

VGI

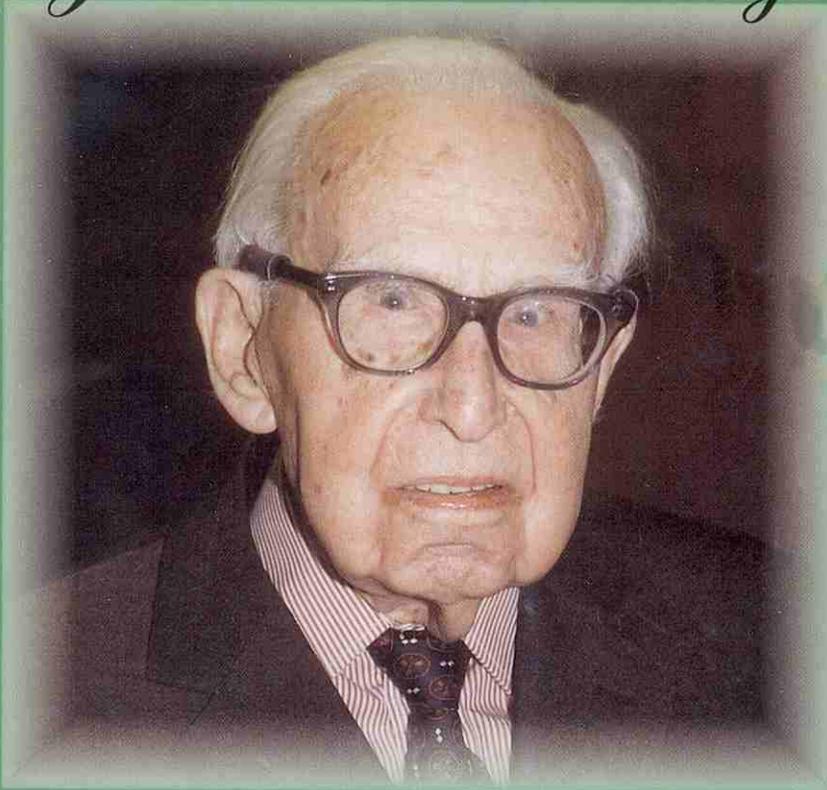
Österreichische Zeitschrift für
**VERMESSUNG &
GEOINFORMATION**

85. Jahrgang 1997

Heft 4/97

Organ der Österreichischen Gesellschaft für Vermessung und Geoinformation und der Österreichischen
Geodätischen Kommission

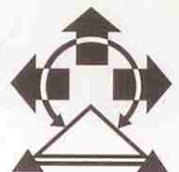
Zum 100. Geburtstag



Präsident i. R.

Prof. Ing. Dr. techn.

Karl Neumaier





Zur optimalen Approximation von Höhenprofilen

Andrzej Borkowski, Wrocław, Dirk Burghardt und Siegfried Meier, Dresden

Zusammenfassung

Gemessene Höhenwerte mit unterschiedlichen Fehlercharakteristiken können mittels linearer Prädiktion, optimaler Splines und aktiver Splines (Snakes) approximiert werden. Die genannten Methoden haben eine gemeinsame Grundlage, die Tichonov-Regularisierung. Obwohl sie Variationsprobleme unterschiedlicher Art lösen, führen die entsprechenden Verfahren auf äquivalente Ergebnisse. Als Ergänzung zu einem robusten Prädiktionsverfahren von Kraus [4] wurden Höhenprofile sowohl mit optimalen als auch mit aktiven Splines approximiert, wobei die berechneten Profile etwa gleichwertig sind.

Abstract

Sampled terrain data with different error characteristics can be approximated by linear prediction, optimal and active splines (snakes) respectively. The mentioned methods are established on the same base, the Tichonov regularization. Although they solve variational problems of different kind the appropriate procedures lead to equivalent results. In addition to a robust procedure of linear prediction by Kraus [4] elevation profiles was approximated by optimal as well as by active splines, whereby the adjusted profiles was found to be nearly the same one.

1. Vorbemerkungen zum Approximationsproblem

Im Fehlerhaushalt moderner, registrierender geodätischer Meßverfahren sind die zufälligen Fehler in der Regel klein, die sog. systematischen mindestens von gleicher Größenordnung, wenn nicht sogar dominierend, grobe Fehler ggf. aufgesetzt. Ein Beispiel sind Laserscannerdaten über baumbestandenem Gelände. Um den Einfluß der Reflexionswerte im Vegetationsbereich auf das approximierende Geländeprofil zu eliminieren (und für verwandte Aufgaben), hat Prof. Kraus in dieser Zeitschrift ein robustes Verfahren der Prädiktionsfilterung vorgeschlagen [4].

Eine Alternative dazu ist die Spline-Approximation. Beide Verfahren sind nach der Regularisierungstheorie von Tichonov [5], vgl. auch [1], [6], Sonderfälle einer globalen Kollokationsaufgabe bzw. Lösungen von Variationsproblemen. Die lineare Prädiktion ergibt sich, wenn im Variationsproblem 1. Ableitungen des Signals auftreten und als sog. Kernfunktion die Autokovarianzfunktion (AKF) benutzt wird. Approximierende Spline-Funktionen erhält man, wenn im Variationsproblem 2. Ableitungen des Signals einbezogen werden.

Ein drittes Verfahren, welches ebenfalls ein Variationsproblem mit 1. und 2. Ableitungen des Signals (energieminimierend) löst und mit der Ti-

chonov-Regularisierung begründet werden kann, ist die Approximation mit energiegeladenen Splines, den sog. Snakes [3]. Vorrangig in der Bildverarbeitung zur Kantendetektion benutzt, neuerdings auch zur formerhaltenden Linienverdrängung vorgeschlagen [2], könnte es ebenfalls zur Approximation von Höhendaten dienen.

Alle genannten Verfahren lösen Variationsprobleme auf der Grundlage von einander gleichwertigen Optimalitätskriterien bzw. Extremalprinzipien. Man wird daher erwarten dürfen, daß sie (wenigstens genähert) gleichwertige Ergebnisse liefern. Nachfolgend werden die Spline-Approximation (Abschnitt 2) und die Snakes-Approximation (Abschnitt 3) als Ergänzung zur Arbeit von Prof. Kraus [4] behandelt: zuerst jeweils der 2D-Ansatz (Relief), dann spezifiziert die 1D-Approximation (Profile) mit Beschreibung und Test der iterativen Verfahren. Als Testbeispiel benutzen wir das gleiche wie in [4], reduziert um einen linearen Trend.

2. Spline – Approximation

2.1 Zweidimensionale Approximation

Auf dem ebenen Bereich B seien n fehlerbehaftete Höhenwerte $h_i = h(x_i, y_i)$; $i = 1, 2, \dots, n$ gegeben. Die Aufgabe ist nun, eine ausgleichende Funktion $z = z(x, y)$ zu finden, welche die h_i approximiert, d.h. das Funktional

$$\Phi_\alpha(z) = \Phi_{1,\alpha}(z) + \sum_{i=1}^n \alpha [z(P_i) - h(P_i)]^2$$

$$\text{mit } \Phi_{1,\alpha}(z) = \iint_B (z_{xx}^2 + 2z_{xy}^2 + z_{yy}^2) dx dy \quad (2.1)$$

und den freien Parameter α mit $\alpha \geq 0$ minimiert [1], [5]. Tiefgestellte x, y bedeuten wie üblich partielle Differentiation nach x, y .

Die analytische Lösung des Variationsproblems $\Phi_\alpha(z) = \text{Min.}$ erfolgt mittels einer „geeigneten“ Regularisierung nach Tichonov [5] und ergibt

$$z(x, y) = \frac{1}{2} \sum_{i=1}^n \lambda_i r_i^2 \ln r_i^2 + d_0 + d_1 x + d_2 y$$

$$\text{mit } \sum_{i=1}^n \lambda_i = 0, \sum_{i=1}^n \lambda_i x_i = 0, \sum_{i=1}^n \lambda_i y_i = 0 \quad (2.2)$$

$$\text{und } r_i^2 = (x-x_i)^2 + (y-y_i)^2,$$

in der englischsprachigen Literatur auch *thin plate spline* genannt. Der (kritische) Parameter α bewichtet die beiden Terme des Funktional (2.1) gegenseitig. Ist α klein, speziell $\alpha \rightarrow 0$, verschwindet der zweite Term in (2.1) und der Approximationsspline (2.2) geht in einen reinen Interpolationsspline über. Ist dagegen α groß genug, speziell $\alpha \rightarrow \infty$, tendiert der Approximationsspline gegen eine ausgleichende Ebene. Somit steuert der Parameter α die Glättungseigenschaften des Approximationssplines. Er hängt nicht nur von der Genauigkeit der Inputdaten, sondern auch von den Daten selbst und ihrer Verteilung in B ab, und muß daher für jeden

gegebenen Datensatz (etwas aufwendig) individuell gewählt werden (vgl. Abschnitt 2.2).

2.2 Eindimensionale Approximation

Die analytische Lösung des eindimensionalen Variationsproblems ist [1]

$$z(x) = \sum_{i=1}^n \lambda_i |x-x_i|^3 + d_0 + d_1 x$$

$$\text{mit } \sum_{i=1}^n \lambda_i = 0, \sum_{i=1}^n \lambda_i x_i = 0. \quad (2.3)$$

Die $n+2$ Parameter $d_0, d_1, \lambda_i (i=1, 2, \dots, n)$ erhält man als Lösung eines Gleichungssystems vom Typ

$$\begin{bmatrix} \alpha & a_{12} & \dots & a_{1n} & 1 & x_1 \\ a_{21} & \alpha & \dots & a_{2n} & 1 & x_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \dots & \alpha & 1 & x_n \\ 1 & 1 & \dots & 1 & 0 & 0 \\ x_1 & x_2 & \dots & x_n & 0 & 0 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \\ d_0 \\ d_1 \end{bmatrix} = \begin{bmatrix} h_1 \\ h_2 \\ \vdots \\ h_n \\ 0 \\ 0 \end{bmatrix} \quad (2.4)$$

Sollen nun die Höhenwerte h_i mit Grobfehlern bzw. asymmetrischer Fehlerverteilung wie im Testbeispiel des Laserscannings über baumbestandenem Relief approximiert werden, braucht man als weiteren Parameter einen Schwellwert ε , proportional zur Standardabweichung $\sigma_{h-a-priori}$ des Abtastfehlers im baumlosen Gelände, etwa $\varepsilon := \gamma \sigma_{h-a-priori}$, $\gamma = 2$ bis 3 . Wenn $v_i := [h(P_i) - z(P_i)] > \varepsilon$, ist die Höhe h_i am Punkt P_i als (wahrscheinlicher) Reflexionswert im Vegetationsbereich zu eliminieren. Im Gegensatz zum Glättungsparameter α

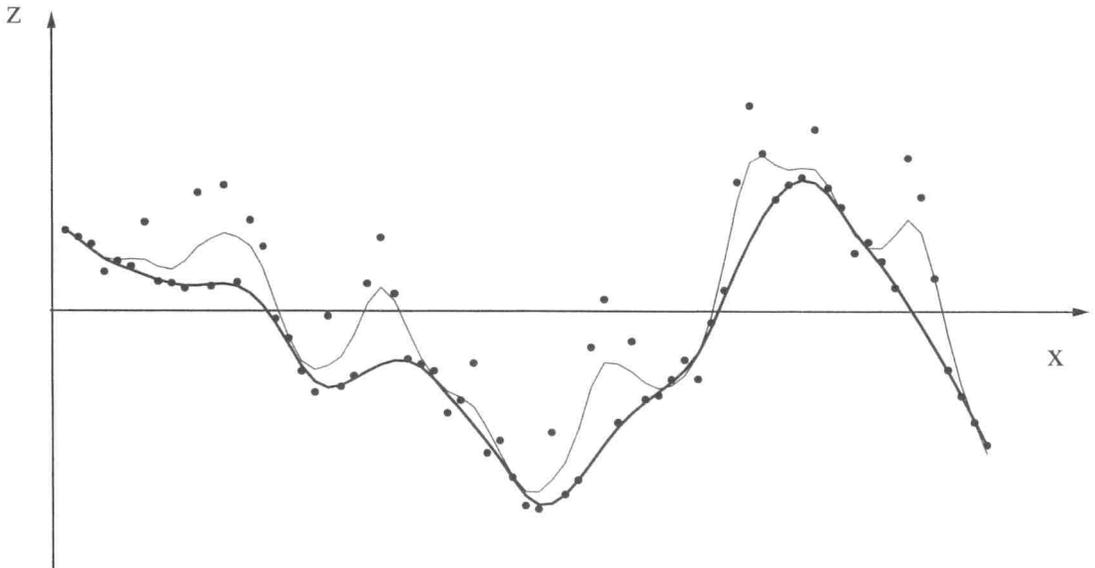


Abb. 1: Spline-Approximation (dicke Linie) einer Meßwertreihe (Punkte) mit asymmetrischer Fehlerverteilung. Erste Annäherung (dünne Linie) nach dem 2. Schritt des Verfahrens (vgl. Abschnitt 2.2)

ist der Auswahlparameter ε rein in der Fehlercharakteristik des Meßverfahrens begründet.

Die 1D-Spline-Approximation mit Elimination der Reflexionswerte im Vegetationsbereich wird nun iterativ in drei Schritten realisiert.

1. *Schritt.* Aus (2.4) berechnet man mit dem Standardwert $\alpha=0$ zunächst den Parametervektor des Interpolationssplines $\underline{\lambda} = \underline{A}^{-1} \underline{h}$, wobei \underline{A} die Koeffizientenmatrix des Gleichungssystems bedeutet. Nun kann man $(\alpha I + \underline{A}) \underline{\lambda} = \underline{h} + d \underline{h}$ mit I als Einheitsmatrix schreiben, woraus $\alpha I \underline{\lambda} = d \underline{h}$ folgt. Die Parameter $\alpha_i = |dh_i / \lambda_i|$; $i=1, 2, \dots, n$ können an jedem Punkt P_i bestimmt werden, indem man Änderungen dh_i annimmt. Als Startwert diene etwa $dh_i \approx \varepsilon$. Zur weiteren Rechnung benutzen wir das Mittel der α_i . Diese Vorgehensweise erlaubt es sofort, zumindest ein annähernd „richtiges“ α zu finden und stellt eine Abhilfe im kritischen Punkt des Verfahrens dar.

2. *Schritt.* Erneute Lösung des Systems mit $\alpha > 0$. Man erhält die neuen Parameter λ_i, d_0, d_1 des Approximationssplines. Damit haben wir einen ersten Verlauf des Splines (2.3) mit den speziellen Höhenwerten z_i in den Stützpunkten P_i erhalten (Beispiel in Abb. 1).

3. *Schritt.* Aus den Approximationswerten z_i des 2. Schrittes und den Meßwerten h_i werden Residuen v_i gebildet. Ist $v_i > \varepsilon$, so wird die zugehörige Höhe h_i eliminiert.

Nach Elimination von k Werten ($1 \leq k < n$) gehen wir zum 2. Schritt zurück und Lösen das Gleichungssystem (2.4) für $n-k$ Stützpunkte. Mit den neuen Approximationswerten z_i wiederholt sich die Auswahl im 3. Schritt. Es wird solange iteriert, bis – entsprechend dem o.a. Kriterium – keine Werte mehr eliminiert werden können.

2.3 Ergebnisse numerischer Tests

Die Experimente mit dem Datensatz Abb. 1 haben gezeigt, daß das Verfahren relativ schnell, nach etwa drei bis vier Schritten konvergiert, und es läßt eine große Variationsbreite des Glättungsparameters a zu. Um einen gewünschten Glättungsgrad zu erzielen muß man ggf. die Iteration für *verschiedene* α wiederholen. Ein sinnvolles Kriterium dafür scheint $\hat{\sigma}_z \approx \sigma_{h-a-priori}$ zu sein, wobei die Standardabweichung aus Residuen v_i , ohne die Ausreißer zu berücksichtigen, berechnet wird.

Ein gewisser Nachteil des Verfahrens liegt darin, daß man die Matrix in (2.4) mehrmals invertieren muß. Darüber hinaus kann das Verfahren numerisch instabil sein, wenn die Daten sehr irregulär verteilt sind. In diesem Fall emp-

fiehl es sich, die Daten gemäß $\bar{x}_i = x_i / (x_{\max} - x_{\min})$ zu skalieren.

Abb. 1 zeigt einen, die systematischen Fehler eliminierenden Approximationsspline mit $\varepsilon = 3\sigma_{h-a-priori}$, $dh = 20\sigma_{h-a-priori}$, $\hat{\sigma}_z = 0,8$.

3. Snakes – Approximation

3.1 Zweidimensionale Approximation

Das zweidimensionale Pendant zu hinreichend glatten Kurvenstücken, denen Energien zugeordnet werden (*Snakes*), sind hinreichend glatte, energiegeladene Flächenstücke. Solche Gebilde nennen wir *Flakes*. Um Punkthöhen $h_i = h(x_i, y_i)$ auf dem ebenen Bereich B durch Flakes zu approximieren, hat man – analog zum Snakes-Konzept – das Funktional

$$I(z) = \iint_B E_{ges}(z; z_x, z_y; z_{xx}, z_{xy}, z_{yy}) dx dy \quad (3.1)$$

zu minimieren. Die Gesamtenergie E_{ges} setzt sich aus äußerer Energie E_{ext} und innerer Energie E_{int} zusammen. E_{ext} allein würde die approximierende Fläche in die Stützpunkte zwingen. Um fehlerausgleichend zu approximieren, benutzt man zusätzlich den Term E_{int} : die bereits in den Meßwerten bzw. in einer geglätteten Version davon enthaltene Vorinformation über Neigung und Krümmung sorgt für einen ausgleichenden, hinreichend glatten Flächenverlauf.

Die zum 2D-Variationsproblem $I(z) = M$ gehörende 2D-Eulergleichung lautet

$$E_z - \frac{\partial}{\partial x} E_{z_x} - \frac{\partial}{\partial y} E_{z_y} + \frac{\partial^2}{\partial x^2} E_{z_{xx}} + \frac{\partial^2}{\partial x \partial y} E_{z_{xy}} + \frac{\partial^2}{\partial y^2} E_{z_{yy}} = 0, \quad (3.2)$$

wobei tiefgestellte Symbole wieder partielle Differentiation bedeuten, z.B. $E_{z_x} = \partial E / \partial z_x$. Um (3.2) weiter zu spezifizieren, müssen die Energieterme *zweckdienlich* definiert werden (s. unten).

Als Alternative zu Flakes bietet sich ein die Oberfläche approximierendes Netz verknoteter, regulär oder irregulär verlaufender Snakes mit Knoten in den Stützpunkten an. Diese Linienapproximation führt auf ein (ggf. sehr großes) System von Eulergleichungen mit erheblichen Konsequenzen für die numerischen Lösungsverfahren. Beide 2D-Konzepte sind noch nicht getestet worden.

3.2 Eindimensionale Approximation

Für ein Höhenprofil $z(x)$ erhält man die 1D-Eulergleichung

$$E_z - \frac{\partial}{\partial x} E_{z_x} + \frac{\partial^2}{\partial x^2} E_{z_{xx}} = 0. \quad (3.3)$$

Die Energieterme werden wie folgt angesetzt. E_{ext} sei proportional zu den Residuen $v_i = (x_i) - z(x_i)$ zwischen gemessenen Höhen $h(x_i)$ und jenen des aktuellen Profilverlaufs $z(x_i)$:

$$E_{ext}(z_i) \sim \begin{cases} v_i^2 & \text{für } v_i \leq \varepsilon \\ 0 & \text{sonst} \end{cases} \quad (3.4)$$

wächst (quadratisch) mit der Entfernung zwischen $h(x_i)$ und $z(x_i)$. Es besteht keine Anziehung- bzw. Verschiebungskraft mehr, wenn $z(x_i) = h(x_i)$ oder wenn $h(x_i)$ ein Reflexionswert aus dem Vegetationsbereich ist: der Schwellwert ε hat die gleiche Bedeutung und wird ebenso bestimmt wie im Abschnitt 2.2. Bei positiven und negativen Grobfehlern sind die v_i in (3.4) durch $|v_i|$ zu ersetzen, und im Sonderfall gleichgewichtiger $h(x_i)$ gilt $E_{ext} \sim v_i^2$ für alle $i=1,2, \dots, n$.

E_{int} soll die (vorgezeichnete) Profilform möglichst gut erhalten bzw. Änderungen in den (gegenseitig bewichteten) 1. und 2. Ableitungen z_x, z_{xx} möglichst minimieren:

$$E_{int} = (\alpha |w_x|^2 + \beta |w_{xx}|^2) / 2$$

mit $w := z - z^t, w_x := z_x - z_x^t, w_{xx} := z_{xx} - z_{xx}^t$, (3.5)

wobei sich die Differenzen immer auf den letzten Iterationsschritt t beziehen.

Nach Einsetzen von $E_{ges} = E_{ext} + E_{int}$ in (3.3) ergibt sich die Eulergleichung 4. Ordnung

$$\frac{\partial E_{ext}}{\partial z} - \alpha (z_{xx} - z_{xx}^t) + \beta (z_{xxxx} - z_{xxxx}^t) = 0. \quad (3.6)$$

Ihre Diskretisierung und iterative Lösung ist in [2], [3] ausführlich beschrieben. Die Iterationsvorschrift lautet

$$\underline{z}_t = \underline{z}_{t-1} - (\underline{A} + \gamma \underline{I})^{-1} \underline{E}_z \quad (3.7)$$

mit

$$\underline{A} := \begin{bmatrix} a & b & c & 0 & 0 & 0 & \dots \\ b & a & b & c & 0 & 0 & \dots \\ c & b & a & b & c & 0 & \dots \\ 0 & c & b & a & b & c & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

$$a := 2\alpha + 6\beta, \quad b := -\alpha - 4\beta, \quad c := \beta,$$

$$E_{z_i} := \partial E_{ext} / \partial z|_i = -2[h(x_i) - z(x_i)].$$

Der Parameter γ (in der Regel $\gamma=1$) bestimmt die „Auflösung“ und damit die Konvergenzgeschwindigkeit des Verfahrens. Es wird in drei Schritten realisiert.

1. *Schritt.* Die Inputdaten werden mit einem Tiefpaßfilter (mehr oder weniger stark) geglättet. Man erhält einen ersten, groben Verlauf von $z(x)$; vgl. Abb. 2. Da zur Lösung des Variationsproblems gewisse Randwerte benutzt werden, ist es sinnvoll, den Datensatz um mindestens zwei Punkte am Anfang und Ende des Profils z.B. periodisch zu erweitern, damit der Iterationsprozeß schneller konvergiert.

2. *Schritt.* Wahl der Parameter $\alpha, \beta, \gamma, \varepsilon$. Berechnung und Inversion der Matrix $\underline{A} + \gamma \underline{I}$.

3. *Schritt.* Berechnung des Vektors \underline{E} und eines neuen Verlaufs von $z(x)$ nach (3.7). Dieser Schritt

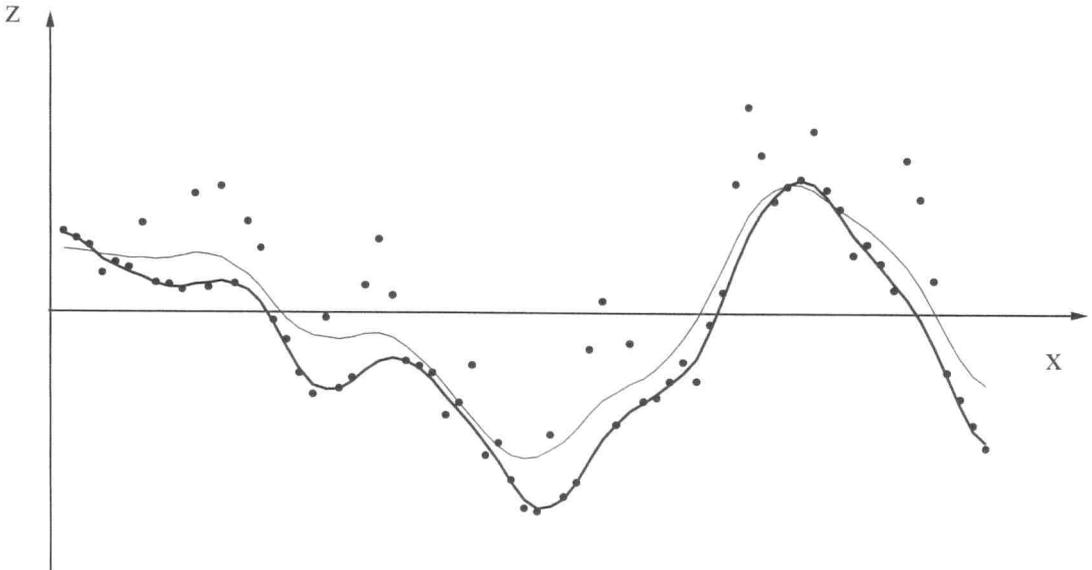


Abb. 2: Snakes-Approximation (dicke Linie) einer Meßwertreihe (Punkte) mit asymmetrischer Fehlerverteilung. Erste Annäherung (dünne Linie) nach dem 1. Schritt des Verfahrens (vgl. Abschnitt 3.2)

wird solange wiederholt, bis eine gewählte Genauigkeitsanpassung, z.B. $\hat{\sigma}_z \approx \sigma_{h-a-priori}$, erreicht ist. Konvergiert der Iterationsprozeß nicht, so ist zum 2. Schritt zurückzukehren; α und β sind neu zu bestimmen.

3.3 Ergebnisse numerischer Tests

Die Snakes-Approximation erfordert mehr Iterationsschritte als die Spline-Approximation 2.2. Trotzdem ist sie relativ schnell, denn man braucht die Koeffizientenmatrix in (3.7) nur einmal zu invertieren, sofern die Steuerparameter α und β schon a-priori die „richtige“ Größenordnung haben. Dem kommt entgegen, daß – wie Experimente mit dem Datensatz Abb. 2 gezeigt haben – das Verfahren eine hohe Stabilität bezüglich Variation von α , β aufweist. In Abb. 2 ist das Ergebnis der Snakes-Approximation mit den Parametern $\alpha=2$, $\beta=20$, $\gamma=1$, $\varepsilon=3\sigma_{h-a-priori}$, $\hat{\sigma}_z=0,7$ dargestellt.

4. Vergleich der Approximationsverfahren

Als Alternativen zu der von Prof. Kraus [4] vorgestellten robusten linearen Prädiktion wurden im 2. Abschnitt eine Spline-Funktion und im 3. Abschnitt eine Lösung mittels Snakes zur Approximation von Höhenwerten mit groben Fehlern angegeben. Ausgehend vom 2D-Ansatz wurden die analytischen 1D-Lösungen samt numerischer Umsetzung diskutiert und an einem typischen Beispiel getestet. Abschließend vergleichen wir die Verfahren nach ihren wichtigsten Merkmalen.

(1) Alle drei Verfahren werden iterativ realisiert, wobei die Snakes-Approximation am schnellsten zu sein scheint, denn man hat dabei eine Bandmatrix nur einmal zu invertieren.

(2) Zur robusten Prädiktion wird eine Schätzung der AKF und zusätzlich ein Histogramm der Residuen benötigt. Spline- und Snakes-Approximation brauchen dagegen kaum Vor- und Zuarbeiten. Dafür sind hier Steuerparameter α bzw. α und β zu wählen, zweifellos der kritische Punkt dieser Verfahren. Allerdings werden für die „ro-

buste Wichtung“ in der Prädiktion auch zwei ziemlich willkürlich wählbare Parameter gebraucht.

(3) Die lineare Prädiktion ist in der Geodäsie seit langem bekannt. Infolge der fehlertheoretischen Begründung paßt sie gut in das klassische geodätische Denken. Dies scheint z.B. auf die Snakes-Approximation vordergründig nicht zuzutreffen, jedoch gewinnt man, namentlich in der Bildverarbeitung, zunehmend an Erfahrung und Sicherheit mit dieser durchaus neuartigen Approximationsmethode.

Unsere Untersuchungen haben ein weiteres Mal die Flexibilität des Snakes-Konzeptes bestätigt; mit ortsabhängigen Parametern $\alpha=\alpha(x)$, $\beta=\beta(x)$ kann die Approximation auch lokal gesteuert werden, z.B. in der Umgebung von Geländekanten und -knicken. Es erscheint daher sinnvoll, den 2D-Ansatz, speziell seine numerische Realisierung weiter zu verfolgen.

Literatur

- [1] Aronov, V.I.: Metody postroeniä kart geologo-gefizičeskich priznakov i geometrizaciä zalezëj nefi i gaza na EVM. Nedra, Moskva 1990.
- [2] Burghardt, D.; Meier, S.: Cartographic Displacement Using the Snakes Concept. In: Förstner, W.; Plümer, L. (eds.): Semantic Modeling for the Acquisition of Topographic Information from Images and Maps, pp. 59–71, Birkhäuser Verlag, Basel 1997.
- [3] Kass, M.; Witkin, A.; Terzopoulos, D.: Snakes: Active Contour Models. Proceedings of the First International Conference of Computer Vision, IEEE Comput. Soc. Press, pp. 259–268, 1987.
- [4] Kraus, K.: Eine neue Methode zur Interpolation und Filterung von Daten mit schiefer Fehlerverteilung, Österreichische Z. f. Vermessung u. Geoinformation, 85 (1997) 1, S. 25–30.
- [5] Tichonov, A.N.; Arsenin, W.J.: Metody reschenija nekorrektnych sadaž. Nauka, Moskva 1979.
- [6] Žurkin, I.G.; Nejman, Ū. M.: Metody vycislenij geodezii. Nedra, Moskva 1988.

Anschriften der Autoren

Dr.-Ing. Andrzej Borkowski, Katedra Geodezji i Fotogrametrii, Akademia Rolnicza we Wroclawiu, ul. Grunwaldzka 53, PL-50-357 Wroclaw
Dipl.-Phys. Dirk Burghardt und o. Univ.-Prof. Dr.-Ing. Siegfried Meier, Institut für Planetare Geodäsie, Technische Universität Dresden, Mommsenstr. 13, D-01062 Dresden