounded*, i. e. bounded on bounded intervals of $\mathbb{R}$. This means that there exist nondecreasing functions ψ^+, $\psi^- : \mathbb{R} \mapsto \mathbb{R}$ with the property

$$\psi = \psi^+ - \psi^-. \tag{2.1}$$

We define on $\mathbb{R}$ the nondecreasing function

$$\psi^{\pm} = \psi^+ + \psi^-. \tag{2.2}$$

Definition 2.1 *We say that the locators φ_i are adapted to the model if*

$$\mathsf{E}\psi(Y_i - \varphi_i(\theta_0)) = 0, \quad i \in \mathbb{N}. \tag{2.3}$$

We say that the estimator $\hat{\theta}_n$ is adapted to the model if the locators are adapted in the sense of (2.3) and the estimator is consistent in the sense of (1.10).

In the rest of paper we consider the following conditions of regularity of the estimator $\hat{\theta}_n$ in the model (1.1).

(R1) The second moments (variances if (2.3) holds)

$$\sigma_i^2 = \mathsf{E}\left[\psi(Y_i - \varphi_i(\theta_0))\right]^2 \tag{2.4}$$

are uniformly bounded in the mean, i. e.,

$$\sup_{n\in\mathbb{N}} \frac{1}{n} \sum_{i=1}^{n} \sigma_i^2 < \infty. \tag{2.5}$$

(R2) The gradients

$$\dot{\varphi}_i(\theta) = \left(\frac{\partial}{\partial\theta_1}, \ldots, \frac{\partial}{\partial\theta_m}\right)' \varphi_i(\theta), \quad \theta \in \Theta,\ i \in \mathbb{N} \tag{2.6}$$

exist and are locally bounded and locally Lipschitz in the sense that one can find a closed ball

$$B = B_\delta(\theta_0) = \{\mathbf{y} \in \mathbb{R}^m : \|\mathbf{y} - \theta_0\| \le \delta\} \tag{2.7}$$

and a constant $\lambda > 0$ possibly depending on B, such that $B \subset \Theta$ and

$$\|\dot{\varphi}_i(\theta)\| \le \lambda, \quad \theta \in B,\ i \in \mathbb{N} \tag{2.8}$$

$$\|\dot{\varphi}_i(\theta) - \dot{\varphi}_i(\tilde{\theta})\| \le \lambda\|\theta - \widetilde{\theta}\|, \quad \theta, \widetilde{\theta} \in B,\ i \in \mathbb{N}. \tag{2.9}$$

(R3) There exists $\tau_0 > 0$ such that the functions

$$H_i(t) = \mathsf{E}\,\psi(Y_i - \varphi_i(\theta_0) + t), \quad i \in \mathbb{N} \tag{2.10}$$

are differentiable on the interval $(-\tau_0, \tau_0)$ and the derivatives

$$h_i(t) = \frac{d}{dt} H_i(t), \quad i \in \mathbb{N} \tag{2.11}$$

satisfy the condition

$$\lim_{\tau\downarrow 0} \sup_{n\in\mathbb{N}} \frac{1}{n} \sum_{i=1}^{n} \omega(h_i, \tau) = 0 \tag{2.12}$$

where $\omega(h_i,\tau) = \sup_{|t|\le\tau} |h_i(0) - h_i(t)|$, $0<\tau<\tau_0$, is the modulus of continuity of $h_i(t)$ in the neighborhood of $t=0$. Further, the variances σ_i^2 from (R1), gradients $\dot\varphi_i$ from (R2) and functions h_i from (R3) satisfy

$$\Sigma_n = \frac{1}{n}\sum_{i=1}^n \sigma_i^2 \dot\varphi_i(\theta_0)\,\dot\varphi_i(\theta_0)' \to \Sigma, \tag{2.13}$$

$$\Phi_n = \frac{1}{n}\sum_{i=1}^n h_i(0)\dot\varphi_i(\theta_0)\,\dot\varphi_i(\theta_0)' \to \Phi \tag{2.14}$$

where the $m\times m$ matrices Σ and Φ are positive definite.

(R4) There exist constants $\tau_0>0$ and κ such that the function (2.2) satisfies for all $0<\tau<\tau_0$ the relation

$$\sup_{n\in\mathbb{N}} \frac{1}{n}\sum_{i=1}^n \mathsf{E}\left[\psi^\pm(X_i+\tau) - \psi^\pm(X_i-\tau)\right]^2 < \kappa \tag{2.15}$$

where $X_i = Y_i - \varphi(\theta_0)$.

(R4+) There exist constants $\tau_0>0$ and $q>0$ and κ such that the function (2.2) satisfies for all $0<\tau<\tau_0$ the relation

$$\sup_{n\in\mathbb{N}} \frac{1}{n}\sum_{i=1}^n \mathsf{E}\left[\psi^\pm(X_i+\tau) - \psi^\pm(X_i-\tau)\right]^2 < \kappa\tau^q \tag{2.16}$$

where $X_i = Y_i - \varphi(\theta_0)$.

Sufficient conditions for (R3), (R4) and (R4+) will be studied in the next section. Here we formulate the main result of the paper. We remind that the asymptotic relations are considered for $n\to\infty$ unless otherwise stated.

Theorem 2.2 *If the estimator $\widehat\theta_n \sim \langle\psi;\varphi_i\rangle$ is adapted to the model (1.1) in the sense of Definition 2.1 and satisfies the regularity conditions (R1) – (R4) then it is $\sqrt{n}$–consistent in the sense of (1.11).*

Theorem 2.3 *Let the estimator $\widehat\theta_n \sim \langle\psi;\varphi_i\rangle$ be adapted to the model (1.1) in the sense of Definition 2.1 and satisfy the regularity conditions (R1) – (R4) and (R4+). If*

$$n^{-1/2}\sum_{i=1}^n \psi(Y_i - \varphi_i(\theta_0))\,\dot\varphi_i(\theta_0) \xrightarrow{\mathcal{L}} N(0,\Sigma) \tag{2.17}$$

then the estimator $\hat\theta_n$ is asymptotically normal in the sense of (1.12) with the variance-covariance matrix

$$V = \Phi^{-1}\,\Sigma\,\Phi^{-1}. \tag{2.18}$$

The proofs of Theorem 2.2 and 2.3 are deferred to Section 6. Here we present a sufficient condition for the condition (2.17) of Theorem 2.3.

Proposition 2.4 *If the assumptions (2.3) and (2.13) hold and for some $\gamma > 0$,*

$$\sup_{i\in\mathbb{N}} \mathsf{E}\|\psi(Y_i - \varphi_i(\theta_0))\,\dot{\varphi}(\theta_0)\|^{2+\gamma} < \infty \tag{2.19}$$

then the asymptotic normality condition (2.17) holds.

Proof: Clear from the Lyapunov central limit theorem. □

3 Results under restricted generality

In this section we restrict in different ways the generality of the model (1.1) and also the generality of the M-estimator $\hat{\theta}_n \sim \langle \psi; \varphi_i \rangle$ studied in the previous section. We study sufficient conditions for the assumptions of Theorems 2.2 and 2.3 under this restricted generality.

Definition 3.1 *The general statistical model with independent observations defined by (1.1) is said to be*

(i) regression model *if there are given sets $\mathcal{X} \subset \mathbb{R}^k$, $T \subset \mathbb{R}$, and a mapping $\phi : \mathcal{X}\times\Theta \mapsto T$, and if for $1 \le i \le n$ are given realizations $\mathbf{x}_i$ of $\mathbf{x} = (x_1, \dots, x_k)' \in \mathcal{X}$ and families of distributions $\mathcal{F}_i = \{F_i(y|\vartheta) : \vartheta \in T\}$, both possibly depending on n, such that*

$$G(y|i,\theta) = F_i(y|\phi(\mathbf{x}_i,\theta)) \quad \textit{for } 1 \le i \le n \quad \textit{and} \quad \theta \in \Theta; \tag{3.1}$$

(ii) homogeneous regression model *if it satisfies (i) and*

$$\mathcal{F}_i = \mathcal{F} = \{F(y|\vartheta) : \vartheta \in T\} \quad \textit{for } 1 \le i \le n \tag{3.2}$$

where the family of distributions $\mathcal{F}$ depends neither on i nor on n;

(iii) linear regression model *if it satisfies (i), $\mathcal{X}$ belongs to the same Euclidean space $\mathbb{R}^m$ as Θ and*

$$\phi(\mathbf{x},\theta) = \mathbf{x}'\theta \quad \textit{for } \mathbf{x} \in \mathcal{X} \textit{ and } \theta \in \Theta; \tag{3.3}$$

(iv) regression model with additive errors *if it satisfies (i) and $\mathcal{F}_i$ are location families not depending on n, i. e. if $T = \mathbb{R}$ and*

$$\mathcal{F}_i = \{F_i(y - \vartheta) : \vartheta \in \mathbb{R}\}, \quad 1 \le i \le n, \tag{3.4}$$

for a sequence of parent distributions $F_1(y)$, $F_2(y), \dots$ not depending on n.

The combinations of properties (ii) – (iv) of regression models are admitted. In this manner we obtain the following important special cases.

Example 3.2 Homogeneous regression with additive errors. This means the *standard nonlinear regression* where the observations are defined by formula

$$Y_i = \phi(\mathbf{x}_i, \theta_0) + \mathcal{E}_i, \quad 1 \le i \le n, \tag{3.5}$$

and the additive errors $\mathcal{E}_i$ are i.i.d. by the parent F of the location family $\mathcal{F} = \{F(y - \vartheta) : \vartheta \in \mathbb{R}\}$ satisfying simultaneously the assumptions (3.2) and (3.4).

Example 3.3 Homogeneous linear regression with additive errors. This means the *standard linear regression* where $\mathcal{X} \subset \mathbb{R}^m$ and

$$Y_i = \mathbf{x}_i'\theta_0 + \mathcal{E}_i, \quad 1 \le i \le n, \tag{3.6}$$

where the additive errors $\mathcal{E}_i$ satisfy the conditions of Example 3.2.

Example 3.4 The general homogeneous regression leads to independent observations

$$Y_i \sim F(y|\phi(\mathbf{x}_i, \theta_0)), \quad 1 \le i \le n, \tag{3.7}$$

specified by a $k \times n$ matrix

$$\mathbf{X}_n = (\mathbf{x}_1, \dots, \mathbf{x}_n) \tag{3.8}$$

of regressors and a family of distributions $\mathcal{F} = \{F(y|\vartheta) : \vartheta \in T\}$. If $\mathcal{F}$ is a location family then we obtain the standard nonlinear regression of Example 3.2.

Example 3.5 The homogeneous linear regression in general differs from the standard linear regression. It has been called *pseudolinear regression* in Liese and Vajda [20]. Here the independent observations

$$Y_i \sim F(y|\mathbf{x}_i'\theta_0), \quad 1 \le i \le n, \tag{3.9}$$

are specified by the matrix (3.8) and by a family of distributions $\mathcal{F} = \{F(y|\vartheta) : \vartheta \in T\}$. If $T = \mathbb{R}$ and $\mathcal{F}$ is a location family then the pseudolinear regression reduces to the standard linear regression of Example 3.3. If $\mathcal{F}$ is an exponential family then the pseudolinear regression model reduces to the *generalized linear model*. As an example of the generalized linear regression we can consider the *Cox model* where $\mathcal{F}$ consists of the exponential distributions $F(y|\vartheta) = 1 - \exp\{\vartheta \ln(1 - F(y))\}$, $\vartheta \in \mathbb{R}$, for a given distribution $F(y) = F(y|1)$ differentiable on the support $(0, \infty)$ (then $\Lambda(y) = -\ln(1 - F(y))$ is a cumulative hazard function).

Next we study the adaptation condition (2.3) in the homogeneous regression models and standard nonlinear regression models introduced above. This condition means in fact that

$$\int \psi(y - \varphi_i(\theta))\, dG(y|i, \theta) = 0 \quad \text{for all} \quad \theta \in \Theta \text{ and } i \in \mathbb{N}. \tag{3.10}$$

In the homogeneous regression models the adaptation (3.10) reduces to evaluation of solutions $a(\vartheta)$ of the system of equations

$$\int \psi(y-a)\, dF(y|\vartheta) = 0, \quad \vartheta \in T, \tag{3.11}$$

in the real variable $a \in \mathbb{R}$. Indeed, by (3.1) and (3.2), (3.10) holds provided

$$\varphi_i(\theta) = a(\phi(\mathbf{x}_i, \theta)) \quad \text{if} \quad \int \psi(y - a(\vartheta))\, dF(y|\vartheta) = 0, \quad \vartheta \in T. \tag{3.12}$$

In the standard nonlinear regression of Example 3.2 with an error distribution $F(y)$, the adaptation condition (3.12) further simplifies into

$$\varphi_i(\theta) = \phi(\mathbf{x}_i, \theta) + b(F) \quad \text{if} \quad b(F) = a(0), \quad \text{i.e.} \quad \int \psi(y - b(F))\, dF(y) = 0. \tag{3.13}$$

An M-estimator $\hat{\theta}_n \sim \langle \psi; \phi(\mathbf{x}_i, \theta) + c\rangle$ with a fixed $c \in \mathbb{R}$ is in fact adapted to all nonlinear regression models (3.5) with error distributions F restricted by the condition $b(F) = c$. However, this condition may not be easily verifiable for some functions ψ. In order to obtain an $M-$estimator adapted to the standard nonlinear regression models (3.5) with an arbitrary error distribution F, it suffices to extend the parameter space Θ into $\Theta^* = \Theta \times \mathbb{R}$ and replace $\varphi_i(\theta) = \phi(\mathbf{x}_i, \theta) + c$ by

$$\varphi_i^*(\theta^*) = \phi(\mathbf{x}_i, \theta) + b \quad \text{for } \theta^* = (\theta, b) \in \Theta^*,$$

i. e. to consider the M-estimator

$$\hat{\theta}_n^* = (\widehat{\theta}_n, \hat{b}_n) \sim \langle \psi; \phi(\mathbf{x}_i, \theta) + b\rangle \tag{3.14}$$

of the extended true parameter $\theta_0^* = (\theta_0, b_0)$ where $b_0 = b(F)$. The validity of (2.3) for $\hat{\theta}_n^*$, i. e. the validity of (2.3) with $\varphi_i(\theta_0)$ replaced by $\varphi_i^*(\theta_0^*) = \phi(\mathbf{x}_i, \theta_0) + b_0$ is obvious.

Now we present simple conditions which imply the assumptions (R4) and (R4+) of Theorems 2.2 and 2.3 for particular versions of the M-estimators (1.9) and general model (1.1).

Proposition 3.6 *If both components ψ^+ and ψ^- of the decomposition (2.1) are Lipschitz on $\mathbb{R}$ then the M-estimator $\hat{\theta}_n \sim \langle \psi; \varphi_i\rangle$ satisfies the regularity condition (R4+) in the general model (1.1).*

Proof: Under the assumptions of this proposition the function $\psi^{\pm}$ defined in (2.2) satisfies the Lipschitz condition

$$|\psi^{\pm}(y_1) - \psi^{\pm}(y_2)| \le C|y_1 - y_2|$$

for some constant C and all y_1, $y_2 \in \mathbb{R}$. Therefore the expression in the brackets of (2.15) is bounded above by $(2\tau)^2$. This means that (2.16) with $\kappa = 4C^2$, $q = 2$ and arbitrary $\tau > 0$ holds for the model (1.1). □

The next result is an alternative to Proposition 3.6. In this result we assume that ψ is absolutely continuous on $\mathbb{R}$. Similarly as in (1.6), this means that there exists a measurable and locally absolutely integrable function $\dot{\psi} : \mathbb{R} \mapsto \mathbb{R}$ satisfying the condition

$$\dot{\psi}(y) = \frac{d\psi(y)}{dy} \quad \text{a.e.} \tag{3.15}$$

with respect to the Lebesgue measure on $\mathbb{R}$. Then, similarly as in (1.7), for every $y \in \mathbb{R}$

$$\psi(y) = \psi(0) + \int_{(0,y]} \dot{\psi}(s)ds \quad \text{(cf. (1.8))}$$

and, moreover,

$$\psi^+(y) = \psi^+(0) + \int_{(0,y]} \dot{\psi}(s)\, I(\dot{\psi}(s) > 0)\, ds$$

and

$$\psi^-(y) = \psi^-(0) - \int_{(0,y]} \dot{\psi}(s)\, I(\dot{\psi}(s) < 0)\, ds$$

for the components of the decomposition (2.1). Therefore (2.2) implies that for every $y \in \mathbb{R}$

$$\psi^\pm(y) = \psi^\pm(0) + \int_{(0,y]} |\dot{\psi}(s)|\, ds \quad \text{(cf. (1.8))}. \tag{3.16}$$

Obviously, if $\dot{\psi}$ is bounded a.e. on $\mathbb{R}$ then it follows from the formulas above that ψ^+ and ψ^- are Lipschitz on $\mathbb{R}$ so that Proposition 3.6 is applicable. Therefore the next result is interesting only in situations where $\dot{\psi}$ is unbounded.

Proposition 3.7 *Let ψ be absolutely continuous on $\mathbb{R}$ with an a. e. derivative $\dot{\psi}$. The M-estimator $\hat{\theta}_n \sim \langle \psi; \varphi_i \rangle$ satisfies the regularity condition* (R4+) *in the general model* (1.1) *if one of the following conditions holds:*

(i) *$\dot{\psi}$ is square integrable on $\mathbb{R}$;*

(ii) *for $X_i = Y_i - \varphi_i(\theta_0)$ and some $\varepsilon > 0$*

$$C := \sup_{n \in \mathbb{N}} \frac{1}{n} \sum_{i=1}^{n} \mathsf{E} \sup_{|s| \le \varepsilon} (\dot{\psi}(X_i + s))^2 < \infty; \tag{3.17}$$

(iii) *$\dot{\psi} = \dot{\psi}_1 + \dot{\psi}_2$ where $\dot{\psi}_1$ satisfies (i) and $\dot{\psi}_2$ satisfies (ii).*

Proof: By (3.16) and Schwarz' inequality, for every $y \in \mathbb{R}$ and $\tau > 0$

$$\begin{aligned} [\psi^\pm(y+\tau) - \psi^\pm(y-\tau)]^2 &\le \left(\int_{(y-\tau,y+\tau]} |\dot{\psi}(s)| ds \right)^2 \\ &\le 2\tau \int_{(y-\tau,y+\tau]} (\dot{\psi}(s))^2 ds \\ &\le 2\tau \int_{\mathbb{R}} (\dot{\psi}(s))^2 ds =: A_1(\dot{\psi}). \end{aligned} \tag{3.18}$$

Therefore, if $\dot\psi$ is square integrable then (2.16) holds for $q = 1$, all $\tau > 0$ and

$$\kappa = 2\int_{\mathbb{R}} (\dot\psi(s))^2\, ds.$$

Further, by (3.18),

$$\left[\psi^{\pm}(y+\tau) - \psi^{\pm}(y-\tau)\right]^2 \leq 4\tau^2 \sup_{|s|\leq\tau} (\dot\psi(y+s))^2 =: A_2(\dot\psi).$$

Therefore, if (ii) holds then (2.16) holds for $q = 2$, $\kappa = 4C^2$ and all $0 < \tau \leq \varepsilon$. Finally, from the above inequalities we see that

$$\left[\int_{y-\tau}^{y+\tau} |\dot\psi(t)| dt\right]^2 \leq \min(A_1(\dot\psi), A_2(\dot\psi)).$$

From here and

$$\left[\int_{y-\tau}^{y+\tau} |\dot\psi_1(t) + \dot\psi_2(t)| dt\right]^2 \leq 2\left[\int_{y-\tau}^{y+\tau} |\dot\psi_1(t)| dt\right]^2 + 2\left[\int_{y-\tau}^{y+\tau} |\dot\psi_2(t)| dt\right]^2$$

we obtain the statement in (iii). □

The following proposition presents similar conditions as Proposition 3.7 for the estimator $\widehat\theta_n \sim \langle\psi, \varphi_i\rangle$ for nonexplosive $\psi^{\pm}$.

Definition 3.8 *We say that a nondecreasing function $\xi : \mathbb{R} \mapsto \mathbb{R}$ is explosive if there exists $\tau > 0$ such that*

$$\sup_{y\in\mathbb{R}} [\xi(y+\tau) - \xi(y-\tau)] = \infty.$$

Thus $\psi^{\pm}$ is nonexplosive if for every $\tau > 0$

$$C(\tau) := \sup_{y\in\mathbb{R}} \left[\psi^{\pm}(y+\tau) - \psi^{\pm}(y-\tau)\right] < \infty. \tag{3.19}$$

Clearly, $C(\tau)$ is nondecreasing in the domain $\tau > 0$ with $C(0) \geq 0$. Nonexplosive $\psi^{\pm}$ satisfies the inequalities

$$\left[\psi^{\pm}(y+\tau) - \psi^{\pm}(y-\tau)\right]^2 \leq C(\tau)\left[\psi^{\pm}(y+\tau) - \psi^{\pm}(y-\tau)\right] \tag{3.20}$$

and

$$\left[\psi^{\pm}(y+\tau) - \psi^{\pm}(y-\tau)\right]^2 \leq (C(\tau))^2. \tag{3.21}$$

Proposition 3.9 *Every $M-$estimator $\widehat\theta_n \sim \langle\psi, \varphi_i\rangle$ with nonexplosive $\psi^{\pm}$ satisfies the regularity assumption* (R4) *in the model* (1.1)*. If there exist constants $\tau_0, q > 0$ and κ such that for $X_i = Y_i - \varphi_i(\theta_0)$ and all $0 < \tau < \tau_0$*

$$\sup_{n\in\mathbb{N}} \frac{1}{n}\sum_{i=1}^{n} \mathsf{E}\left[\psi^{\pm}(X_i+\tau) - \psi^{\pm}(X_i-\tau)\right] < \kappa\tau^q \tag{3.22}$$

then it satisfies also (R4+).

Proof: The first assertion is clear from (2.15) and (3.21). The second assertion follos from (2.16), (3.20) and (3.22). □

Proposition 3.10 *Let $\psi^{\pm}$ of the estimator of Proposition 3.9 be piecewise constant with finitely many jumps of sizes $\Delta_k > 0$ at points t_k, and let for some fixed $\varepsilon > 0$ the neighborhoods $N_k(\tau) = (t_k - \tau, t_k + \tau)$, $0 < \tau < \varepsilon$, be disjoint for different k. If the distribution functions $F_i(y)$ of X_i in (3.22) have densities in the union $U(\varepsilon) = \cup_k N_k(\tau)$ and*

$$C := \sup_{y \in \mathbb{N}} \frac{1}{n} \sum_{i=1}^{n} \sup_{y \in U(\varepsilon)} f_i(y) < \infty, \tag{3.23}$$

then (3.22) holds for $\tau_0 = \varepsilon/2$, $q = 1$, and $\kappa = 4C \sum_k \Delta_k$.

Proof: If $\tau \le \varepsilon/2$ then

$$\psi^{\pm}(y+\tau) - \psi^{\pm}(y-\tau) = \begin{cases} \Delta_k & \text{if} \quad y \in N_k(\tau) \\ 0 & \text{otherwise} \end{cases}$$

and

$$\int_{N_k(\tau)} dF_i(y) \le F(t_k + 2\tau) - F(t_k - 2\tau) \le 4\tau \sup_{y \in U(\varepsilon)} f_i(y).$$

Therefore

$$\mathsf{E}[\psi^{\pm}(X_i + \tau) - \psi^{\pm}(X_i - \tau)] \le \sum_k \Delta_k \int_{N_k(\tau)} dF_i(y).$$

The desired result follows from here. □

Our last result is concerning estimators $\widehat{\theta}_n \sim \langle \psi, \varphi_i \rangle$ with nonexplosive ψ in the general regression models where $G(y|i,\theta) = F_i(y|\phi(\mathbf{x}_i,\theta))$ and $\varphi_i(\theta) = a(\phi(\mathbf{x}_i,\theta))$, see (3.1) and (3.12). We use the notation

$$\vartheta_i = \phi(\mathbf{x}_i, \theta) \qquad \text{and} \qquad a_i = a(\vartheta_i) \tag{3.24}$$

In this notation the functions $H_i(t)$ of (2.10) are given by the formula

$$H_i(t) = \int \psi(y - a_i + t) dF(y|\vartheta_i), \quad t \in \mathbb{R}, \tag{3.25}$$

in the general regression model (3.1). In the simplified notation

$$F_i(y) = F(y + a_i|\vartheta_i), \qquad F_{i,s}(y) = F(y + a_i - s|\vartheta_i) \tag{3.26}$$

it holds

$$H_i(t) = \int \psi(y) dF_i(y - t), \quad t \in \mathbb{R}, \tag{3.27}$$

so that, for $s \neq 0$,

$$\frac{1}{s}[H_i(t+s) - H_i(t)] = \int \psi(y) d\Phi_{i,s,t}(y) \tag{3.28}$$

where

$$\Phi_{i,s,t}(y) = \frac{F_{i,s}(y-t) - F_i(y-t)}{s}, \quad y \in \mathbb{R}.$$

Let us consider $\psi^{\pm} = \psi^+ + \psi^-$ and suppose that for some $\tau > 0$

$$\psi^+,\ \psi^- \in L_1(F_{i,t}) \quad \text{for all } i \in \mathbb{N} \text{ and all } |t| \leq \tau. \tag{3.29}$$

Here and in the sequel, $L_1(G)$ denotes the Banach space of functions absolutely integrable with respect to the measure defined on $\mathbb{R}$ by a nondecreasing and right continuous function $G : \mathbb{R} \mapsto \mathbb{R}$. We assume nonexplosive $\psi^{\pm}$ defined by the condition (3.19).

Proposition 3.11 *Let an M-estimator $\widehat{\theta}_n \sim \langle \psi, a(\mathbf{x}_i, \theta) \rangle$ with non-explosive $\psi^{\pm}$ be adapted to the general regression model (3.1). Further, let ψ^+, $\psi^- \in L_1(F_{i,t})$ for some $\tau > 0$ and all $|t| \leq \tau$ and $i \in \mathbb{N}$, let all distributions $F_{i,s}$, $i \in \mathbb{N}$, $s \in \mathbb{R}$, be differentiable on $\mathbb{R}$ with derivatives $f_{i,s}$, and put $f_i = f_{i,0}$.*

(I) *If*

$$\sup_{|s| \leq \tau} f_{i,s} \in L_1(\psi^{\pm}) \tag{3.30}$$

then the convolutions $H_i(t)$ are absolutely continuous on $(-\tau/2, \tau/2)$, with a. e. derivatives

$$h_i(t) = -\int f_i(y-t) d\psi(y), \quad i \in \mathbb{N}. \tag{3.31}$$

(II) *If f_i are locally Lipschitz in sense that for every $y \in \mathbb{R}$*

$$|f_i(y-t) - f_i(y)| \leq \lambda_i(y)\,|t|, \quad t \in (-\tau, \tau), \tag{3.32}$$

and both f_i and λ_i belong to $L_1(\psi^{\pm})$ then the previous condition (3.30) is satisfied. If, moreover,

$$\limsup_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \lambda_i \in L_1(\psi^{\pm}) \tag{3.33}$$

then $\widehat{\theta}_n$ satisfies the regularity condition (R3) *for $\tau_0 = \tau/2$.*

Proof: Let $|s| \leq \tau/2$, $|t| \leq \tau/2$ and i be arbitrary fixed. Then

$$\int d\Phi_{i,s,t}(y) = 0, \tag{3.34}$$

$$\Phi_{i,s,t}(\infty) = \lim_{y\to\infty} \Phi_{i,s,t}(y) = 0 \tag{3.35}$$

and, by (3.29),

$$\int |\psi(y)| d\Phi_{i,s,t}(y) < \infty. \tag{3.36}$$

Hence, by (3.27) and the Fubini theorem,

$$\begin{aligned}
\frac{1}{s}[H_i(t+s) - H_i(t)] &= \int\int I(0 < x \le y) d\psi(x) d\Phi_{i,s,t}(y) \\
&= \int\int I(0 < x \le y) d\Phi_{i,s,t}(y) d\psi(x) \\
&= \int\int I(x \le y < \infty) d\Phi_{i,s,t}(y) d\psi(x) \\
&= \int (\Phi_{i,s,t}(\infty) - \Phi_{i,s,t}(x)) d\psi(x).
\end{aligned}$$

Therefore, by (3.34)–(3.36),

$$\frac{1}{s}[H_i(t+s) - H_i(t)] = -\int \Phi_{i,t,s}(y) d\psi(y). \tag{3.37}$$

Since

$$\lim_{s\to 0} \Phi(i,t,s)(y) = f_i(y-t) \quad \text{a.e.}$$

and since (3.30) justifies interchange of the integral and $\lim_{s\to 0}$ in (3.37), assertion (I) is proved. The first part of assertion (II) follows from the inequality

$$f_i \le \sup_{|s|\le\tau} f_{i,s} \le f_i + \lambda_i \tau,$$

and the second part follows from the first part and from the fact that, under (3.31) and (3.32),

$$\frac{1}{n}\sum_{i=1}^{n} \sup_{|t|\le\tau} |h_i(t) - h_i(0)| \le \frac{\tau}{n}\sum_{i=1}^{n}\int \lambda_i(y) d\psi(y).$$

Indeed, under (3.33) the $\limsup_n$ of the right-hand side tends to zero as $\tau \downarrow 0$. □

For bounded sensitivities ψ the assumptions of Proposition 3.9 simplify in sense that (3.29) is automatically satisfied.

For the standard nonlinear regression model with an absolutely continuous error distribution F and the same $b(F)$ as in (3.13), the condition (3.29) simplifies into

$$\psi^+, \psi^- \in L_1(F(y - b(F))) \quad \text{and} \quad \lim_{a\to\infty} \psi^{\pm}(a) \sup_{|y|\ge a} f(y) = 0, \tag{3.38}$$

where f is the derivative of F. Further, (3.30) takes on the form

$$\sup_{|s|\le\tau} f(y - b(F) - s) \in L_1(\psi^{\pm}), \tag{3.39}$$

the Lipschitz condition (3.32) is in this case

$$|f(y-b(F)-t)-f(y-b(F))| \leq \lambda(y)\,|t|, \tag{3.40}$$

and the remaining conditions of assertion (II) reduce to $f,\,\lambda \in L_1(\psi^{\pm})$.

Applicability of the results of this section is illustrated in the next sections.

4 $L_{1+\alpha}$–estimators

Let us start with two examples.

Example 4.1 Perhaps the best known of all M-estimators is the L_2-*estimator*

$$\hat{\theta}_n \sim \langle \psi(y)=y;\, \varphi_i \rangle. \tag{4.1}$$

Here $\rho(y) = y^2/2$ and the decomposition (2.1) and formula (2.2) are trivial in the sense that $\psi^- \equiv 0$ and $\psi^+(y) = \psi^{\pm}(y) = \psi(y) = y$. Since $\rho(y) = y^2/2$, it follows from the definition of $\hat{\theta}_n$ that, in any model (1.1), $\hat{\theta}_n$ minimizes the L_2-distance between observations $\mathbf{Y}_n = (Y_1,\ldots,Y_n)'$ and locators $\boldsymbol{\varphi}_n(\theta) = (\varphi_1(\theta),\ldots,\varphi_n(\theta))'$,

$$\hat{\theta}_n = \arg\min_{\Theta} \|\mathbf{Y}_n - \boldsymbol{\varphi}_n(\theta)\|_2, \tag{4.2}$$

where $\|\cdot\|_2$ denotes the L_2-norm. The rule (2.3) for adaptation of locators reduces to a *mean value rule* $\varphi_i(\theta_0) = \mathsf{E}Y_i$, i. e. the formula (3.10) for locators takes on the form

$$\varphi_i(\theta) = \int y\,dG(y|i,\theta), \quad \theta \in \Theta, \tag{4.3}$$

where $G(y|i,\theta)$ are the distributions of model (1.1). Similarly, the particular adaptation rules (3.12) and (3.13) reduce to

$$\varphi_i(\theta) = a(\phi(\mathbf{x}_i,\theta)) \quad \text{for} \quad a(\vartheta) = \int y\,dF(y|\vartheta)$$

and

$$\varphi_i(\theta) = \phi(\mathbf{x}_i,\theta) + \int y\,dF(y),$$

respectively.

In the regression models with additive errors, (4.2) represents a *least squared error* criterion. Due to the simplicity of both, the criterion function $\|\mathbf{Y}_n - \boldsymbol{\varphi}_n(\theta)\|_2$ and the universal adaptation rule (4.3), the L_2-estimators play a fundamental role in the statistical practice as well as in the theory. The linearity of $\psi(y)$, placing these estimators into the center of interest of the linear statistics, makes the asymptotic theory of these estimators relatively easy. This theory has been developed into considerable details, see e. g. Rao [25].

Example 4.2 Another well known M-estimator is the L_1-*estimator*

$$\hat{\theta}_n \sim \langle \psi(y) = 1 - 2I(y < 0);\, \varphi_i \rangle. \tag{4.4}$$

Here, as before, $I(\cdot)$ denotes the indicator of events. The ψ-function of the L_1-estimator is an example where the decompositions (2.1), (2.2) are trivial in the sense that $\psi^- \equiv 0$ and $\psi(y) = \psi^+(y) = \psi^{\pm}(y)$ has a jump of size 2 at $y = 0$. Since $\rho(y) =| y|$, it follows from the definition 3.1 that, in any model (1.1), $\hat{\theta}_n$ minimizes the L_1-distance between observations $\mathbf{Y}_n = (Y_1, \ldots, Y_n)$ and locators $\boldsymbol{\varphi}_n(\theta) = (\varphi_1(\theta), \ldots, \varphi_n(\theta))$,

$$\hat{\theta}_n = \arg\min_{\Theta} \|\mathbf{Y}_n - \boldsymbol{\varphi}_n(\theta)\|_1, \tag{4.5}$$

where $\|\cdot\|_1$ denotes the L_1-norm. The general rule (2.3) for adaptation of locators reduces to the *median rule*, $\varphi_i(\theta_0) = \mathrm{med} Y_i$, i. e. (3.10) takes on the form

$$\varphi_i(\theta) = \mathrm{med} G(y|i, \theta), \quad \theta \in \Theta, \tag{4.6}$$

where

$$\mathrm{med} G(y|i, \theta) = \inf\{y \in \mathbb{R} : G(y|i, \theta) \geq 1/2\}$$

denotes the median of $G(y|i, \theta)$. Similarly, the special adaptation rules (3.12) and (3.13) reduce to

$$\varphi_i(\theta) = a(\phi(\mathbf{x}_i, \theta)) \quad \text{for} \quad a(\vartheta) = \mathrm{med} F(y|\vartheta)$$

and

$$\varphi_i(\theta) = \phi(\mathbf{x}_i, \theta) + \mathrm{med} F(y).$$

In the regression models with additive errors, (4.2) represents a *least absolute error* criterion. Due to the relative simplicity of both, the criterion function $\|\mathbf{Y}_n - \boldsymbol{\varphi}_n(\theta)\|_1$ and the universal adaptation rule (4.6), the L_1-estimators play an important role in the statistical practice as well as in the theory (see e. g. Serfling [27], Dodge [4], Farenbrother [6], Ronchetti [26], Pollard [22], Knight [15] and references therein).

The L_1-or L_2-estimators $\hat{\theta}_n$ can be embedded into various families of estimators $\hat{\theta}_n^{(\alpha)}$ with a parameter $\alpha \in \mathbb{R}$ controlling finite-sample-size properties, such as rejection regions and variances-covariances of deviations $\hat{\theta}_n^{(\alpha)} - \theta_0$, or asymptotic properties like influence curves and relative efficiencies.

In this section we study the family of *quantile* $L_{1+\alpha}$-*estimators*

$$\hat{\theta}_n^{(\alpha)} \sim \langle \psi(y) = 1 + \alpha - 2I(y < 0);\, \varphi_i \rangle, \quad -1 < \alpha < 1, \tag{4.7}$$

where $\hat{\theta}_n^{(0)}$ is the L_1-estimator of Example 4.2. The ψ-functions of (4.7) differ from the ψ-function of (4.4) by a constant shift α : if $\alpha > 0$ then the sensitivity is suppressed in the domain $y < 0$ and enhanced in the domain $y \geq 0$, while for $\alpha < 0$ the opposite is true. Note that for the extremal $\alpha = 1$ or $\alpha = -1$ we obtain in (4.7) sensitivities concentrated only on $y \geq 0$ or $y < 0$, respectively. The corresponding quantile L_2- and L_0-estimators are legitimate particular cases of the M-estimators studied in this paper (one of them is studied at the end of this section). Notice that the quantile L_2-estimator differs from the usual L_2-estimator of Example 4.1.

Since $\rho(y) = y\,\psi(y) = \rho_\alpha(y)$ where

$$\rho_\alpha(y) = (1+\alpha)\,y\,I(y>0) - (1-\alpha)\,yI(y<0),$$

the definition of M-estimator implies that

$$\hat{\theta}_n^{(\alpha)} = \arg\min_{\Theta} \left((1+\alpha)\,\|\mathbf{Y}_n - \boldsymbol{\varphi}_n(\theta)\|_1^+ + (1-\alpha)\,\|\mathbf{Y}_n - \boldsymbol{\varphi}_n(\theta)\|_1^-\right), \tag{4.8}$$

where

$$\|\mathbf{Y}_n - \boldsymbol{\varphi}_n(\theta)\|_1^{+(-)} = \sum_{i=1}^n |Y_i - \varphi_i(\theta)|^{+(-)}$$

and

$$|y|^+ = |y|\,I(y>0) \quad \text{and} \quad |y|^- = |y|\,I(-y>0).$$

Thus we see that the criterion (4.8) differs from (4.5) in that the criterion function takes the values $|Y_i - \varphi_i(\theta)|$ with different weights $1+\alpha$ or $1-\alpha$, depending on whether $Y_i - \varphi_i(\theta)$ is positive or negative. Since the above defined $\rho_\alpha(y)$ is twice larger than $\rho_\beta(y)$ of (1.12) for $\beta = (1+\alpha)/2$, the quantile L_1-estimators $\widehat{\theta}_n^{(\alpha)} \sim \langle \rho_\alpha; \varphi_i \rangle$ coincide with the estimators $\widehat{\theta}_n^{(\beta)} \sim \langle \rho_\beta; \varphi_i \rangle$ where ρ_β is given by (1.13) for $\beta = (1+\alpha)/2$. If these estimators are applied in the regression models then they are called *regression quantiles*.

For the ψ-function defined in (4.7), and for arbitrary $\varphi \in \mathbb{R}$ and arbitrary distribution function $G(y)$,

$$\int \psi(y-\varphi)dG(y) = 1 + \alpha - 2G(\varphi).$$

Consequently, the general rule (3.10) for adaptation of locators reduces into the $(1+\alpha)/2$-*quantile rule*

$$\varphi_i(\theta) = G^{-1}\left((1+\alpha)/2\,|\,i,\theta\right), \quad \theta \in \Theta, \tag{4.9}$$

where

$$G^{-1}(\beta) = \inf\left\{y \in \mathbb{R} : G(y) \geq \beta\right\}, \quad 0 < \beta < 1, \tag{4.10}$$

is the quantile function of $G(y)$. From (3.12) or (3.13) we obtain the special adaptation rules

$$\varphi_i(\theta) = F^{-1}\left((1+\alpha)/2\,|\,\phi(\mathbf{x}_i,\theta)\right) \quad \text{or} \quad \varphi_i(\theta) = \phi(\mathbf{x}_i,\theta) + F^{-1}((1+\alpha)/2).$$

The second of these rules is used in the standard nonlinear regression model (3.5). It cannot be used if the error distribution $F(y)$ is unknown. By (3.14), in this case one can consider an extended $L_{1+\alpha}$-estimator $(\hat{\theta}_n^{(\alpha)}, \hat{b}_n^{(\alpha)})$ of the extended true parameter $(\theta_0,\, b_0 = F^{-1}((1+\alpha)/2))$ adapted by the rule

$$\varphi_i(\theta, b) = \phi(\mathbf{x}_i, \theta) + b, \quad (\theta, b) \in \widetilde{\Theta}, \quad \widetilde{\Theta} = \Theta \times \mathbb{R}. \tag{4.11}$$

But

$$Y_i = \phi(\mathbf{x}_i, \theta_0) + b_0 + \tilde{\mathcal{E}}_i, \quad \tilde{\mathcal{E}}_i \sim \tilde{F}(y) = F(y + b_0),$$

where $\tilde{F}^{-1}((1+\alpha)/2) = 0$, and $\phi(\mathbf{x}_i, \theta) + b$ is a special case of a general function $\tilde{\phi}(\mathbf{x}_i, \tilde{\theta})$ of $(m+1)$-dimensional parameter $\tilde{\theta} \in \tilde{\Theta}$, $\tilde{\Theta} \subset \mathbb{R}^{m+1}$ open. Therefore $(\hat{\theta}_n^{(\alpha)}, \hat{b}_n^{(\alpha)})$ is a special case of a general $L_{1+\alpha}$-estimator $\widehat{\tilde{\theta}}_n$ of true $\tilde{\theta}_0 \in \tilde{\Theta}$ in the model

$$Y_i = \tilde{\phi}(\mathbf{x}_i, \widetilde{\theta}_0) + \tilde{\mathcal{E}}_i, \quad \tilde{\mathcal{E}}_i \sim \tilde{F}(y), \quad \tilde{F}^{-1}((1+\alpha)/2) = 0. \tag{4.12}$$

All conditions imposed in this model on $\tilde{F}(y)$ and $\tilde{\phi}(\mathbf{x}_i, \widetilde{\theta})$ easily transform into conditions on distribution $F(y) = \tilde{F}(y - F^{-1}((1-\alpha)/2))$ of the errors $\mathcal{E}_i$ in the model (3.5) and on $\phi(\mathbf{x}_i, \theta) + b$. Similarly, all properties of the estimator $\hat{\tilde{\theta}}_n^{(\alpha)}$ straightforward transform into properties of the particular version $(\hat{\theta}_n^{(\alpha)}, \hat{b}_n^{(\alpha)})$. Hence, in the standard nonlinear (and linear) regression with an unknown error distribution, it suffices to investigate the estimators (4.7) under the assumption

$$F^{-1}((1+\alpha)/2) = 0, \tag{4.13}$$

using the adaptation rule

$$\varphi_i(\theta) = \phi(\mathbf{x}_i, \theta), \quad \theta \in \Theta, \tag{4.14}$$

for $\Theta \subset \mathbb{R}^m$ open and $m \geq 2$.

The estimators $(\hat{\theta}_n^{(\alpha)}, \hat{b}_n^{(\alpha)})$, $\alpha \in (-1, 1)$, with the adaptation rule (4.11), have been introduced into the literature by Koenker and Basset [16]. As said above, these estimators, called regression quantiles, coincide with $(\theta_n^{(\beta)}, \hat{b}_n^{(\beta)})$ defined by the criterion functions (1.13) for $\beta = (1+\alpha)/2 \in (0, 1)$. Koenker and Basset established the asymptotic normality of these estimators in the standard linear regression (3.6) with an unknown distribution $F(y)$. Jurečková and Procházka [13] extended their result to the standard nonlinear regression (3.5) with an unknown $F(y)$. In this section we study the estimators (4.7) under the restrictions (4.13), (4.14). As argued above, our study covers as particular cases the estimators $(\hat{\theta}_n^{(\alpha)}, \hat{b}_n^{(\alpha)})$ of $(m-1)$-dimensional parameter θ_0 and $b_0 = F^{-1}((1-\alpha)/2)$ in the model (3.5) free of the restriction (4.13).

We shall obtain asymptotic normality of the estimators $\hat{\theta}_n^{(\alpha)}$, $\alpha \in (-1, 1)$, defined by (4.7) and (4.13), from Theorem 2.1 under the assumption (4.13). To this end we assume the following.

(a) $\hat{\theta}_n^{(\alpha)}$ is consistent in the sense of (1.10).

(b) One can find a closed ball $B \subset \Theta$ of a radius $\delta > 0$ centered at θ_0 on which there exist the gradients

$$\dot{\phi}(\mathbf{x}_i, \theta) = \left(\frac{\partial}{\partial \theta_1}, \dots \frac{\partial}{\partial \theta_m} \right)' \phi(\mathbf{x}_i, \theta), \qquad i \in \mathbb{N},$$

and a constant λ possibly depending on B, such that

$$\|\dot{\phi}(\mathbf{x}_i, \theta)\| \le \lambda \quad \text{and} \quad \|\dot{\phi}(\mathbf{x}_i, \theta) - \dot{\phi}(\mathbf{x}_i, \tilde{\theta})\| \le \lambda \|\theta - \tilde{\theta}\|$$

for all $\theta, \tilde{\theta} \in B,\ i \in \mathbb{N}$, i.e. the regularity condition (R2) holds.

(c) It holds

$$\Psi_n = \frac{1}{n} \sum_{i=1}^{n} \dot{\phi}(\boldsymbol{x}_i, \theta)\, \dot{\phi}(\boldsymbol{x}_i, \theta)' \to \Psi,$$

where the $m \times m$ matrix Ψ is positive definite.

(d) The error distribution function $F(y)$ is differentiable on an interval $(-\tau, \tau)$ and the derivative $f(y)$ of $F(y)$ is continuous at $y = 0$ with $f(0) > 0$.

Theorem 4.3 *If the conditions (a) – (d) hold and the error distribution satisfies* (4.13) *then the estimators* $\hat{\theta}_n^{(\alpha)}$ *defined by* (4.7) *and* (4.14) *are asymptotically normal in the sense*

$$\sqrt{n} \left(\hat{\theta}_n^{(\alpha)} - \theta_0 \right) \xrightarrow{\mathcal{L}} N \left(0, \frac{1 - \alpha^2}{4 f^2(0)} \Psi^{-1} \right), \tag{4.15}$$

where $f(0) > 0$ *is defined by (d) and the positive definite matrix* Ψ *is defined by (c).*

Proof: Let $\alpha \in (-1, 1)$ and $F(y)$ satisfying (4.13) fulfil assumptions (a) – (d). We shall verify that $\hat{\theta}_n^{(\alpha)}$ satisfies all assumptions of Theorem 2.3. By Propositions 3.9, 3.10, and (d), $\hat{\theta}_n^{(\alpha)}$ satisfies the regularity condition (R4+). By (2.10), (4.13) and (4.14), if $t \in \mathbb{R}$ then

$$H_i(t) = \int \psi(y + t) dF(y) = 1 + \alpha - 2F(-t), \quad i \in \mathbb{N}.$$

Consequently, by (d), the estimators $\hat{\theta}_n^{(\alpha)}$ satisfy the regularity condition (R3) of Theorem 2.3 for $h_i(t) = 2f(-t)$ and $\tau_0 = \tau$. As to the remaining conditions, (2.3) was clarified above, the consistency was assumed in (a), (R2) was assumed in (b) and (R1) holds because

$$\begin{aligned} \sigma_i^2 &= \int \psi^2(y) dF(y) = (1 + \alpha)^2 \int_0^\infty dF(y) + (1 - \alpha)^2 \int_{-\infty}^0 dF(y) \\ &= (1 + \alpha)^2 (1 - (1 + \alpha)/2) + (1 - \alpha)^2 (1 + \alpha)/2 \\ &= 1 - \alpha^2 \quad \text{for all } i \in \mathbb{N}. \end{aligned}$$

Conditions (2.13), (2.14) of condition (R3) follow from (c) for

$$\Sigma = (1-\alpha^2)\,\Psi \quad \text{and} \quad \Phi = 2f(0)\,\Psi.$$

Since the functions $\psi(t)$ as well as the gradients $\dot{\varphi}_i$ are bounded (see (b)), the remaining condition (2.17) of Theorem 2.3 holds by Proposition 2.4. The desired relation (4.15) thus follows from Theorem 2.3. □

By what has been said above, the following assertion about an arbitrary error distribution $F(y)$ follows from Theorem 4.3. In this assertion, and in the rest of section, we put

$$\beta = \frac{1+\alpha}{2}, \quad \beta \in (0,1). \tag{4.16}$$

Corollary 4.4 *Let $\alpha \in (-1,1)$ be arbitrary, and let β be given by (4.16). If conditions (a) – (d) hold with $F(y)$ replaced by $\tilde{F}(y) = F(y - F^{-1}(\beta))$ then the above specified $L_{1+\alpha}$-estimator $(\hat{\theta}_n^{(\alpha)}, \hat{b}_n^{(\alpha)})$ is asymptotically normal in the sense*

$$\sqrt{n}\left[(\hat{\theta}_n^{(\alpha)}, \hat{b}_n^{(\alpha)}) - (\theta_0, F^{-1}(\beta))\right] \xrightarrow{\mathcal{L}} N\left(0, \frac{\beta(1-\beta)}{f^2(F^{-1}(\beta))}\tilde{\Psi}^{-1}\right) \quad \text{as } n \to \infty \tag{4.17}$$

for $f(y) = dF(y)/dy$ and the matrix

$$\tilde{\Psi} = \begin{pmatrix} \Psi & , & 0 \\ 0 & , & 1 \end{pmatrix},$$

where Ψ is given by (c).

The asymptotic laws (4.15), (4.17) have been established for the L_1-estimator, where $\beta = 1/2$, as well as for the general $L_{1+\alpha}$-estimator under various conditions, see e. g. Pollard [22], Jurečková and Procházka [13] and other cited there. Let us compare the present conditions for these laws with the conditions assumed in the two cited papers.

Pollard [22] assumed (6.7) so that his conditions can be compared with those of Theorem 4.3. He studied the L_1-estimator $\hat{\theta}_n^{(0)}$ in the standard linear regression, where (b) is automatically fulfilled and the matrices considered in (c) are

$$\Psi_n = \frac{1}{n}\sum_{i=1}^{n} \mathbf{x}_i\,\mathbf{x}_i'.$$

For these matrices, (c) is a classical condition of regression analysis. As shown on p. 189 of Pollard [22], this condition is somewhat stronger than what is assumed in his Theorem 1. On the other hand, our condition (d) is slightly weaker than the assumption that $F(y)$ is continuously differentiable in an interval $(-\tau, \tau)$ with the derivative $f(y)$ positive on $(-\tau, \tau)$, which appears in the mentioned Theorem 1. The consistency of $\hat{\theta}_n^{(0)}$ assumed in (a) takes place under (c) and (d). This can be proved by applying the Convexity Lemma on p. 187 of Pollard [22].

Thus, as to the L_1-estimators in linear models, the conditions obtained from Theorem 2.3 are comparable with previously published ones, obtained by methods tailor-designed for these estimators and models. In this sense the comparison with [22] demonstrates that Theorem 2.3 is not trivial.

Jurečková and Procházka [13] studied the same estimator and model as Corollary 4.4. The conditions (b), (c) of this corollary are the same as (b), (c) in Theorem 4.3. The condition (d) is changed in the sense that $F(y)$ is differentiable in an interval $(F^{-1}(\beta) - \tau, F^{-1}(\beta) + \tau)$ with the derivative $f(y)$ continuous at $y = F^{-1}(\beta)$ and $f(F^{-1}(\beta)) > 0$. The consistency of $(\hat{\theta}_n^{(\alpha)}, \hat{b}_n^{(\alpha)})$ required in (a) follows under (b), (c), (d) by the same method as used above for the consistency of $\hat{\theta}_n^{(0)}$. Jurečková and Procházka assumed, in addition to (b), (c), (d), that $\phi(\mathbf{x}, \theta)$ is strictly monotone in each component of θ, twice differentiable in each of these components, with the first and second derivatives uniformly bounded on $\mathcal{X} \times \Theta$, and that the above mentioned $f(y)$ is symmetric about $y = 0$, bounded on $\mathbb{R}$ and differentiable on $(F^{-1}(\beta) - \tau, F^{-1}(\beta) + \tau)$. Moreover, they assumed that $\mathcal{X} \subset \mathbb{R}^k$ and $\Theta \subset \mathbb{R}^m$ are compact, and that the regression functions $\phi(\mathbf{x}_i, \theta)$ and gradients $\dot{\phi}(\mathbf{x}_i, \theta)$ satisfy some additional conditions.

Obviously, here one can deduce a stronger conclusion in favour of Theorem 2.3 than formulated in the context of the simpler L_1-estimator above. On the other hand, it is clear that the results obtained from Theorem 2.3 cannot always be as strong as the results achievable for special M-estimators and models. This can be illustrated by a reference to [15], where the L_1-estimator is studied in a standard linear regression with error distribution $F(y)$. The author proved an asymptotic law similar to (4.15) even in situations where the derivative $f(y)$ of $F(y)$ is discontinuous at the median of $F(y)$. To this end, by exploiting special features of the ψ-function defined in (4.4), and special properties of linear models, he formulated asymptotic normality conditions different from (c), (d) in Proposition 4.1, and also from the conditions considered in the previous literature. Example 4.6 below illustrates that a similar non-applicability of our theory may take place also for other M-estimators.

Remark 4.5 By (4.17), the asymptotic relative efficiency in the class of quantile $L_{1+\alpha}$-estimators depends on the function $\Gamma(\beta) = \beta(1-\beta)/f^2(F^{-1}(\beta))$; if $\beta_0 = \arg\min_{\beta \in (0,1)} \Gamma(\beta)$ then the estimator with $\alpha = 2\beta_0 - 1$ is relatively most efficient (cf. (4.16)). By the l'Hospital rule, if f has differentiable tails with a derivative $\dot{f}$ then, for $\beta \to 0$ and $\beta \to 1$,

$$\lim \Gamma(\beta) = \lim \frac{1 - 2\beta}{2\dot{f}(F^{-1}(\beta))} = \infty$$

provided $\dot{f}(y) \uparrow 0$ for $y \to \infty$ and $\dot{f}(y) \downarrow 0$ for $y \to -\infty$. In this typical case the indices α of all relatively most efficient estimators are bounded away from -1 and 1. If $\dot{f}$ is continuous on $\mathbb{R}$ then at least one such relatively most efficient $L_{1+\alpha}$-estimator exists.

Example 4.6 Let the error distribution be exponential, $F(y) = (1 - e^{-y})\, I(y > 0)$. Then $f(F^{-1}(\beta)) = 1 - \beta$. In this case $\Gamma(\beta) = \beta/(1-\beta)$ is increasing on $(0, 1)$, so that one can expect that the extremal quantile L_0-estimator $\hat{\theta}_n^{(-1)}$ maximizes the asymptotic relative efficiency in the class of estimators $\hat{\theta}_n^{(\alpha)}$, $\alpha \in [-1, 1]$. According to (4.7), the adapted version of this estimator is defined by

$$\hat{\theta}_n^{(-1)} = \underset{\theta \in \Theta}{\arg\min} \sum_{i=1}^{n} |Y_i - \phi(\mathbf{x}_i, \theta)|\, I(Y_i < \phi(\mathbf{x}_i, \theta)).$$

Here

$$H(t) = 2(e^t - 1)\, I(t < 0),$$

and

$$h(t) = 2e^t\, I(t < 0).$$

We see that the regularity condition (R4) does not hold. Consequently, Theorem 2.3 is not applicable to $\hat{\theta}_n^{(-1)}$, i.e. (4.17) is not guaranteed for $\alpha = -1$ ($\beta = 0$). In fact, since $\Gamma(0) = 0$, one can expect in this case a higher rate of consistency than $\sqrt{n}$ obtained in (4.17). The higher rate of consistency can be easily verified if $\Theta = \mathbb{R}$ and $\phi(\mathbf{x}, \theta) = \theta$, i.e. if $Y_i = \theta_0 + \mathcal{E}_i$ where $\mathcal{E}_i$ are exponentially distributed errors. Then

$$\hat{\theta}_n^{(-1)} = \min\{Y_1, \ldots, Y_n\}$$

so that

$$\mathsf{P}\left(n(\hat{\theta}_n^{(-1)} - \theta_0) > t\right) = e^{-t}, \quad t \in \mathbb{R},$$

i.e. $\hat{\theta}_n^{(-1)}$ is consistent of the order n.

5 $L_{2+\alpha}$–estimators

In the statistical literature, the classical L_2-estimator (4.1) has been embedded to many families of M-estimators. These can usually be interpreted as families of $L_{2+\alpha}$-estimators

$$\langle \psi_\alpha; \varphi_i \rangle, \quad \alpha \in \mathbb{R}, \tag{5.1}$$

with $\psi_\alpha(y)$ continuous at $\alpha = 0$ and ψ_0 coinciding with $\psi(y)$ of (4.1), i.e. satisfying for all $y \in \mathbb{R}$ the relations

$$\lim_{\alpha \to 0} \psi_\alpha(y) = \psi_0(y) \quad \text{and} \quad \psi_0(y) = y. \tag{5.2}$$

In other words, the family of estimators can be rearranged so that $\alpha = 0$ leads to the L_2-estimator.

Example 5.1 The *Huber estimators* (see e. g. [12]) form a family of the type (5.1) with

$$\psi_\alpha(y) = \int_0 y I(-|\alpha|^{-1} < s < |\alpha|^{-1}) ds \quad \text{for } \alpha \neq 0, \tag{5.3}$$

extended to $\alpha = 0$ in accordance with (5.2). Here the decomposition (2.1) and formula (2.2) are trivial in the sense that $\psi_\alpha^- \equiv 0$ and $\psi_\alpha^+ = \psi_\alpha^\pm = \psi_\alpha$. The *skipped mean* is defined by

$$\psi_\alpha(y) = y\, I(-|\alpha|^{-1} < y < |\alpha|^{-1}) \quad \text{for} \quad \alpha \neq 0$$

and extended by (5.2). If $\alpha \neq 0$ then $\psi_\alpha^+(y)$ coincide with Huber's (5.3), $\psi_\alpha^-(y) = I(y \geq |\alpha|^{-1}) - I(y < -|\alpha|^{-1})$ and

$$\rho_\alpha(y) = \alpha^{-2} - (\alpha^{-2} - y^2)\, I(-|\alpha|^{-1} < y < \alpha).$$

For more details about this and the next example we refer to [8]. The *Tukey biweight* is defined by

$$\psi_\alpha(y) = y(\alpha^{-2} - y^2)^2\, I(-|\alpha|^{-1} < y < |\alpha|^{-1}) \quad \text{for } \alpha \neq 0,$$

where, for $\alpha \neq 0$,

$$\begin{aligned} \psi_\alpha^+ &= \int_0^y I\left(-\left(\sqrt{3}|\alpha|\right)^{-1} < s < \left(\sqrt{3}|\alpha|\right)^{-1}\right) ds, \\ \psi_\alpha^-(y) &= \int_0^y I\left(s > \left(\sqrt{3}|\alpha|\right)^{-1}\right) d\psi(s) - \int_0^y I\left(s < -\left(\sqrt{3}|\alpha|\right)^{-1}\right) d\psi(s), \end{aligned}$$

and

$$\rho_\alpha(y) = \frac{1}{6|\alpha|^6} - \frac{(\alpha^{-2} - y^2)}{6}\, I\left(-|\alpha|^{-1} < y < |\alpha|^{-1}\right).$$

Portnoy [23] and independently Vajda [29] studied the family of $L_{2+\alpha}$ estimators defined by

$$\psi_\alpha(y) = y\, e^{-(\alpha y)^2} \quad \text{for} \quad \alpha \neq 0 \tag{5.4}$$

with

$$\rho_\alpha(y) = \frac{1}{2\alpha^2}\left(1 - e^{-(\alpha y)^2}\right) \quad \text{for} \quad \alpha \neq 0.$$

As is shown in the second reference, the estimators defined by (5.4) can be obtained from a minimum distance rule applied to α^2-divergences of theoretical and empirical distributions.

For the $L_{2+\alpha}$-estimators with $\alpha \neq 0$ studied in this example, there is no universal adaptation rule similar to the $(1+\alpha)/2$-quantile rule (4.9) of previous section, or to the mean value rule (4.3) applicable when $\alpha = 0$. One general adaptation rule applicable to these estimators is given in the next proposition.

Proposition 5.2 *Consider an M-estimator $\hat{\theta}_n \sim \langle\psi;\varphi_i\rangle$ with a monotone $\psi(y)$, skew-symmetric about $y = 0$, in the standard nonlinear regression model (3.5) with an error distribution $F(y)$ satisfying the condition $\psi \in L_1(F)$. If $F(y) - F(0)$ is skew-symmetric about $y = 0$ (i. e. if the errors are symmetrically distributed about zero) then the locators are adapted by the rule*

$$\varphi_i(\theta) = \phi(\mathbf{x}_i, \theta), \quad i \in \mathbb{R}. \tag{5.5}$$

This adaptation is unique unless there exists a constant $b \in \mathbb{R}$ such that

$$\psi(y - b) = \psi(y) \quad F - a.s. \tag{5.6}$$

Proof: The skew-symmetries of ψ and F imply that

$$\int \psi(y)dF(y) = 0.$$

By (3.13), this means that (5.5) is an adaptation rule. If $b \neq 0$ then the monotonicity of ψ implies that $\psi(y - b) - \psi(y)$ does not change sign on $\mathbb{R}$. Therefore

$$\int \psi(y - b)dF(y) \neq \int \psi(y)dF(y) = 0$$

unless (5.6) holds. By (3.13), this implies the uniqueness of the rule (5.5). □

The skew-symmetry of the above considered sensitivity functions ψ about 0 means that the sensitivity of the corresponding estimators to errors in data is symmetrically distributed about 0. In the rest of this section we study one class of $L_{2+\alpha}$-estimators with sensitivity functions ψ_α skew-asymmetric about 0. Such estimators are convenient when errors in data are asymmetrically distributed. As an example we may consider the situation when nonnegative data X_i are transformed into $Y_i = \ln X_i$ for fitting a symmetric location model on $\mathbb{R}$. Then an error ε in data X_i leads to an error $\varepsilon\, e^{-Y_i}$ in data Y_i, which is exponentially decreasing with increasing values of Y_i. This partially motivates the following steps.

Let us study the family of *exponential $L_{2+\alpha}$-estimators*

$$\hat{\theta}_n^{(\alpha)} \sim \langle \psi(y) = y\, e^{\alpha y};\, \varphi_i \rangle\,, \quad \alpha \in \mathbb{R}, \tag{5.7}$$

where $\hat{\theta}_n^{(0)}$ is the L_2-estimator of Example 4.1. Here

$$\rho(y) = \begin{cases} \dfrac{e^{\alpha y}(\alpha y - 1) + 1}{\alpha^2} & \text{if} \quad \alpha \neq 0 \\[2ex] y^2 & \text{if} \quad \alpha = 0. \end{cases} \tag{5.8}$$

A strong additional motivation for the estimators (5.7) is a relatively simple adaptation, in the sense of (3.13), to the generalized regression models (3.5) with exponential parent families $\mathcal{F}$, in particular to the generalized linear models mentioned in Example 3.5. The only estimator with this property studied so far in the literature seems to be the classical MLE. Thus the class (5.7) deserves to be investigated in detail.

By (3.13), the adaptation of $\widehat{\theta}_n^{(\alpha)}$ to the regression model (3.1) with a parent family $\mathcal{F} = \{F(y|\vartheta) : \vartheta \in y\}$ reduces to solution of equations (3.12), which are now of the form

$$\int (y-a)\, e^{\alpha(y-a)} dF(y|\vartheta) = 0, \quad \vartheta \in T. \tag{5.9}$$

We restrict ourselves to the homogeneous regression model (3.2) with exponential families $\mathcal{F}$ in the natural form (cf. Brown [3]), i.e. with densities

$$f(y|\vartheta) = e^{\vartheta y - c(\vartheta)} \sim F(y|\vartheta), \quad \vartheta \in T, \tag{5.10}$$

with respect to a σ-finite measure ν on $\mathbb{R}$, where

$$T = \left\{\vartheta \in \mathbb{R} : 0 < \int e^{\vartheta y} d\nu(y) < \infty \right\} \quad \text{and} \quad c(\vartheta) = \ln \int e^{\vartheta y} d\nu(y). \tag{5.11}$$

Here T is convex, and $c(\vartheta)$ is a *cumulant generating function* convex on T.

For families $\mathcal{F}$ in a natural form, the distributions figuring in (1.1) are given by

$$G(y|i,\theta) \sim g(y|i,\theta) = f(y)|\phi(\mathbf{x}_i,\theta) = e^{\phi(\mathbf{x}_i,\theta)\, y - c(\phi(\mathbf{x}_i,\theta))} \tag{5.12}$$

for all y from the support of ν, and all $i \in \mathbb{N}$ and $\theta \in \Theta$. If $\phi(\mathbf{x},\theta) = \mathbf{x}'\theta$ then we obtain generalized linear models with *natural link functions* (see e. g. Fahrmeir and Kaufmann [5]) where

$$G(y|i,\theta) \sim g(y|i,\theta) = f(y|\mathbf{x}_i'\theta) = e^{\mathbf{x}_i'\theta y - c(\mathbf{x}_i'\theta)} \tag{5.13}$$

for all y from the support of ν and all $i \in \mathbb{N}$ and $\theta \in \Theta$. The exponential families are assumed to be nontrivial in the sense that ν is not concentrated in one point, that T has a nonempty interior, and that all values $\mathbf{x}_i'\theta$ or $\phi(\mathbf{x}_i',\theta)$ are in this interior.

In a nontrivial exponential family $\mathcal{F}$, the cumulant generating function $c(\vartheta)$ is strictly convex and infinitely differentiable on the interior T^0 of T, with derivatives

$$\dot{c}(\vartheta) = \frac{dc(\vartheta)}{d\vartheta} \quad \text{and} \quad \ddot{c}(\vartheta) = \frac{d^2 c(\vartheta)}{d\vartheta^2} \tag{5.14}$$

satisfying for all $\vartheta \in T^0$ the equalities

$$\int (y - \dot{c}(\vartheta))\, e^{\vartheta y - c(\vartheta)} d\nu(y) = 0 \tag{5.15}$$

and

$$\int (y - \dot{c}(\vartheta))^2\, e^{\vartheta y - c(\vartheta)} d\nu(y) = \ddot{c}(\vartheta). \tag{5.16}$$

The derivative $\dot{c}(\vartheta)$ of the strictly convex function $c(\vartheta)$ is increasing on the interior T^0, and the second derivative $\ddot{c}(\vartheta)$ is positive on T^0. By (5.15) and (5.16), $\dot{c}(\vartheta)$ is the mean in $\mathcal{F}$,

$$\mu(\vartheta) = \int y f(y|\vartheta) d\nu(y), \tag{5.17}$$

and $\ddot{c}(\vartheta)$ is the variance or, equivalently, the *Fisher information* of $\mathcal{F}$, i. e.

$$\mu(\vartheta) = \dot{c}(\vartheta) \quad \text{and} \quad \mathcal{I}(\vartheta) = \ddot{c}(\vartheta), \quad \vartheta \in T^0. \tag{5.18}$$

Moreover, for each $\vartheta \in T^0$, $a = \dot{c}(\vartheta)$ is the unique solution of the equations

$$\int (y-a)\, e^{\vartheta y - c(\vartheta)} d\nu(y) = 0 \tag{5.19}$$

and

$$\int (y-a)^2\, e^{\vartheta y - c(\vartheta)} d\nu(y) = \ddot{c}(\vartheta). \tag{5.20}$$

For simplicity, we study the important particular case where $T = \mathbb{R}$. Then in the homogeneous regression models (3.1) under consideration, equations in (5.9) reduce to

$$\int (a-y)\, e^{\alpha(y-a)+\vartheta y - c(\vartheta)} d\nu(y) = 0, \quad \vartheta \in \mathbb{R}, \tag{5.21}$$

which can be obtained from equations (5.19) with ϑ replaced by $\vartheta + \alpha$. Therefore, given any $\alpha \in \mathbb{R}$,

$$a(\vartheta) = \dot{c}(\vartheta + \alpha), \quad \vartheta \in \mathbb{R}, \tag{5.22}$$

are the unique solutions of equations (5.21). According to (3.12), this means that the *pseudoadditive rule*

$$\varphi_i(\theta) = \dot{c}\left(\phi(\mathbf{x}_i, \theta) + \alpha\right), \quad \theta \in \Theta, \tag{5.23}$$

leads to the adaptation of exponential $L_{2+\alpha}$-estimators to the exponential homogeneous regression models under consideration in the sense of (3.12), i. e. the adapted versions of the estimators (5.7) are

$$\widehat{\theta}_n^{(\alpha)} \sim \langle y\, e^{\alpha y};\, \dot{c}(\phi(\mathbf{x}_i, \theta) + \alpha) \rangle, \quad \alpha \in \mathbb{R}. \tag{5.24}$$

Replacing $\phi(\mathbf{x}_i, \theta)$ by the scalar product $\mathbf{x}_i'\theta$ we obtain from (5.23), (5.24) corresponding formulas for the exponential $L_{2+\alpha}$-estimators adapted to generalized linear models with natural link functions.

Let us look at the restrictions which Theorem 2.3 imposes on the estimators (5.24) and the respective exponential regression models. We start with sufficient conditions for (R3) and (R4+). The decomposition of the ψ-function figuring in (5.24) is as follows

$$\psi^+(y) = \begin{cases} y\, e^{\alpha y}\, I(\alpha y + 1 \geq 0) + \dfrac{1}{2} I(\alpha y + 1 \geq 0) & \text{if} \quad \alpha \neq 0, \\ y & \text{if} \quad \alpha = 0 \end{cases} \tag{5.25}$$

and

$$\psi^-(y) = \begin{cases} -y\, e^{\alpha y}\, I(\alpha y + 1 < 0) + \dfrac{1}{2}\, I(\alpha y + 1 \geq 0) & \text{if} \quad \alpha \neq 0, \\ 0 & \text{if} \quad \alpha = 0. \end{cases} \tag{5.26}$$

The part $\psi^-(y)$ is non-explosive, Lipschitz and bounded and square integrable on $\mathbb{R}$. The other part $\psi^+(y)$ is not so nice – it is explosive, non-Lipschitz, unbounded and square non-integrable on $\mathbb{R}$. To satisfy (R4+) we shall need Proposition 3.7.

Proposition 5.3 *The estimator $\widehat{\theta}_n$ defined by (5.7) fulfils in the model under consideration for all $\alpha \in \mathbb{R}$ the condition 2.12 in the regularity condition* (R3)*. If the expectations $\mu(\vartheta)$ and Fisher informations $\mathcal{I}(\vartheta)$ defined in* (5.18) *satisfy, for*

$$\phi_i = \phi(\mathbf{x}_i, \theta_0), \quad i \in \mathbb{N}, \tag{5.27}$$

and some $\alpha \in \mathbb{R}$, the inequalities

$$\sup_{i \in \mathbb{N}} \left[\mu(\phi_i + 2\alpha) - \mu(\phi_i + \alpha)\right]^2 < \infty \tag{5.28}$$

and

$$\sup_{n \in \mathbb{N}} \frac{1}{n} \sum_{i=1}^{n} \sup_{|t| \leq 2|\alpha|} \mathcal{I}(\phi_i + t) < \infty, \tag{5.29}$$

then the corresponding estimator $\widehat{\theta}_n^{(\alpha)}$ defined by (5.7) *fulfills also the regularity conditions* (R1) *and* (R4+)*.*

Proof: (I) For every $\alpha \in \mathbb{R}$, the derivative

$$\dot{\psi}^+(y) = (\alpha y + 1)\, e^{\alpha y}\, I(\alpha y + 1 \geq 0) \tag{5.30}$$

of $\psi^+(y)$ is nondecreasing on $\mathbb{R}$ if $\alpha \geq 0$, and nonincreasing if $\alpha < 0$. Consequently,

$$\Psi_\tau(y) := \sup_{|s| \leq \tau} \left|\dot{\psi}^+(y + s)\right| = \dot{\psi}^+(y + \tau\, \mathrm{sgn}\alpha), \quad y \in \mathbb{R}, \ \tau > 0, \tag{5.31}$$

where

$$\mathrm{sgn}\alpha = \begin{cases} 1 & \text{if} \quad \alpha \geq 0, \\ -1 & \text{if} \quad \alpha < 0. \end{cases}$$

By using the relation

$$|\dot{\psi}^+(y) - \dot{\psi}(y)| \leq \sup_{t \in \mathbb{R}} |\dot{\psi}^-(t)| \leq 1/e^2, \quad y \in \mathbb{R},$$

we find that (3.17) is equivalent to the condition

$$\sup_{n \in \mathbb{N}} \frac{1}{n} \sum_{i=1}^{n} \mathsf{E}\left[\dot{\psi}(Y_i - \dot{c}(\phi_i + \alpha) + \tau\, \mathrm{sgn}\alpha)\right]^2 < \infty \quad \text{for some } \tau > 0, \tag{5.32}$$

where $\dot{\psi}(y) = (\alpha y + 1)\, e^{\alpha y}$. By the Taylor Theorem the difference in the brackets of (5.28) equals $\alpha \ddot{c}(\phi_i + \alpha_i)$ for some $\alpha_i \in \mathbb{R}$. Using (5.19), (5.20), we obtain that the expectation of (5.31) is equal to

$$\begin{aligned} & e^{2|\alpha|\tau + b_i(\alpha)} \left[\alpha^2 \ddot{c}(\phi_i + 2\alpha) + (\alpha^2 \ddot{c}(\phi_i + \alpha_i) + |\alpha|\,\tau + 1)^2\right] \\ \leq \; & e^{2|\alpha|\tau} \left[\alpha^2 \ddot{c}(\phi_i + 2\alpha) + (\alpha^2 \ddot{c}(\phi_i + \alpha_i) + |\alpha|\,\tau + 1)^2\right], \end{aligned}$$

where

$$\begin{aligned} 0 \;&\leq\; b_i(\alpha) = c(\phi_i + 2\alpha) - c(\phi_i) - 2\alpha \dot{c}(\phi_i + 2\alpha) \\ &\leq\; 2|\alpha| \sup_{|t| \leq 2|\alpha|} \mathcal{I}(\phi_i + t) \end{aligned} \tag{5.33}$$

because $c(\vartheta)$ is convex and $\ddot{c}(\vartheta) = \mathcal{I}(\vartheta) > 0$. Therefore, if (5.28) and (5.29) hold then (5.32) holds too, and Proposition 3.7 implies that (R4+) holds.

(II) Using (5.33) and the notation of part (I), we get from the definition of σ_i^2 in (2.4) and from (5.23),

$$\begin{aligned} \sigma_i^2 \;&=\; \mathsf{E}\left[\psi(Y_i - \dot{c}(\phi_i + \alpha))\right]^2 \\ &=\; e^{b_i(\tau)} \left[\ddot{c}(\phi_i + 2\alpha) + (\dot{c}(\phi_i + 2\alpha) - \dot{c}(\phi_i + \alpha))^2\right]. \end{aligned} \tag{5.34}$$

By (5.33), the assumptions (5.28) and (5.29) imply the inequality (2.5) required in (R1).

(III) Using (5.33), we get from the formula for $H_i(t)$ in (2.10) and from (5.23),

$$\begin{aligned} H_i(t) \;&=\; \mathsf{E}\, \psi(Y_i - \dot{c}(\phi_i + \alpha) + t) \\ &=\; t\, e^{\alpha t + \tilde{b}_i(\alpha)}, \quad t \in \mathbb{R}, \end{aligned}$$

where (cf. (5.33))

$$\tilde{b}_i(\alpha) = c(\phi_i + \alpha) + c(\phi_i) - \alpha \dot{c}(\phi_i + \alpha) \leq 0 \tag{5.35}$$

due to the convexity of $c(\vartheta)$. This function is differentiable on $\mathbb{R}$ with the derivative

$$h_i(t) = (\alpha t + 1)\, e^{\alpha t + \tilde{b}_i(\alpha)}, \quad t \in \mathbb{R}. \tag{5.36}$$

Since $\tilde{b}_i(\alpha) \leq 0$, it holds for all $i \in \mathbb{N}$

$$\begin{aligned} |h_i(t) - h_i(0)| \;&\leq\; \left|(\alpha t + 1)\, e^{\alpha t} - 1\right| e^{\tilde{b}_i(\alpha)} \leq \left|(\alpha t + 1)\, e^{\alpha t} - 1\right| \\ &=\; \left|\alpha t + (\alpha t + 1)\,(e^{\alpha t} - 1)\right|. \end{aligned}$$

Therefore the condition 2.12 is satisfied on the infinite interval $(-\tau_0, \tau_0) = \mathbb{R}$ even if the conditions (5.28) and/or (5.29) fail to hold. □

Combining Proposition 5.3 and Theorem 2.3, we obtain the following assertion, in which we use for $\alpha \in \mathbb{R}$ and $i \in \mathbb{N}$ the constants $b_i(\alpha)$, $\tilde{b}_i(\alpha)$ defined by (5.33), (5.35) and

$$\sigma_i^2(\alpha) = e^{b_i(\alpha)} \left[\mathcal{I}(\phi_i + 2\alpha) + (\mu(\phi_i + 2\alpha) - \mu(\phi_i + \alpha))^2\right] \tag{5.37}$$

for ϕ_i and $\mu(\vartheta)$, $\mathcal{I}(\vartheta)$ defined by (5.27) and (5.18). We also use the formulas

$$\dot{\varphi}_i(\theta_0) = \mathcal{I}(\phi_i + \alpha)\, \dot{\phi}_i, \quad \text{for} \quad \dot{\phi}_i = \dot{\phi}(\mathbf{x}_i, \theta_0) \tag{5.38}$$

for the gradients $\dot{\varphi}_i(\theta_0)$ considered in conditions (2.8), (2.9) of (R2), provided the derivatives

$$\dot{\phi}(\mathbf{x}_i, \theta) = \left(\frac{\partial}{\partial\theta_1}, \dots, \frac{\partial}{\partial\theta_m}\right)' \phi(\mathbf{x}_i, \theta), \quad i \in \mathbb{N}, \tag{5.39}$$

exist in an open ball $B \subset \Theta$ centered at θ_0. We restrict ourselves to the exponential models (3.7) which satisfy (R2), i.e. for which the last condition holds and the gradients (5.39) satisfy (2.8) and (2.9).

Theorem 5.4 *Let for some $\alpha \in \mathbb{R}$ the estimator $\widehat{\theta}_n \sim \langle\psi, \varphi_i\rangle$ defined by (5.7) and (5.23) satisfy (R2) and the conditions (5.27) and (5.28) in a homogeneous regression model with exponential parent family (3.2). Further, let*

$$\Sigma_n := \frac{1}{n}\sum_{i=1}^{n} \sigma_i^2(\alpha)\, (\mathcal{I}(\phi_i + \alpha))^2\, \phi_i \dot{\phi}_i' \to \Sigma \tag{5.40}$$

and

$$\Phi_n := \frac{1}{n}\sum_{i=1}^{n} e^{\tilde{b}_i(\alpha)} (\mathcal{I}(\phi_i + \alpha))^2 \phi_i\, \dot{\phi}_i' \to \Phi, \tag{5.41}$$

where the matrices Σ and Φ are positive definite. Finally, let

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n} \psi(Y_i - \mu(\phi_i + \alpha))\, \mathcal{I}(\phi_i + \alpha)\, \dot{\phi}_i \xrightarrow{\mathcal{L}} N(0, \Sigma). \tag{5.42}$$

If $\widehat{\theta}_n^{(\alpha)}$ is consistent then it is asymptotically normal in the sense

$$\sqrt{n}(\widehat{\theta}_n^{(\alpha)} - \theta_0) \xrightarrow{\mathcal{L}} N(0,\, \Phi^{-1}\Sigma\, \Phi^{-1}). \tag{5.43}$$

Proof: Since $\widehat{\theta}_n^{(\alpha)}$ is consistent and satisfies (5.23), it is adapted to the model under consideration. By Proposition 5.3, (5.40) and (5.41), it satisfies the regularity conditions (R1), (R3) and (R4+). The remaining regularity condition (R2) is assumed. Since (5.42) means in the present situation the same as (2.17) all assumptions of Theorem 2.3 are satisfied. Therefore (5.43) follows from Theorem 2.3. □

Let us look at the special case

$$\mathcal{X} \subset \mathbb{R}^m, \quad \phi(\mathbf{x}, \theta) = \mathbf{x}'\theta \quad \text{and} \quad \alpha = 0, \tag{5.44}$$

i. e. the L_2-estimator $\widehat{\theta}_n^{(0)}$ of a true parameter $\theta_0 \in \Theta = \mathbb{R}$ in a generalized linear model with natural link function. Then (3.1) reduces to

$$G(y|i,\theta) \sim g(y|i,\theta) = e^{\mathbf{x}_i'\theta - c(\mathbf{x}_i'\theta)}, \quad \theta \in \mathbb{R}^m, \ i \in \mathbb{N}, \tag{5.45}$$

further $\phi_i = \mathbf{x}_i'\theta_0$ in (5.27), the gradients of (5.39) are given by formula

$$\dot{\phi}_i = \dot{\phi}(\mathbf{x}_i, \theta_0) = \mathbf{x}_i, \quad \theta_0 \in \mathbb{R}^m, \ i \in \mathbb{N}, \tag{5.46}$$

and the ψ-function is linear, $\psi(y) = y$. The conditions (5.28), (5.29) and (R2) take place if

$$\sup_{i \in \mathbb{N}} \|\mathbf{x}_i\| < \infty. \tag{5.47}$$

Further, (5.37) implies that

$$e^{\tilde{b}_i(0)} = 1 \quad \text{and} \quad \sigma_i^2(0) = \mathcal{I}(\mathbf{x}_i'\theta_n), \quad i \in \mathbb{N},$$

in the conditions (5.40), (5.41) of Theorem 5.4 so that they reduce to

$$\Sigma_n = \frac{1}{n}\sum_{i=1}^n (\mathcal{I}(\mathbf{x}_i'\theta_0))^3 \mathbf{x}_i\mathbf{x}_i' \to \Sigma \tag{5.48}$$

and

$$\Phi_n = \frac{1}{n}\sum_{i=1}^n (\mathcal{I}(\mathbf{x}_i'\theta_0))^2 \mathbf{x}_i\mathbf{x}_i' \to \Phi \tag{5.49}$$

for some positive definite matrices Σ and Φ. The remaining condition of Theorem 5.4 takes on the form

$$\frac{1}{\sqrt{n}}\sum_{i=1}^n (Y_i - \mu(\mathbf{x}_i'\theta_0))\, \mathcal{I}(\mathbf{x}_i'\theta_0)\mathbf{x}_i \xrightarrow{\mathcal{L}} N(0, \Sigma) \tag{5.50}$$

for Σ figuring in (5.48). We shall show that (5.50) follows from (5.47) and (5.48). Indeed, then $\vartheta_i = \mathbf{x}_i'\theta_0$ and $\mathcal{I}_i = \mathcal{I}(\vartheta_i)$ are uniformly bounded for $i \in \mathbb{N}$. Hence if $t \to 0$ then, uniformly for $i \in \mathbb{N}$,

$$c(\vartheta_i + t\,\mathcal{I}_i) = c(\vartheta_i) + \mu(\vartheta_i)\, t\,\mathcal{I}_i + \mathcal{I}_i^2\,\frac{t^2}{2} + o(t^2).$$

Further, for every $\xi \in \mathbb{R}$,

$$\mathsf{E}\exp\{(Y_i - \mu(\vartheta_i))\,\mathcal{I}_i\,\xi/\sqrt{n}\} = c(\vartheta_i + \xi\,\mathcal{I}_i/\sqrt{n}) - c(\vartheta_i)\,\xi\,\mathcal{I}_i/\sqrt{n}.$$

It follows from here that the moment generating functions $M_n(\tau) = \mathsf{E}\exp\{Z_n'\tau\}$, $\tau = (\tau_1, \ldots, \tau_m) \in \mathbb{R}^m$, of the random variables

$$Z_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (Y_i - \mu(\varphi_i))\, \mathcal{I}_i\, \mathbf{x}_i, \quad n \in \mathbb{N},$$

converge under (5.47) and (5.48) pointwise to

$$M(\tau) = \exp\left\{\frac{1}{2}\tau\, \Sigma\, \tau'\right\},$$

which suffices for (5.50). Therefore the following statement holds.

Corollary 5.5 *Let a generalized linear model* (5.45) *satisfy* (5.47) – (5.49). *If the* L_2-*estimator* $\widehat{\theta}_n^{(0)}$ *of a true parameter* $\theta_0 \in \mathbb{R}^m$ *is consistent, then it is asymptotically normal in sense of* (5.43), *where* Σ *and* Φ *are the matrices appearing in* (5.48) *and* (5.49).

Note that under the weak convergence of probability measures

$$\frac{1}{n} \sum_{i=1}^{n} \delta_{\mathbf{x}_i} \Rightarrow \mu$$

of Dirac's probability measures $\delta_{\mathbf{x}_i}$ on the regressor space $\mathcal{X}$, the conditions (5.48), (5.49) hold for

$$\Sigma = \int_{\mathcal{X}} \mathcal{I}(\mathbf{x}'\theta_0)^2 \mathbf{x}\mathbf{x}' \mu(d\mathbf{x}), \qquad \Phi = \int_{\mathcal{X}} \mathcal{I}(x'\theta_0) \mathbf{x}\mathbf{x}' \mu(d\mathbf{x}).$$

Similarly, the conditions (5.40), (5.41) hold but the formulas for the limit matrices are more complicated. Let us also note that in the generalized linear models of Corollary 5.5, none of the estimators $\widehat{\theta}_n^{(\alpha)}$, $\alpha \in \mathbb{R}$, is in general the MLE. Below is studied a special where $\widehat{\theta}_n^{(0)}$ is the MLE.

A similar asymptotic normality result as presented by Corollary 5.5 has been proved for the MLE in generalized linear models with natural link functions in Theorem 3 of Fahrmeir and Kaufmann [5]. The conditions of that theorem are weaker but less easily verifiable than the conditions (5.47) – (5.49) of Corollary 5.5, and the theorem does not provide the asymptotic variance-covariance matrix. Therefore the two results are not directly comparable.

The power of Theorem 2.3 has been verified in Section 4 by an application to the linear and nonlinear regression models. Another verification of this power can be obtained by an application to the model with observations Y_i with distributions from a natural exponential family. In this special case the $L_{2+\alpha}$-estimators

$$\hat{\vartheta}_n^{(\alpha)} \sim \langle t\, e^{\alpha t};\, \dot{c}(\vartheta + \alpha) \rangle, \quad \alpha \in \mathbb{R}, \tag{5.51}$$

estimate a true value $\vartheta_0 \in \mathbb{R}$ of the parameter of these distributions and $\hat{\vartheta}_n^{(0)}$ is the MLE. The estimators (5.51) are special cases of estimators (5.24) obtained for $\Theta = \mathbb{R}$ and $\phi(\mathbf{x}, \theta) = \theta$, so that $\phi_i = \vartheta_0$ and $\dot{\phi}_i = 1$ in the formulas above. Consequently, in (5.33) and (5.35).

$$\begin{aligned} b_i(\alpha) &= c(\vartheta_0 + 2\alpha) - c(\vartheta_0) - 2\alpha\dot{c}(\vartheta_0 + \alpha) =: b(\alpha), \\ \tilde{b}_i(\alpha) &= c(\vartheta_0 + \alpha) - c(\vartheta_0) - \alpha\dot{c}(\vartheta_0 + \alpha) =: \tilde{b}(\alpha), \end{aligned}$$

and in (5.37)

$$\sigma_i^2(\alpha) = e^{b(\alpha)} \left[\mathcal{I}(\vartheta_0 + 2\alpha) + (\mu(\vartheta_0 + 2\alpha) - \mu(\vartheta_0 + \alpha))^2\right] =: \sigma^2(\alpha).$$

Therefore (5.48) and (5.49) hold for

$$\Sigma_n = \Sigma = \sigma^2(\alpha)\,(\mathcal{I}(\vartheta_0 + \alpha))^2$$

and

$$\Phi_n = \Phi = e^{\tilde{b}(\alpha)}(\mathcal{I}(\vartheta_0 + \alpha))^2,$$

so that in (5.43) we have

$$\Phi^{-1}\Sigma\,\Phi^{-1} = \frac{e^{c(\vartheta_0+2\alpha)-c(\vartheta_0)}[\mathcal{I}(\vartheta_0 + 2\alpha) + (\mu(\vartheta_0 + 2\alpha) - \mu(\vartheta_0 + \alpha))^2]}{[e^{c(\vartheta_0+\alpha)-c(\vartheta_0)}\mathcal{I}(\vartheta_0 + \alpha)]^2} = s^2(\alpha). \tag{5.52}$$

The assumptions of Theorem 5.4 hold except the consistency which is clarified in the next proposition where we assume $T = \mathbb{R}$ for simplicity.

Proposition 5.6 *For every exponential family under consideration, the estimators $\hat{\vartheta}_n^{(\alpha)}$ defined by (5.51) are consistent, with values uniquely given by the formula*

$$\mu(\tilde{\vartheta}_n^{(\alpha)} + \alpha) = \frac{\sum_{i=1}^n Y_i\, e^{\alpha Y_i}}{\sum_{i=1}^n e^{\alpha Y_i}}, \quad n \in \mathbb{N}, \tag{5.53}$$

for $\mu(t) = \dot{c}(t)$ strictly increasing on $\mathbb{R}$.

Proof: Let $\alpha \in \mathbb{R}$ and $n \in \mathbb{N}$ be arbitrary fixed. By definition, $\hat{\vartheta}_n^{(\alpha)}$ minimizes

$$M_n(\vartheta) = \sum_{i=1}^n \rho(Y_i - \mu(\vartheta)), \tag{5.54}$$

where $\rho(t)$ is given by (5.8). If $\alpha = 0$ then the assertion is obvious. Suppose that $\alpha \neq 0$. Since $\mu(t)$ is infinitely differentiable on $\mathbb{R}$, we can consider the derivatives

$$\begin{aligned} \dot{M}_n(\vartheta) = \frac{d}{d\vartheta}M_n(\vartheta) &= \dot{\mu}(\vartheta)\sum_{i=1}^n \psi(Y_i - \mu(\vartheta)) \\ &= \dot{\mu}(\vartheta)\, e^{-\alpha\mu(\vartheta)} \sum_{i=1}^n (Y_i - \mu(\vartheta))\, e^{-\alpha Y_i} \end{aligned}$$

and

$$\ddot{M}_n(\vartheta) = \frac{d^2}{d\vartheta^2} M_n(\vartheta) = \ddot{\mu}(\vartheta) \sum_{i=1}^{n} \psi(Y_i - \mu(\vartheta)) + (\dot{\mu}(\vartheta))^2 Z_n(\vartheta),$$

where

$$\begin{aligned} Z_n(\vartheta) &= \sum_{i=1}^{n} \dot{\psi}(Y_i - \mu(\vartheta)) \\ &= \alpha \sum_{i=1}^{n} \psi(Y_i - \mu(\vartheta)) + \sum_{i=1}^{n} e^{\alpha(Y_i - \mu(\vartheta))}. \end{aligned}$$

By (5.18), $\dot{\mu}(\vartheta)$ is the Fisher information $\mathcal{I}(\vartheta) > 0$ for all $\vartheta \in \mathbb{R}$. Therefore $\hat{\vartheta}_n^{(\alpha)}$ given by (5.53) is the only solution of the equation $\dot{M}_n(\vartheta) = 0$. Further,

$$\begin{aligned} \ddot{M}_n(\hat{\vartheta}_n^{(\alpha)}) &= (\mathcal{I}(\hat{\vartheta}_n^{(\alpha)}))^2 Z_n(\hat{\vartheta}_n^{\alpha}) \\ &= (\mathcal{I}(\hat{\vartheta}_n^{(\alpha)}))^2 \sum_{i=1}^{n} e^{\alpha(Y_i - \mu(\hat{\vartheta}_n^{(\alpha)}))} > 0, \end{aligned}$$

so that $\hat{\vartheta}_n^{(\alpha)}$ is a unique local minimum of $M_n(\vartheta)$ on $\mathbb{R}$. We shall prove the relation

$$M_n(\vartheta) \geq M_n(\hat{\vartheta}_n^{(\alpha)}), \quad \vartheta \in \Theta, \tag{5.55}$$

which implies that $\hat{\vartheta}_n^{(\alpha)}$ is a unique global minimum of $M_n(\vartheta)$ on $\mathbb{R}$, i. e. that the second half of Proposition 5.6 is valid. By (5.54) and (5.8), for every $\vartheta \in \mathbb{R}$,

$$M_n(\vartheta) = \frac{1}{\alpha^2} \left(1 - \Gamma_n(\vartheta) \sum_{i=1}^{n} e^{\alpha Y_i} \right)$$

where

$$\Gamma_n(\vartheta) = \left(1 + \alpha \left[\mu(\vartheta + \alpha) - \mu(\hat{\vartheta}_n^{(\alpha)} + \alpha)\right]\right) e^{-\alpha\mu(\vartheta+\alpha)}.$$

Therefore (5.55) holds if

$$\Gamma_n(\vartheta) \leq \Gamma_n(\hat{\vartheta}_n^{(\alpha)}) = e^{-\alpha\mu(\hat{\vartheta}_n^{(\alpha)}+\alpha)}, \quad \theta \in \Theta,$$

i. e. if $\Delta_n(\vartheta) = \mu(\vartheta + \alpha) - \mu(\hat{\vartheta}_n^{(\alpha)} + \alpha)$ satisfies the relation

$$1 + \alpha\Delta_n(\vartheta) \leq e^{\alpha\Delta_n(\vartheta)}, \quad \vartheta \in \Theta.$$

This completes the proof of the second half of Proposition 5.6. The first part (consistency of $\hat{\vartheta}_n^{(\alpha)}$) follows, via the strict monotonicity and continuity of $\mu(t)$, from the fact that, by (5.53) and the law of large numbers,

$$\mu(\hat{\vartheta}_n^{(\alpha)} + \alpha) \xrightarrow{P} \mu(\vartheta + \alpha) \quad \text{as } n \to \infty.$$

□

By combining Proposition (5.6) with what has been said before, we obtain the following result.

Proposition 5.7 *Assume $T = \mathbb{R}$ for the exponential family (5.10). The estimators (5.51) of a true parameter $\vartheta_0 \in \mathbb{R}$ are explicitly given by formula (5.53). They are consistent and asymptotically normal in the sense that, for all $\alpha \in \mathbb{R}$,*

$$\sqrt{n}(\hat{\vartheta}_n^{(\alpha)} - \vartheta_0) \xrightarrow{\mathcal{L}} N(0, s^2(\alpha)) \quad \text{as } n \to \infty, \tag{5.56}$$

where $s^2(\alpha)$ is given by (5.52).

The asymptotic normality result (5.56) was obtained from the theory of Section 2 under the same generality as it can be obtained by a direct analysis of the concrete class of estimators

$$\hat{\vartheta}_n^{(\alpha)} = \mu^{-1}\left(\frac{\sum_{i=1}^n Y_i e^{\alpha Y_i}}{\sum_{i=1}^n e^{\alpha Y_i}}\right) - \alpha, \quad \alpha \in \mathbb{R}. \tag{5.57}$$

None of the assumptions of this theory imposed a superfluous restriction on the model or α. Again, this verifies in some sense that the general theory is strong enough to deal with concrete situations.

In the rest of section we study the exponential $L_{2+\alpha}$-estimators of parameters of two well known exponential families.

Example 5.8 Let the family (3.2) be standard normal with a location parameter $\vartheta \in \mathbb{R}$. Then

$$c(\vartheta) = \frac{\vartheta^2}{2}, \quad \mu(\vartheta) = \vartheta, \quad \mathcal{I}(\vartheta) = 1, \quad \vartheta \in \mathbb{R}, \tag{5.58}$$

and the dominating measure ν is the standard normal probability measure. By (5.57), the exponential $L_{2+\alpha}$-estimates are given by the formula

$$\hat{\vartheta}_n^{(\alpha)} = \frac{\sum_{i=1}^n Y_i\, e^{\alpha Y_i}}{\sum_{i=1}^n e^{\alpha Y_i}} - \alpha, \quad \alpha \in \mathbb{R}, \tag{5.59}$$

and, by (5.52) and (5.58), $s^2(\alpha) = 1 + \alpha^2$. By Proposition 5.7, the estimators (5.59) are asymptotically normal with asymptotic mean 0 and asymptotic variances $1 + \alpha^2$. If instead of the standard normal law $f(y|\vartheta_0)$ under consideration, the observations are governed by

$$(1-\varepsilon)\, f(y|\vartheta_0) + \varepsilon\, f(y|\vartheta_0, \sigma), \quad 0 < \varepsilon < 1, \tag{5.60}$$

where $f(y|\vartheta_0, \sigma)$ is a normal density with location ϑ_0 and scale $\sigma > 0$, then

$$b(\vartheta_0|\alpha, \varepsilon) = \frac{\varepsilon\, \vartheta_0 \sigma(\sigma - 1) \exp\left\{\frac{1}{2}\left[(\vartheta_0\sigma + \alpha)^2 - \vartheta_0^2\right]\right\}}{(1-\varepsilon)\exp\{2\vartheta_0\alpha\} + \varepsilon\sigma \exp\left\{\frac{1}{2}\left[(\vartheta_0\sigma + \alpha)^{2-\vartheta_0^2}\right]\right\}}$$

is the asymptotic bias of $\hat{\vartheta}_n^{(\alpha)}$. By a suitable choice of $\alpha \neq 0$, this bias can be held at a considerably lower levels over an a priori expected domain of ϑ_0 than is the level due to the MLE $\hat{\vartheta}_n^{(0)}$.

Example 5.9 Let the family (3.2) be Poisson with a parameter $\vartheta = \ln\lambda \in \mathbb{R}$. Then

$$c(\vartheta) = \mu(\vartheta) = \mathcal{I}(\vartheta) = e^{\vartheta}, \quad \vartheta \in \mathbb{R}, \tag{5.61}$$

and the dominating measure ν on $\mathbb{R}$ is finite and discrete,

$$\nu = \sum_{k=0}^{\infty} \frac{\delta_k}{k!},$$

where δ_k is the Dirac measure concentrating the mass 1 at the point $k \in \mathbb{R}$. In this case, by (5.57), the exponential $L_{2+\alpha}$-estimates are given by the formula

$$\hat{\vartheta}_n^{(\alpha)} = \ln \frac{\sum_{i=1}^n Y_i\, e^{\alpha Y_i}}{\sum_{i=1}^n e^{\alpha Y_i}} - \alpha \tag{5.62}$$

and, by (5.52) and (5.58),

$$s^2(\alpha) = \exp\left\{e^{\vartheta_0+\alpha}(e^{\alpha}-1) - \vartheta_0\right\}\left[1 + e^{\vartheta_0}(e^{\alpha}-1)^2\right]. \tag{5.63}$$

Therefore, by Proposition 5.7, the estimators defined by (5.62) are asymptotically normal with asymptotic mean 0 and asymptotic variances (5.63). This means that

$$\sqrt{n}\left(e^{\hat{\vartheta}_n^{(\alpha)}} - e^{\vartheta_0}\right) = e^{\vartheta_0}\sqrt{n}\left(e^{\hat{\vartheta}_n - \vartheta_0} - 1\right)$$

tends in law to

$$N\left(0, \exp\left\{e^{\vartheta_0+\alpha}(e^{\alpha+1}) + \vartheta_0\right\}\left[1 + e^{\vartheta_0}(e^{\alpha}-1)^2\right]\right),$$

i. e., that the exponential $L_{2+\alpha}$-estimators $\hat{\lambda}_n^{(\alpha)}$ of $\lambda_0 = e^{\vartheta_0}$ are asymptotically normal in the sense

$$\sqrt{n}\left(\hat{\lambda}_n^{(\alpha)} - \lambda_0\right) \xrightarrow{\mathcal{L}} N\left(0, \lambda_0 \exp\left\{\lambda_0 e^{\alpha}(e^{\alpha}-1)\right\}\left[1 + \lambda_0(e^{\alpha}-1)^2\right]\right).$$

If instead of the Poisson distribution $F(y|\vartheta)$ under consideration the observations are distributed by

$$(1-\varepsilon)\, F(y|\vartheta) + \varepsilon G(y), \quad 0 < \varepsilon < 1, \tag{5.64}$$

where

$$G(y) = \zeta(2)^{-1} \sum_{k=1}^{\infty} \frac{1}{k^2}\, I(y > k)$$

and $\zeta(s)$, $s > 1$, is the Riemann function, then the asymptotic bias of the MLE $\hat{\vartheta}_n^{(0)}$ is infinite for arbitrarily small ε. Indeed, if $\tilde{Y}_i$ are observations i.i.d. by (5.64) then

$$E\,\tilde{Y}_i = (1-\varepsilon)\, e^{\vartheta_0} + \varepsilon\zeta(2)^{-1} \sum_{k=1}^{\infty} \frac{k}{k^2} = \infty.$$

On the other hand, the asymptotic bias $b(\vartheta_0|\alpha,\varepsilon)$ of every estimator $\hat{\vartheta}_n^{(\alpha)}$ with $\alpha < 0$ satisfies the relation

$$\lim_{\varepsilon\downarrow 0} b(\vartheta_0|\alpha,\varepsilon) = 0 \quad \text{for every } \vartheta_0 \in \mathbb{R}.$$

The results in the last two examples demonstrate that in the class of $L_{2+\alpha}$-estimators one can find more robust alternatives to the L_2-estimator (MLE). The price payed for the robustness is a larger asymptotic variance when the observations are not contaminated.

6 Proof of Theorems 2.2 and 2.3

Unless otherwise explicitly stated, we consider in this section arbitrary model (1.1) and M-estimator $\hat{\theta}_n \sim \langle \psi; \varphi_i \rangle$ where ψ can be decomposed as the difference (2.1) of two nondecreasing functions ψ^+ and ψ^-. We suppose for simplicity that both these functions are right-continuous. Then also their sum $\psi^\pm$ introduced in (2.2) and ψ itself are right continuous. We shall formulate a series of auxiliary statements leading to the proofs of the Theorems 2.2 and 2.3. All statements refer to the concepts and conditions introduced in Sections 1 and 2. Most of these statements are technical but some of them are interesting also from the statistical point of view.

If $\xi : \mathbb{R} \mapsto \mathbb{R}$ is nondecreasing and right continuous then there exists unique measure μ_ξ on the Borel subsets of $\mathbb{R}$ associated with ξ and satisfying relation $\mu(a,b]) = \xi(b) - \xi(a)$ for all real numbers $a < b$. If $\phi : \mathbb{R} \mapsto \mathbb{R}$ is measurable then the Lebesgue-Stieltjes integral is defined as the Lebesgue integral for the associated measure, e. g.

$$\int_{(a,b]} \phi(s)\, d\xi(s) = \int_{(a,b]} \phi(s)\, \mu_\xi(ds).$$

If η is another monotone right continuous function then the bivariate Lebesgue–Stieltjes integral

$$\int_{(a,b]^2} \phi(s,t)\, d\xi(s)\, d\eta(t) \tag{6.1}$$

can be defined by means of the associated measure $\mu_\xi \otimes \mu_\eta$ on the Borel subsets of $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$. For locally bounded functions $\phi(s)$ (e. g. for linear combinations of monotone functions), and for differences $\xi(s) = \xi^+(s) - \xi^-(s)$ of two nondecreasing right-continuous functions, one can define the Lebesgue–Stieltjes integral

$$\int_{(a,b]} \phi(s)\, d\xi(s) = \int_{(a,b]} \phi(s)\, d\xi^+(s) - \int_{(a,b]} \phi(s)\, d\xi^-(s).$$

If $\eta = \eta^+ - \eta^-$ is a similar difference then one can similarly extend the bivariate Lebesgue–Stieltjes integrals (6.1). Using the bounded measurable function

$$\phi(s,t) = I(a < t \le b)\, I(a < s \le t) = I(a < s \le b) I(s \le t \le b)$$

defined by means of the indicator function $I(\cdot)$, and employing equalities of the type

$$\int I(a < s \le t)\, d\xi(s) = \xi(t) - \xi(a),$$

one obtains from the Fubini theorem the *per partes* rule

$$\int_{(a,b]} \eta(s)\, d\xi(s) + \int_{(a,b]} \xi(s-)\, d\eta(s) = \xi(b)\, \eta(b) - \xi(a)\, \eta(a) \tag{6.2}$$

for Lebesgue–Stieltjes integrals. In this rule, $\xi(s-)$ denotes the left continuous version of $\xi(s)$.

Our first statement is concerning the criterion function ρ satisfying according to (1.7) for all $y \in \mathbb{R}$ the relation

$$\rho(y) = \rho(0) + \int_{(0,y]} \psi(s) ds. \tag{6.3}$$

Proposition 6.1 *For all y, $t \in \mathbb{R}$ holds the generalized Taylor formula*

$$\rho(y+t) = \rho(y) + \psi(y)\, t + R(y,t) \tag{6.4}$$

where the remainder is

$$R(y,t) = \int_{y}^{y+t} (y+t-s)\, d\psi(s). \tag{6.5}$$

Proof: By (1.7) and the *per partes* rule (6.2),

$$\begin{aligned} \rho(y+t) - \rho(y) &= \int_{(y,y+t]} \psi(s)\, ds \\ &= \psi(y)\, t + \int_{(y,y+t]} (y+t-s)\, d\psi(s). \end{aligned}$$

□

By applying the generalized Taylor formula (6.4) in (1.3) we obtain

$$\begin{aligned} M_n(\theta) - M_n(\theta_n) &= \frac{1}{n} \sum_{i=1}^{n} [\rho(Y_i - \varphi_i(\theta)) - \rho(Y_i - \varphi_i(\theta_0))] \\ &= \frac{1}{n} \sum_{i=1}^{n} \psi(X_i)\, t_i + \frac{1}{n} \sum_{i=1}^{n} R(X_i, t_k) \end{aligned} \tag{6.6}$$

where $X_i = Y_i - \varphi_i(\theta_0)$ and $t_i = \varphi_i(\theta) - \varphi_i(\theta_0)$. The first sum in the last row is linear in t_i. Therefore we are interested in the behavior of the expected remainders $\mathsf{E}R(X_i, t)$ in a neighborhood of $t = 0$.

Proposition 6.2 *Let the regularity condition* (R3) *hold and let $X_i = Y_i - \varphi_i(\theta_0)$. Then the expectations $\mathsf{E}R(X_i, t)$ are locally quadratic in the sense that, for the functions $h_i : (-\tau_0, \tau_0) \mapsto \mathbb{R}$ introduced in* (R3) *and all $0 < \tau < \tau_0$*

$$\sup_{|t| \le \tau} \left| \mathsf{E}R(X_i,\, t) - h_i(0)\, \frac{t^2}{2} \right| \le \frac{t^2}{2} \omega(h_i, \tau). \tag{6.7}$$

Proof: Consider $t \in (-\tau_0, \tau_0)$. If (R3) holds then, by the Fubini theorem and (6.4),

$$\begin{aligned}
\mathsf{E}R(X_i,t) &= \int_0^t \mathsf{E}\psi(X_i+s)\,ds - t\mathsf{E}\psi(X_i) \\
&= \int_0^t [H_i(s)-H_i(0)]ds \\
&= \int_0^t \int_0^s h_i(u)duds = \int_0^t (t-u)h_i(u)du \\
&= \int_0^t (t-u)h_i(0)du + \int_0^t (t-u)[h_i(u)-h_i(0)]du.
\end{aligned}$$

The rest is clear from here and from the definition of $\omega(h_i,\tau)$. □

The next result estimates fluctuations of the remainders $R(X_i,t)$ around $\mathsf{E}\,R(X_i,t)$.

Proposition 6.3 *If the regularity condition* (R4+) *holds then for τ_0, q and κ considered in* (R4+)*, and for $X_i = Y_i - \varphi_i(\theta_0)$ and all $0<\tau<\tau_0$,*

$$\sup_{n\in\mathbb{N}} \frac{1}{n}\sum_{i=1}^n \mathsf{E}\sup_{|t|\le\tau}(R(X_i,t))^2 < \kappa\,\tau^{2+q}. \tag{6.8}$$

Proof: Let $y\in\mathbb{R}$ be arbitrary fixed. By substitution $y+t\mapsto t$ and the convention (1.8), it follows from (6.5)

$$R(y,t) = \int_{(0,t]} (t-s)\,d\psi(s+y) = \int_{(t^-,t^+]} |t-s|\,d\psi(s+y)$$

where $t^- = \min\{0,t\}$ and $t^+ = \max\{0,t\}$. Hence for every $t\in\mathbb{R}$

$$\begin{aligned}
|R(y,t)| &= \left|\int_{(t^-,t^+]} |t-s|\,d\psi^+(s) - \int_{(t^-,t^+]} |t-s|\,d\psi^-(s)\right| \\
&\le \left|\int_{(t^-,t^+]} |t-s|\,d\psi^\pm(y+s)\right| \quad \text{(cf. 2.2)} \\
&\le |t|\left[\psi^\pm(y+t^+) - \psi^\pm(y-t^-)\right] \\
&\le |t|\left[\psi^\pm(y+|t|) - \psi^\pm(y-|t|)\right].
\end{aligned}$$

Consequently,

$$\sup_{|t|\le\tau}(R(y,t))^2 \le t^2\left[\psi^\pm(y+\tau) - \psi^\pm(y-\tau)\right]^2$$

and (6.8) follows from (R4+). □

Next follows an important technical result which is sharper than a similar result in [21] and which is proved by a different method. Consider closed balls $B_\gamma \subset \mathbb{R}^m$ of diameters

$0 < \gamma \leq \delta$, $\delta \leq \infty$, centered at $0 \in \mathbb{R}^m$, and a sequence $S_1(\mathbf{u})$, $S_2(\mathbf{u}), \ldots$ of continuous independent zero-mean random processes $(S_i(\mathbf{u}) : \mathbf{u} \in B_\delta)$ with $S_1(0) = S_2(0) = \cdots = 0$. For given $0 < \gamma \leq \delta$ and $n \in \mathbb{N}$, we shall estimate the expected modulus of continuity

$$\Omega_n(\gamma) = \mathsf{E} \sup_{u \in B_\gamma} \left| \frac{1}{\sqrt{n}} \sum_{i=1}^n S_i(u) \right| \tag{6.9}$$

of the normalized sum at $\mathbf{u} = 0$. A useful estimate will be obtained by means of the theory of empirical processes, in particular by the results in Chapter 2 of [30]. We suppose that for some $\delta > 0$

$$|S_i(\mathbf{u}) - S_i(\tilde{\mathbf{u}})| \leq \Lambda_i \|\mathbf{u} - \tilde{\mathbf{u}}\| \quad \text{for all} \quad \mathbf{u}, \tilde{\mathbf{u}} \in B_\delta,\ i \in \mathbb{N}, \tag{6.10}$$

and

$$\sup_{n \in \mathbb{N}} \mathsf{E} \frac{1}{n} \sum_{i=1}^n L_n^2 < \infty \quad \text{for} \quad L_n = \left(\frac{1}{n} \sum_{i=1}^n \Lambda_i^2 \right)^{1/2} \tag{6.11}$$

Set

$$\Gamma_n(\gamma) = \sup_{\mathbf{u} \in B_\gamma} \left[\frac{1}{n} \sum_{i=1}^n S_i^2(\mathbf{u}) \right]^{1/2}.$$

Proposition 6.4 *Suppose that $S_1(\mathbf{u}), S_2(\mathbf{u}), \ldots$ are continuous , independent zero-mean stochastic processes continuous on B_δ with $S_i(0) = 0$. If the condition (6.10) and (6.11) hold then there exists a universal constants K and $\kappa(d)$ such that for $\gamma \leq \delta$*

$$\mathsf{E} \sup_{\|\mathbf{u}\| \leq \gamma} \left| \frac{1}{\sqrt{n}} \sum_{i=1}^n S_i(\mathbf{u}) \right| \leq \gamma \mathsf{E} \left[L_n \Gamma \left(d, \frac{2\Gamma_n(\gamma)}{\delta L_n} \right) \right] \tag{6.12}$$

and

$$\mathsf{E} \sup_{\|\mathbf{u}\| \leq \delta} \left| \frac{1}{\sqrt{n}} \sum_{i=1}^n S_i(\mathbf{u}) \right| \leq \delta \Gamma(d, 2)\, \mathsf{E} L_n \tag{6.13}$$

where

$$\Gamma(d, s) = 2K \int_0^s \sqrt{|\ln(\kappa(d) t^d)|} dt.$$

For every $0 < \alpha < 1$ there exists a constant $C(\alpha, d)$ such that

$$\mathsf{E} \sup_{\|\mathbf{u}\| \leq \gamma} \left| \frac{1}{\sqrt{n}} \sum_{i=1}^n S_i(\mathbf{u}) \right| \leq C(\alpha, d) \gamma^\alpha \Gamma^\alpha(d, \gamma) \mathsf{E} L_n^{1-\alpha}. \tag{6.14}$$

Proof: Suppose $\varepsilon_1, ..., \varepsilon_n$ are independent binary random variables taking on the values 1 and -1 with equal probability 1/2. Assume that for $n = 1, 2, ...$ the set $A_n \subset \mathbb{R}^n$ is bounded with respect to the Euclidean distance $\|\cdot\|_n$ on $\mathbb{R}^n$. Denote by $\mathbf{N}(\varepsilon, A_n)$ the minimal number of balls of radius $\varepsilon > 0$ covering A_n. Then by Corollary 2.2.8 in [30] there is a universal constant K such that

$$\mathsf{E} \sup_{(a_1,\ldots,a_n)\in A_n} \left|\sum_{i=1}^{n} a_i \varepsilon_i\right| \le K \int_0^\infty \sqrt{\ln \mathbf{N}\left(\frac{\varepsilon}{2}, A_n\right)} d\varepsilon. \tag{6.15}$$

The symmetrization Lemma 2.3.6 of [30] yields

$$\mathsf{E} \sup_{\|\mathbf{u}\|\le\gamma} \left|\frac{1}{\sqrt{n}} \sum_{i=1}^{n} S_i(\mathbf{u})\right| \le \mathsf{E}\left(\mathsf{E}_\varepsilon \sup_{\|\mathbf{u}\|\le\gamma} \left|\frac{1}{\sqrt{n}} \sum_{i=1}^{n} S_i(\mathbf{u})\varepsilon_i\right|\right) \tag{6.16}$$

where the $\varepsilon_1, \ldots, \varepsilon_n$ are independent Bernoulli variables which are independent of the processes $S_1(\mathbf{u}), \ldots, S_n(\mathbf{u})$ and take on the values 1 and -1 with probability 1/2. The symbol E_ε denotes the expectation w.r.t. $\varepsilon_1, \ldots, \varepsilon_n$. To estimate the right hand term we suppose that the processes $S_1(\mathbf{u}), \ldots, S_n(\mathbf{u})$ and random variables $\varepsilon_1, \ldots, \varepsilon_n$ are defined on a product space, say $(\Omega_1 \times \Omega_2, \mathfrak{F}_1 \otimes \mathfrak{F}_2, \mathbb{P}_1 \times \mathbb{P}_2)$ where the processes depend on $\omega_1 \in \Omega_1$ and the binary variables depend on $\omega_2 \in \Omega_2$. Fix $\omega_1 \in \Omega_1$ and introduce

$$A_{n,\gamma}(\omega_1) = \left\{\frac{1}{\sqrt{n}} S_1(\mathbf{u}, \omega_1), \ldots, \frac{1}{\sqrt{n}} S_n(\mathbf{u}, \omega_1), \mathbf{u} \in B_\gamma\right\} \subseteq \mathbb{R}^n.$$

For fixed ω_1 we estimate the entropy number appearing in (6.15). The Lipschitz condition (6.10) implies that for every ε-net for B_γ there is an $L_n\varepsilon-$net for $A_{n,\gamma}(\omega_1)$. For $\gamma \le \delta$ the entropy number of B_γ does not exceed $\kappa(d)(\frac{\gamma}{\varepsilon})^d$ where $\kappa(d)$ is a constant depending on d only. As the diameter of $A_{n,\gamma}$ does not exceed, $2\Gamma_n(\gamma)$ we have

$$\mathbf{N}\left(\frac{\varepsilon}{2}, A_{n,\gamma}\right) \le \begin{cases} \kappa(d)\left[\frac{\gamma}{\varepsilon}\right]^d (2L_n)^d & \text{for} \quad \varepsilon \le 4\Gamma_n(\gamma) \\ 1 \quad \text{for} & \varepsilon > 4\Gamma_n(\gamma) \end{cases} \tag{6.17}$$

and

$$\begin{aligned} \mathsf{E} \sup_{(a_1,\ldots,a_n)\in A_{n,\delta}} \left|\sum_{i=1}^{n} a_i \varepsilon_i\right| &\le K \int_0^{4\Gamma_n(\gamma)} \sqrt{\left|\ln\left[\kappa(d)\left(\frac{2L_n\gamma}{\varepsilon}\right)^d\right]\right|} d\varepsilon \\ &= K 2\delta L_n \int_0^{2\Gamma_n(\gamma)/(\delta L_n)} \sqrt{|\ln(\kappa(d)t^d)|} dt. \end{aligned}$$

To complete the proof we set

$$\Gamma(d, s) = 2K \int_0^s \sqrt{|\ln(\kappa(d)t^d)|} dt,$$

and obtain (6.12). To prove (6.13) it suffices to observe that the assumption (6.10) yields $\Gamma_n(\gamma)/(\delta L_n) \le 1$. Using the inequality

$$\ln x \le \frac{x^{1-\alpha}}{1-\alpha} \quad \text{for} \quad x \ge 1 \text{ and } 0 < \alpha < 1$$

we find a constant $C(d,\alpha)$ such that

$$2K\int_0^s \sqrt{|\ln(\kappa(d)t^d)|}dt \le C(d,\alpha)s^{1-\alpha},$$

which proves the statement (6.14). □

In the next proposition we assume that the adaptation condition (2.3) and the regularity conditions (R1), (R2) and (R4) hold. We introduce the local parameter

$$\mathbf{u} = \theta - \theta_0 \in B_\delta = \{\mathbf{u} \in \mathbb{R}^m : \|\mathbf{u}\| \le \delta\}$$

where $\delta > 0$ is the same as in the definition (2.7) of the ball B in (R2). In the proposition we study the zero-mean version $D_n(\mathbf{u}) - \mathsf{E}\, D_n(\mathbf{u})$ of the random process $(D_n(\mathbf{u}) : \mathbf{u} \in B_\delta)$ defined by

$$D_n(\mathbf{u}) = \sqrt{n}\left(M_n(\theta_0 + \mathbf{u}) - M_n(\theta_0)\right). \tag{6.18}$$

To simplify formulas we use the notations

$$\varphi_i = \varphi_i(\theta_0), \quad \xi_i(\mathbf{u}) = \varphi_i(\theta_0 + \mathbf{u}) - \varphi_i, \quad X_i = Y_i - \varphi_i, \quad \dot\varphi_i = \dot\varphi_i(\theta_0). \tag{6.19}$$

By (1.3) and (6.4),

$$\begin{aligned} D_n(\mathbf{u}) &= \frac{1}{\sqrt{n}}\left(\sum_{i=1}^n [\rho(X_i - \xi_i(\mathbf{u})) - \rho(X_i)]\right) \\ &= \frac{1}{\sqrt{n}}\sum_{i=1}^n \psi(X_i)\,\xi_i(\mathbf{u}) + \frac{1}{\sqrt{n}}\sum_{i=1}^n R_i(\mathbf{u}) \end{aligned}$$

where we put for simplicity

$$R_i(\mathbf{u}) = R(X_i, \xi_i(\mathbf{u})), \quad \mathbf{u} \in B_\delta.$$

It follows that

$$D_n(\mathbf{u}) = \mathcal{L}_n(\mathbf{u}) + \mathcal{D}_n(\mathbf{u}) + \mathcal{R}_n(\mathbf{u}), \quad \mathbf{u} \in B_\delta,$$

where the linear term $\mathcal{L}_n$, deviation $\mathcal{D}_n$ and remainder $\mathcal{R}_n$ are given by

$$\begin{aligned} \mathcal{L}_n(\mathbf{u}) &= \frac{1}{\sqrt{n}}\sum_{i=1}^n \psi(X_i)\,(\dot\varphi_i'\mathbf{u}), \qquad (6.20) \\ \mathcal{D}_n(\mathbf{u}) &= \frac{1}{\sqrt{n}}\sum_{i=1}^n \psi(X_i)[\xi_i(\mathbf{u}) - \dot\varphi_i'\mathbf{u}], \\ \mathcal{R}_n(\mathbf{u}) &= \frac{1}{\sqrt{n}}\sum_{i=1}^n R_i(\mathbf{u}). \end{aligned}$$

Since $\widehat{\theta}_n$ is adapted, (2.3) implies $\mathsf{E}\mathcal{L}_n(\mathbf{u}) = \mathsf{E}\mathcal{D}_n(\mathbf{u}) = 0$, so that

$$D_n(\mathbf{u}) - \mathsf{E}\, D_n(\mathbf{u}) = \mathcal{L}_n(\mathbf{u}) + \mathcal{D}_n(\mathbf{u}) + \mathcal{S}_n(\mathbf{u}), \tag{6.21}$$

where

$$\mathcal{S}_n(\mathbf{u}) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} S_i(\mathbf{u}) \quad \text{for} \quad S_i(\mathbf{u}) = R_i(\mathbf{u}) - \mathsf{E}\, R_i(\mathbf{u}). \tag{6.22}$$

Proposition 6.5 *Let B_γ be a zero centered ball of radius γ and let the adaption condition* (2.3) *and the regularity conditions* (R1), (R2) *and* (R4) *hold. Then for every $0 < \alpha < 1$ and the above considered processes $\mathcal{L}_n(\mathbf{u})$, $\mathcal{D}_n(\mathbf{u})$, $\mathcal{S}_n(\mathbf{u})$ defined on B_δ there exist constants c_0, c_1, c_2 such that, for all $n \in \mathbb{N}$,*

$$\mathsf{E} \sup_{\mathbf{u}\in B_\gamma} |\mathcal{L}_n(\mathbf{u})| \le c_0\gamma \quad \textit{if} \quad 0 < \gamma \le \delta, \tag{6.23}$$

$$\mathsf{E} \sup_{\mathbf{u}\in B_\gamma} |\mathcal{D}_n(\mathbf{u})| \le c_1\gamma^2 \quad \textit{if} \quad 0 < \gamma \le \delta, \tag{6.24}$$

$$\mathsf{E} \sup_{\mathbf{u}\in B_\gamma} |\mathcal{S}_n(\mathbf{u})| \le c_2\gamma \quad \textit{if} \quad 0 < \gamma \le \delta. \tag{6.25}$$

If in addition (R4+) *holds then for every $0 < \alpha < 1$ there exist a constants c_3 and $q > 0$ such that, for all $n \in \mathbb{N}$,*

$$\mathsf{E} \sup_{\mathbf{u}\in B_\gamma} |\mathcal{S}_n(\mathbf{u})| \le c_3\gamma^{1+\alpha q/2} \quad \textit{if} \quad 0 < \gamma \le \delta.$$

Proof: The regularity condition (2.8) implies

$$|\varphi_i(\theta_0 + \mathbf{u}) - \varphi_i(\theta_0)| \le \lambda \|\mathbf{u}\| \le \lambda\gamma \qquad \text{for} \qquad \gamma \le \delta \tag{6.26}$$

and

$$\varphi_i(\theta_0 + \mathbf{u}) - \varphi_i(\theta_0) - \dot{\varphi}_i'\mathbf{u} = \int_0^1 [\dot{\varphi}_i'(\theta_0 + s\mathbf{u}) - \dot{\varphi}_i'(\theta_0)]\mathbf{u} ds. \tag{6.27}$$

Hence by the Lipschitz continuity of $\dot{\varphi}_i$ required in (2.9)

$$\begin{aligned} |\varphi_i(\theta_0 + \mathbf{u}) - \varphi_i - \dot{\varphi}_i'\mathbf{u}| &\le \|\mathbf{u}\| \int_0^1 \lambda \|\mathbf{u}\|\, s ds \\ &= \frac{\lambda}{2} \|\mathbf{u}\|^2. \end{aligned} \tag{6.28}$$

By the definition of $\mathcal{L}_n(\mathbf{u})$, for every $\gamma \le \delta$

$$\mathsf{E} \sup_{\mathbf{u}\in B_\gamma} |\mathcal{L}_n(\mathbf{u})| \le \gamma \mathsf{E} \left\| \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \psi(X_i)\, \dot{\varphi}_i \right\|.$$

By assumption (2.3) we have $\mathsf{E}\psi(X_i) = 0$. Hence it follows from the independence of X_i that

$$\left(\mathsf{E}\left\|\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\psi(X_i)\,\dot{\varphi}_i\right\|\right)^2 \leq \frac{1}{n}\sum_{i=1}^{n}\|\dot{\varphi}_i\|^2\,\mathsf{E}\psi^2(X_i).$$

We see from (3.10) and (2.5) that (6.23) holds for $c_0 = \sqrt{\lambda C}$, where λ is the constant figuring in (3.10) and C is the supremum in (2.5). Similarly, by the definition of $\mathcal{D}_n(\mathbf{u})$ and (6.28),

$$\left(\mathsf{E}\sup_{\mathbf{u}\in B_\gamma}|\mathcal{D}_n(\mathbf{u})|\right)^2 \leq \frac{\lambda\gamma^2}{2n}\sum_{i=1}^{n}\mathsf{E}(\psi(X_i))^2.$$

To prove (6.25) we shall apply Proposition 6.4. The independent zero-mean processes $(S_i(\mathbf{u}) : \mathbf{u} \in B_\gamma)$, $i \in \mathbb{N}$, satisfy all assumptions of Proposition 6.4. Indeed, since $\xi_i(\mathbf{u}) = \varphi_i(\theta_0 + \mathbf{u}) - \varphi_i$, it holds $S_i(0) = 0$. Using similar arguments as in the proof of Proposition 6.3 we get that the modulus of the function

$$|t-s|I(t^- < s \leq t^+) - |\widetilde{t}-s|I(\widetilde{t}^- < s \leq \widetilde{t}^+)$$

is for every $t, \widetilde{t} \in \mathbb{R}$ bounded above by $|t - \widetilde{t}|$. Therefore using similar arguments as in the mentioned proof, we obtain that for all $t, \widetilde{t} \in (-\tau, \tau)$

$$\left|R(y,t) - R(y,\widetilde{t})\right| \leq |t-\widetilde{t}|\left[\psi^{\pm}(y+\tau) - \psi^{\pm}(y-\tau)\right].$$

It follows from here that the processes $(R_i(\mathbf{u}) : \mathbf{u} \in B_\gamma)$ satisfy for all $\mathbf{u}$, $\widetilde{\mathbf{u}} \in B_\gamma$ the inequalities

$$|R_i(\mathbf{u}) - R_i(\widetilde{\mathbf{u}})| \leq \left[\psi^{\pm}(X_i - \varphi_i + \tau_i) - \psi^{\pm}(X_i - \varphi_i - \tau_i)\right]|\xi_i(\mathbf{u}) - \xi_i(\widetilde{\mathbf{u}})|$$

where

$$\tau_i = \tau_i(\mathbf{u}, \widetilde{\mathbf{u}}) = \max\{|\xi_i(\mathbf{u})|,\, |\xi_i(\widetilde{\mathbf{u}})|\}.$$

We get from (6.26) $\tau_i(\mathbf{u}, \widetilde{\mathbf{u}}) \leq \lambda\delta$, so that the monotonicity of $\psi^{\pm}$ implies

$$\begin{aligned}|R_i(\mathbf{u}) - R_i(\widetilde{\mathbf{u}})| &\leq Z_i|\xi_i(\mathbf{u}) - \xi_i(\widetilde{\mathbf{u}})| \\ &= Z_i\left|\int_0^1[\dot{\varphi}_i(\theta_0 + \widetilde{\mathbf{u}} + s(\mathbf{u}-\widetilde{\mathbf{u}}))]'[\mathbf{u}-\widetilde{\mathbf{u}}]ds\right|\end{aligned}$$

where

$$Z_i = \psi^{\pm}(X_i - \varphi_i + \lambda\delta_0) - \psi^{\pm}(X_i - \varphi_i - \lambda\delta_0).$$

Hence by (6.10)

$$|R_i(\mathbf{u}) - R_i(\widetilde{\mathbf{u}})| \leq \lambda Z_i\,\|\mathbf{u} - \widetilde{\mathbf{u}}\|.$$

Thus the zero-mean versions $S_i(\mathbf{u}) = R_i(\mathbf{u}) - \mathsf{E}\,R_i(\mathbf{u})$ satisfy the inequalities

$$|S_i(\mathbf{u}) - S_i(\widetilde{\mathbf{u}})| \leq \widetilde{Z}_i\,\|\mathbf{u} - \widetilde{\mathbf{u}}\| \qquad \text{where} \qquad \widetilde{Z}_i = \lambda Z_i + \lambda\mathsf{E}\,Z_i.$$

Note that

$$\mathsf{E}(\widetilde{Z}_i)^2 \leq 4\lambda^2 \mathsf{E}(Z_i)^2. \tag{6.29}$$

The statement (6.25) now follows from (6.13) with

$$\mathsf{E}L_n = \sup_n \mathsf{E}\left(\frac{1}{n}\sum_{i=1}^n (\widetilde{Z}_i)^2\right)^{1/2} \leq 2\lambda^2 \left(\frac{1}{n}\sum_{i=1}^n \mathsf{E}Z_i{}^2\right)^{1/2}$$

because (R4) guarantees that the right-hand terms are bouded by a constant. □

In the following result we use the above considered ball B_δ, and also similar balls B_γ centered at $0 \in \mathbb{R}^m$ with arbitrary $\gamma > 0$.

Proposition 6.6 *(van der Vaart and Wellner). Let $\widehat{\theta}_n$ be consistent. If there exist constants $0 < \delta_0 \leq \delta$ and $\kappa_1,\, \kappa_2 > 0$ such that*

$$\liminf_{n\to\infty} \inf_{\mathbf{u}\in B_{\delta_0}} \left(\frac{1}{\sqrt{n}}\,\mathsf{E}\,D_n(\mathbf{u}) - \kappa_1\|\mathbf{u}\|^2\right) \geq 0 \tag{6.30}$$

and

$$\limsup_{n\to\infty} \mathsf{E} \sup_{u\in B_\gamma} |D_n(\mathbf{u}) - \mathsf{E}\,D_n(\mathbf{u})| \leq \kappa_2\gamma \quad \textit{for all} \quad 0 < \gamma \leq \delta_0 \tag{6.31}$$

then the estimator $\hat{\theta}_n$ under consideration is $\sqrt{n}$-consistent in the sense of (1.11).

Proof: See Theorem 3.2.5 of [30]. □

Proposition 6.7 *Let the estimator satisfy the adaption condition 2.3, the regularity conditions (R2)-(R4) and the condition (2.14) of Theorem (2.3). Then for every $\mathbf{u} \in B_\delta$ and the matrix Φ_n defined in (2.14),*

$$\sup_{\mathbf{u}\in B_\delta} \left|\frac{1}{\sqrt{n}}\,\mathsf{E}\,D_n(\mathbf{u}) - \frac{1}{2}\,\mathbf{u}'\Phi_n\mathbf{u}\right| \leq \frac{(\lambda\delta)^2}{2n}\sum_{i=1}^n \omega(h_i, \lambda\delta), \tag{6.32}$$

where λ is the constant from the regularity condition (2.9). Furthermore, (6.30) holds for some δ_0 and some $\kappa_1 > 0$.

Proof: By (6.21),

$$\frac{1}{\sqrt{n}}\,\mathsf{E}\,D_n(\mathbf{u}) - \frac{1}{2}\sum_{i=1}^n h_i(0)\,\xi_i^2(\mathbf{u}) = \frac{1}{n}\sum_{i=1}^n \left[\mathsf{E}\,R(Y_i - \varphi_i,\, \xi_i(\mathbf{u})) - \frac{1}{2}h_i(0)\xi_i^2(\mathbf{u})\right],$$

where $\|\xi_i(\mathbf{u})\| \leq \lambda\delta$. Relation (6.32) follows from here and from Proposition 6.2. To complete the proof we note that by (6.28) and $(a-b)^2 \leq 2a^2 + 2b^2$ it holds

$$\xi_i^2(\mathbf{u}) \geq (\dot{\varphi}_i(\mathbf{u}))^2 - \frac{1}{2}\left(\frac{\lambda}{2}\,\|\mathbf{u}\|^2\right)^2.$$

□

Proof of Theorem 2.2 Clear from Propositions 6.5-6.7. □

Introduce

$$\tilde{D}_n(\mathbf{v}) = \sqrt{n}\, D_n(\mathbf{v}/\sqrt{n}), \quad \widetilde{\mathcal{L}}_n(\mathbf{v}) = \sqrt{n}\, \mathcal{L}_n(\mathbf{v}/\sqrt{n})$$

and, similarly, also $\tilde{\mathcal{D}}_n(\mathbf{v})$, $\tilde{\mathcal{R}}_n(\mathbf{v})$ and $\widetilde{\mathcal{S}}_n(\mathbf{v}) = \tilde{\mathcal{R}}_n(\mathbf{v}) - \mathsf{E}\tilde{\mathcal{R}}_n(\mathbf{v})$, for $\mathbf{v} \in B_r$ and all sufficiently large n.

Proposition 6.8 *If all assumptions of Proposition 6.5 hold then for every closed ball B_r,*

$$\lim_{n\to\infty} \mathsf{E} \sup_{\mathbf{v}\in B_r} |\tilde{\mathcal{D}}_n(\mathbf{v})| = \lim_{n\to\infty} \mathsf{E} \sup_{\mathbf{v}\in B_r} |\widetilde{\mathcal{S}}_n(\mathbf{v})| = 0. \tag{6.33}$$

Consequently,

$$\sup_{\mathbf{v}\in B_r} \left| \tilde{D}_n(\mathbf{v}) - \mathsf{E}\, \tilde{D}_n(\mathbf{v}) - \widetilde{\mathcal{L}}_n(\mathbf{v}) \right| \xrightarrow{P} 0 \quad \text{as} \quad n \to \infty. \tag{6.34}$$

Proof: By Proposition 6.5, for all $r > 0$

$$\mathsf{E} \sup_{\mathbf{v}\in B_r} |\tilde{\mathcal{D}}_n(\mathbf{v})| \le \sqrt{n}\, c_1 \left(\frac{r}{\sqrt{n}}\right)^2 \quad \text{and} \quad \mathsf{E} \sup_{\mathbf{v}\in B_r} |\widetilde{\mathcal{S}}_n(\mathbf{v})| \le \sqrt{n}\, c_2 \left(\frac{r}{\sqrt{n}}\right)^{1+\alpha q/2}.$$

(6.33) is clear from here. □

In the following lemma we consider

$$Z = (Z_1, \ldots, Z_m)' \sim N(0, \Sigma), \tag{6.35}$$

where Σ is the matrix defined by (2.13).

Proposition 6.9 *If the assumptions of Theorem 2.3 hold then for every $r > 0$, the distribution of the process $(\widetilde{\mathcal{L}}_n(\mathbf{v}) : \mathbf{v} \in B_r)$ tends weakly to the distribution of $(\mathbf{v}'Z : \mathbf{v} \in B_r)$.*

Proof: For a fixed $\mathbf{v} \in B_r$, $\mathbf{v}'\Sigma_n\mathbf{v}$ is the variance of the vector $\widetilde{\mathcal{L}}_n(\mathbf{v})$, where Σ_n is defined in (2.13). By (2.17),

$$\widetilde{\mathcal{L}}_n(\mathbf{v}) \xrightarrow{\mathcal{L}} N(0, \mathbf{v}'\Sigma\mathbf{v}) \quad \text{as} \quad n \to \infty.$$

The stated convergence follows from the fact that $\widetilde{\mathcal{L}}_n(\mathbf{v})$ is linear in $\mathbf{v}$. □

In the next lemma and its proof, we consider the matrices Φ and Φ_n defined in (2.14) and the random vector Z defined by (6.35).

Proposition 6.10 *If the assumptions of Proposition 6.7 hold then for every closed ball B_r,*

$$\lim_{n\to\infty} \sup_{\mathbf{v}\in B_r} \left| \mathsf{E}\, \tilde{D}_n(\mathbf{v}) - \frac{1}{2}\, \mathbf{v}'\Phi\, \mathbf{v} \right| = 0 \tag{6.36}$$

and the process

$$\tilde{D}(\mathbf{v}) = \frac{1}{2}\,\mathbf{v}'\Phi\,\mathbf{v} - \mathbf{v}'Z\,, \quad \mathbf{v} \in \mathbb{R}^m, \tag{6.37}$$

is minimized at the unique $\Phi^{-1}Z$, *i.e.*

$$\Phi^{-1}Z = arg\ min_{\mathbf{v}\in\mathbb{R}^m}\ \tilde{D}(\mathbf{v}). \tag{6.38}$$

Proof: By Proposition 6.7, for every B_r under consideration

$$\sup_{\mathbf{v}\in B_r}\left|\mathsf{E}\,\tilde{D}_n(\mathbf{v}) - \frac{1}{2}\,\mathbf{v}'\Phi_n\mathbf{v}\right| \le \frac{(\lambda r)^2}{2n}\sum_{i=1}^{n}\omega(h_i, r/\sqrt{n})$$

and, by (2.12), the right hand side tends to zero as $n \to \infty$. The relation (6.37) follows from the easily verifiable formula

$$\tilde{D}(\mathbf{v}) = \frac{1}{2}\left\|\Phi^{1/2}\mathbf{v} - \Phi^{-1/2}Z\right\| - Z'\Phi Z,$$

where $\Phi^{1/2}$ is the symmetric root of the matrix Φ. □

Proof of Theorem 2.3. Define a random sequence

$$\widehat{\mathbf{v}}_n = \sqrt{n}(\widehat{\theta}_n - \theta_0), \quad n \in \mathbb{N}.$$

By definition of $\tilde{D}_n(\mathbf{v})$, for each $n \in \mathbb{N}$,

$$\widetilde{\mathbf{v}}_n = \arg\min_{\mathbf{v}\in\mathbb{R}^m}\ \tilde{D}_n(\mathbf{v}). \tag{6.39}$$

By Propsition 6.7, $\widehat{\theta}_n$ is $\sqrt{n}$-consistent, so that the sequence of distributions of $\widetilde{\mathbf{v}}_n$ is tight. By (6.34) and (6.36), for every closed ball B_r, the distribution of the process $(\tilde{D}_n(v) : v \in B_r)$ converges weakly to the distribution of $(\tilde{D}(v) : v \in B_r)$ defined by (6.37) and satisfying (6.38). By the *argmax continuous mapping* Theorem 3.2.2 of [30], this implies

$$\hat{v}_n \xrightarrow{\mathcal{L}} \Phi^{-1}Z = N(0, \Phi^{-1}\Sigma\Phi) \quad \text{as } n \to \infty,$$

which proves (1.12) and (2.18). □

References

[1] **Arcones, M. A. :** *M-estimators converging to a stable limit.* J. Multivariate Anal. 74, 193–221 (2000)

[2] **Arcones, M. A. :** *Asymptotic distribution of regression M-estimators.* J. Statist. Plann. Inference 97, 235–261 (2001)

[3] **Brown, L. D. :** *Fundamentals of Statistical Exponential Families.* IMS Lecture Notes Vol. 9 (1986)

[4] **Dodge, Y. :** *An introduction to statistical data analysis L_1-norm based.* In: *Statistical Data Analysis Based on the L_1 norm and Related Methods*, (Y. Dodge, ed.), Amsterdam, North Holland, pp. 1–22 (1988)

[5] **Fahrmeier, L.**, and **Kaufmann, H. :** *Consistency and asymptotic normality of the maximum likelihood estimator in generalized linear models .* Ann. Statist. 13, 342–368 (1985)

[6] **Farenbrother, R. W. :** *The historical development of the L_1 and L_∞ estimation procedures.* In: *Statistical Data Analysis Based on the L_1-norm and Related Methods*, (Y. Dodge, ed.), Amsterdam, North Holland, pp. 37–64 (1987)

[7] **Hallin, M.**, and **Jurečková, J. :** *Optimal tests for autoregressive models based on autoregression rank scores.* Ann. Statist. 27, 1385–1414 (1999)

[8] **Hampel, F. R., Rousseeuw, P. J., Ronchetti, E. M.**, and **Stahel, W. A. :** *Robust Statistics. The Approach Based on Influence Functions.* Wiley, New York (1986)

[9] **Hewitt, H.**, and **Stromberg, K. :** *Real and Abstract Analysis.* Springer–Verlag, Berlin, Heidelberg, New York (1965)

[10] **Hjort, N. L.**, and **Pollard, D. :** *Asymptotic for Minimizers of Convex Processes.* Statistical Research Report, University of Oslo (1993)

[11] **Huber, P. J. :** *Robust estimation of a location parameter.* Ann. Math. Statist. 35, 73–101 (1964)

[12] **Huber, P. J. :** *Robust Statistics.* Wiley, New York (1981)

[13] **Jurečková, J.**, and **Procházka, B. :** *Regression quantiles and trimmed squares estimator in nonlinear regression model.* Nonparametric Statistics 3, 201–222 (1994)

[14] **Jurečková, J.**, and **Sen, P. K. :** *Robust Statistical Inference: Asymptotics and Interrelations.* Wiley, New York (1996)

[15] **Knight, K. :** *Limiting Distributions for L_1-regression estimators under general conditions.* Ann. Statist. 26, 755–770 (1998)

[16] **Koenker, R.**, and **Basset, G. :** *Regression quantiles.* Econometrica 46, 33–500 (1978)

[17] **Koul, H. L.**, and **Saleh, A. K. E. :** *Autoregression quantiles and related rank scores processes.* Ann. Statist. 23, 670–689 (1995)

[18] **Liese, F.**, and **Vajda, I. :** *Consistency of M-estimates in general regression models.* J. Multivariate Anal. 50, 93–114 (1994)

[19] **Liese, F.**, and **Vajda, I. :** *Necessary and sufficient conditions for consistency of generalized M-estimates.* Metrika 42, 291–324 (1995)

[20] **Liese, F.**, and **Vajda, I. :** *M-estimators of structural parameters in pseudolinear models.* Appl. Statist. 20, 1514–1533 (1999)

[21] **Liese, F.**, and **Vajda, I. :** *A general asymptotic theory of M-estimators.* Math. Methods Statist. (submitted) (2002)

[22] **Pollard, D. :** *Asymptotics for least absolute deviation regression estimates.* Econometric Theory 7, 186–199 (1991)

[23] **Portnoy, S. :** *Robust estimation in dependent situations.* Ann. Statist. 5, 22–43 (1977)

[24] **Portnoy, S. :** *Asymptotic behavior of regression quantiles in nonstationary cases.* J. Multivariate Anal. 38, 100–113 (1991)

[25] **Rao, C. R. :** *Linear Statistical Inference and Its Appliations.* 2nd ed., Wiley, New York (1973)

[26] **Ronchetti, E. :** *Bounded influence in regression: A review.* In: Statistical Data Analysis Based on the L_1 norm and Related Methods, (Y. Dodge, ed.), Amsterdam, North Holland, pp. 65–80 (1988)

[27] **Serfling, R. J. :** *Approximation Theorems of Mathematical Statistics.* Wiley, New York (1980)

[28] **Sen, P. K.**, and **Singer, J. M. :** *Large Sample Methods in Statistics.* Chapman & Hall, London (1993)

[29] **Vajda, I. :** *Efficiency and robustness control via distorted maximum likelihood estimation.* Kybernetika 22, 47–67 (1986)

[30] **van der Vaart, A.**, and **Wellner, J. A. :** *Weak Convergence and Empirical Processes.* Springer, New York (1996)

[31] **Yohai, V. J.**, and **Maronna, R. A. :** *Asymptotic behaviour of M-estimates for the linear model.* Ann. Statist. 7, 258–268 (1979)

[32] **Zhao, L. C., Rao, C. R.**, and **Chen, X. :** *A note on the consistency of M-estimates in linear models.* Collection: Stochastic Processes, 359–367, Springer, New York (1993)

[33] **Zhao, L. C. :** *Some contributions to M-estimators in linear models.* J. of Statist. Planning and Inference 88, 189-203 (2000)

received: September 16, 2002

Authors:

Prof. Dr. Friedrich Liese
Universität Rostock
Fachbereich Mathematik
18051 Rostock
Germany

e-mail: friedrich.liese@mathematik.uni-rostock.de

Prof. Dr. Igor Vajda
Institute of Information Theory and Automation
Academy of Sciences of the Czech Republic
Pod vodàrenskou věži 4
CZ-182 08 Praha
Czech Republic

e-mail: vajda@utia.cas.cz

Rostock. Math. Kolloq. **57**, 53–70 (2003)

Subject Classification (AMS)
35K55,35B40,35B65

Fouzi Zaouch

Global Existence and Boundedness of Solutions of the Time-Dependent Ginzburg-Landau Equations with a Time-Dependent Magnetic Field

ABSTRACT. This paper is concerned with existence, uniqueness and long-time asymptotic behavior of the solutions of the time-dependent Ginzburg-Landau equations of superconductivity, in the case where the applied magnetic field $\mathbf{H}$ is time-dependent. We first prove existence and uniqueness of solutions with H^1-initial data. This result is obtained under the "$\phi = -\omega(\nabla \cdot \mathbf{A})$" gauge with $\omega > 0$. These solutions become then uniformly bounded in time for the H^1-norm, by assuming time-uniform boundedness on $\mathbf{H}$ and its time derivative.

KEY WORDS. Superconductivity, Ginzburg-Landau equation, gauge, initial boundary value problems, global existence and uniqueness.

1 Introduction

In this paper we consider the Ginzburg-Landau model for superconductivity in the nonstationary case. Based on an averaging method of the BCS theory, a time-dependent Ginzburg-Landau model was derived by Gor'kov and Eliashberg in 1968 [1]. The study of this model for superconductivity may give a better understanding of the physical state of a superconductor, especially for the high-temperature superconductors. It is known from the physics literature that the realization of this physical phenomena and then the validation of this model is only possible under temperatures near the critical temperature. The equations describing the state of a superconducting material near the critical temperature are nonlinear differential equations for the *order-parameter* ψ, the *vector potential* $\mathbf{A}$ and the *electric potential* ϕ, whose evolution in presence of a *magnetic field* $\mathbf{H}$ is governed by the following system

$$\eta\left(\frac{\partial}{\partial t} + i\kappa\phi\right)\psi = -\left(\frac{i}{\kappa}\nabla + \mathbf{A}\right)^2\psi + \left(1-|\psi|^2\right)\psi \quad \text{in} \quad \Omega\times(0,\infty), \tag{1.1}$$

$$\frac{\partial \mathbf{A}}{\partial t} + \nabla\phi = -\nabla\times\nabla\times\mathbf{A} + \mathbf{J}_s + \nabla\times\mathbf{H} \quad \text{in} \quad \Omega\times(0,\infty), \tag{1.2}$$

where $\mathbf{J}_s$ is given by

$$\mathbf{J}_s \equiv \mathbf{J}_s(\psi, \mathbf{A}) = \frac{1}{2i\kappa}(\psi^*\nabla\psi - \psi\nabla\psi^*) - |\psi|^2\mathbf{A} = -\mathrm{Re}\left[\psi^*\left(\frac{i}{\kappa}\nabla + \mathbf{A}\right)\psi\right]. \tag{1.3}$$

Equations (1.1)-(1.3) are satisfied everywhere in a domain Ω, which is the region occupied by the superconducting material and at all times $t > 0$. The associated boundary conditions are

$$\mathbf{n}\cdot\left(\frac{i}{\kappa}\nabla + \mathbf{A}\right)\psi + \frac{i}{\kappa}\gamma\psi = 0 \quad \text{and} \quad \mathbf{n}\times(\nabla\times\mathbf{A} - \mathbf{H}) = \mathbf{0} \quad \text{on} \quad \partial\Omega, \tag{1.4}$$

where $\partial\Omega$ is the boundary of Ω and $\mathbf{n}$ the local outer unit normal to $\partial\Omega$. They must be satisfied at all times $t > 0$. Henceforth, the term „TDGL Equations“ refers to the system of equations (1.1)-(1.4).

We assume that Ω is a bounded domain in $\mathbb{R}^n$ ($n = 2$ or 3) with a boundary $\partial\Omega$ of class $C^{1,1}$. The parameters appearing in the TDGL equations are dimensionless physical constants; η is the friction coefficient and κ is the Ginzburg-Landau parameter. Here η measures the temporal rate of change and the value of κ determines the type of superconductor: $\kappa \leq 1/\sqrt{2}$ describes what is known as a *type I* superconductor and $\kappa \geq 1/\sqrt{2}$ as a *type II*. The function γ is defined and Lipschitz continuous on $\partial\Omega$ and $\gamma(x) \geq 0$ for $x \in \partial\Omega$. We use the following common notation: $\nabla \equiv$ grad, $\nabla\cdot \equiv$ div, $\nabla\times \equiv$ curl and $\nabla^2 = \nabla\cdot\nabla \equiv \Delta$, i is the imaginary unit and a superscript* denotes the complex conjugation.

The order parameter ψ is a complex-valued function, it describes the center-of-mass motion of the "superelectron", whose density is $n_s = |\psi|^2$ and whose flux is $\mathbf{J}_s$. $\psi = 0$ corresponds to the normal state, and in a perfect superconducting state $|\psi| = 1$. The vector potential $\mathbf{A}$ takes its values in $\mathbb{R}^n$, it represents the *magnetic potential*, i.e. $\mathbf{B} = \nabla\times\mathbf{A}$. The scalar potential ϕ determines the *electric field* $\mathbf{E} = -\frac{\partial\mathbf{A}}{\partial t} - \nabla\phi$. The vector $\mathbf{H}$ represents the (externally) *applied magnetic field*; it is a given function of space and time, which is divergence free, $\nabla\cdot\mathbf{H} = 0$ at all time. The difference $\mathbf{M} = \mathbf{B} - \mathbf{H}$ is known as the *magnetisation*. The trivial solution ($\psi = 0$, $\mathbf{B} = \mathbf{H}$, $\mathbf{E} = 0$) represents the normal state, where all superconducting properties have been lost. For further physics details about Ginzburg-Landau equations, one may consult [1] or [2].

Several works have been devoted recently to questions of existence, uniqueness and long time asymptotic behavior of the solutions of equations (1.1)-(1.4) when the applied magnetic field is stationary, i.e. $\mathbf{H}(t) = \mathbf{H}_0$; as a bibliographical review, we refer to [3], [4], [5], [6], [7] and [8]. To overcome the uniqueness deficiency in equations (1.1)-(1.4), the authors in the mentioned references adopted some *gauge* transformation like the zero-electric gauge ($\phi = 0$), the London gauge ($\nabla\cdot\mathbf{A} = 0$) or the Lorentz gauge ($\phi = -\nabla\cdot\mathbf{A}$). On the other hand, it is known that in presence of an applied time-independent magnetic field $\mathbf{H}$, the TDGL equations enjoy the free energy functional, whose advantage is the getting of some estimates on the solutions.

In contrast to the above situation, Fleckinger, Kaper and Takáč considered in [9] equations (1.1)-(1.4) with a time-dependent magnetic field $\mathbf{H}(t)$. They established in the general context of "$\phi = -\omega(\nabla \cdot \mathbf{A})$" gauge ($\omega > 0$) the existence of a dynamical process. However, some regularities of the solutions obtained are lost in the limit case $\omega = 0$. When $\mathbf{H}$ is stationary, this process becomes a dynamical system enjoying the existence of a global attractor. Subsequently Kaper and Takáč [11] proved that in the special case where the applied magnetic field is asymptotically stationary, the dynamical process generated by the TDGL equations is asymptotically autonomous, i.e. its large-time asymptotic limit is a dynamical system, whose attractor coincides with the one of the dynamical process.

In this paper, we present new, more general results concerning existence, uniqueness and regularity of solutions to the TDGL equations when the applied magnetic field $\mathbf{H}$ exhibits strong temporal fluctuations. In practice $\mathbf{H}$ is either time-independent or time-periodic. For instance, we are able to show global existence for all times $t \geq 0$ if $\mathbf{H}$ is time-periodic. The Lyapunov functional method applied in [9], [10] and [11] is not suitable for treating other than weak temporal fluctuations that disappear for large time with certain convergence rate. Our method of proving global existence and boundedness of solutions for all times $t \geq 0$ significantly improves and extends the classical Lyapunov functional method. Our discussion will rely on the choice of the "$\phi = -\omega(\nabla \cdot \mathbf{A})$" gauge ($\omega > 0$), introduced in [10]. We omit the degenerate case $\omega = 0$. The outline of the paper is as follows. In section 2, we introduce preliminary material, gauge invariance among others, and recall basic results for use in subsequent sections. In section3, we first homogenize the boundary conditions, give definitions of the function spaces we are going to use and assumptions on the data, and after we reformulate the problem into an equivalent abstract initial value problem. Section 4 contains results concerning existence, uniqueness and regularity of solutions to the original equations, the proof of local existence is based on the contraction mapping principle, while global existence is derived from estimates on the energy type functional. In our existence result, we obtain solutions of the TDGL equations from H^1-initial data and without requiring L^∞-bound of the initial order parameter ψ_0. In section 5, we establish that the solutions obtained become uniformly bounded with respect to $t \geq 0$, this will lead to the existence of an absorbing set for the process.

2 Preliminaries

The TDGL equations are not mathematically well posed unless some gauge fixing has been done. It is known in [12] that the solutions of equations (1.1)-(1.4) are unique up to a gauge transformation $\mathcal{G}_\chi$

$$\mathcal{G}_\chi : (\psi,\ \mathbf{A},\ \phi) \longrightarrow \left(\psi \mathrm{e}^{i\kappa\chi},\ \mathbf{A} + \nabla\chi,\ \phi - \frac{\partial \chi}{\partial t}\right),$$

here χ is a given real-valued function (sufficiently smooth) of position and time. In our investigation we adopt the "$\phi = -\omega(\nabla \cdot \mathbf{A})$" gauge. We restrict ourselves to the case $\omega > 0$. Formaly we determine this gauge by taking $\chi \equiv \chi_\omega(x,t)$ as a solution of the following boundary value problem

$$\begin{aligned} \frac{\partial \chi}{\partial t} - \omega\Delta\chi &= \phi + \omega(\nabla \cdot \mathbf{A}) \quad \text{in} \quad \Omega \times (0,\infty), \\ \mathbf{n} \cdot (\nabla\chi) &= -\mathbf{n} \cdot \mathbf{A} \quad \text{on} \quad \partial\Omega \times (0,\infty). \end{aligned}$$

The initial condition $\chi(\cdot,0) = \chi_0$ can be chosen arbitrarily. By virtue of the current gauge, $\mathbf{A}$ and ϕ satisfy the identities

$$\phi + \omega(\nabla \cdot \mathbf{A}) = 0 \quad \text{in} \quad \Omega \times (0,\infty), \tag{2.1}$$

$$\mathbf{n} \cdot \mathbf{A} = 0 \quad \text{on} \quad \partial\Omega \times (0,\infty). \tag{2.2}$$

On the other hand the TDGL equations may be given as

$$\eta\frac{\partial \psi}{\partial t} = -\left(\frac{i}{\kappa}\nabla + \mathbf{A}\right)^2 \psi + i\eta\kappa\omega\psi(\nabla \cdot \mathbf{A}) + \left(1 - |\psi|^2\right)\psi \text{ in } \Omega \times (0,\infty), \tag{2.3}$$

$$\frac{\partial \mathbf{A}}{\partial t} = -\nabla \times \nabla \times \mathbf{A} + \omega\nabla(\nabla \cdot \mathbf{A}) + \mathbf{J}_s + \nabla \times \mathbf{H} \quad \text{in} \quad \Omega \times (0,\infty), \tag{2.4}$$

where $\mathbf{J}_s$ is given by (1.3) and the boundary conditions become

$$\mathbf{n} \cdot \nabla\psi + \gamma\psi = 0,\ \mathbf{n} \cdot \mathbf{A} = 0,\ \mathbf{n} \times (\nabla \times \mathbf{A} - \mathbf{H}) = \mathbf{0} \text{ on } \partial\Omega \times (0,\infty). \tag{2.5}$$

For the initial condition, we put

$$\psi(\cdot,0) = \psi_0 \quad \text{and} \quad \mathbf{A}(\cdot,0) = \mathbf{A}_0 \ \text{ in } \ \Omega, \tag{2.6}$$

where ψ_0 and $\mathbf{A}_0$ are given.

Now we introduce notations conventions concerning functional spaces, in order to reformulate the gauged TDGL equations (2.3)-(2.6) as an abstract evolution equation in a real Banach space. Throughout, for $p \geq 1$, $L^p(\Omega)$ will denote the usual Lebesgue space, with the norm $\|\cdot\|_p$, $(\cdot,\cdot)$ is the usual inner-product in $L^2(\Omega)$. For nonnegative integer m, we will denote by $H^m(\Omega)$ the usual Sobolev space, with norm $\|\cdot\|_{H^m}$. In the case of nonintegers m, $H^m(\Omega)$ is the fractional Sobolev space defined by interpolation, see [12]. The corresponding spaces of complex-valued functions will be denoted by $\mathcal{L}^p(\Omega)$ and $\mathcal{H}^m(\Omega)$ and the corresponding spaces of vector valued functions will be denoted by $\mathbf{L}^p(\Omega)$ and $\mathbf{H}^m(\Omega)$. Without any possible ambiguity, we use the same symbol $\|\cdot\|_p$ to indicate the norms in $\mathcal{L}^p(\Omega)$ and $\mathbf{L}^p(\Omega)$, and the inner-product for $p=2$ is defined in the usual way. We sometimes use $\|\cdot\|_X$ to denote the norm defined on a Banach space X. To fix the time-dependence of the functions entering equations (2.3)-(2.5), we define the following spaces: For any given $T>0$, $p\geq 1$ and any given Banach space X,

$$L^p(0,T;X) = \left\{ u : t\in(0,T)\to u(\cdot,t)\in X \text{ measurable, and } \int_0^T \|u(\cdot,t)\|_X^p \,\mathrm{d}t < \infty \right\},$$

$$L^\infty(0,T;X) = \left\{ u : t\in(0,T)\to u(\cdot,t)\in X \text{ measurable, and } \sup_{0<t<T} \|u(\cdot,t)\|_X < \infty \right\},$$

$$W^{1,p}(0,T;X) = \left\{ u\in L^p(0,T;X) \text{ absolutely continuous } : \frac{\partial u}{\partial t}\in L^p(0,T;X) \right\}.$$

The spaces $W^{m,p}(0,T;X)$ are defined in similar ways. $C(0,T;X)$ denotes the space of continuously X-valued functions defined in $[0,T]$.

For later purpose we recall some known inequalities and formulas concering vector-valued functions, details and proofs are contained in [13] and [14].

Poincaré inequality: For all $\mathbf{A}\in\mathbf{H}^1(\Omega)$, with $\mathbf{n}\cdot\mathbf{A}=0$ on $\partial\Omega$

$$\lambda_0\|\mathbf{A}\|_{\mathbf{H}^1}^2 \leq \|\nabla\times\mathbf{A}\|_2^2 + \|\nabla\cdot\mathbf{A}\|_2^2, \tag{2.7}$$

λ_0 is a positive constant.

Green′s formulas:

(i) For any $\mathbf{A}\in\mathbf{H}(\mathrm{div};\Omega) := \{\mathbf{A}\in\mathbf{L}^2(\Omega) : \nabla\cdot\mathbf{A}\in L^2(\Omega)\}$ and $\varphi\in H^1(\Omega)$

$$\int_\Omega (\nabla\cdot\mathbf{A})\varphi\,\mathrm{d}x + \int_\Omega \mathbf{A}\cdot(\nabla\varphi)\,\mathrm{d}x = \int_{\partial\Omega} (\mathbf{n}\cdot\mathbf{A})\varphi\,\mathrm{d}\sigma(x). \tag{2.8}$$

(ii) For any $\mathbf{A}\in\mathbf{H}(\mathrm{curl};\Omega) := \{\mathbf{A}\in\mathbf{L}^2(\Omega) : \nabla\times\mathbf{A}\in\mathbf{L}^2(\Omega)\}$ and $\mathbf{B}\in\mathbf{H}^1(\Omega)$

$$\int_\Omega (\nabla\times\mathbf{A})\cdot\mathbf{B}\,\mathrm{d}x - \int_\Omega \mathbf{A}\cdot(\nabla\times\mathbf{B})\,\mathrm{d}x = \int_{\partial\Omega} \mathbf{B}\cdot(\mathbf{A}\times\mathbf{n})\,\mathrm{d}\sigma(x). \tag{2.9}$$

Gronwall′s inequality: Let $\eta(t)$ be a positive, absolutely continuous function on $[0,T]$, $T>0$, satisfying $\eta'(t)\leq\mu(t)\eta(t)+\nu(t)$ a.e. $t\in[0,T]$, where μ and ν are integrable on

$[0, T]$, then

$$\eta(t) \leq \mathrm{e}^{\int_0^t \mu(s)\mathrm{d}s} \left[\eta(0) + \int_0^t \mathrm{e}^{-\int_0^s \mu(r)\mathrm{d}r} \nu(s)\ \mathrm{d}s \right] \quad \text{for all} \quad t \in [0, T]. \tag{2.10}$$

3 Abstract Equation

Before we start to reformulate the gauged TDGL equations (2.3)-(2.6) into an equivalent abstract initial-value problem, we turn the boundary condition in the right hand side of (2.5) into a homogenous one. This is achieved at each fixed instant. At each time t, assume $\mathbf{H} \in \mathbf{L}^2(\Omega)$ and consider $\mathbf{A_H}$ the unique weak solution of the strongly elliptic boundary-value problem

$$\nabla \cdot \mathbf{A_H} = 0 \qquad \text{and} \qquad \nabla \times \nabla \times \mathbf{A_H} = \nabla \times \mathbf{H} \quad \text{in} \quad \Omega, \tag{3.1}$$

$$\mathbf{n} \cdot \mathbf{A_H} = 0 \qquad \text{and} \qquad \mathbf{n} \times (\nabla \times \mathbf{A_H} - \mathbf{H}) = 0 \quad \text{on} \quad \partial\Omega. \tag{3.2}$$

The existence of $\mathbf{A_H}$ is guaranted by the Lax-Milgram theorem applied to the continuous and coercive bilinear form

$$Q(\mathbf{A}, \mathbf{B}) = \int_\Omega (\nabla \times \mathbf{A}) \cdot (\nabla \times \mathbf{B})\ \mathrm{d}x + \omega \int_\Omega (\nabla \cdot \mathbf{A})(\nabla \cdot \mathbf{B})\ \mathrm{d}x,$$

on the space $\{\mathbf{A} \in \mathbf{H}^1(\Omega) : \mathbf{n} \cdot \mathbf{A} = 0 \text{ on } \partial\Omega\}$.

The mapping $\mathbf{H} \in \mathbf{L}^2(\Omega) \longmapsto \mathbf{A_H} \in \mathbf{H}^1(\Omega)$ is linear, time independent and continuous, see [9].

The gauged TDGL equations (2.3)-(2.4) are equivalent to a problem in terms of ψ and the reduced vector potential $\tilde{\mathbf{A}} := \mathbf{A} - \mathbf{A_H}$

$$\eta \frac{\partial \psi}{\partial t} = -\left(\frac{i}{\kappa} \nabla + \tilde{\mathbf{A}} + \mathbf{A_H} \right)^2 \psi + i\eta\kappa\omega\psi(\nabla \cdot \tilde{\mathbf{A}}) + \left(1 - |\psi|^2\right) \psi \quad \text{in} \quad \Omega \times (0, \infty), \tag{3.3}$$

$$\frac{\partial \tilde{\mathbf{A}}}{\partial t} = -\nabla \times \nabla \times \tilde{\mathbf{A}} + \omega \nabla(\nabla \cdot \tilde{\mathbf{A}}) + \tilde{\mathbf{J}}_s - |\psi|^2 \mathbf{A_H} - \frac{\partial \mathbf{A_H}}{\partial t} \quad \text{in} \quad \Omega \times (0, \infty), \tag{3.4}$$

where $\tilde{\mathbf{J}}_s = \mathbf{J}_s(\psi, \tilde{\mathbf{A}})$ is given by the expression in (1.3), and the boundary condition (2.5) reduces to

$$\mathbf{n} \cdot \nabla\psi + \gamma\psi = 0,\ \ \mathbf{n} \cdot \tilde{\mathbf{A}} = 0 \quad \text{and} \quad \mathbf{n} \times (\nabla \times \tilde{\mathbf{A}}) = \mathbf{0} \quad \text{on} \quad \partial\Omega \times (0, \infty). \tag{3.5}$$

The supplemented initial condition is

$$\psi(\cdot, 0) = \psi_0 \ \text{ and } \ \tilde{\mathbf{A}}(\cdot, 0) = \tilde{\mathbf{A}}_0 = \mathbf{A}_0 - \mathbf{A_H}(0) \ \text{ in } \ \Omega. \tag{3.6}$$

We come now to introduce a convenient abstract frame for the system of equations (3.3)-(3.6). In the sequel we will consider the solutions ψ and $\tilde{\mathbf{A}}$ of the system of equations (3.3)-(3.6) as a vector representing the pair

$$\tilde{u} = (\psi, \tilde{\mathbf{A}}) = (\psi, \mathbf{A} - \mathbf{A_H}), \tag{3.7}$$

so we adopt the notations

$$\mathbb{L}^p(\Omega) = \mathcal{L}^p(\Omega) \times \mathbf{L}^p(\Omega) \quad \text{and} \quad \mathbb{H}^s(\Omega) = \mathcal{H}^s(\Omega) \times \mathbf{H}^s(\Omega),$$

and indicate, without any possible confusion, the norm in $\mathbb{L}^p(\Omega)$ by $\|\cdot\|_p$. We set $X = \mathbb{L}^2(\Omega)$ and define some suitable operators related to the dissipative terms in (3.3) and (3.4), we define two linear operators L_1 and L_2 respectively from $\mathcal{H}^1(\Omega)$ and $\mathbf{H}^1(\Omega)$ to their dual spaces by

$$(L_1\psi, \phi) = \int_\Omega \nabla\psi \cdot \nabla\phi^* \, \mathrm{d}x + \int_{\partial\Omega} \gamma\psi\phi^* \, \mathrm{d}\sigma(x), \tag{3.8}$$

$$(L_2\mathbf{A}, \mathbf{B}) = \int_\Omega (\nabla \times \mathbf{A}) \cdot (\nabla \times \mathbf{B}) \, \mathrm{d}x + \omega \int_\Omega (\nabla \cdot \mathbf{A})(\nabla \cdot \mathbf{B}) \, \mathrm{d}x. \tag{3.9}$$

Operators L_1 and L_2 are selfadjoint and positive definite. Moreover the classical theory of second order differential operators allows the extension of L_1 and L_2 as unbounded linear selfadjoint operators respectively on $\mathcal{L}^2(\Omega)$ and $\mathbf{L}^2(\Omega)$, in which case $L_1\psi = -\Delta\psi$ and $L_2\mathbf{A} = \nabla \times \nabla \times \mathbf{A} - \omega\nabla(\nabla \cdot \mathbf{A})$ in Ω, with

$$\begin{aligned} \mathcal{D}(L_1) &= \{\psi \in \mathcal{H}^2(\Omega) : \mathbf{n} \cdot \nabla\psi + \gamma\psi = 0 \ \text{ on } \ \partial\Omega\}, \\ \mathcal{D}(L_2) &= \{\mathbf{A} \in \mathbf{H}^2(\Omega) : \mathbf{n} \cdot \mathbf{A} = 0 \ \text{ on } \ \partial\Omega\}. \end{aligned}$$

Let $\mathcal{A}$ be the linear selfadjoint operator in X defined by

$$\begin{aligned} &\mathcal{D}(\mathcal{A}) = \mathcal{D}(L_1) \times \mathcal{D}(L_2), \\ &\mathcal{A}v = \left(-\frac{1}{\eta\kappa^2}\Delta\psi \ , \ \nabla \times \nabla \times \mathbf{A} - \omega\nabla(\nabla \cdot \mathbf{A}) \right), \ v = (\psi, \mathbf{A}) \in \mathcal{D}(\mathcal{A}) \, . \end{aligned} \tag{3.10}$$

Since $\mathcal{A}$ is positive definite on X, it is then a sectorial operator. It follows that $-\mathcal{A}$ is the infinitesimal generator of an holomorphic semigroup $(\mathrm{e}^{-\mathcal{A}t})_{t\geq 0}$, see [15] and [16], Fractional powers $\mathcal{A}^\alpha$ are well defined for $\alpha \in \mathbb{R}$, they are unbounded for $\alpha > 0$ and $X^\alpha := \mathcal{D}(\mathcal{A}^\alpha)$ is a closed linear subspace of $\mathbb{H}^{2\alpha}(\Omega)$ for $0 < \alpha < 1$ and contains the range of $\mathrm{e}^{-\mathcal{A}t}$ for $\alpha \geq 0$ and $t > 0$. In particular we have

$$X^{1/2} = \{v = (\psi, \mathbf{A}) \in \mathbb{H}^1(\Omega) : \mathbf{n} \cdot \mathbf{A} = 0 \ \text{ on } \ \partial\Omega\}. \tag{3.11}$$

In general it is possible to consider $\mathcal{A}$ as a linear operator in $\mathbb{L}^p(\Omega)$ with $1 < p < \infty$, we will use the same symbol $\mathcal{A}$ if no confusion is possible. In this case the L^p-theory for elliptic differential operators proves that $-\mathcal{A}$ generates an holomorphic semigroup $(\mathrm{e}^{-\mathcal{A}t})_{t\geq 0}$ in $\mathbb{L}^p(\Omega)$.

On the other hand we consider the initial value problem for the transformed solution $\tilde{u} = (\psi, \tilde{\mathbf{A}})$

$$\frac{\mathrm{d}\tilde{u}}{\mathrm{d}t} + \mathcal{A}\tilde{u} = \mathcal{F}(t, \tilde{u}(t)) \quad \text{for} \quad t > 0 \quad \text{and} \quad \tilde{u}(0) = \tilde{u}_0, \tag{3.12}$$

in X, where $\mathcal{F}(t, \tilde{u}) = (\varphi, \mathbf{F})$, $\tilde{u}_0 = (\psi_0, \tilde{\mathbf{A}}_0)$, φ and $\mathbf{F}$ are given by the following

$$\begin{aligned}\varphi \equiv \varphi(t, \psi, \tilde{\mathbf{A}}) = \frac{1}{\eta}\Big[-\frac{2i}{\kappa}(\nabla\psi)\cdot(\tilde{\mathbf{A}} + \mathbf{A_H}) - \frac{i}{\kappa}(1 - \eta\kappa^2\omega)\psi(\nabla\cdot\tilde{\mathbf{A}}) \\ -\psi|\tilde{\mathbf{A}} + \mathbf{A_H}|^2 + \left(1 - |\psi|^2\right)\psi\Big],\end{aligned} \tag{3.13}$$

$$\mathbf{F} \equiv \mathbf{F}(t, \psi, \tilde{\mathbf{A}}) = \tilde{\mathbf{J}}_s - |\psi|^2\mathbf{A_H} - \frac{\partial \mathbf{A_H}}{\partial t}. \tag{3.14}$$

Let $\tilde{u}_0 \in \mathbb{H}^1(\Omega)$, we say that $\tilde{u}$ is a *mild solution* of equation (3.12) on the interval $[0, T]$, for some $T \in (0, \infty)$, if $\tilde{u} : [0, T] \longrightarrow \mathbb{H}^1(\Omega)$ is continuous and

$$\tilde{u}(t) = \mathrm{e}^{-\mathcal{A}t}\tilde{u}_0 + \int_0^t \mathrm{e}^{-\mathcal{A}(t-s)}\mathcal{F}(s, \tilde{u}(s))\,\mathrm{d}s \quad \text{for} \quad 0 \leq t \leq T. \tag{3.15}$$

In particular a mild solution plays the role of a weak solution $(\psi, \tilde{\mathbf{A}})$ for the system of equation (3.3)-(3.5). Of course the existence of a weak solution $u = (\psi, \mathbf{A})$ to the gauged TDGL equations (2.3)-(2.5) requires some regularity about $\mathbf{A_H}$; this suggests that some control should be imposed on the time-dependence of $\mathbf{H}$. Clearly, in definition (3.15) of mild solution, the action of the semigroup $(\mathrm{e}^{-\mathcal{A}t})$ on $\mathcal{F}$ is in $\mathbb{L}^{3/2}(\Omega)$, this is because $\mathcal{F}$ maps $[0, T] \times \mathbb{H}^1(\Omega)$ in $\mathbb{L}^{3/2}(\Omega)$, so it is to distinguish that the operator $\mathcal{A}$ appearing under the symbol integral in (3.15) is considered in $\mathbb{L}^{3/2}(\Omega)$. Furthermore we see that the regularity of the integral in (3.15) introduced by the term $\frac{\partial \mathbf{A_H}}{\partial t}$, namely $\int_0^t \mathrm{e}^{-L_2(t-s)}\frac{\partial \mathbf{A_H}}{\partial s}(s)\,\mathrm{d}s$, determines the regularity of the mild solution $\tilde{u}$ of equation (3.12).

4 Existence and Uniqueness

In this section, we study the existence and uniqueness of a mild solution of the initial value problem (3.12). We assume the applied magnetic field $\mathbf{H}(t)$ in $\mathbf{L}^2(\Omega)$ at each $t \geq 0$ and

$(\mathbf{H_0})$ $\quad \mathbf{H} \in L^\infty\left(0, T; \mathbf{L}^2(\Omega)\right) \cap W^{1,2}\left(0, T; \mathbf{L}^2(\Omega)\right), \quad 0 < T < \infty.$

Note that by virtue of [9], $(\mathbf{H}_0)$ implies

$$t \in [0, T] \longrightarrow \int_0^t \mathrm{e}^{-L_2(t-s)}\frac{\partial \mathbf{A_H}}{\partial t}(s)\,\mathrm{d}s \in \mathbf{H}^1(\Omega) \quad \text{is Hölder continuous.} \tag{4.1}$$

Theorem 1 *For every initial data* $\tilde{u}_0 = (\psi_0, \tilde{\mathbf{A}}_0) \in X^{1/2}$ *the initial value problem* (3.12) *has a unique mild solution* $\tilde{u} = (\psi, \tilde{\mathbf{A}})$ *such that*

$$\tilde{u} \in C\left(0, T; \mathbb{H}^1(\Omega)\right) \cap W^{1,2}\left(0, T; \mathbb{L}^2(\Omega)\right).$$

Proof: The proof of local existence and uniqueness is based on the contraction mapping principle. To this goal, we construct a Banach space $C\left(0, \tau; \mathbb{H}^1(\Omega)\right)$ (τ *small enough*) such that the mapping $\mathcal{G}$ defined from the integral equation in (3.15), namely

$$\mathcal{G}\tilde{u}(t) = \mathrm{e}^{-\mathcal{A}t}\tilde{u}_0 + \int_0^t \mathrm{e}^{-\mathcal{A}(t-s)}\mathcal{F}(s, \tilde{u}(s))\,\mathrm{d}s, \tag{4.2}$$

acts as a contraction map on some closed subset. We need to prove the following properties

$$\mathcal{F}(t,\cdot) : \mathbb{H}^1(\Omega) \longrightarrow \mathbb{L}^{3/2}(\Omega) \text{ is locally Lipschitz for each } t \in [0, T], \tag{4.3}$$

$$\mathrm{e}^{-\mathcal{A}t} : \mathbb{L}^{3/2}(\Omega) \longrightarrow \mathbb{H}^1(\Omega) \text{ for } t > 0 \text{ and } \int_0^\tau \|\mathrm{e}^{-\mathcal{A}t}\|_{\mathcal{L}(\mathbb{L}^{3/2},\mathbb{H}^1)}\,\mathrm{d}t < \infty. \tag{4.4}$$

Given (4.3) and (4.4), the standard proof of [15, theorem 3.3.3] can be used; we show that there are some positive constants τ and ε both small enough such that if we denote $\mathcal{X} = \left\{v \in C\left(0, \tau; X^{1/2}\right) : v(0) = \tilde{u}_0, \|v(t) - \tilde{u}_0\|_{\mathbb{H}^1} \le \varepsilon\right\}$, then $\mathcal{G} : \mathcal{X} \to \mathcal{X}$ is a contraction map and hence possesses a unique fixed point.

In order to establish (4.3), we need to estimate each term separately. Let two elements $\tilde{u}_1 = (\psi_1, \tilde{\mathbf{A}}_1)$ and $\tilde{u}_2 = (\psi_2, \tilde{\mathbf{A}}_2)$ of $\mathbb{H}^1(\Omega)$, we have for example

$$\begin{aligned}\|\nabla\psi_2 \cdot \tilde{\mathbf{A}}_2 - \nabla\psi_1 \cdot \tilde{\mathbf{A}}_1\|_{3/2} &\le \|\nabla(\psi_2 - \psi_1)\|_2\|\tilde{\mathbf{A}}_2\|_6 + \|\nabla\psi_1\|_2\|\tilde{\mathbf{A}}_2 - \tilde{\mathbf{A}}_1\|_6 \\ &\le C\|\tilde{u}_2 - \tilde{u}_1\|_{\mathbb{H}^1},\end{aligned}$$

where C is a positive constant depending only on the norm of $\tilde{u}_1$ and $\tilde{u}_2$ in $\mathbb{H}^1(\Omega)$. Here we have used the continuous Sobolev imbedding of $H^1(\Omega)$ in $L^6(\Omega)$. For the other terms in $\mathcal{F}$, we argue analogously. It follows that if B_R denotes the ball of radius R centered at the origin in $\mathbb{H}^1(\Omega)$, then

$$\|\mathcal{F}(t, \tilde{u}_1) - \mathcal{F}(t, \tilde{u}_2)\|_{3/2} \le C\|\tilde{u}_1 - \tilde{u}_2\|_{\mathbb{H}^1} \quad \text{for all} \quad \tilde{u}_1, \tilde{u}_2 \in B_R, \tag{4.5}$$

C is the Lipschitz constant, it depends on R but not on t.

The proof of the claim in (4.4) uses the smoothing action of the semigroup $\mathrm{e}^{-\mathcal{A}t}$ and some imbedding theorems established for second-order elliptic differential operators. More precisely, rather than (4.4) we can check

$$\|e^{-\mathcal{A}t}\|_{\mathcal{L}(\mathbb{L}^{3/2},\mathbb{H}^1)} \leq Ct^{-\gamma}e^{-\delta t} \quad \text{for all} \quad t > 0, \tag{4.6}$$

for some positive constants C, δ and $\gamma > 3/4$ independent on t. We refer to [14, theorem 1.6.1] for the proof of this, see also [16].

Next, to show the solution $\tilde{u} = (\psi, \tilde{\mathbf{A}})$ of equation (3.12) is global, some estimates on the energy type functional defined in $\mathbb{H}^1(\Omega)$ by

$$\begin{aligned} E_\omega[\psi, \mathbf{A}] = \int_\Omega \Big[\Big| \Big(\frac{i}{\kappa}\nabla + \mathbf{A} \Big) \psi \Big|^2 + \frac{1}{2}\left(1 - |\psi|^2\right)^2 + 2\omega(\nabla \cdot \mathbf{A})^2 \\ + |\nabla \times \mathbf{A} - \mathbf{H}|^2 \Big] \, dx + \frac{1}{\kappa^2}\int_{\partial\Omega} \gamma |\psi|^2 \, d\sigma(x), \end{aligned} \tag{4.7}$$

are needed. In fact, from the consideration on $\mathbf{H}$ stated in ($\mathbf{H_0}$), it can be shown (see [9]), that the pair $u = (\psi, \mathbf{A})$ related to $\tilde{u} = (\psi, \tilde{\mathbf{A}})$ by (3.7) satisfies

$$u \in L^\infty\left(0, T; \mathbb{H}^1(\Omega)\right) \cap W^{1,2}\left(0, T; \mathbb{L}^2(\Omega)\right) \quad \text{and} \quad \nabla \cdot \mathbf{A} \in L^2\left(0, T; \mathbb{H}^1(\Omega)\right).$$

Thus, again by ($\mathbf{H_0}$), we obtain

$$\tilde{u} \in L^\infty\left(0, T; \mathbb{H}^1(\Omega)\right) \cap W^{1,2}\left(0, T; \mathbb{L}^2(\Omega)\right).$$

However, this regularity result concerning $\tilde{u}$ can be improved by the smoothness of the action of $e^{-\mathcal{A}t}$ to prove continuity of $\tilde{u}$. In fact, we have as claimed in (4.1) the map $t \in [0, T] \longrightarrow \int_0^t e^{-L_2(t-s)} \frac{\partial \mathbf{A_H}}{\partial t}(s) \, ds \in \mathbf{H}^1(\Omega)$ is continuous, it suffices then to show that

$$t \in [0, T] \longrightarrow \int_0^t e^{-\mathcal{A}(t-s)} \mathcal{F}'\left(s, \tilde{u}(s)\right) \, ds \in \mathbb{H}^1(\Omega) \quad \text{is continuous,}$$

where $\mathcal{F}'\left(t, \tilde{u}(t)\right) = \mathcal{F}\left(t, \tilde{u}(t)\right) + \left(0, \frac{\partial \mathbf{A_H}}{\partial t}(t)\right)$. At first, we remark that

$$\left(t \longrightarrow \mathcal{F}'(t, \tilde{u}(t))\right) \in L^\infty\left(0, T; \mathbb{L}^{3/2}(\Omega)\right). \tag{4.8}$$

To check this, we shall estimate each term in $\mathcal{F}'$ separately. For example

$$\|\nabla\psi(t) \cdot \tilde{\mathbf{A}}(t)\|_{3/2} \leq \|\nabla\psi(t)\|_2 \|\tilde{\mathbf{A}}(t)\|_6 \leq C\|\psi(t)\|_{\mathcal{H}^1} \|\tilde{\mathbf{A}}(t)\|_{\mathbf{H}^1},$$

where C is the Sobolev constant relative to the continuous imbedding of $H^1(\Omega)$ in $L^6(\Omega)$. The other remaining terms can be estimated in the similar way, which confirm (4.8). In the sequel, we define

$$\begin{aligned} \mathcal{F}_\lambda(t) &= \int_0^{t-\lambda} e^{-\mathcal{A}(t-s)} \mathcal{F}'(s, \tilde{u}(s)) \, ds \quad \text{for } \lambda \leq t \leq T, \\ \mathcal{F}_\lambda(t) &= 0 \quad \text{for } 0 \leq t \leq \lambda. \end{aligned}$$

For $\lambda > 0$ small, $\mathcal{F}_\lambda$ is well defined and continuous. Indeed we write for $\lambda < t < T$

$$\mathcal{F}_\lambda(t+h) - \mathcal{F}_\lambda(t) = I_1 + I_2$$

where

$$\begin{aligned} I_1 &= \int_0^{t-\lambda} \mathrm{e}^{-\mathcal{A}(t-s)}(\mathrm{e}^{-\mathcal{A}h} - I)\mathcal{F}'(s, \tilde{u}(s))\, \mathrm{d}s \\ I_2 &= \int_{t-\lambda}^{t+h-\lambda} \mathrm{e}^{-\mathcal{A}(t+h-s)}\mathcal{F}'(s, \tilde{u}(s))\, \mathrm{d}s. \end{aligned}$$

By using (4.6), we have

$$\|I_1\|_{\mathbb{H}^1} \leq C \int_0^{t-\lambda} (t-s)^{-\gamma}\mathrm{e}^{-\delta(t-s)}\|(\mathrm{e}^{-\mathcal{A}h} - I)\mathcal{F}'(s, \tilde{u}(s))\|_{3/2}\, \mathrm{d}s.$$

Furthermore thanks to (4.8), we can apply Lebesgue theorem to obtain

$$\|I_1\|_{\mathbb{H}^1} \longrightarrow 0 \quad \text{as} \quad h \to 0$$

On the other hand

$$\begin{aligned} \|I_2\|_{\mathbb{H}^1} &\leq C \int_{t-\lambda}^{t+h-\lambda} (t+h-s)^{-\gamma}\mathrm{e}^{-\delta(t+h-s)}\|\mathcal{F}'(s, \tilde{u}(s))\|_{3/2}\, \mathrm{d}s \\ &\leq C \sup_{0\leq t\leq T} \|\mathcal{F}'(s, \tilde{u}(s))\|_{3/2} \int_\lambda^{h+\lambda} s^{-\gamma}\mathrm{e}^{-\delta s}\, \mathrm{d}s \end{aligned}$$

and we obtain

$$\|I_2\|_{\mathbb{H}^1} \longrightarrow 0 \quad \text{as} \quad h \to 0$$

When $h \to 0^-$, we obtain a similar estimate and the remaining case $0 \leq t \leq \lambda$ is trivial. Therefore $\mathcal{F}_\lambda \in C\left(0, T; \mathbb{H}^1(\Omega)\right)$.

Now for $t \in [t_0, t_1] \subset (0, T)$, we estimate

$$\begin{aligned} \left\| \mathcal{F}_\lambda(t) - \int_0^t \mathrm{e}^{-\mathcal{A}(t-s)}\mathcal{F}'(s, \tilde{u}(s))\, \mathrm{d}s \right\|_{\mathbb{H}^1} &\leq \int_{t-\lambda}^t \|\mathrm{e}^{-\mathcal{A}(t-s)}\mathcal{F}'(s, \tilde{u}(s))\|_{\mathbb{H}^1}\, \mathrm{d}s \\ &\leq C \int_0^\lambda s^{-\gamma}\mathrm{e}^{-\delta s}\, \mathrm{d}s. \end{aligned}$$

Passing to limit $\lambda \to 0^+$, uniformly for $t_0 \leq t \leq t_1$ (t_0 and t_1 are arbitrary), we obtain that the map $\left(t \in (0, T) \longrightarrow \int_0^t \mathrm{e}^{-\mathcal{A}(t-s)}\mathcal{F}'(s, \tilde{u}(s))\, \mathrm{d}s\right)$ is continuous. It remains to show continuity for t=0 and t=T and this is achieved analogously. Therefore

$$\tilde{u} = (\psi, \tilde{\mathbf{A}}) \in C\left(0, T; \mathbb{H}^1(\Omega)\right).$$

Remark 1 It is not hard to see that the order parameter ψ satisfies moreover the "*maximum principle*": if $\psi_0 \in \mathcal{L}^\infty(\Omega)$ then

$$|\psi(x,t)| \leq \max(1, \|\psi_0\|_\infty) \quad \text{for all} \quad (x,t) \in \bar{\Omega} \times [0,T]. \tag{4.9}$$

As a consequence of theorem 1, we obtain that the pair $(\psi, \tilde{\mathbf{A}})$ is a weak solution of equations (3.3) and (3.4), while the boundary condition (3.5) is satisfied in some sense of traces. Observe that theorem 1 includes a comparable result for the pair $u = (\psi, \mathbf{A})$, providing that continuity of $\mathbf{A_H}$ in time occurs. Such a regularity is completely controlled by the continuity of $\mathbf{H}$ in time and the hypothesis $(\mathbf{H_0})$ seems to be only a natural minimal condition for the existence and uniqueness result in theorem 1. However condition $(\mathbf{H_0})$ may be strengthened by requiring that $\mathbf{H} \in C(0,T;\mathbf{L}^2(\Omega))$, in this case we obtain the solution $u = (\psi, \mathbf{A}) \in C(0,T;\mathbb{H}^1(\Omega))$ and satisfies the gauged TDGL equations (2.3)-(2.4) in a weak sense.

We now concentrate on the regularity of the dependence of the solution $\tilde{u}$ on the initial data $\tilde{u}_0$. As in [9], we can verify as well that the map $\tilde{u}_0 \in X^{1/2} \longrightarrow \tilde{u} \in C(0,T;\mathbb{H}^1(\Omega))$ is uniformly Lipschitz continuous on bounded subsets of $X^{1/2}$. This implies the following

Theorem 2 *The solutions of the abstract initial-value problem* (3.12) *generate a dynamical process $U = \{U(t,s) : 0 \leq s \leq t \leq T\}$ on $X^{1/2}$ by the definition*

$$\tilde{u}(t) = U(t,s)\tilde{u}(s) \quad \text{for} \quad 0 \leq s \leq t \leq T. \tag{4.10}$$

Also, for $0 \leq s < t \leq T$, each map $U(t,s) : X^{1/2} \to X^{1/2}$ is compact.

We omit the proof since the arguments are similar.

Remark 2 Let us mention that in the particular case, where the magnetic field $\mathbf{H}$ is time constant, the result obtained in [10] concerning asymptotic behavior of the mild solution as $t \to \infty$ remains true, namely the process U becomes a dynamical system $S = \{S(t) : t \geq 0\}$ on $X^{1/2}$, by the definition

$$S(t-s) = U(t,s) \quad \text{for} \quad t \geq s \geq 0.$$

Moreover the dynamical system enjoys the following properties

(i) The functional E_ω defined in (4.7) is a Liapunov functional for S.

(ii) Each $\tilde{u}_0 \in X^{1/2}$ has a relatively compact orbit in $\mathbb{H}^1(\Omega)$.

(iii) The ω-limit set of each $\tilde{u}_0 \in X^{1/2}$ is a nonempty compact connected set of divergence-free equilibria.

(iv) There is a global attractor for S.

Here the sense of definitions is borrowed from [17].

5 Global Boundedness

In the sequel, we would like to show that in a special case of a smooth magnetic field $\mathbf{H}$, the solutions ψ and $\mathbf{A}$ of the gauged TDGL equations (2.3)-(2.5) become bounded uniformly with respect to $t \geq 0$. In what follows C will denote various constants depending only on the data κ, η, $\mathbf{H}$ and the constants entering the equations (2.3)-(2.4), but not on t. Also we use the symbol ∂_t to denote the time derivative $\frac{d}{dt}$. Throughout this section, we shall assume that $\mathbf{H}(t) \in \mathbf{H}^1(\Omega)$ for $t \geq 0$ with $\mathbf{H} \in C\left(0, T; \mathbf{L}^2(\Omega)\right)$ for all $T > 0$, $u_0 = (\psi_0, \mathbf{A}_0) \in X^{1/2}$, with $\psi_0 \in \mathcal{L}^\infty(\Omega)$ and $\|\psi_0\|_\infty \leq 1$. Let $u = (\psi, \mathbf{A})$ the corresponding solution of the TDGL equations starting from u_0. Remark that since $\mathbf{H}$ is time continuous, it is also the case for the solution u. We have the following estimate on the L^2-norm of ψ and $\mathbf{A}$.

Lemma 1 *Assume $\mathbf{H} \in L^\infty\left(0, \infty; \mathbf{L}^2(\Omega)\right)$, then there exists $C > 0$ such that*

$$\|\psi(t)\|_2^2 + \|\mathbf{A}(t)\|_2^2 \leq C\left[\mathrm{e}^{-\lambda_0 \omega_0 t}\left(\|\psi_0\|_2^2 + \|\mathbf{A}_0\|_2^2\right) + 1\right] \quad \textit{for all} \quad t \geq 0, \tag{5.1}$$

where $\omega_0 = \min(1, \omega)$.

Proof: Multiplying the equation (2.3) by the complex conjugate ψ^*, integrating over Ω and taking the real part, we obtain

$$\frac{\eta}{2}\partial_t\|\psi\|_2^2 = -\frac{1}{\kappa^2}\int_{\partial\Omega} \gamma|\psi|^2 \, \mathrm{d}\sigma(x) - \left\|\left(\frac{i}{\kappa}\nabla + \mathbf{A}\right)\psi\right\|_2^2 + \|\psi\|_2^2 - \|\psi\|_4^4. \tag{5.2}$$

On the other hand taking the inner product of (2.4) with $\mathbf{A}$, it yields from (2.8) and (2.9)

$$\frac{1}{2}\partial_t\|\mathbf{A}\|_2^2 = -\|\nabla \times \mathbf{A}\|_2^2 - \omega\|\nabla \cdot \mathbf{A}\|_2^2 + \int_\Omega \mathbf{A} \cdot \mathbf{J}_s \, \mathrm{d}x + \int_\Omega \mathbf{H} \cdot (\nabla \times \mathbf{A}) \, \mathrm{d}x. \tag{5.3}$$

The last two terms in the right-hand side of (5.3) can be majorized as follows: let $\varepsilon > 0$, replace $\mathbf{J}_s$ in (1.3), so we can apply (4.9) and standard Hölder's and Young's inequalities to obtain

$$\left|\int_\Omega \mathbf{A} \cdot \mathbf{J}_s \, \mathrm{d}x\right| \leq \frac{\varepsilon}{2}\|\mathbf{A}\|_2^2 + \frac{1}{2\varepsilon}\left\|\left(\frac{i}{\kappa}\nabla + \mathbf{A}\right)\psi\right\|_2^2,$$

$$\left|\int_{\Omega}\mathbf{H}\cdot(\nabla\times\mathbf{A})\ \mathrm{d}x\right|\leq\frac{\varepsilon}{2}\|\nabla\times\mathbf{A}\|_2^2+\frac{1}{2\varepsilon}\|\mathbf{H}\|_2^2.$$

Thanks to (2.7), we get

$$\begin{aligned}\frac{1}{2}\partial_t\left(\eta\|\psi\|_2^2+\varepsilon\|\mathbf{A}\|_2^2\right) &\leq -\frac{1}{2}\left\|\left(\frac{i}{\kappa}\nabla+\mathbf{A}\right)\psi\right\|_2^2-\varepsilon\lambda_0\omega_0\|\mathbf{A}\|_{\mathbf{H}^1}^2+\varepsilon^2\|\mathbf{A}\|_{\mathbf{H}^1}^2\\ &\quad+\|\psi\|_2^2+\frac{1}{2}\|\mathbf{H}\|_2^2. \end{aligned}\tag{5.4}$$

Set $\zeta(t)=\eta\|\psi(t)\|_2^2+\varepsilon\|\mathbf{A}(t)\|_2^2$. Since $\mathbf{H}\in L^\infty\left(0,\infty;\mathbf{L}^2(\Omega)\right)$, it follows by choosing $0<\varepsilon<\frac{\lambda_0\omega_0}{2}$ that

$$\partial_t\zeta(t)+\lambda_0\omega_0\zeta(t)\leq C\quad\text{for all}\quad t\geq 0.$$

Thus after substituting in inequality (2.10), we obtain

$$\zeta(t)\leq \mathrm{e}^{-\lambda_0\omega_0 t}\zeta(0)+\frac{C}{\lambda_0\omega_0}\quad\text{for all}\quad t\geq 0.$$

This concludes the proof of the lemma.

The next theorem establishes the H^1-norm global boundedness of the solutions ψ and $\mathbf{A}$ of the TDGL equations (2.3)-(2.6).

Theorem 3 *Provided $\mathbf{H}\in W^{1,\infty}\left(0,\infty;\mathbf{L}^2(\Omega)\right)$, there exists $C>0$ such that*

$$\|\psi(t)\|_{\mathcal{H}^1}^2+\|\mathbf{A}(t)\|_{\mathbf{H}^1}^2\leq C\left[\mathrm{e}^{-\varepsilon t}\left(\|\psi_0\|_{\mathcal{H}^1}^2+\|\mathbf{A}_0\|_{\mathbf{H}^1}^2\right)+1\right]\ \text{for all } t\geq 0,\tag{5.5}$$

where $\varepsilon>0$ is small enough.

Proof: First we estimate the H^1-norm of $\mathbf{A}$. Taking the inner product of (2.4) with $\partial_t\mathbf{A}$, we have

$$\begin{aligned}\frac{1}{2}\partial_t\left(\|\nabla\times\mathbf{A}\|_2^2+\omega\|\nabla\cdot\mathbf{A}\|_2^2\right) = &-\int_{\Omega}\partial_t\mathbf{H}\cdot(\nabla\times\mathbf{A})\ \mathrm{d}x+\partial_t\left(\int_{\Omega}\mathbf{H}\cdot(\nabla\times\mathbf{A})\ \mathrm{d}x\right)\\ &-\|\partial_t\mathbf{A}\|_2^2+\int_{\Omega}\mathbf{J}_s\cdot\partial_t\mathbf{A}\ \mathrm{d}x. \end{aligned}\tag{5.6}$$

Using similar arguments as above, we get

$$\begin{aligned}\left|\int_{\Omega}\partial_t\mathbf{H}\cdot(\nabla\times\mathbf{A})\ \mathrm{d}x\right| &\leq \frac{\varepsilon}{2}\|\nabla\times\mathbf{A}\|_2^2+\frac{1}{2\varepsilon}\|\partial_t\mathbf{H}\|_2^2,\\ \left|\int_{\Omega}\mathbf{J}_s\cdot\partial_t\mathbf{A}\ \mathrm{d}x\right| &\leq \frac{1}{2}\|\partial_t\mathbf{A}\|_2^2+\frac{1}{2}\left\|\left(\frac{i}{\kappa}\nabla+\mathbf{A}\right)\psi\right\|_2^2. \end{aligned}$$

Thus

$$\frac{1}{2}\partial_t\left(\|\nabla\times\mathbf{A}\|_2^2+\omega\|\nabla\cdot\mathbf{A}\|_2^2-2\int_\Omega\mathbf{H}\cdot(\nabla\times\mathbf{A})\,\mathrm{d}x\right)$$
$$\leq\frac{1}{2}\left\|\left(\frac{i}{\kappa}\nabla+\mathbf{A}\right)\psi\right\|_2^2+\frac{\varepsilon}{2}\|\nabla\times\mathbf{A}\|_2^2+\frac{1}{2\varepsilon}\|\partial_t\mathbf{H}\|_2^2. \tag{5.7}$$

Multiplying (5.7) by ε, $0<\varepsilon<1$ and adding estimate (5.4) yield

$$\frac{1}{2}\partial_t\left[\eta\|\psi\|_2^2+\varepsilon(\|\mathbf{A}\|_2^2+\|\nabla\times\mathbf{A}\|_2^2+\omega\|\nabla\cdot\mathbf{A}\|_2^2)-2\varepsilon\int_\Omega\mathbf{H}\cdot(\nabla\times\mathbf{A})\,\mathrm{d}x\right]$$
$$\leq-\varepsilon\lambda_0\omega_0\|\mathbf{A}\|_{\mathbf{H}^1}^2+\frac{3}{2}\varepsilon^2\|\mathbf{A}\|_{\mathbf{H}^1}^2+\|\psi\|_2^2+\frac{1}{2}\left(\|\mathbf{H}\|_2^2+\|\partial_t\mathbf{H}\|_2^2\right), \tag{5.8}$$

so by putting

$$\begin{aligned}\vartheta(t)\;&=\;\eta\|\psi(t)\|_2^2+\varepsilon\left(\|\mathbf{A}(t)\|_2^2+\|\nabla\times\mathbf{A}(t)\|_2^2+\omega\|\nabla\cdot\mathbf{A}(t)\|_2^2\right)\\&\quad-2\varepsilon\int_\Omega\mathbf{H}(t)\cdot(\nabla\times\mathbf{A}(t))\,\mathrm{d}x,\end{aligned}$$

we deduce

$$\begin{aligned}\partial_t\vartheta(t)+\varepsilon\vartheta(t)\;&\leq\;-2\varepsilon\lambda_0\omega_0\|\mathbf{A}\|_{\mathbf{H}^1}^2+\varepsilon^2(4+\omega_1)\|\mathbf{A}\|_{\mathbf{H}^1}^2+(2+\varepsilon\eta)\|\psi\|_2^2\\&\quad+2\|\mathbf{H}\|_2^2+\|\partial_t\mathbf{H}\|_2^2,\end{aligned} \tag{5.9}$$

with $\omega_1=\max(1,\omega)$, which with the assumption $\mathbf{H}\in W^{1,\infty}(0,\infty;\mathbf{L}^2(\Omega))$ implies

$$\partial_t\vartheta(t)+\varepsilon\vartheta(t)\leq C\quad\text{for all}\quad t\geq0,$$

provided $0<\varepsilon<\frac{2\lambda_0\omega_0}{4+\omega_1}$. Hence Gronwall's inequality (2.10) shows

$$\vartheta(t)\leq\mathrm{e}^{-\varepsilon t}\vartheta(0)+\frac{C}{\epsilon}\quad\text{for all}\quad t\geq0.$$

Therefore

$$\|\psi(t)\|_2^2+\|\mathbf{A}(t)\|_{\mathbf{H}^1}^2\leq C\left[\mathrm{e}^{-\varepsilon t}(\|\psi_0\|_2^2+\|\mathbf{A}_0\|_{\mathbf{H}^1}^2)+1\right]\text{ for all }t\geq0. \tag{5.10}$$

On the other hand, to estimate the H^1-norm of ψ, we make use of the energy type functional E_ω introduced in (4.7). Since ψ and $\mathbf{A}$ satisfy equations (2.3)-(2.4), the time derivative of E_ω is

$$\begin{aligned}\partial_tE_\omega(t)\;&=\;-2\int_\Omega\left[\eta|\partial_t\psi-i\kappa\omega\psi(\nabla\cdot\mathbf{A})|^2+|\partial_t\mathbf{A}|^2+\omega^2|\nabla(\nabla\cdot\mathbf{A})|^2\right]\,\mathrm{d}x\\&\quad-2\int_\Omega\partial_t\mathbf{H}\cdot(\nabla\times\mathbf{A}-\mathbf{H})\,\mathrm{d}x.\end{aligned}$$

This implies

$$\partial_t E_\omega(t) \le -2\int_\Omega \partial_t \mathbf{H}\cdot(\nabla\times\mathbf{A}-\mathbf{H})\,\mathrm{d}x. \tag{5.11}$$

Now adding estimates (4.7) and (5.11), thanks to Hölder's and Young's inequalities, so it follows

$$\partial_t E_\omega(t) + E_\omega(t) \le \left\|\left(\frac{i}{\kappa}\nabla + \mathbf{A}\right)\psi\right\|_2^2 + C\left(\|\mathbf{A}\|_{\mathbf{H}^1}^2 + \|\mathbf{H}\|_2^2 + \|\partial_t\mathbf{H}\|_2^2 + 1\right), \tag{5.12}$$

therefore by putting $\xi(t) = E_\omega(t) + \eta\|\psi(t)\|_2^2$, we derive from (5.2) and (5.12)

$$\partial_t\xi(t) + \xi(t) \le C(\|\mathbf{A}\|_{\mathbf{H}^1}^2 + 1) \quad \text{for all} \quad t \ge 0.$$

Once more, Gronwall's inequality (2.10) yields

$$\xi(t) \le \mathrm{e}^{-t}\left[\xi(0) + C\int_0^t \mathrm{e}^s(\|\mathbf{A}(s)\|_{\mathbf{H}^1}^2 + 1)\,\mathrm{d}s\right] \quad \text{for all} \quad t \ge 0,$$

and by (5.10), we infer that

$$\xi(t) \le C\left[\mathrm{e}^{-\varepsilon t}\left(\|\psi_0\|_{\mathcal{H}^1}^2 + \|\mathbf{A}_0\|_{\mathbf{H}^1}^2\right) + 1\right] \quad \text{for all} \quad t \ge 0.$$

Consequently by replacing E_ω in (4.7) and taking in mind (5.10), we conclude

$$\|\nabla\psi(t)\|_2^2 \le C\left[\mathrm{e}^{-\varepsilon t}(\|\psi_0\|_{\mathcal{H}^1}^2 + \|\mathbf{A}_0\|_{\mathbf{H}^1}^2) + 1\right],$$

which proves theorem 2.

Remark 3 Theorem 3 remains true also for the pair $\tilde{u} = (\psi, \tilde{\mathbf{A}})$ of solutions of the reduced homogeneous problem (3.3)- (3.4). On the other hand, we can use equation (3.15) to improve the regularity of the dependence of $\tilde{u}$ on the initial data $\tilde{u}_0$; that is the set $\{U(t,0)\tilde{u}_0 : t \ge 0,\ \|\tilde{u}_0\|_{\mathbb{H}^1} \le R\}$ $(R > 0)$, is relatively compact in $\mathbb{H}^1(\Omega)$.

References

[1] **Gor'kov, L. P.**, and **Eliashberg, G. M. :** *Generalizations of the Ginzburg-Landau equations for non-stationary problems in the case of alloys with paramagnetic impurities.* Zh. Eksp. Teor. Fiz., **54**, 612-626 (1968); Soviet Phys. JETP **27**, 328-334 (1968)

[2] **Tinkham, M. :** *Introduction to Superconductivity.* McGraw-Hill, Inc., New York, 1975

[3] **Tang, M. Q. :** *On an evolutionary system of Ginzburg-Landau equations with fixed total magnetic flux.* Comm. PDEs **20**, 1-36 (1995)

[4] **Liang, J.**, and **Tang, Q. :** *Asymptotic behaviour of the solutions of an evolutionary Ginzburg-Landau model.* J. Math. Anal. Appl. **195**, 92-107 (1995)

[5] **Tang, Q.**, and **Wang, S. :** *Time dependent Ginzburg-Landau equations of superconductivity.* Physica D **88**, 139-166 (1995)

[6] **Du, Q. :** *Global existence and uniqueness of solutions of the time-dependent Ginzburg-Landau model for superconductivity.* Appl. Anal. **53**, 1-18 (1994)

[7] **Chen, Z. M., Hoffmann, K.-H.**, and **Liang, J. :** *On a nonstationary Ginzburg-Landau superconductivity model.* Math. Methods Appl. Sci. **16**, 855-875 (1993)

[8] **Rodriguez-Bernal, A., Wang, B.**, and **Willie, R. :** *Asymptotic behavior of time-dependent Ginzburg-Landau equations of superconductivity.* Math. Meth. Appl. Sci., **22**, 1647-1669 (1999)

[9] **Takáč, P. :** *On the dynamical system generated by a superconductivity model.* Proc. ICIAM' 95, Z. Angew. Math. Mech., **76** (S2), 349-352 (1996)

[10] **Fleckinger-Pellé, J., Kaper, H. G.**, and **Takáč, P. :** *Dynamics of the Ginzburg-Landau equations of superconductivity.* Nonl. Anal. TMA, **32**, 647-665 (1998)

[11] **Kaper, H. G.**, and **Takáč, P. :** *Ginzburg-Landau dynamics with a time-dependent magnetic field.* Nonlinearity, **11**, 291-305 (1998)

[12] **Fleckinger-Pellé, J.**, and **Kaper, H. G. :** *Gauges for the Ginzburg-Landau equations of superconductivity.* Proc. ICIAM' 95, Z. Angew. Math. Mech., **76**, 345-348 (1996)

[13] **Adams, R. A. :** *Sobolev Spaces.* Academic Press, New York 1975

[14] **Girault, V.**, and **Raviart, P.-A. :** *Finite Element Methods for Navier-Stokes Equations.* Springer-Verlag, New York 1986

[15] **Henry, D. :** *Geometric Theory of Semilinear Parabolic Equations.* Lecture Notes in Math. **840**, Springer-Verlag, New York 1981

[16] **Pazy, A. :** *Semigroups of Linear Operators and Applications to Partial Differential Equations.* Springer-Verlag 1983

[17] **Temam, R. :** *Infinite-Dimensional Dynamical Systems in Mechanics and Physics.* Springer-Verlag, New York 1988

received: July 8, 2002

Author:

Fouzi Zaouch
Universität Rostock
Fachbereich Mathematik
18051 Rostock
Germany

e-mail: fouzi.zaouch@mathematik.uni-rostock.de

Rostock. Math. Kolloq. **57**, 71–91 (2003) Subject Classification (AMS)
81S10,53C80

Katharina Habermann, Andreas Klein

Lie derivative of symplectic spinor fields, metaplectic representation, and quantization

ABSTRACT. In the context of Riemannian spin geometry it requires skilful handling to define a Lie derivative of (Riemannian) spinor fields.

A Lie derivative of symplectic spinor fields in the direction of Hamiltonian vector fields can be defined in a very natural way. It is the aim of this note to present this construction. Furthermore, an immediate interpretation of this Lie derivative in the language of natural ordering quantization is given.

Introduction

In the context of Riemannian spin geometry, the general question of constructing a Lie derivative for spinor fields has been studied by several authors. Yvette Kosmann, for instance, gave a geometric construction of a so-called metric Lie derivative of spinor fields in [12]. This approach was extended by Jean–Pierre Bourguignon and Paul Gauduchon in [2]. The problem with it is to compare spinor fields for different metrics, since a diffeomorphism ϕ transforms the metric tensor g to ϕ^*g and the (Riemannian) spinor fields over (M, g) will be transformed into spinor fields over (M, ϕ^*g). Other studies focussed on relations between Killing vector fields and Killing spinors such as [14] by Andrei Moroianu and [1] Dmitri Alekseevsky *et al.* A further result in this direction was the finding of Katharina Habermann that conformal vector fields act by a certain kind of conformal Lie derivative on the space of solutions of the twistor equation. In [7] she discussed the relevant $\mathbb{Z}_2$-graded algebra.

Studing the problem in the symplectic setting, one deals with symplectic spinor fields over (M, ω) and $(M, \phi^*\omega)$, respectively. In the case of a Hamiltonian vector field all spinor fields live over the same symplectic manifold and a definition of a Lie derivative for symplectic spinor fields in the direction of a Hamiltonian vector field in the classical way of defining a Lie derivative for geometrical objects is possible. It is the aim of this note to present this construction.

Furthermore, an immediate interpretation of this Lie derivative in the language of natural ordering quantization is given. This interpretation was inspired by Theorem 1 in the book [6] of Maurice de Gosson. The observation is that there is a one-parameter group of metaplectic operators, which is associated to a quadratic Hamiltonian and gives solutions of a Schrödinger equation. A similar Schrödinger equation but without any spinorial context was established in the book [5] of Victor Guillemin and Shlomo Sternberg. Moreover, a detailed discussion of this Schrödinger equation can be found in the mentioned book of Maurice de Gosson. In this paper, we put the Schrödinger equation in the context of symplectic spin geometry and give a new and completely self-contained proof. Finally, the Schrödinger equation gives the Lie derivative of constant symplectic spinor fields on $\mathbb{R}^{2n}$ in the direction of the Hamiltonian vector field associated to the quadratic Hamiltonian.

Altogether, our computations also illustrate a remark of Bertram Kostant in his paper on symplectic spinors. There, symplectic spinor fields were introduced in order to give the construction of the half-form bundle and the half-form pairings in the context of geometric quantization. These half–densities are related to a certain line subbundle of the symplectic spinor bundle, which sometimes is also known as metaplectic correction. And Bertram Kostant notices that Hamiltonian vector fields clearly operate as Lie differentiation on smooth symplectic spinor fields ([13] 5.5).

Acknowledgement. This work was partially supported by the Gerhard Hess – programme of the Deutsche Forschungsgemeinschaft (DFG). Additionally, parts of this paper were written during the stay of the first author at the Blekinge Technical University in Karlskrona (Sweden). She wishes to thank in particular Professor Maurice de Gosson for the kind hospitality and many useful discussions.

1 Preparations

1.1 Some Notations

We consider the standard space $\mathbb{R}^{2n}$ with the Euklidean product $\langle\ ,\ \rangle$. Further, let J be the $2n \times 2n$-matrix given by

$$J = \begin{pmatrix} 0 & -\mathbf{1} \\ \mathbf{1} & 0 \end{pmatrix} ,$$

where $\mathbf{1}$ denotes the $n \times n$-matrix $\mathbf{1} = diag(1, \dots, 1)$. Then the standard symplectic structure ω_0 on $\mathbb{R}^{2n}$ is defined to be

$$\omega_0(\ ,\ ) = \langle J\ ,\ \rangle .$$

We remark that for local coordinates $(p, q) = (p_1, \ldots, p_n, q_1, \ldots, q_n)$ on $\mathbb{R}^{2n}$ the standard symplectic structure ω_0 writes as

$$\omega_0 = \sum_{j=1}^{n} dp_j \wedge dq_j \ .$$

For the canonical standard basis $\{a_1, \ldots, a_n, b_1, \ldots, b_n\}$ of $\mathbb{R}^{2n}$ one computes readily

$$\omega_0(a_j, a_k) = 0 \ , \quad \omega_0(b_j, b_k) = 0 \ , \quad \text{and} \quad \omega_0(a_j, b_k) = \delta_{jk} \qquad \text{for} \quad j, k = 1, \ldots, n \ .$$

This says that $\{a_1, \ldots, a_n, b_1, \ldots, b_n\}$ is a symplectic basis of the symplectic vector space $(\mathbb{R}^{2n}, \omega_0)$.

The symplectic group $Sp(2n, \mathbb{R})$ is the group of real $2n \times 2n$-matrices leaving the standard symplectic structure ω_0 on $\mathbb{R}^{2n}$ invariant, i.e. the group $Sp(2n, \mathbb{R})$ consists of those real $2n \times 2n$-matrices A satisfying the relation

$$A^\top J A = J \ . \tag{1.1}$$

Thus, the Lie algebra $\mathfrak{sp}(2n, \mathbb{R})$ of the symplectic group is given by the space of all real $2n \times 2n$-matrices B with

$$B^\top J + J B = 0 \ . \tag{1.2}$$

Moreover, let

$$\begin{array}{cc} B_{jk} = \begin{pmatrix} 0 & \vdots & 0 \\ \ldots & 1 & \ldots \\ 0 & \vdots & 0 \end{pmatrix} & \leftarrow j\text{-th row} \\ \qquad\quad \uparrow k\text{-th column} & \end{array}$$

be the $n \times n$-matrix with a 1 as the only nonvanishing entry at the j-th row and the k-th column for $j, k = 1, \ldots, n$. Using these $n \times n$-matrices, we introduce the following $2n \times 2n$-matrices

$$X_{jk} = \begin{pmatrix} B_{jk} & 0 \\ 0 & -B_{kj} \end{pmatrix} \ , \quad Y_{jk} = \begin{pmatrix} 0 & B_{jk} + B_{kj} \\ 0 & 0 \end{pmatrix} \ , \quad \text{and} \quad Z_{jk} = \begin{pmatrix} 0 & 0 \\ B_{jk} + B_{kj} & 0 \end{pmatrix}$$

for $j, k = 1, \ldots, n$. Now, it is a well known fact that the set

$$\{Y_{jk} \text{ and } Z_{jk} \text{ for } 1 \le j \le k \le n \ , \ X_{jk} \text{ for } 1 \le j, k \le n\}$$

of $2n \times 2n$-matrices is a basis of the symplectic Lie algebra $\mathfrak{sp}(2n, \mathbb{R})$.

1.2 The Metaplectic Representation and symplectic Clifford multiplication

This section recalls well known basics on the metaplectic group and its representation. See also [10, 15].

For the symplectic group, the subgroup $Sp(2n,\mathbb{R}) \cap O(2n,\mathbb{R}) \cong U(n)$ is maximal compact. This implies $\pi_1(Sp(2n,\mathbb{R})) \cong \mathbb{Z}$ for the fundamental group of $Sp(2n,\mathbb{R})$. Consequently, the symplectic group has a – up to isomorphism – uniquely determined covering group of order 2. The metaplectic group $Mp(2n,\mathbb{R})$ is defined to be this two-fold covering group of $Sp(2n,\mathbb{R})$, giving the exact sequence

$$1 \quad \to \quad \mathbb{Z}_2 \quad \to \quad Mp(2n,\mathbb{R}) \ \overset{\rho}{\to} \ Sp(2n,\mathbb{R}) \quad \to \quad 1$$

with double covering map ρ. For our computations, it is sufficient to know the differential $\rho_* : \mathfrak{mp}(2n,\mathbb{R}) \to \mathfrak{sp}(2n,\mathbb{R})$ of this double covering. Due to Crumeyrolle [3], the Lie algebra of the metaplectic group is given by the set of all symmetric homogeneous polynomials of degree 2 in the elements of $\mathbb{R}^{2n}$. Thus, the set

$$\{a_j \cdot a_k \text{ and } b_j \cdot b_k \text{ for } 1 \le j \le k \le n \ , \ a_j \cdot b_k + b_k \cdot a_j \text{ for } 1 \le j,k \le n\}$$

is a basis of the metaplectic Lie algebra $\mathfrak{mp}(2n,\mathbb{R})$. This Lie algebra may be represented as a Lie subalgebra of the symplectic Clifford algebra. So we write formally $v \cdot w$ for the polynomial given by the two vectors v and w. Later, this notation will be consistent with the Clifford multiplication of vectors and functions.

Then one proves (cf. [9] Proposition 1.2)

Lemma 1.1 *The differential $\rho_* : \mathfrak{mp}(2n,\mathbb{R}) \to \mathfrak{sp}(2n,\mathbb{R})$ is given by $\rho_*(a_j \cdot a_k) = -Y_{jk}$, $\rho_*(b_j \cdot b_k) = Z_{jk}$, and $\rho_*(a_j \cdot b_k + b_k \cdot a_j) = 2\,X_{jk}$ for $j,k = 1,\dots,n$.* □

The Schrödinger quantization prescription

$$\begin{array}{lcll} 1 \in \mathbb{R} & \mapsto & \sigma(1) := \text{multiplication by } i, & \\ a_j \in \mathbb{R}^{2n} & \mapsto & \sigma(a_j) := \text{multiplication by } i x_j \ , & \text{and} \\ b_j \in \mathbb{R}^{2n} & \mapsto & \sigma(b_j) := \dfrac{\partial}{\partial x_j} & \text{for} \quad j = 1,\dots,n \ , \end{array}$$

where the operators $\sigma(1)$, $\sigma(a_j)$, and $\sigma(b_j)$ for $j = 1,\dots,n$ are continuous operators acting on the Schwartz space $\mathcal{S}(\mathbb{R}^n)$ of rapidly decreasing smooth functions on $\mathbb{R}^n$, gives the symplectic

Clifford multiplication

$$\begin{array}{rcl}\mu : \mathbb{R}^{2n} \times \mathcal{S}(\mathbb{R}^n) & \to & \mathcal{S}(\mathbb{R}^n) \\ (v, f) & \mapsto & \mu(v, f) = v \cdot f := \sigma(v) f \ .\end{array}$$

It is an elementary computation to prove the relation

$$v \cdot w \cdot f - w \cdot v \cdot f = -i\omega_0(v, w) f$$

for vectors $v, w \in \mathbb{R}^{2n}$ and functions $f \in \mathcal{S}(\mathbb{R}^n)$.

The metaplectic group has a natural representation acting on the Hilbert space $L^2(\mathbb{R}^n)$. A concrete realization of this representation is given by the following specification (cf. [10]).

Consider $g(a) = \left(\sqrt{\det(a)}, \begin{pmatrix} a & 0 \\ 0 & (a^\top)^{-1} \end{pmatrix} \right)$ where $a \in GL(n, \mathbb{R})$. Choosing a square root of $\det(a)$, one has $g(a) \in Mp(2n, \mathbb{R})$ and

$$(L(g(a))f)(x) = \sqrt{\det(a)} f(a^\top x) \ , \quad x \in \mathbb{R}^n \ . \tag{1.3}$$

The set of all matrices $\tau(b) = \begin{pmatrix} \mathbf{1} & b \\ 0 & \mathbf{1} \end{pmatrix}$, where $b^\top = b$ and $\mathbf{1}$ denotes the $n \times n$-matrix $\mathbf{1} = diag(1, \ldots, 1)$ is simply connected. Thus, $\tau(b)$ can be understood as an element of $Mp(2n, \mathbb{R})$, such that $t(0)$ is the unit element in $Mp(2n, \mathbb{R})$. For $\tau(b)$ it is

$$(L(\tau(b))f)(x) = \mathrm{e}^{-\frac{i}{2}\langle bx, x \rangle} f(x) \ , \quad x \in \mathbb{R}^n \ . \tag{1.4}$$

Choosing a square root $i^{1/2}$, the element $\sigma = (i^{1/2}, J)$ can be considered as an element $\sigma \in Mp(2n, \mathbb{R})$. Here, one obtains

$$(L(\sigma)f)(x) = \left(\frac{i}{2\pi} \right)^{\frac{n}{2}} \int_{\mathbb{R}^n} \mathrm{e}^{i\langle x, y \rangle} f(y) dy \ , \quad x \in \mathbb{R}^n \ . \tag{1.5}$$

That gives $L(\sigma) = i^{\frac{n}{2}} \mathcal{F}^{-1}$, where $\mathcal{F} : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$ denotes the usual Fourier transform. Finally, we remark that the metaplectic group is generated by all these types of elements, since the corresponding matrices in $Sp(2n, \mathbb{R})$ already give the whole symplectic group.

With respect to this representation, the symplectic Clifford multiplication is $Mp(2n, \mathbb{R})$-equivariant, i.e. we have the relation

$$\mu(\rho(g)v, L(g)f) = L(g)\mu(v, f)$$

for all $g \in Mp(2n, \mathbb{R})$, $v \in \mathbb{R}^{2n}$, and $f \in \mathcal{S}(\mathbb{R}^n)$.

The differential of the metaplectic representation is more interesting for our computations. In order to be able to give precise calculations, we are going to deduce the differential detailly.

Proposition 1.2 *The differential $L_* : \mathfrak{mp}(2n,\mathbb{R}) \to \mathfrak{u}(\mathcal{S}(\mathbb{R}^n))$ of the metaplectic representation L is given by*

$$\begin{aligned} L_*(a_j \cdot a_k)(f) &= -ia_j \cdot a_k \cdot f \\ L_*(b_j \cdot b_k)(f) &= -ib_j \cdot b_k \cdot f \\ L_*(a_j \cdot b_k + b_k \cdot a_j)(f) &= -i(a_j \cdot b_k + b_k \cdot a_j) \cdot f \end{aligned}$$

for $j, k = 1, \ldots, n$.

Proof: Generally, the differential $L_* : \mathfrak{mp}(2n,\mathbb{R}) \to \mathfrak{u}(\mathcal{S}(\mathbb{R}^n))$ may be computed via

$$L_*(X)f = \frac{d}{dt}(L(\exp(tX))f_{|t=0} \; .$$

We will make use of this formula in the progress of this proof.

First, the relation $\rho(\exp(tX)) = \exp(t\rho_*(X))$ gives

$$\begin{aligned} \rho(\exp(t\,a_j \cdot a_k)) &= \exp(t\rho_*(a_j \cdot a_k)) \quad = \quad \exp(-tY_{jk}) \\ &= \exp\begin{pmatrix} 0 & -t(B_{jk}+B_{kj}) \\ 0 & 0 \end{pmatrix} \quad = \quad \begin{pmatrix} 1 & -t(B_{jk}+B_{kj}) \\ 0 & 1 \end{pmatrix} \; , \end{aligned}$$

$$\begin{aligned} \rho(\exp(t\,b_j \cdot b_k)) &= \exp(t\rho_*(b_j \cdot b_k)) \quad = \quad \exp(tZ_{jk}) \\ &= \exp\begin{pmatrix} 0 & 0 \\ t(B_{jk}+B_{kj}) & 0 \end{pmatrix} \quad = \quad \begin{pmatrix} 1 & 0 \\ t(B_{jk}+B_{kj}) & 1 \end{pmatrix} \\ &= J\begin{pmatrix} 1 & -t(B_{jk}+B_{kj}) \\ 0 & 1 \end{pmatrix} J^{-1} \; , \end{aligned}$$

and

$$\begin{aligned} \rho(\exp(t(a_j \cdot b_k + b_k \cdot a_j))) &= \exp(t\rho_*(a_j \cdot b_k + b_k \cdot a_j)) \quad = \quad \exp(2t\,X_{jk}) \\ &= \exp\begin{pmatrix} 2tB_{jk} & 0 \\ 0 & -2tB_{kj} \end{pmatrix} \\ &= \begin{pmatrix} \exp(2tB_{jk}) & 0 \\ 0 & \exp(-2tB_{jk})^\top \end{pmatrix} \; . \end{aligned}$$

Thus,

$$\exp(t\,a_j \cdot a_k) = \tau(-t(b_{jk} + B_{kj})) \; ,$$

$$\exp(t\,b_j \cdot b_k) = \sigma\,\tau(-t(B_{jk}+B_{kj}))\,\sigma^{-1},$$

and

$$\exp(t(a_j \cdot b_k + b_k \cdot a_j)) = g(\exp(2tB_{jk})) \; .$$

Finally, this gives

$$\begin{aligned}
(L_*(a_j \cdot a_k)f)(x) &= \frac{d}{dt}(L(\exp(t\, a_j \cdot a_k))f)(x)_{|t=0} \\
&= \frac{d}{dt} e^{\frac{i}{2}t\langle (B_{jk}+B_{kj})x,x\rangle} f(x)_{|t=0} \\
&= \frac{i}{2}\langle (B_{jk}+B_{kj})x,x\rangle f(x) \\
&= i x_j x_k\, f(x) \quad = \quad -i a_j \cdot a_k \cdot f(x)\ ,
\end{aligned}$$

$$\begin{aligned}
(L_*(b_j \cdot b_k)f)(x) &= \frac{d}{dt}(L(\exp(t\, b_j \cdot b_k))f)(x)_{|t=0} \\
&= \frac{d}{dt}(L(\sigma)\circ L(\tau(-t(B_{jk}+B_{kj})))\circ L(\sigma)^{-1}(f))(x)_{|t=0} \\
&= i\mathcal{F}^{-1}(x_j x_k \mathcal{F}(f))(x) \\
&= -i\frac{\partial^2 f}{\partial x_j \partial x_k}(x) \quad = \quad -i b_j \cdot b_k \cdot f(x)\ ,
\end{aligned}$$

and

$$\begin{aligned}
(L_*(a_j \cdot b_k + b_k \cdot a_j)f)(x) &= \frac{d}{dt}(L(\exp(t(a_j \cdot b_k + b_k \cdot a_j)))f)(x)_{|t=0} \\
&= \frac{d}{dt}\sqrt{\det(\exp(2tB_{jk}))}\, f(\exp(2tB_{jk})^\top x)_{|t=0} \\
&= \frac{1}{2}\frac{d}{dt}\det(\exp(2tB_{jk}))_{|t=0}\ f(x) + \frac{d}{dt} f(\exp(2tB_{jk})^\top x)_{|t=0} \\
&= \frac{1}{2}Tr\left(\frac{d}{dt}\exp(2tB_{jk})_{|t=0}\right)\ f(x) + df\left(\frac{d}{dt}\exp(2tB_{jk})^\top_{|t=0} x\right) \\
&= \frac{1}{2}Tr(2B_{jk})f(x) + df(2B_{kj}x) \\
&= \delta_{jk} f(x) + 2x_j \frac{\partial f}{\partial x_k}(x) \\
&= x_j \frac{\partial f}{\partial x_k}(x) + \frac{\partial}{\partial x_k} x_j f(x) \quad = \quad -i(a_j \cdot b_k + b_k \cdot a_j)\cdot f(x)\ ,
\end{aligned}$$

which are the asserted relations. □

1.3 Symplectic Spinor Fields

Let (M,ω) be a $2n$-dimensional symplectic manifold and R the $Sp(2n,\mathbb{R})$-principal fibre bundle of all symplectic frames over M. A metaplectic structure on (M,ω) is a principal fibre bundle P over M having $Mp(2n,\mathbb{R})$ as structure group together with a bundle morphism $f : P \to R$ which is equivariant with respect to the homomorphism $\rho : Mp(2n,\mathbb{R}) \to Sp(2n,\mathbb{R})$.

That is, we have the following commutative diagram

$$\begin{tikzcd}
P \times Mp(2n,\mathbb{R}) \arrow[r] \arrow[d, "f \times \rho"'] & P \arrow[d, "f"'] \arrow[dr] & \\
 & & M \\
R \times Sp(2n,\mathbb{R}) \arrow[r] & R \arrow[ur] &
\end{tikzcd}$$

such that a metaplectic structure can be understood as a lift of the symplectic frame bundle R with respect to the double covering ρ.

Generally, one has a cohomological obstruction to lifting the structure group of a principal fibre bundle. The topological condition to the existence of a metaplectic structure is given by $c_1(M) \equiv 0 \bmod 2$.

If (M, ω) is a $2n$-dimensional symplectic manifold with fixed metaplectic structure P then the symplectic spinor bundle is defined to be the associated Hilbert bundle

$$\mathcal{Q} = P \times_L L^2(\mathbb{R}^n).$$

Furthermore, we need the subbundle

$$\mathcal{S} = P \times_L \mathcal{S}(\mathbb{R}^n).$$

Observing that the symplectic Clifford multiplication is $Mp(2n, \mathbb{R})$-equivariant, it lifts to the bundle level to a symplectic Clifford multiplication

$$\begin{array}{rcl}
\mu : TM \otimes \mathcal{S} & \rightarrow & \mathcal{S} \\
X \otimes \varphi & \mapsto & \mu(X, \varphi) = X \cdot \varphi
\end{array}$$

on the symplectic spinor bundle $\mathcal{S}$. Obviously, we have the relation

$$X \cdot Y \cdot \varphi - Y \cdot X \cdot \varphi = -i\omega(X, Y)\varphi$$

for vector fields X, Y and spinor fields φ.

Furthermore, the $L^2(\mathbb{R}^n)$-scalar product on the fibres gives a canonical Hermitian scalar product $\langle\ ,\ \rangle$ on $\mathcal{Q}$. $\Gamma(\mathcal{Q}) = \Gamma(\mathcal{S})$ denotes the space of all smooth symplectic spinor fields.

Moreover, any symplectic covariant derivative on the tangent bundle TM of (M, ω) induces a covariant derivative on the symplectic spinor bundle $\mathcal{Q}$, the spinor derivative

$$\nabla : \Gamma(\mathcal{Q}) \rightarrow \Gamma(T^*M \otimes \mathcal{Q}),$$

which in the following will also be denoted by ∇. If $e_1, \ldots, e_n, f_1, \ldots, f_n$ denotes any local symplectic frame on (M, ω) then the spinor derivative writes as

$$\nabla_X \varphi = X(\varphi) + \frac{i}{2} \sum_{j=1}^{n} \{e_j \cdot \nabla_X f_j - f_j \cdot \nabla_X e_j\} \cdot \varphi\ . \tag{1.6}$$

Here a covariant derivative $\nabla : \Gamma(TM) \to \Gamma(T^*M \otimes TM)$ on a symplectic manifold (M, ω) is called symplectic if and only if $\nabla\omega = 0$. The torsion of such a connection is defined to be

$$T^\nabla(X, Y) = \nabla_X Y - \nabla_Y X - [X, Y].$$

Then the connection is said to be torsionfree, if and only if $T^\nabla \equiv 0$.

Finally, for the Clifford multiplication, the spinor derivative, and the Hermitian scalar product we have the following relations

$$\begin{aligned}
(X \cdot Y - Y \cdot X) \cdot \varphi &= -i\omega(X, Y)\varphi \\
\langle X \cdot \varphi, \psi\rangle &= -\langle \varphi, X \cdot \psi\rangle \\
\nabla_X(Y \cdot \varphi) &= (\nabla_X Y) \cdot \varphi + Y \cdot \nabla_X \varphi \\
X\langle \varphi, \psi\rangle &= \langle \nabla_X \varphi, \psi\rangle + \langle \varphi, \nabla_X \psi\rangle \\
\langle \varphi, \psi\rangle &= \overline{\langle \psi, \varphi\rangle}.
\end{aligned}$$

1.4 Symplectic Spinor Fields and Diffeomorphisms

In order to define the Lie derivative of symplectic spinor fields we first illustrate how symplectic spinor fields behave under diffeomorphisms. In Riemannian spin geometry, the problem of transforming a spinor field under diffeomorphisms of the manifold is studied in detail in the paper [4] of Dabrowski and Percacci. This method can be carried over to our situation of symplectic spinor fields.

Let (M, ω) be a $2n$-dimensional symplectic manifold and let ϕ be any orientation preserving diffeomorphism of M. Then ϕ induces an isomorphism ϕ_* of the $Sp(2n, \mathbb{R})$-principal frame bundles R^ϕ and R according to the symplectic structures $\phi^*\omega$ and ω

$$\begin{aligned}
\phi_* \quad : \quad & R^\phi \to R \\
& (e_1, \ldots, e_n, f_1, \ldots, f_n) \mapsto (\phi_* e_1, \ldots, \phi_* e_n, \phi_* f_1, \ldots, \phi_* f_n)\ .
\end{aligned}$$

This isomorphism maps symplectic frames with respect to $\phi^*\omega$ to symplectic frames for the symplectic structure ω.

Let (P, f) be a fixed metaplectic structure for (M, ω). Moreover, (P^ϕ, f^ϕ) denotes the metaplectic structure for $(M, \phi^*\omega)$ such that ϕ_* lifts to an isomorphism $\tilde{\phi}_* : P^\phi \to P$, i.e. such that the following diagramm commutes

$$\begin{array}{ccccc}
 & & \tilde{\phi}_* & & \\
 & P^\phi & \to & P & \\
f^\phi & \downarrow & & \downarrow & f\ . \\
 & R^\phi & \to & R & \\
 & & \phi_* & &
\end{array}$$

Let $\mathcal{Q} = P \times_L L^2(\mathbb{R}^n)$ and $\mathcal{Q}^\phi = P^\phi \times_L L^2(\mathbb{R}^n)$ denote the corresponding symplectic spinor bundles. A symplectic spinor field over (M, ω) is a section of the symplectic spinor bundle $\mathcal{Q}$, or, equivalently, an L-equivariant map $\varphi : P \to L^2(\mathbb{R}^n)$. Now, we define the transformed symplectic spinor field $(\phi^{-1})_*\varphi$ by the equation

$$(\phi^{-1})_*\varphi = \varphi \circ \tilde{\phi}_* : P^\phi \to L^2(\mathbb{R}^n) \ ,$$

where this spinor field also is regarded as an L-equivariant map. Then $(\phi^{-1})_*\varphi$ is a symplectic spinor field over $(M, \phi^*\omega)$ with respect to the metaplectic structure (P^ϕ, f^ϕ).

Obviously, ϕ is a symplectomorphism between the symplectic manifolds (M, ω) and $(M, \phi^*\omega)$. Thus, if ∇ is any symplectic covariant derivative on (M, ω) then ∇^ϕ defined by

$$\nabla^\phi_{(\phi^{-1})_*X}(\phi^{-1})_*Y = (\phi^{-1})_*(\nabla_X Y)$$

for vector fields X and Y gives a symplectic covariant derivative for $(M, \phi^*\omega)$. This implies that the induced spinor derivative in $\mathcal{Q}^\phi$ which we also denote by ∇^ϕ satisfies

$$\nabla^\phi_{(\phi^{-1})_*X}(\phi^{-1})_*\varphi = (\phi^{-1})_*(\nabla_X \varphi).$$

Furthermore,

$$((\phi^{-1})_*X) \cdot ((\phi^{-1})_*\varphi) = (\phi^{-1})_*(X \cdot \varphi)$$

holds true for the symplectic Clifford multiplication.

2 The Lie Derivative of Symplectic Spinor Fields

In this section, we will define the Lie derivative of symplectic spinor fields in the direction of Hamiltonian vector fields. This can be done in a very natural way.

Let (M, ω) be a symplectic manifold. A vector field X over M is called Hamiltonian vector field if there is a smooth function $h : M \to \mathbb{R}$ such that

$$\omega(X, \quad) = dh \ .$$

The Hamiltonian vector field given by a function h often is denoted also by X_h. Further, let $\mathcal{L}_X$ denote the Lie derivative in the direction of X. Then, the well known relation

$$\mathcal{L}_X = d \circ i_X + i_X \circ d$$

gives

$$\mathcal{L}_X\omega = d \circ i_X\omega + i_X \circ d\omega = d(\omega(X, \quad)) = ddh = 0.$$

For the sake of simplicity, we assume M to be a closed manifold. For the Hamiltonian vector field X, let $\{\phi_t : M \to M\}_{t\in\mathbb{R}}$ be the one-parameter transformation group of diffeomorphisms induced by X, i.e. we have

$$X(x) = \frac{d}{dt}\phi_t(x)_{|t=0} \qquad \text{for } x \in M\ .$$

Then, $\mathcal{L}_X\omega = 0$ gives

$$\phi_t^*\omega = \omega \qquad \text{for } t \in \mathbb{R}.$$

Let $\mathcal{Q}$ and $\mathcal{S}$ denote the symplectic spinor bundles with respect to a fixed metaplectic structure P over (M,ω).

In section 1.4 we gave a description how a diffeomorphism $\phi : M \to M$ for a given symplectic spinor field $\varphi \in \Gamma(\mathcal{Q})$ over (M,ω) induces a symplectic spinor field $(\phi^{-1})_*\varphi$ over $(M,\phi^*\omega)$. Since $\phi_t^*\omega = \omega$, in our situation each $({\phi_t}^{-1})_*\varphi$ is a symplectic spinor field over (M,ω), i.e. lies in $\Gamma(\mathcal{Q})$. This allows the following definition.

Definition 2.1 *The Lie derivative of the symplectic spinor field $\varphi \in \Gamma(\mathcal{Q})$ in the direction of the Hamiltonian vector field X is defined to be*

$$\mathcal{L}_X\varphi \;=\; \frac{d}{dt}(\phi_t^{-1})_*\varphi_{|t=0},$$

where $\{\phi_t\}_{t\in\mathbb{R}}$ denotes the one-parameter transformation group induced by X.

Recalling the construction of $(\phi^{-1})_* : \Gamma(\mathcal{Q}) \to \Gamma(\mathcal{Q}^\phi)$ in section 1.4, one sees that $(\phi^{-1})_*$ is determined only up to sign. For this reason we additionally require $(\phi_0^{-1})_* = \mathrm{id}_{\Gamma(\mathcal{Q})}$ for the smooth family of mappings $(\phi_t^{-1})_* : \Gamma(\mathcal{Q}) \to \Gamma(\mathcal{Q})$.

Proposition 2.2 *Let ∇ be any torsionfree symplectic connection on (M,ω) and let X be any fixed Hamiltonian vector field. Then the Lie derivative of symplectic spinor fields in the direction of X can be expressed in the following form*

$$\mathcal{L}_X\varphi = \nabla_X\varphi + \frac{i}{2}\sum_{j=1}^{n}\{\nabla_{e_j}X\cdot f_j - \nabla_{f_j}X\cdot e_j\}\cdot\varphi \qquad \textit{for } \varphi \in \Gamma(\mathcal{Q}),$$

where $e_1,\dots,e_n,f_1,\dots,f_n$ denotes any local symplectic frame on (M,ω).

Proof: First, one has the relation

$$(\mathcal{L}_X\omega)(Y,Z) = X(\omega(Y,Z)) - \omega([X,Y],Z) - \omega(Y,[X,Z])$$

for vector fields X, Y, Z. Let $e_1, \dots, e_n, f_1, \dots, f_n$ be a local symplectic frame on (M, ω) and $\varphi \in \Gamma(\mathcal{Q})$ a symplectic spinor field. Then one obtains for the torsionfree symplectic connection ∇

$$\begin{aligned}
\sum_{j=1}^{n}\{\omega(\nabla_{e_j}X, f_j) + \omega(e_j, \nabla_{f_j}X)\} &= \sum_{j=1}^{n}\{\omega(\nabla_X e_j, f_j) - \omega([X, e_j], f_j) \\
&\qquad + \omega(e_j, \nabla_X f_j) - \omega(e_j, [X, f_j])\} \\
&= \sum_{j=1}^{n}\{X(\omega(e_j, f_j)) - \omega([X, e_j], f_j) - \omega(e_j, [X, f_j])\} \\
&= \sum_{j=1}^{n}(\mathcal{L}_X\omega)(e_j, f_j) \\
&= 0, \qquad (2.7)
\end{aligned}$$

by $\mathcal{L}_X\omega = 0$.

Let $\overline{s} : U \to P$ be a lift of the local symplectic frame $s = (e_1, \dots, e_n, f_1, \dots, f_n) : U \to R$ into the metaplectic structure. We consider the sections

$$s_t = ((\phi_t^{-1})_* e_1, \dots, (\phi_t^{-1})_* e_n, (\phi_t^{-1})_* f_1, \dots, (\phi_t^{-1})_* f_n) : \phi_t^{-1}(U) \to R \quad \text{for } t \in \mathbb{R}$$

and lifts $\overline{s}_t : \phi_t^{-1}(U) \to P$ of s_t, such that $\overline{s}_t$ gives a smooth family satisfying $\overline{s}_0 = \overline{s}$. If φ is locally given by $\varphi_{|U} = [\overline{s}, u]$ then

$$(\phi_t^{-1})_* \varphi_{|\phi_t^{-1}(U)} = [\overline{s}_t, u \circ \phi_t].$$

Furthermore, we have mappings $g_t : U \cap \phi_t^{-1}(U) \to Mp(2n, \mathbb{R})$ given by

$$\overline{s}_t = \overline{s} g_t \, .$$

With

$$\begin{aligned}
&((\phi_t^{-1})_* e_1, \dots, (\phi_t^{-1})_* e_n, (\phi_t^{-1})_* f_1, \dots, (\phi_t^{-1})_* f_n) = \\
&\qquad = (e_1, \dots, e_n, f_1, \dots, f_n) \begin{pmatrix} \omega((\phi_t^{-1})_* e_l, f_k) & \omega((\phi_t^{-1})_* f_l, f_k) \\ \omega(e_k, (\phi_t^{-1})_* e_l) & \omega(e_k, (\phi_t^{-1})_* f_l) \end{pmatrix}_{k,l=1,\dots,n}
\end{aligned}$$

one derives

$$\rho(g_t) = \begin{pmatrix} \omega((\phi_t^{-1})_* e_l, f_k) & \omega((\phi_t^{-1})_* f_l, f_k) \\ \omega(e_k, (\phi_t^{-1})_* e_l) & \omega(e_k, (\phi_t^{-1})_* f_l) \end{pmatrix}_{k,l=1,\dots,n},$$

where $\rho : Mp(2n, \mathbb{R}) \to Sp(2n, \mathbb{R})$ denotes the double covering. With

$$\mathcal{L}_X Y = \frac{d}{dt}(\phi_t^{-1})_* Y = [X, Y]$$

for all vector fields Y on M, one sees

$$\begin{aligned}
\frac{d}{dt}\rho(g_t)_{|t=0} &= \begin{pmatrix} \omega(\mathcal{L}_X e_l, f_k) & \omega(\mathcal{L}_X f_l, f_k) \\ \omega(e_k, \mathcal{L}_X e_l) & \omega(e_k, \mathcal{L}_X f_l) \end{pmatrix}_{k,l=1,\dots,n} \\
&= \begin{pmatrix} \omega([X,e_l], f_k) & 0 \\ 0 & \omega(e_k, [X, f_l]) \end{pmatrix}_{k,l=1,\dots,n} \\
&\qquad + \begin{pmatrix} 0 & \omega([X,f_l], f_k) \\ 0 & 0 \end{pmatrix}_{k,l=1,\dots,n} \\
&\qquad + \begin{pmatrix} 0 & 0 \\ \omega(e_k, [X, e_l]) & 0 \end{pmatrix}_{k,l=1,\dots,n} .
\end{aligned}$$

Having $\mathcal{L}_X \omega = 0$, we conclude

$$\omega(e_k, [X, f_l]) = X(\omega(e_k, f_l)) - \omega([X, e_k], f_l) = -\omega([X, e_k], f_l)$$

as well as

$$\omega([X, f_l], f_k) = \omega([X, f_k], f_l) \qquad \text{and} \qquad \omega([X, e_k], e_l) = \omega([X, e_l], e_k) \ .$$

We obtain

$$\begin{aligned}
\frac{d}{dt}\rho(g_t)_{|t=0} &= \sum_{k,l=1}^{n} \Bigg\{ \omega([X,e_l], f_k) \begin{pmatrix} B_{kl} & 0 \\ 0 & -B_{lk} \end{pmatrix} \\
&\qquad + \frac{1}{2}\omega([X, f_l], f_k) \begin{pmatrix} 0 & B_{kl} + B_{lk} \\ 0 & 0 \end{pmatrix} \\
&\qquad + \frac{1}{2}\omega(e_k, [X, e_l]) \begin{pmatrix} 0 & 0 \\ B_{kl} + B_{lk} & 0 \end{pmatrix} \Bigg\} \\
&= \sum_{k,l=1}^{n} \left\{ \omega([X,e_l], f_k) X_{kl} + \frac{1}{2}\omega([X, f_l], f_k) Y_{kl} + \frac{1}{2}\omega(e_k, [X, e_l]) Z_{kl} \right\} \\
&= \frac{1}{2}\sum_{k,l=1}^{n} \rho_*\big(\omega([X,e_l], f_k)(a_k \cdot b_l + b_l \cdot a_k) + \omega(f_k, [X, f_l]) a_k \cdot a_l \\
&\qquad + \omega(e_k, [X, e_l]) b_k \cdot b_l\big).
\end{aligned}$$

With

$$\frac{d}{dt}\rho(g_t)_{|t=0} = \rho_*\left(\frac{d}{dt} g_{t|t=0}\right),$$

the definition of the Clifford multiplication, Proposition 1.2, equation (2.7), and relation

(1.6), we compute on U

$$
\begin{aligned}
\mathcal{L}_X\varphi &= \frac{d}{dt}[s_t, u\circ\phi_t]_{|t=0} \\
&= \frac{d}{dt}[sg_t, u\circ\phi_t]_{|t=0} \\
&= \frac{d}{dt}[s, L(g_t)(u\circ\phi_t)]_{|t=0} \\
&= [s, L_*(\frac{d}{dt}g_{t|t=0})u + \frac{d}{dt}u\circ\phi_{t|t=0}] \\
&= X(\varphi) - \frac{i}{2}\sum_{k,l=1}^{n}\{\omega([X,e_l],f_k)(e_k\cdot f_l + f_l\cdot e_k) \\
&\qquad +\omega(f_k,[X,f_l])e_k\cdot e_l + \omega(e_k,[X,e_l])f_k\cdot f_l\}\cdot\varphi \\
&= X(\varphi) - \frac{i}{4}\sum_{k=1}^{n}\{[X,e_k]\cdot f_k + f_k\cdot[X,e_k] - [X,f_k]\cdot e_k - e_k\cdot[X,f_k]\}\cdot\varphi \\
&= X(\varphi) - \frac{i}{4}\sum_{k=1}^{n}\{\nabla_X e_k\cdot f_k - \nabla_{e_k}X\cdot f_k + f_k\cdot\nabla_X e_k - f_k\cdot\nabla_{e_k}X \\
&\qquad -\nabla_X f_k\cdot e_k + \nabla_{f_k}X\cdot e_k - e_k\cdot\nabla_X f_k + e_k\cdot\nabla_{f_k}X\}\cdot\varphi \\
&= X(\varphi) + \frac{i}{2}\sum_{k=1}^{n}\{e_k\cdot\nabla_X f_k - f_k\cdot\nabla_X e_k\}\cdot\varphi \\
&\qquad +\frac{i}{4}\sum_{k=1}^{n}\{i\omega(e_k,\nabla_X f_k) - i\omega(f_k,\nabla_X e_k)\}\varphi \\
&\qquad +\frac{i}{2}\sum_{k=1}^{n}\{\nabla_{e_k}X\cdot f_k - \nabla_{f_k}X\cdot e_k\}\cdot\varphi \\
&\qquad +\frac{i}{4}\sum_{k=1}^{n}\{i\omega(\nabla_{e_k}X,f_k) - i\omega(\nabla_{f_k}X,e_k)\}\varphi \\
&= \nabla_X\varphi - \frac{1}{4}\sum_{k=1}^{n}X(\omega(e_k,f_k))\varphi \\
&\qquad +\frac{i}{2}\sum_{k=1}^{n}\{\nabla_{e_k}X\cdot f_k - \nabla_{f_k}X\cdot e_k\}\cdot\varphi \\
&= \nabla_X\varphi + \frac{i}{2}\sum_{k=1}^{n}\{\nabla_{e_k}X\cdot f_k - \nabla_{f_k}X\cdot e_k\}\cdot\varphi\,,
\end{aligned}
$$

which proves the proposition. □

As it is well known, the commutator of two Hamiltonian vector fields is a Hamiltonian vector field, too. Ideed, if $X = X_h$ is given by the function h and $Y = X_g$ by a function g then the commutator is the Hamiltonian vector field defined by the Poisson bracket of g and h, i.e.

$$[X_h, X_g] = -X_{\{h,g\}} \ . \tag{2.8}$$

For the Lie derivative in the direction of the commutator one has the following relation.

Corollary 2.3 *Let $\varphi \in \Gamma(\mathcal{Q})$ a symplectic spinor field and let X, Y are Hamiltonian vector fields on (M, ω), then*

$$\mathcal{L}_{[X,Y]}\varphi = [\mathcal{L}_X, \mathcal{L}_Y]\varphi.$$

Proof: Using (2.8) and Proposition 2.2, this proof is immediate. □

In case that M is not closed, all considerations hold true locally.

3 The Lie Derivative as Schrödinger Equation

This section illustrates how the Schrödinger equation for a quadratic Hamiltonian function relates to the Lie derivative of a constant symplectic spinor field over $\mathbb{R}^{2n}$.

3.1 The Schrödinger equation for quadratic Hamiltonians

We consider quadratic Hamiltonians H of the form $H(z) = z^\top Qz$ for $z \in \mathbb{R}^{2n}$, where Q is any real $2n \times 2n$-matrix. In general, one could add an additional absolut real term. But, this is completely inessential, because it does not play any role for the dynamics of the system. Or, physically speaking, the choice of the zero-energy-level is arbitrary.

Lemma 3.1 *Let $H : \mathbb{R}^{2n} \to \mathbb{R}$ be a quadratic Hamiltonian on $\mathbb{R}^{2n}$, which is given by $H(z) = z^\top Qz$, where Q is any $2n \times 2n$-matrix. Then, there exists a $2n \times 2n$-matrix $A \in \mathfrak{sp}(2n, \mathbb{R})$ such that the Hamiltonian vector field X_H of H is given by $X_H(z) = Az$ for $z \in \mathbb{R}^{2n}$.*

Proof: Let $\gamma(t)$ be a curve in $\mathbb{R}^{2n}$ with $\gamma(0) = z$ and $\dot\gamma(0) = w$. Then

$$dH(w)_z = \frac{d}{dt}H(\gamma(t))_{|t=0} = \frac{d}{dt}(\gamma(t))^\top Q(\gamma(t))_{|t=0} = w^\top Qz + z^\top Qw = w^\top(Q + Q^\top)z \ .$$

On the other hand, the Hamiltonian vector field X_H of H is given by

$$dH(w) = \omega_0(X_H, w) = -\langle Jw, X_H\rangle = -(Jw)^\top X_H = -w^\top J^\top X_H = w^\top JX_H \ .$$

Thus, at any point $z \in \mathbb{R}^{2n}$ we have $JX_H(z) = (Q + Q^\top)z$, and consequently

$$X_H(z) = -J(Q + Q^\top)z = J^\top(Q + Q^\top)z = ((Q + Q^\top)J)^\top z \ .$$

Taking $A = -J(Q + Q^\top)$, we have $X_H(z) = Az$ for $z \in \mathbb{R}^{2n}$. Furthermore, this A satisfies

$$A^\top J = -(Q + Q^\top)J^\top J = -(Q^\top + Q)$$

as well as

$$-JA = -(Q + Q^\top) \ ,$$

which gives $A^\top J + JA = 0$, or equivalently, $A \in \mathfrak{sp}(2n, \mathbb{R})$. □

Each quadratic Hamiltonian on $\mathbb{R}^{2n}$ can be written as a linear combination, i.e. as a sum of multiples of the functions on $\mathbb{R}^{2n}$ given by the expressions

$$\begin{aligned} H^1_{jk}(p,q) &= p_j p_k \ , \\ H^2_{jk}(p,q) &= q_j q_k \ , && \text{and} \\ H^3_{jk}(p,q) &= p_j q_k = \frac{1}{2}(p_j q_k + q_k p_j) && \text{for} \quad j,k = 1,\ldots,n \ . \end{aligned}$$

We call these functions *generating quadratic Hamiltonians.*

Lemma 3.2 *For the generating quadratic Hamiltonians the corresponding elements in* $\mathfrak{sp}(2n, \mathbb{R})$ *due to Lemma 3.1 are given in the following way.*

(1) *If* $H = H^1_{jk}$, *then* $A = -Z_{jk} = -Y^\top_{jk}$.

(2) *If* $H = H^2_{jk}$, *then* $A = Y_{jk} = Z^\top_{jk}$.

(3) *If* $H = H^3_{jk}$, *then* $A = X^\top_{jk}$.

Proof:

(1) $H = H^1_{jk}$ is given by $Q = \begin{pmatrix} B_{jk} & 0 \\ 0 & 0 \end{pmatrix}$. Thus

$$A = -\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} B_{jk} + B_{kj} & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ -B_{jk} - B_{kj} 0 & \end{pmatrix} = -Z_{jk} = -Y^\top_{jk} \ .$$

(2) $H = H^2_{jk}$ is given by $Q = \begin{pmatrix} 0 & 0 \\ 0 & B_{jk} \end{pmatrix}$, which implies

$$A = -\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 0 & B_{jk} + B_{kj} \end{pmatrix} = \begin{pmatrix} 0 & B_{jk} + B_{kj} \\ 0 & 0 \end{pmatrix} = Y_{jk} = Z^\top_{jk} \ .$$

(3) Finally, $H = H^3_{jk}$ is given by $Q = \frac{1}{2}\begin{pmatrix} 0 & B_{jk} \\ B_{kj} & 0 \end{pmatrix}$. This yields

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & B_{jk} \\ B_{kj} & 0 \end{pmatrix} = \begin{pmatrix} B_{kj} & 0 \\ 0 & -B_{jk} \end{pmatrix} = X^\top_{jk} \, .$$

$\square$

Let $\mathcal{H}$ denote the Hamilton operator which is given by H via normal ordering quantization, i.e. one obtains $\mathcal{H}$ by replacing in H formally the variable p_j by the multiplication operator ix_j and q_k by the operator $\frac{\partial}{\partial x_k}$. Thereby "normal ordering" means that the expression $p_j q_k = \frac{1}{2}(p_j q_k + q_k p_j)$ is replaced be the operator $\frac{i}{2}\left(x_j \frac{\partial}{\partial x_k} + \frac{\partial}{\partial x_k} x_j\right)$.

Corollary 3.3 *For the quadratic Hamiltonian H let A be given by Lemma 3.1 and let $\mathcal{H}$ be the Hamilton operator given via normal ordering quantization. Then, one has the relation*

$$L_* \circ \rho_*^{-1}(A^\top) = -i\mathcal{H} \, .$$

Proof: Since L_* and ρ_*^{-1} are linear and H is a linear combination of the generating quadratic Hamiltonians, it suffices to prove the assertion for the generating quadratic Hamiltonians. Then, by Lemma 1.1, Lemma 3.2, and Proposition 1.2 one has

$$\begin{aligned} L_* \circ \rho_*^{-1}(-Y_{jk}) &= ix_j x_k = -i(ix_j)(ix_k) = -i\mathcal{H} && \text{for } H = H^1_{jk}, \\ L_* \circ \rho_*^{-1}(Z_{jk}) &= -i\frac{\partial^2}{\partial x_j \partial x_k} = -i\mathcal{H} && \text{for } H = H^2_{jk}, \quad \text{and} \\ L_* \circ \rho_*^{-1}(X_{jk}) &= -\frac{i}{2}\left(ix_j \frac{\partial}{\partial x_k} + \frac{\partial}{\partial x_k} ix_j\right) = -i\mathcal{H} && \text{for } H = H^3_{jk} \, . \end{aligned}$$

$\square$

We consider a quadratic Hamiltonian H on $\mathbb{R}^{2n}$ with $A \in \mathfrak{sp}(2n, \mathbb{R})$ given according to Lemma 3.1. Then, we consider the family $S_t \in Sp(2n, \mathbb{R})$ of symplectic matrices defined by $S_t = \exp(tA^\top)$ for $t \in \mathbb{R}$. We lift this family of symplectic matrices into the double covering of $Sp(2n, \mathbb{R})$. That is, we consider the family $M_t \in Mp(2n, \mathbb{R})$ given by $\rho(M_t) = S_t$ such that M_0 is the unit element in $Mp(2n, \mathbb{R})$.

Definition 3.4 *For fixed $\psi_0 \in \mathcal{S}(\mathbb{R}^n)$ we define $\psi(t, x) := L(M_t)(\psi_0)(x)$, where $t \in \mathbb{R}$ and $x \in \mathbb{R}^n$.*

Furthermore let $\psi(t)$ be the curve in $\mathcal{S}(\mathbb{R}^n)$ given by $\psi(t)(x) := \psi(t, x)$ for $x \in \mathbb{R}^n$, i.e. $\psi(t) = L(M_t)\psi_0$.

Proposition 3.5 *$\psi(t)$ satisfies the Schrödinger equation*

$$\frac{d}{dt}\psi(t)_{|t=0} = -i\mathcal{H}(\psi_0)\ .$$

Proof: We have

$$\frac{d}{dt}\psi(t)_{|t=0} = \frac{d}{dt}L(M_t)(\psi_0)_{|t=0} = L_*\left(\frac{d}{dt}M_{t|t=0}\right)(\psi_0)\ .$$

The definition of M_t gives

$$\rho_*\left(\frac{d}{dt}M_{t|t=0}\right) = \frac{d}{dt}\rho(M_t)_{|t=0} = \frac{d}{dt}S_{t|t=0} = \frac{d}{dt}\exp(tA^\top)_{|t=0} = A^\top\ .$$

Hence

$$\frac{d}{dt}M_{t|t=0} = \rho_*^{-1}(A^\top)$$

and finally

$$\frac{d}{dt}\psi(t)_{|t=0} = L_* \circ \rho_*^{-1}(A^\top)(\psi_0) = -i\mathcal{H}(\psi_0)$$

by the previous Lemma. □

Let us now give the announced interpretation of the Lie derivative.

3.2 Interpretation as Lie derivative

In fact, the Schrödinger equation above gives the Lie derivative of a constant symplectic spinor field φ_0 on $\mathbb{R}^{2n}$ in the direction of the Hamiltonian vector field X.

First observe that the symplectic standard basis $\{a_1, \ldots, a_n, b_1, \ldots, b_n\}$ gives a global section s of the symplectic frame bundle R of $\mathbb{R}^{2n}$. Then $\overline{s}$ denotes a lift of s into the canonical metaplectic structure P of $\mathbb{R}^{2n}$.

Now, if ψ_0 is any fixed function in $\mathcal{S}(\mathbb{R}^n)$ the symplectic spinor field φ_0 over $\mathbb{R}^{2n}$ is defined to be

$$\varphi_0 = [\overline{s}, \psi_0]\ .$$

Further, we consider the family $\{\phi_t : \mathbb{R}^{2n} \to \mathbb{R}^{2n}\}_{t\in\mathbb{R}}$ given by

$$\phi_t(z) := \exp(tA)z \qquad \text{for } z \in \mathbb{R}^{2n}\ ,$$

where A denotes the matrix according to Lemma 3.1. Then,

$$\frac{d}{dt}\phi_t(z)_{|t=0} = Az = X(z)\ ,$$

which says that $\{\phi_t\}_{t\in\mathbb{R}}$ is exactly the one-parametergroup of diffeomorphisms induced by the Hamiltonian vector field X. Recalling the computations in the proof of Proposition 2.2, one has

$$(\phi_t^{-1})_*\varphi_0 = [\overline{s}_t, \psi_0]$$

with $s_t = s\exp(-tA)$ and $\overline{s}_t$ its lift to P. Since P is an $Mp(2n,\mathbb{R})$-principal fibre bundle, we obtain a family of elements $N_t \in Mp(2n,\mathbb{R})$ such that

$$\overline{s}_t = \overline{s}N_t \qquad \text{with} \qquad \rho(N_t) = \exp(-tA)\ .$$

Hence,

$$(\phi_t^{-1})_*\varphi_0 = [\overline{s}_t, \psi_0] = [\overline{s}N_t, \psi_0] = [\overline{s}, L(N_t)\psi_0]\ .$$

For a fixed element $\tilde{J} \in \rho^{-1}(J)$ the metaplectic representation was given by $L(\tilde{J}) = i^{\frac{n}{2}}\mathcal{F}^{-1}$ (cf. equation (1.5)). Thus, $i^{-\frac{n}{2}}L(\tilde{J}) \circ \mathcal{F} = id$. Using relation (1.1) we obtain

$$\rho(N_t)\rho(\tilde{J}) = \exp(-tA)J = (\exp(tA))^{-1}J = J(\exp(tA))^\top = J\exp(tA^\top) = \rho(\tilde{J})\rho(M_t)\ ,$$

where M_t is given above. Consequently,

$$L(N_t) \circ L(\tilde{J}) = L(\tilde{J}) \circ L(M_t)\ .$$

Altogether, we arrive at

$$(\phi_t^{-1})_*\varphi_0 = i^{-\frac{n}{2}}[\overline{s}, L(N_t) \circ L(\tilde{J}) \circ \mathcal{F}\psi_0] = i^{-\frac{n}{2}}[\overline{s}, L(\tilde{J}) \circ L(M_t) \circ \mathcal{F}\psi_0]\ .$$

Finally, we compute the Lie derivative of φ_0 in the direction of X and obtain, by Definition 2.1,

$$\mathcal{L}_X\varphi_0 = i^{-\frac{n}{2}}\left[\overline{s}, L(\tilde{J})\left(\frac{d}{dt}L(M_t)(\mathcal{F}\psi_0)_{|t=0}\right)\right] = -i[\overline{s}, \mathcal{F}^{-1} \circ \mathcal{H} \circ \mathcal{F}(\psi_0)]\ .$$

Here, the Fourier transform $\mathcal{F}$ means the transition between position and momentum representations.

Concluding Remarks

Fixing a compatible almost complex structure for (M,ω), Andreas Klein introduced a globally defined Fourier transform acting on symplectic spinor fields. See [11]. If one would define a Hamilton operator $\hat{\mathcal{H}}$ acting on symplectic spinor fields in the way that

$$\hat{\mathcal{H}}[\overline{s}, \psi] := [\overline{s}, \mathcal{H}\psi]\ ,$$

however, this does not work in general. The reason is that $\hat{\mathcal{H}}$ is not well defined by this relation.

But, setting formally

$$\mathfrak{q}(h)\varphi := i\mathcal{L}_{X_h}\varphi \ , \tag{3.9}$$

equation (2.3) gives

$$\begin{aligned} \mathfrak{q}(\{h,g\})\varphi &= i\mathcal{L}_{X_{\{h,g\}}}\varphi = -i\mathcal{L}_{[X_h,X_g]}\varphi = -i\mathcal{L}_{X_h}\circ\mathcal{L}_{X_g}\varphi + i\mathcal{L}_{X_h}\circ\mathcal{L}_{X_h}\varphi \\ &= i\mathfrak{q}(h)\circ\mathfrak{q}(g)\varphi - i\mathfrak{q}(g)\circ\mathfrak{q}(h)\varphi = i[\mathfrak{q}(h),\mathfrak{q}(g)]\varphi \ , \end{aligned}$$

which is in fact the "magic" Heisenberg relation

$$[\mathfrak{q}(h),\mathfrak{q}(g)]\varphi = -i\mathfrak{q}(\{h,g\})\varphi \ .$$

We do not claim that (3.9) gives a quantization procedure for arbitrary Hamiltonians over any symplectic manifold, although this expression makes sense in the general situation. We deduced the Heisenberg relation purely formal.

References

[1] **Alekseevsky, D. V., Cortes, V., Devchand, C.**, and **Semmelmann, U. :** *Killing spinors are Killing vector fields in Riemannian supergeometry.* Journal of Geometry and Physics **26**, 37-50 (1998)

[2] **Bourguignon, J.-P.**, and **Gauduchon, P. :** *Spineurs, Opérateurs de Dirac et Variations de Métriques.* Commun. Math. Phys. **144**, 581-599 (1992)

[3] **Crumeyrolle, A. :** *Orthogonal and Symplectic Clifford Algebras.* Kluwer Academic Publishers 1990

[4] **Dabrowski, L.**, and **Percacci, R. :** *Spinors and diffeomorphisms.* Commun. Math. Phys. **106**, 691-704 (1986)

[5] **Guillemin, V.**, and **Sternberg, S. :** *Symplectic techniques in physics.* Cambridge University Press 1990

[6] **de Gosson, M. :** *Maslov classes, metaplectic representation and Lagrangian quantization.* Akademie Verlag Berlin 1997

[7] **Habermann, K. :** *The graded algebra and the derivative $\overline{\mathcal{L}}$ of spinor fields related to the twistor equation.* Journal of Geometry and Physics **18**, 131-146 (1996)

[8] **Habermann, K. :** *The Dirac Operator on Symplectic Spinors.* Annals of Global Analysis and Geometry **13**, 155-168 (1995)

[9] **Habermann, K. :** *Basic Properties of Symplectic Dirac Operators.* Commun. Math. Phys. **184**, 629-652 (1997)

[10] **Kashiwara, M.**, and **Vergne, M. :** *On the Segal-Shale-Weil Representations and Harmonic Polynomials.* Inventiones math. **44**, 1-47 (1978)

[11] **Klein, A. :** *Eine Fourietransformation für symplektische Spinoren und Anwendungen in der Quantisierung.* diploma thesis, Technical University Berlin 2000

[12] **Kosmann, Y. :** *Dérivées de Lie des spineurs.* Ann. Mat. Pura ed Appl. **91**, 317-395 (1972)

[13] **Kostant, B. :** *Symplectic Spinors.* Symposia Mathematica. vol XIV (1974)

[14] **Moroianu, A. :** *On the Infinitesimal Isometries of Manifolds with Killing Spinors.* Journal of Geometry and Physics **35**, 63-74 (2000)

[15] **Wallach, N. R. :** *Symplectic Geometry and Fourier Analysis.* MATH SCI PRESS 1977

received: September 16, 2002

Authors:

Katharina Habermann
Ernst-Moritz-Arndt-Universität Greifswald
Institut für Mathematik und Informatik
Jahnstrasse 15a
17487 Greifswald
Germany

e-mail: khaberma@mail.uni-greifswald.de

Andreas Klein
Humboldt-Universität zu Berlin
Institut für Mathematik
Unter den Linden 6
10099 Berlin
Germany

e-mail: klein@mathematik.hu-berlin.de

Rostock. Math. Kolloq. **57**, 93–98 (2003)

Subject Classification (AMS)
39A13,15A57

Lothar Berg

On the Representation of Continuous Solutions of Two-Scale Difference Equations at Dyadic Points

ABSTRACT. The paper gives some insight into the structure of continuous solutions of two-scale difference equations at dyadic points. An example is given in which the solution is estimated.

KEY WORDS. Two-scale difference equations, 2^l-slanted matrices, recursions

Let φ be a continuous compactly supported solution of the two-scaled difference equation (cf. [3])

$$\varphi\left(\frac{t}{2}\right) = \sum_{n=0}^{N} c_n \varphi(t-n) \tag{1}$$

($t \in \mathbb{R}$) with $N \in \mathbb{N}$ (in fact it must be $N \geq 2$), $c_n \in \mathbb{C}$, $c_0 c_N \neq 0$ and

$$\sum_{n=0}^{N} c_n = 2^M ,$$

($M \in \mathbb{N}$). In [2, Corollary 2.5] it was shown that the restriction of φ to $[0,1]$ possesses at dyadic points the representation

$$\varphi\left(\frac{k}{2^l}\right) = c_0^l \sum_{j=1}^{N-1} y_{N+k-j} \varphi(j) \tag{2}$$

($k, l \in \mathbb{N}_0$, $0 \leq k \leq 2^l$) where the coefficients are defined by the initial values

$$y_1 = \cdots = y_{N-1} = 0, \quad y_N = 1 \tag{3}$$

and the recursions

$$c_0 y_k = \sum_{j=\lceil \frac{k}{2} \rceil}^{\lfloor \frac{N+k}{2} \rfloor} c_{N+k-2j} y_j \,. \tag{4}$$

Here $\lfloor . \rfloor$ denotes the floor, and $\lceil . \rceil$ the ceiling function, cf. [4, p. 52]. It is suitable to use the extensions $y_j = 0$ for $j < 0$ and $c_n = 0$ for both $n < 0$ and $n > N$, respectively, and to introduce the infinite two-scale matrix

$$\mathbf{A} = (c_{2j-k}) \qquad (1 \le j, k)\,.$$

Then, for $l \in \mathbb{N}$, the matrix $\mathbf{A}^l$ possesses the entries

$$c_0^l y_{2^l+N-1},\ c_0^l y_{2^l+N-2},\ c_0^l y_{2^l+N-3},\ \dots \tag{5}$$

in its first row, cf. [2, Theorem 2.4]. It can easily be seen that $\mathbf{A}^l$ is a 2^l-slanted matrix, i.e.

$$\mathbf{A}^l = \left(c^{(l)}_{2^l j-k}\right) \qquad (1 \le j, k) \tag{6}$$

where $c_n^{(1)} = c_n$ and

$$c^{(l+m)}_{2^{l+m}j-k} = \sum_{i=1}^{\infty} c^{(l)}_{2^l j-i} c^{(m)}_{2^m i-k}\,,$$

in particular $c_0^{(l)} = c_0^l$, $c^{(l)}_{(2^l-1)N} = c_N^l$, and $c_n^{(l)} = 0$ for both $n < 0$ and $n > (2^l-1)N$, respectively.

For our next considerations we need the following submatrices of $\mathbf{A}$:

$$A_l = (c_{2j-k}) \qquad (1 \le j, k \le 2^l + N - 1)$$

with $l \in \mathbb{N}_0$. If A_0 is diagonalizable then there exist matrices $\Lambda = \mathrm{diag}(\lambda_1, \dots, \lambda_N)$ and E with

$$A_0 = E^{-1} \Lambda E\,, \tag{7}$$

where the j-th row $(e_{j1}, \dots, e_{jN})$ of E is a left eigenvector of A_0 to the eigenvalue $\lambda_i (j \in \{1, \dots, N\})$. This eigenvector can be continued to a left eigenvector $(e_{j1}, e_{j2}, \dots)$ of $\mathbf{A}$ to the same eigenvalue. The matrix of these eigenvectors we denote by

$$\mathbf{E} = (e_{jk}) \qquad (1 \le j \le N,\ 1 \le k)\,,$$

and we also need the finite submatrices

$$G_l = (e_{jk}) \qquad (1 \le j \le N,\ 1 \le k \le 2^l + N - 1)\,. \tag{8}$$

Theorem *Let A_0 be diagonalizable. Then with the foregoing notations the first $2^l + N - 1$ terms of (5) can be represented as*

$$c_0^l y_{2^l+N-k} = \sum_{i=1}^{N} \lambda_i^l f_{1i} e_{ik} \qquad (k = 1, \dots, 2^l + N - 1, l \in \mathbb{N}_0) \tag{9}$$

where $(f_{11}, \dots, f_{1N})$ is the first row of $E^{-1} = (f_{jk})$.

Proof: The right-hand sides of (9) for $k = 1, \ldots, 2^l + N - 1$ are the entries of the first row of the matrix $E^{-1}\Lambda G_l$. We have to show that they coincide with the first $2^l + N - 1$ entries of the first row of $\mathbf{A}^l$. For $l = 0$ this is clear. For $l \geq 1$ the matrices $\mathbf{A}$ and A_l can be splitted into the following block forms

$$\mathbf{A} = \begin{pmatrix} A_l & * \\ O & * \end{pmatrix}, \quad A_l = \begin{pmatrix} A_0 & * \\ O & * \end{pmatrix},$$

where the asterisks indicate suitable submatrices and O suitable zero matrices. Hence,

$$\mathbf{A}^l = \begin{pmatrix} A_l^l & * \\ O & * \end{pmatrix}, \quad A_l^l = \begin{pmatrix} A_0^l & * \\ O & * \end{pmatrix}, \tag{10}$$

where $A_l^l = (c_{2^l j-k}^{(l)})$ $(1 \leq j, k \leq 2^l + N - 1)$ using the notation (6). Since $c_{2^l j-k}^{(l)} = 0$ for $2^l j - k > (2^l - 1)N$, and therefore for both $N + 1 \leq j$ and $1 \leq k \leq 2^l + N - 1$, we have in fact

$$A_l^l = \begin{pmatrix} A_0^l & * \\ O & O \end{pmatrix}. \tag{11}$$

Comparison of the Jordan normal form

$$A_l = E_l^{-1} \begin{pmatrix} \Lambda & O \\ O & J \end{pmatrix} E_l \tag{12}$$

with (7) and (8) shows that the outer factors must have the block forms

$$E_l^{-1} = \begin{pmatrix} E^{-1} & * \\ O & * \end{pmatrix}, \; E_l = \begin{pmatrix} E & * \\ O & * \end{pmatrix} = \begin{pmatrix} G_l \\ * \end{pmatrix}.$$

Comparison of (11) with (12) implies that $J^l = 0$ and therefore

$$A_l^l = \begin{pmatrix} E^{-1} & * \\ O & * \end{pmatrix} \begin{pmatrix} \Lambda^l G_l \\ O \end{pmatrix} = \begin{pmatrix} E^{-1}\Lambda^l G_l \\ O \end{pmatrix}.$$

Now, the assertion follows from (10) $\quad \square$

Remarks $\quad 1^\circ$. Choosing in (2) $k = 2^l - m$ then by means of (9) with $k = m + j$ we get some insight into the structure of $\varphi\left(1 - \frac{m}{2^l}\right)$, $0 \leq m \leq 2^l$. Though the result can be used for explicit calculations of φ, this is not recommended.

2°. The entries of the eigenvectors $(e_{i1}, e_{i2}, \ldots)$ satisfy analogous recursions as in (4), namely

$$\lambda_i e_{ik} = \sum_{j=\lceil \frac{k}{2} \rceil}^{\lfloor \frac{N+k}{2} \rfloor} c_{2j-k} e_{ij}\,.$$

3°. In the case $N=2$ formula (9) was already set up (with other notations) in [1, (3.1)].

4°. The case that A_0 is non-diagonalizable can be treated with some more effort, cf. [1, (3.3)] in the case $N=2$.

5°. In [2, Proposition 2.7] it must be $m_0=0$.

6°. The first column of E^{-1} is a right eigenvector of A_0 to the eigenvector 1. This implies $\varphi(j)=f_{j1}$ (up to a constant factor), cf. [2, (2.4)] with $t=1$.

7°. Formula (9) can be simplified if the entries f_{1k} of the first row of E^{-1} are normlized according to $f_{1k}=1$ so far as $f_{1k}\neq 0$. But it is also possible that $f_{1k}=0$ for a fixed k as in the folowing

Example Choosing $c_0=\frac{1}{4}$, $c_3=1$, $c_4=\frac{3}{4}$ and $c_n=0$ otherwise, so that $N=4$, then $\Lambda=\operatorname{diag}\left(1\,\frac{1}{2}\,-\frac{1}{2}\,\frac{3}{4}\right)$ and

$$E^{-1}=\begin{pmatrix}0&1&1&1\\0&2&-2&3\\1&-3&1&-9\\0&0&0&5\end{pmatrix},\qquad E=\begin{pmatrix}1&1&1&1\\\frac{1}{2}&\frac{1}{4}&0&-\frac{1}{4}\\\frac{1}{2}&-\frac{1}{4}&0&\frac{1}{20}\\0&0&0&\frac{1}{5}\end{pmatrix}.$$

Hence, (9) yields in particular

$$y_{2^l+3}=2^{l-1}(1+(-1)^l)\,,\quad y_{2^l+2}=(1-(-1)^l)2^{l-2}\,,$$
$$y_{2^l+1}=0\,,\quad y_{2^l}=\frac{1}{5}\left(3^l-(5-(-1)^l)2^{l-2}\right)$$

$(l\in\mathbb{N}_0)$. Formula (2) with $\varphi(3)=1$ and $\varphi(j)=0$ otherwise specializes to

$$\varphi\left(\frac{k}{2^l}\right)=\frac{1}{4^l}y_{k+1} \tag{13}$$

for $0\le k\le 2^l$. But (13) is even valid for $0\le k\le 3\cdot 2^l$, since $\varphi(t)=0$ for $t\le 0$ and (1) imply $\varphi(\frac{t}{2})=\frac{1}{4}\varphi(t)$ for $0\le t\le 3$. The recursions (4) specialize to

$$y_{2j}=3y_j+y_{j+2},\ y_{2j-1}=4y_j \tag{14}$$

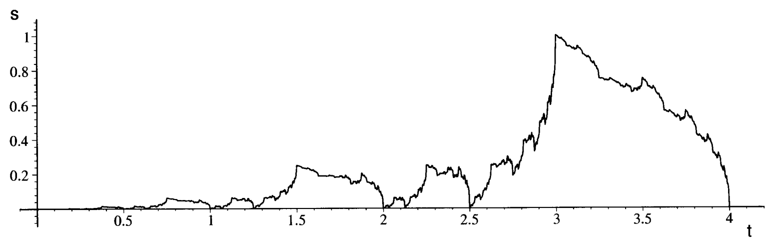
for $j\in\mathbb{N}$, and with the initial values (3) with $N=4$ we obtain for the first values

j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
y_j	0	0	0	1	0	0	4	3	0	4	0	3	16	12	12	13	0	0	16	15	0	16	12	21	64	60

where it easily follows by induction that

$$y_{(2^{2n}+1)2^m+1}=0 \tag{15}$$

for all $m,n\in\mathbb{N}_0$. The solution $s=\varphi(t)$ is plotted in the following picture:



Introducing the notations

$$x_n = y_{3\cdot 2^n+1},\ z_n = y_{3\cdot 2^n+3},\ u_n = y_{3\cdot 2^n+2},\ v_n = y_{3\cdot 2^n},\ w_n = y_{3\cdot 2^n-1}$$

$(n \in \mathbb{N}_0)$ and using (14) we find the recursions

$$\begin{aligned} &x_n = 4x_{n-1},\ z_n = 12x_{n-2} + 4z_{n-2},\\ &u_n = 3x_{n-1} + z_{n-1},\ v_n = 3v_{n-1} + u_{n-1},\ w_n = 12v_{n-2} + 4u_{n-2}, \end{aligned}$$

and by means of the initial values from the forgoing table their solutions

$$\begin{aligned} &x_n = 4^n,\ z_n = 4^n + ((-1)^n - 3)2^{n-1},\ u_n = 4^n - ((-1)^n + 3)2^{n-2},\\ &v_n = 4^n + 3\cdot 2^{n-2} + \frac{1}{5}((-2)^{n-2} - 3^{n+2}),\ w_n = 4^n + 3\cdot 2^{n-1} + \frac{1}{5}((-2)^{n-1} - 4\cdot 3^{n+1})\,. \end{aligned}$$

Proposition *The solutions y_k of (14) with (3) for $N = 4$ satisfy the estimates*

$$0 \le y_k \le \left(\frac{k-1}{3}\right)^2 \tag{16}$$

$(k \in \mathbb{N})$ *where both bounds are sharp for infinitely many k.*

Proof: The first inequality of (16) follows from (14) and the initial values (3) with $N = 4$, the sharpness from (15). For $k = 3\cdot 2^n + 1$ $(n \in \mathbb{N}_0)$ the second inequality is in fact an equality in view of $x_n = 4^n$. For $1 \le k \le 3$ it is trivial. For $k \ne 3\cdot 2^n + 1$ and $k \ge 5$ we shall prove the better inequality

$$y_k \le \frac{1}{9}k(k-2)\,. \tag{17}$$

For $k = 3\cdot 2^n + 2$ $(n \in \mathbb{N}_0)$ we have

$$y_k = u_n \le 4^n - 2^{n-1} = \frac{1}{9}(k-2)\left(k - \frac{7}{2}\right)$$

and (17) is valid. For $k = 3\cdot 2^n - 2$ we have

$$y_k = 3w_{n-1} + x_{n-1} \le 4^n + \frac{12}{5}(2^n - 3^n)$$

and (17) is valid when $n \geq 2$ ($n = 1$ corresponds to $y_4 = 1$).

In order to complete the proof we introduce the sets $M_n = \{3 \cdot 2^n + 2, \dots, 3 \cdot 2^{n+1}\}$ ($n \in \mathbb{N}_0$). The inequality (17) is valid for $k \in M_0 = \{5, 6\}$. If (17) is valid for $k \in M_n$ then by means of the recursions (14) it follows that (17) is valid for the odd k from M_{n+1}. Analogously, we see that (17) is also valid for the even k from M_{n+1} if we simultaneously take into account the already treated two special cases. Hence by induction, (17) is valid for all $k \in \bigcup\limits_{n=0}^{\infty} M_n$ $\square$

In view of (13) and the continuity of φ we immediately get the

Corollary *For $0 \leq t \leq 3$ the solution of our example for (1) with $\varphi(3) = 1$ satisfies the estimates*

$$0 \leq \varphi(t) \leq \frac{1}{9} t^2$$

where both bound are sharp for infinitely many t.

References

[1] **Berg, L.**, and **Krüppel, M. :** *On a simple system of discrete two-scale difference equations.* Z. Anal. Anwend. **19**, 999-1016 (2000)

[2] **Berg, L.**, and **Krüppel, M. :** *On discrete solutions of two-scale difference equations.* Rostock. Math. Kolloq. **56**, 63-79 (2002)

[3] **Daubechies, I.**, and **Lagarias, J. :** *Two-scale difference equations I. Existence and global regularity of solutions.* SIAM J. Math. Anal. **22**, 1388-1410 (1991)

[4] **Wirsching, G. J. :** *The Dynamical System Generated by the $3n + 1$ Function.* Lect. Notes Math. **1681**, 1-158 (1998)

received: October 18, 2002

Author:

Prof. i.R. Dr. Lothar Berg
Universität Rostock
Fachbereich Mathematik
18051 Rostock
Germany

e-mail: lothar.berg@mathematik.uni-rostock.de

Rostock. Math. Kolloq. **57**, 99–110 (2003)

Subject Classification (AMS)
54E05, 54E15, 54B20, 54A10, 54D30

Somashekhar Naimpally

Proximity and Hyperspace Topologies

Dedicated to my friend Professor Dr. Harry Poppe on his 70^{th} birthday.

ABSTRACT. In this paper we give a survey of the use of proximities in hyperspace topologies. A proximal hypertopology corresponding to a LO-proximity is a g eneralization of the well known Vietoris topology. In case we start with an EF-proximity, the proximal hypertopology equals the Hausdorff uniform topology corresponding to the totally bounded uniformity and, being contained in both the Vietoris and Hausdorff uniform topologies, serves as a bridge between the two. Wattenberg and Beer-Himmelberg-Prickry-Van Vleck showed that the locally finite hypertopology induced by a metrizable space is the sup of the Hausdorff metric topologies induced by all compatible metrics. Naimpally-Sharma showed that this follows from the fact that a Tychonoff space is normal iff its fine uniformity induces the locally finite hypertopology. Di Concilio-Naimpally-Sharma showed that in a Tychonoff space the fine uniformity induces the proximal locally finite hypertopology.

We study DELTA topologies introduced by Poppe, and their proximal variations. We show that a short proof can be given of the Beer-Tamaki result concerning the uniformizability of (proximal) DELTA hypertopologies via the Attouch-Wets approach used by Beer in dealing with the Fell topology. Finally we present a result concerning (Proximal) DELTA-U-hypertopolgies. Several new hypertopologies are introduced.

KEY WORDS AND PHRASES. proximity, hyperspace, Δ-topology, proximal Δ-topology, U-topology, ΔU-topology, proximal ΔU-topology, Function space, Vietoris topology, Fell topology, Hausdorff uniformity.

1 Introduction

Suppose (X, T) (respectively (X, V)) is a T_1 topological space (respectively a uniform space). Then it is well known that on $\mathrm{CL}(X)$, the hyperspace of all non-empty closed subsets of X, one can define Vietoris topology $\tau(\mathrm{V})$ (respectively a Hausdorff uniformity V_{H}) such that X

is topologically (respectively, uniformly) embedded in $\mathrm{CL}(X)$. But it is not known how one can define directly a proximity on the hyperspace of a given proximity space (X, δ). Nachman ([21]) tackled this problem in the case of an EF-proximity δ on X via Hausdorff uniformities associated with compatible uniformities on X. An attempt was made to use proximity in hyperspaces in [16] and a little later in [4]. Since the paper [16] remains unpublished and the paper [4] dealt with proximities in the context of metric spaces, an impression continues in the literature that proximal hypertopologies exist only in metric spaces. The aim of this paper is to correct this impression and show that proximal topologies can be defined using LO-proximities in any T_1 space. Recently there has been some work done in the general case. (See e. g. [9], [10]) See [1] for the latest results on compactness in function spaces via hyperspaces and (uniform) convergence structures.

(X, T) denotes a T_1-topological space and δ denotes any compatible LO-proximity on X. The symbol δ_0 denotes the fine LO-proximity and it is well known (Urysohn Theorem) that it is EF iff X is normal. If (X, T) is Tychonoff, then we generally choose δ to be EF. $\mathrm{CL}(X)$ denotes the family of all non-empty closed subsets of X and $\mathrm{K}(X)$ denotes the family of all non-empty compact subsets. We use the symbol Δ to denote a subfamily of $\mathrm{CL}(X)$ and we assume, without any loss of generality, that it is a cover of X and is closed under finite unions and contains all singletons.

For any set $E \subset X$ and $\mathsf{E} \subset \mathsf{T}$ we use the following notation:

$E^- = \{A \in \mathrm{CL}(X) : A \cap E \neq \emptyset\}$

$\mathsf{E}^- = \{A \in \mathrm{CL}(X) : A \cap E \neq \emptyset\}$ for each $E \in \mathsf{E}\}$

$E^{++} = \{A \in \mathrm{CL}(X) : A \ll E$ w. r. t. δ i. e. $A\ \underline{\delta}\ E^c\}$

$E^+ = \{A \in \mathrm{CL}(X) : A \subset E$ i. e. $A \ll E$ w. r. t. $\delta_0\}$

The **Δ-topology** $\tau(\Delta)$ is generated by a basis of the form $E^+ \vee \mathsf{E}^-$, where $E^c \in \Delta$ and $\mathsf{E} \subset \mathsf{T}$ is finite. ([26], [27])

The **proximal Δ-topology** (w. r. t. δ) $\sigma(\delta\Delta)$ is generated by a basis of the form $E^{++} \vee \mathsf{E}^-$, where $E^c \in \Delta$ and $\mathsf{E} \subset \mathsf{T}$ is finite. We omit δ if it is obvious from the context.

If in the above, the family E is locally finite, then we have the **locally finite Δ-topology** $\tau(\mathrm{LF}\Delta)$ and the **proximal locally finite Δ-topology** (w. r. t. δ) $\sigma(\mathrm{LF}\delta\Delta)$.

The **ΔU-topology** $\tau(\Delta\mathrm{U})$ is generated by a basis of the form $E^+ \vee \mathsf{E}^-$, where $E^c \in \Delta$ or $\mathrm{cl}E \in \Delta$ and $\mathsf{E} \subset \mathsf{T}$ is finite.

The **proximal ΔU-topology** (w. r. t. δ) $\sigma(\delta\Delta\mathrm{U})$ is generated by a basis of the form $E^{++} \vee \mathsf{E}^-$, where $E^c \in \Delta$ or $\mathrm{cl}E \in \Delta$ and $\mathsf{E} \subset \mathsf{T}$ is finite.

If in the above, the family E is locally finite, then we have the **locally finite ΔU-topology** $\tau(\mathrm{LF}\Delta\mathrm{U})$ and the **proximal locally finite ΔU-topology** (w. r. t. δ) σ $(\mathrm{LF}\delta\Delta\mathrm{U})$.

Well known special cases are:

(a) when $\Delta = \mathrm{CL}(X)$, $\tau(\Delta) = \tau(\mathrm{V})$ the **Vietoris or finite topology** ([20])
$\sigma(\delta\Delta) = \sigma(\delta)$ the **proximal topology** ([16])
$\tau(\mathrm{LF}\Delta) = \tau(\mathrm{LF})$ the **locally finite topology** ([19])
$\sigma(\mathrm{LF}\delta\Delta) = \sigma(\mathrm{LF}\delta)$ the **proximal locally finite topology** ([16])

To make the notation simpler, we'll omit δ from all proximal topologies whenever it is understood from the context: thus we'll use σ for $\sigma(\delta)$, $\sigma(\mathrm{LF})$ for $\sigma(\mathrm{LF}\delta)$ etc.

(b) When $\Delta = \mathrm{K}(X)$, $\tau(\Delta) = \tau(\mathrm{F})$ the **Fell topology** ([17])
and we define three new ones
$\sigma(\delta\Delta) = \sigma(\delta\mathrm{F})$ the **proximal Fell topology**
$\tau(\mathrm{LF}\Delta) = \tau(\mathrm{LFF})$ the **locally finite Fell topology**
$\sigma(\mathrm{LF}\delta\Delta) = \sigma(\mathrm{LF}\delta\mathrm{F})$ the **proximal locally finite Fell topology**
$\tau(\Delta\mathrm{U}) = \tau(\mathrm{U})$ the **U-topology** ([8])
and we define three new ones
$\sigma(\delta\Delta\mathrm{U}) = \sigma(\delta\mathrm{U})$ the **proximal U-topology**
$\tau(\mathrm{LF}\Delta\mathrm{U}) = \tau(\mathrm{LFU})$ the **locally finite U-topology**
$\sigma(\mathrm{LF}\delta\Delta\mathrm{U}) = \sigma(\mathrm{LF}\delta\mathrm{U})$ the **proximal locally finite U-topology**

Of course, if the proximity δ is EF or R (and so X is Tychonoff or regular respectively) then $\tau(\mathrm{F}) = \sigma(\mathrm{F})$, $\tau(\mathrm{LFF}) = \sigma(\mathrm{LFF})$, $\tau(\mathrm{U}) = \sigma(\mathrm{U})$ and $\tau(\mathrm{LFU}) = \sigma(\mathrm{LFU})$.

Many interesting properties of the Fell topology stem from the fact that it is also a proximal topology!. In generalizing results from the Fell topology to Δ-topologies, we find that some hold for $\tau(\Delta)$ and others for $\sigma(\Delta)$!!

(c) If (X, d) is a metric space, δ is the metric proximity induced by d and Δ denotes the ring generated by closed balls of non-negative radii, then

$\tau(\Delta) = \tau(\mathrm{B})$ the **Ball topology** ([2])
$\sigma(\Delta) = \sigma(\mathrm{B})$ the **proximal Ball topology** ([14])
and we introduce two new ones
$\tau(\mathrm{LF}\Delta) = \tau(\mathrm{LFB})$ the **locally finite Ball topology**,
$\sigma(\mathrm{LF}\Delta) = \sigma(\mathrm{LFB})$ the **proximal locally finite Ball topology**

In addition we have the well known **Hausdorff metric** d_{H} and the **Hausdorff metric topology** $\tau(d_{\mathrm{H}})$.

If (X, V) is a uniform space, then we have the **Hausdorff uniformity** V_{H} and the **Hausdorff uniform topology** $\tau(\mathsf{V}_{\mathrm{H}})$.

[2] is a standard reference on hyperspace topologies and we give below other relevant bibliography for the interested reader.

2 VIETORIS, PROXIMAL AND (PROXIMAL) LOCALLY FINITE TOPOLOGIES

Suppose (X, T) is a T_1 topological space, δ any compatible LO-proximity on X and δ_0 the fine LO-proximity. If (X, T) is Tychonoff, the fine EF-proximity is denoted by $\delta^{\#}$ (the functionally indistinguishable EF-proximity), the fine uniformity is denoted by $\mathsf{V}^{\#}$, and the finest totally bounded uniformity is denoted by V^{*}. If δ is a compatible EF-proximity on X, then $\Pi(\delta)$ is the family of all uniformities which induce δ and V_ω denotes the coarsest totally bounded member of $\Pi(\delta)$ ([25]). We note that since, in general, there are many proximities compatible with (X, T), proximal hypertopologies provide us with a large number of hypertopologies. For further details see [16].

Theorem 2.1 ([16]) (a) $\tau(\mathrm{V}) = \sigma(\delta_0)$

(b) $\tau(\mathrm{LF}) = \sigma(\mathrm{LF}\delta_0)$

(c) $\sigma \subset \sigma(\mathrm{LF})$ *and* $\tau(\mathrm{V}) \subset \tau(\mathrm{LF})$.

In each case $\subset$ *is replaced by* $=$ *if and only if* X *is feebly compact (i. e. every locally finite family of open sets in* X *is finite).*

(d) *In general* $\tau(\mathrm{V})$ *and* σ *are independent.*

(e) *If* $\delta < \delta'$ *and* δ *is* EF, *then* $\sigma(\delta) \subset \sigma(\delta')$ *and* $\sigma(\mathrm{LF}\delta) \subset \sigma.(\mathrm{LF}\delta')$.

Consequently, $\sigma \subset \tau(\mathrm{V})$.

(f) *If* δ *is EF and* $\delta \neq \delta_0$, *then* $\sigma \neq \tau(\mathrm{V})$ *and* $\sigma(\mathrm{LF}) \neq \tau(\mathrm{LF})$.

Corollary 2.2 ([16]) *If* δ *is* EF, *then* (a), (b) *and* (c) *are mutually equivalent and each implies* (d)*:*

(a) $\tau(\mathrm{V}) = \sigma$

(b) $\tau(\mathrm{LF}) = \sigma(\mathrm{LF})$

(c) $\delta = \delta_0$

(d) (X, T) *is normal.*

Corollary 2.3 ([4]) *If (X,d) is a metric space and δ is the metric proximity, then the following are equivalent:*

(a) (X,d) *is Atsuji or* UC

(i. e. every real valued continuous function on X is uniformly continuous.)

(b) $\tau(\mathrm{V}) \subset \tau(d_{\mathrm{H}})$

(c) $\delta = \delta_0$

(d) $\tau(\mathrm{V}) = \sigma$

Theorem 2.4 ([16]) *Suppose (X,δ) is an* EF*-proximity space and V and V_ω are in $\Pi(\delta)$. Then*

(a) $\sigma = \tau(\mathsf{V}_{\omega\mathrm{H}}) \subset \tau(\mathsf{V}_{\mathrm{H}}) \subset \sigma(\mathrm{LF}\delta^{\#}) \subset \tau(\mathrm{LF})$ *and* $\sigma = \tau(\mathsf{V}_{\mathrm{H}})$ *implies* $\mathsf{V} = \mathsf{V}_\omega$.

It follows that if (X,T) is normal, then $\tau(\mathrm{V}) = \tau(\mathsf{V}^*_{\mathrm{H}}) \subset \tau(\mathrm{LF})$.

(b) $\sigma = \tau(\mathsf{V}_{\omega\mathrm{H}}) \subset \tau(\mathrm{V}) \subset \tau(\mathrm{LF})$

(c) $\sigma(\mathrm{LF}\delta^{\#}) = \tau(\mathsf{V}^*_{\mathrm{H}})$

(d) $\sigma(\delta^{\#}) = \tau(\mathsf{V}^*_{\mathrm{H}})$

Corollary 2.5 ([24]) *The following are equivalent:*

(a) (X,T) *is normal.*

(b) δ_0 *is* EF.

(c) $\tau(\mathsf{V}^{\#}_{\mathrm{H}}) = \tau(\mathrm{LF})$.

(d) $\sigma(\delta^{\#}) = \tau(\mathrm{LF})$.

The following Hesse diagram shows the various relationships:

$$\begin{tikzcd}
& \tau(\mathrm{LF}) = \sigma(\mathrm{LF}\delta_0)\,. \arrow[dl, no head] \arrow[dr, no head] & \\
\sigma(\mathrm{LF}\delta^{\#}) = \tau(\mathsf{V}^{\#}_{\mathrm{H}}) \arrow[dr, no head] & & \tau(\mathrm{V}) = \sigma(\delta_0) \arrow[dl, no head] \\
& \sigma(\delta^{\#}) = \tau(\mathsf{V}^*_{\mathrm{H}}) \arrow[d, no head] & \\
& \sigma = \tau(\mathsf{V}_{\omega\mathrm{H}}) &
\end{tikzcd}$$

Remark 2.6 In [5] it was shown that if X is a metrizable space, the locally finite topology $\tau(\mathrm{LF})$ on $\mathrm{CL}(X)$ is the sup of the Hausdorff metric topologies $\{\tau(d_{\mathrm{H}})\}$ corresponding to equivalent compatible metrics $\{d\}$ on X. This result was generalized in ([24]) to : a Tychonoff space X is normal if and only if the locally finite topology $\tau(\mathrm{LF})$ equals $\tau(\mathsf{V}^{\#}_{\mathrm{H}})$, the topology induced by the Hausdorff uniformity corresponding to the fine uniformity $\mathsf{V}^{\#}_{\mathrm{H}}$. The question then arises: in a non-normal Tychonoff space what is $\tau(\mathsf{V}^{\#}_{\mathrm{H}})$? The answer was provided in ([16]) that it is precisely the proximal locally finite topology $\sigma(\mathrm{LF}\delta^{\#})$ induced by the fine proximity $\delta^{\#}$ on X. This shows the importance of proximal topologies in this problem.

3 (PROXIMAL) DELTA TOPOLOGIES

Poppe ([26], [27]) first studied Δ-topologies as generalizations of the Fell topology. On the other hand, many workers in this area have used, in the context of metric spaces, *bounded* sets to study new hyperspace topologies e.g. the bounded Vietoris (proximal) topology ([12]), Attouch-Wets topology ([3]) etc. It is not widely known that boundedness can also be defined in general topological spaces in an abstract way ([18]) and this provides a technique to give simple proofs and generalizations of several results in this area. In this section we give a glimpse of this approach and refer the interested readers to ([15]) for further information.

A boundedness $\mathbf{H}$ in a T_1-space (X, T) is a non-empty family of subsets of X which contains finite unions and subsets of its members. Well known examples of $\mathbf{H}$ include

(a) metrically bounded subsets of a metric space

(b) the family of subsets of compact sets in a topological space

(c) the family of all totally bounded subsets of a uniform space etc.

In what follows we'll usually take $\Delta\mathbf{H} \cap \mathrm{CL}(X)$. Then $\tau(\Delta)$ is a generalization of the **bounded Vietoris topology**. We note here that the **upper Δ-topology** $\tau^+(\Delta)$ is generated by $\{E^+ : E^c \in \Delta\}$ and we have similar definitions of other "upper" topologies.

Theorem 3.1 (cf. [12]) *Suppose (X, T) is a T_1-topological space and Δ, Δ' are two subrings of $\mathrm{CL}(X)$.*

Then the following are equivalent on $\mathrm{CL}(X)$:

(a) $\tau(\Delta) \subset \tau(\Delta')$

(b) $\tau^+(\Delta) \subset \tau^+(\Delta')$

(c) *For each* $B \in \Delta - \{X\}$*,* $B \subset U \in \mathsf{T}$ *implies the existence of* $B' \in \Delta'$ *such that* $B \subset B' \subset U$*.*

Corollary 3.2 *Suppose* H *is a boundedness in a is a* T_1*-topological space* (X, T) *such that* $\mathrm{K}(X) \subset \mathsf{H}$ *and* $\Delta\mathsf{H} \cap \mathrm{CL}(X)$*. Then the following are equivalent:*

(a) $\tau(\mathrm{F}) = \tau(\Delta)$ *on* $\mathrm{CL}(X)$*,*

(b) $\Delta - \{X\} \subset \mathrm{K}(X)$*,*

(c) X *is boundedly compact.*

We have the well known results:

Corollary 3.3 (a) *If* (X, d) *is a metric space, then on* $\mathrm{CL}(X)$ *the Fell topology equals the bounded Vietoris topology if and only if* X *is boundedly compact.*

Replacing the metric d *by the equivalent bounded metric* $d' = \min\{d, 1\}$*, we get the result:*

(b) $\tau(\mathrm{F}) = \tau(\mathrm{V})$ *on* $\mathrm{CL}(X)$ *if and only if* X *is compact.*

There are analogous results using abstract boundedness in proximity and uniform spaces and we refer to ([15])*.*

We now prove some simple results that will generalize relations between $\tau(\mathrm{F})$*,* $\tau(\mathrm{V})$ *as well as between* $\tau(\mathrm{F})$*,* $\tau(\mathsf{V}_H)$*. Proofs in the following are similar to those in* (3.1)*.*

Theorem 3.4 *Suppose* (X, T) *is a* T_1*-space and suppose* $\Delta \subset \mathrm{CL}(X)$ *is a ring containing* $\mathrm{K}(X)$*. Then the following are equivalent:*

(a) $\tau(\Delta) = \tau(\Delta\mathrm{U})$

(b) $\tau^+(\Delta) = \tau^+(\Delta\mathrm{U})$

(c) $X \in \Delta$*.*

(d) $\tau(\Delta) = \tau(\mathrm{V})$*.*

Theorem 3.5 *Suppose* (X, V) *is a Hausdorff uniform space with a compatible* EF*-proximity* δ*. Suppose* $\Delta \subset \mathrm{CL}(X)$ *is a ring containing* $\mathrm{K}(X)$*. Then the following are equivalent:*

(a) $\sigma(\Delta) = \sigma(\Delta\mathrm{U})$

(b) $\sigma^+(\Delta) = \sigma^+(\Delta\mathrm{U})$

(c) $X \in \Delta$.

(d) $\sigma(\Delta) = \tau(\mathsf{V}_{\mathrm{H}})$

Remark 3.6 As we noted above, $\tau(\mathrm{F}) = \sigma(\mathrm{F})$ and so the two above results generalize the well known result that the following are equivalent:

(a) $\tau(\mathrm{F}) = \tau(\mathrm{V})$

(b) $\tau(\mathrm{F}) = \tau(\mathsf{V}_{\mathrm{H}})$

(c) X is compact.

4 UNIFORMIZATION OF (PROXIMAL) DELTA TOPOLOGIES

Beer and Tamaki ([6], [7]) investigated necessary and sufficient conditions for the unformizability of (proximal) Δ-topologies. Their proof involves construction of special Urysohn functions. In ([22]) we study these and (proximal) ΔU-topologies and provide an alternate approach.

Suppose (X, T) is a Tychonoff space with a compatible EF-proximity δ and suppose $\Delta \subset \mathrm{CL}(X)$ is a ring.

(a) Δ is called **proximally Urysohn** iff whenever $D \in \Delta$ and $A \in \mathrm{CL}(X)$ are far w. r. t. δ then there exists an $S \in \Delta$ such that $D \ll S \subset A^c$.

It is easy to see that the above definition is equivalent to one where the last relation is replaced by $D \ll S \ll A^c$.

Δ is called **Urysohn** if Δ is **proximally Urysohn** w. r. t. the LO-proximity δ_0.

(b) Δ is called a **local family** iff for each $x \in X$ and $V \in \mathsf{T}$ with $x \in V$, implies the existence of a $D \in \Delta$ such that $x \in D^o \subset D \subset V$.

(c) For each $K \in \Delta$ and $W \in \mathsf{V}$, we set

$$[K, W] = \{(A, B) \in \mathrm{CL}(X) \times \mathrm{CL}(X) : A \cap K \subset W(B) \text{ and } B \cap K \subset W(A)\}.$$

The family $\{[K, W] : K \in \Delta \text{ and } W \in \mathsf{V}\}$ is a base for a uniformity on $\mathrm{CL}(X)$ called the **Attouch-Wets uniformity** V_Δ.

The proof of the following result due to Beer and Tamaki uses the Attouch-Wets technique developed by Beer in ([7]) for studying the Fell topology.

Theorem 4.1 *Suppose (X, T) is a Tychonoff space with a compatible EF-proximity δ, V_ω the unique totally bounded uniformity compatible with δ. Suppose Δ is a local proximally Urysohn family. Then the proximal Δ-topology $\sigma(\Delta)$ on $\mathrm{CL}(X)$ is the topology $\tau(\mathsf{V}_{\omega\Delta})$ induced by the Δ-Attouch-Wets uniformity $\mathsf{V}_{\omega\Delta}$ and hence is Tychonoff.*

Conversely, if $\sigma(\Delta)$ is Tychonoff then Δ is a local proximally Urysohn family.

Corollary 4.2 *Suppose (X, T) is a Tychonoff space. Then Δ is a local Urysohn family if and only if $(\mathrm{CL}(X), \tau(\Delta))$ is Tychonoff.*

We conclude with a characterization of completely regular proximal $\Delta\mathrm{U}$-topology $\sigma(\Delta\mathrm{U})$.

Theorem 4.3 *Suppose (X, T) is a Tychonoff space with a compatible LO-proximity δ and Δ is a proximally Urysohn family. Then δ' defined by*

$$A\underline{\delta}'B \quad \textit{iff} \quad \mathrm{cl}\, A \in \Delta \quad \textit{or} \quad \mathrm{cl}\, B \in \Delta \quad \textit{and} \quad A\underline{\delta}B \tag{$*$}$$

is a compatible EF-proximity on X. Further $\delta' \leq \delta$ and $\sigma(\Delta\mathrm{U}) = \sigma(\delta')$.

Corollary 4.4 *Suppose (X, T) is a Tychonoff space and Δ is a Urysohn family. Then δ' defined by*

$$A\underline{\delta}'B \quad \textit{iff} \quad \mathrm{cl}\, A \in \Delta \quad \textit{or} \quad \mathrm{cl}\, B \in \Delta \quad \textit{and} \quad A\underline{\delta_0}B \tag{$**$}$$

is a compatible EF-proximity on X . Further $\delta' \leq \delta_0$ and $\tau(\Delta\mathrm{U}) = \sigma(\delta')$.

Thus $\tau(\Delta\mathrm{U})$ is completely regular.

Corollary 4.5 *Suppose (X, T) is a locally compact Hausdorff space. Then the U-topology $\tau(\mathrm{U})$ is the proximal topology corresponding to the EF-proximity δ_1 induced by the one-point-compactification of X viz:*

$$A\underline{\delta_1}B \quad \textit{iff} \quad \mathrm{cl}\, A \quad \textit{or} \quad \mathrm{cl}\, B \quad \textit{is compact and} \quad A\underline{\delta_0}B\,. \tag{$***$}$$

Remark 4.6 (a) It is interesting to note that in (4.5) if (X, T) is not normal, then δ_0 is not EF but the proximity δ_1 induced by $\Delta = \mathrm{K}(X)$ is indeed EF!

(b) If (X, T) is not locally compact, then the proximity δ_1 as defined by ($***$) is not EF but is a compatible LO-proximity on X. In this case for any compatible EF-proximity δ on X, $\delta_1 < \delta$ but $\sigma(\delta_1) \not\subset \sigma(\delta)$. (Cf. (2.1)(e)).

Acknowledgement: It is a pleasure to dedicate this paper to Harry Poppe, who introduced Δ–topologies and did a lot of research in graph topologies and Ascoli-Arzela type results. I am grateful to my friends Anna Di Concilio, Giuseppe Di Maio, Enrico Meccariello.

References

[1] **Bartsch, R. :** *Compactness properties for some hyperspaces and function spaces.* Shaker Verlag, Aachen 2002

[2] **Beer, G. :** *Topologies on closed and closed convex sets.* Kluwer Academic Pub. 1993

[3] **Beer, G.**, and **di Concilio, A. :** *Uniform continuity on bounded sets and the Attouch-Wets topology.* Proc. Amer. Math. Soc. **112**, 235-243 (1991)

[4] **Beer, G., Lechicki, A., Levi, S.**, and **Naimpally, S. :** *Distance functionals and suprema of hyperspace topologies.* Ann. Mat. Pura. Appl. **162**, 367-381 (1992)

[5] **Beer, G., Himmelberg, C., Prikry, C.**, and **van Vleck, F. :** *The locally finite topology on 2^X.* Proc. Amer. Math. Soc. **101**, 168-171 (1987)

[6] **Beer, G.**, and **Tamaki, R. :** *The infimal value functional and the uniformization of hit-and-miss topologies.* Proc. Amer. Math. Soc. **122**, 601-611 (1994)

[7] **Beer, G.**, and **Tamaki, R. :** *On hit-and-miss hyperspace topologies.* C. M. U. C. **34, 4**, 717-728 (1993)

[8] **Costantini, C.**, and **Vitolo, P. :** *On the infimum of the Hausdorff metric topologies.* Proc. London Math. Soc. **70**, 441-480 (1995)

[9] **di Caprio, D.**, and **Meccariello, E. :** *Notes on Separation Axioms in HyperspacesI.* Q & A in General Topology **18**, 65-86 (2000)

[10] **di Caprio, D.**, and **Meccariello, E. :** *G-uniformities LR-uniformities and hyper-topologies.* Acta Math. Hungarica **88 (1-2)**, 73-93(2000)

[11] **di Maio, G.**, and **Hola, L. :** *A hypertopology determined by the family of totally bounded sets is the minimum of upper Wijsman topologies.* Q & A in Gen. Top. **15**, 51-66 (1997)

[12] **di Maio, G.**, and **Hola, L. :** *On hit-and-miss topologies.* Rend. Acc. Sc. Fis. Mat. Napoli **57**, 103-124 (1995)

[13] **di Maio, G., Hola, L.**, and **Meccariello, E. :** *Notes on hit-and-miss topologies.* Rostock. Math. Kolloq. **52**, 19-32 (1998)

[14] **di Maio, G.**, and **Naimpally, S. :** *Comparison of hypertopologies.* Rend. Istit. Mat. Univ. Trieste **22**, 140-161 (1990)

[15] **di Maio, G.**, and **Naimpally, S. :** *Some notes on hyperspace topologies.* Rend. Acc. Sc. Fis. Mat. Napoli (to appear)

[16] **di Concilio, A., Naimpally, S.**, and **Sharma, P. :** *Proximal Hypertopologies.* Sixth Brazilian Topology Meeting, Campinas, Brazil (1988) [unpublished]

[17] **Fell, J. :** *A Hausdorff topology for the closed subsets of a locally compact non-Hausdorff spaces.* Proc. Amer. Math. Soc. **13**, 472-476 (1962)

[18] **Hu, S. T. :** *Boundedness in a topological space.* J. Math. Pures et Appl. **28**, 287-340 (1949)

[19] **Marjanovic, M. :** *Topologies on collections of closed subsets.* Publ. Inst. Math. (Beograd) **20**, 125-130 (1966)

[20] **Michael, E. :** *Topologies on spaces of subsets.* Trans. Amer. Math. Soc. **71**, 152-182 (1951)

[21] **Nachman, L. :** *Hyperspaces of proximity spaces.* Math. Scand. **23**, 201-213 (1968)

[22] **Naimpally, S. :** *Uniformizing delta (proximal) hypertopologies.* [to be completed with Giuseppe Di Maio and Enrico Meccariello]

[23] **Naimpally, S. :** *Hyperspaces and function spaces.* Q & A in General Topology **9**, 33-59 (1991)

[24] **Naimpally, S.**, and **Sharma, P. :** *Fine uniformity and the locally finite hyperspace topology on 2^X.* Proc. Amer. Math. Soc. **103**, 641-646 (1988)

[25] **Naimpally, S. A.**, and **Warrack, B. D. :** *Proximity Spaces.* Cambridge Tracts in Mathematics no. 59, Cambridge University Press 1970

[26] **Poppe, H. :** *Eine Bemerkung uber Trennungaxiome im Raum der abgeschlossenen Teilmengen eines topologischen Raumes.* Arch. Math. **16** , 197-199 (1965)

[27] **Poppe, H. :** *Einige Bemerkungen uber den Raum der abgeschlossen Mengen.* Fund. Math. **59**, 159-169 (1966)

[28] **Wattenberg, F. :** *Topologies on the set of closed subsets.* Pacific J Math. **68**, 537-557 (1977)

received: October 18, 2002

Author:

Somashekhar Naimpally
96 Dewson Street
Toronto, Ontario
M6H 1H3 CANADA

e-mail: sudha@accglobal.net

Hinweise für Autoren

Um die redaktionelle Bearbeitung und die Herstellung der Druckvorlage zu erleichtern, wären wir den Autoren dankbar, sich betreffs der Form der Manuskripte an den in **Rostock. Math. Kolloq.** veröffentlichten Beiträgen zu orientieren.
Insbesondere beachte man:

1. Manuskripte sollten grundsätzlich **maschinengeschrieben** (Schreibmaschine, PC) in **deutscher oder englischer Sprache** abgefaßt sein. Falls LaTeX - Manuskripte eingesandt werden, sollten eigene definierte Befehle möglichst vermieden werden.

2. Zur inhaltlichen Einordnung der Arbeit sind **1-2 Klassifizierungsnummern** anzugeben.

3. **Literaturzitate** sind im Text durch laufende Nummern (vgl. [3], [4]; [7, 8, 10]) zu kennzeichnen und am Schluß der Arbeit unter der Zwischenüberschrift **Literatur** bzw. **References** zusammenzustellen. Hierbei ist die durch die nachfolgenden Beispiele veranschaulichte Form einzuhalten.

   [3] **Zariski, O.**, and **Samuel, P.:** *Commutative Algebra.* Princeton 1958

   [4] **Steinitz, E.:** *Algebraische Theorie der Körper.* J. Reine Angew. Math. **137**, 167-309 (1920)

   [8] **Gnedenko, B.W.:** *Über die Arbeiten von C.F. Gauß zur Wahrscheinlichkeitsrechnung.* In: Reichardt, H. (Ed.): C.F. Gauß, Gedenkband anläßlich des 100. Todestages. S. 193-204, Leipzig 1957

   Die Angaben erfolgen in Originalsprache; bei kyrillischen Buchstaben sollte die (bibliothekarische) Transliteration verwendet werden.

4. Die aktuelle, vollständige Adresse des Autors sollte enthalten: Vornamen Name / Institution / Struktureinheit / Straße Hausnummer / Postleitzahl Ort / Land / e-mail-Adresse (falls vorhanden).

Es besteht die Möglichkeit, mit dem **Satzsystem LaTeX** oder unter anderen Textprogrammen erstellte Manuskripte auf unter **MS-DOS** formatierten Disketten (**3.5″, 1.44MB**) einzureichen oder per e-mail an folgende Adresse zu schicken:

`romako@mathematik.uni-rostock.de`