Letters from a SteelTown

Calibration problems – An inverse problems view

Heinz W. Engl, heinz.engl@oeaw.ac.at

Introduction

When pricing structured or derivative financial instruments, the typical steps a quant has to do are the following:

1. Choose a model for the movement of the underlying(s)
2. Identify ("calibrate") the model parameters from market prices of liquid instruments
3. Calculate the fair value of the structured instrument by appropriate numerical techniques

We will concentrate on step 2 in this article. The techniques we will present have been applied to various types of inverse problems in science and engineering and should be relevant also in computational finance. At my research groups, we applied advanced inverse problems techniques successfully e.g. to the following problems

- Reconstruct reinforcement bars in concrete from measurements of a scattered magnetic field [Engl, Neubauer]
- Determine optimal cooling strategies in continuous casting and hot rolling of steel (inverse heat conduction problems), e.g. [Binder, Engl, Vessella], [Binder et al.], [Chen et al.]
- Inferring the structure of ion channels from measurements (an inverse problem in a coupled system of partial differential equations similar to the semiconductor equations): [Burger, Eisenberg, Engl], [Burger et al.]

Let us continue with an example from finance:

A first example: Hull-White

Assume that the task is to calculate the value of an annually callable bond under a one factor Hull-White-model

$$dr(t) = (\alpha(t) - \beta(t)r(t))dt + \sigma(t)dW$$

with r(t) being the short rate at time t, dW the increment of a Wiener process, σ (t) the volatility of the Hull-White process, β(t) the mean reversion speed, and α(t) the part of the drift term which has to be adjusted to fit the yield curve of the currency under consideration.

Following Shreve, we can write down closed-form solutions for the fair value at time t of a (credit-risk-free) zero coupon bond B(t,T), which matures at a future time T

$$B(t, T) = \exp\{-r(t)C(t, T) - A(t, T)\}.$$

when the short rate at time t is assumed to be r(t) where

$$K(t) = \int_o^t \beta(u)du,$$

$$A(t, T) = \int_t^T \left[e^{K(v)}\alpha(v) \left(\int_t^T e^{-K(y)}dy \right) - \frac{1}{2}e^{2K(v)}\sigma^2(v) \left(\int_v^T e^{-K(y)}dy \right)^2 \right] dv,$$

$$C(t, T) = e^{K(t)} \int_t^T e^{-K(y)}dy,$$

Assume now that – for some reason – the reversion speed and the volatility are known or given functions. Then K(t) is known, C(t,T) is known, and A(t,T) is a linear well-defined integral operator acting on α(t) plus the known function containing the volatility. Thus, the determination of alpha amounts to solving an integral equation. If we use market prices of zero coupon bonds (or swap rates) for different maturities as data, the determination of α(t) as a, say, piecewise constant function should not be a difficult problem. But there are surprises:

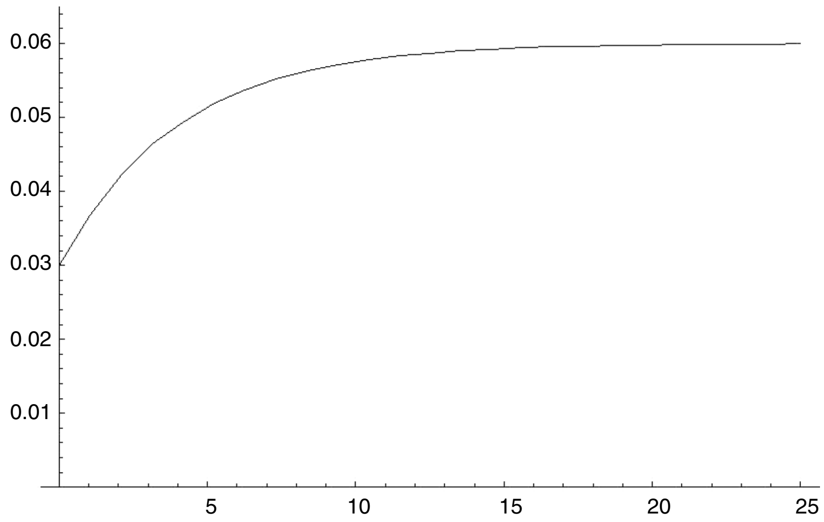
Curve fitting: A naïve approach

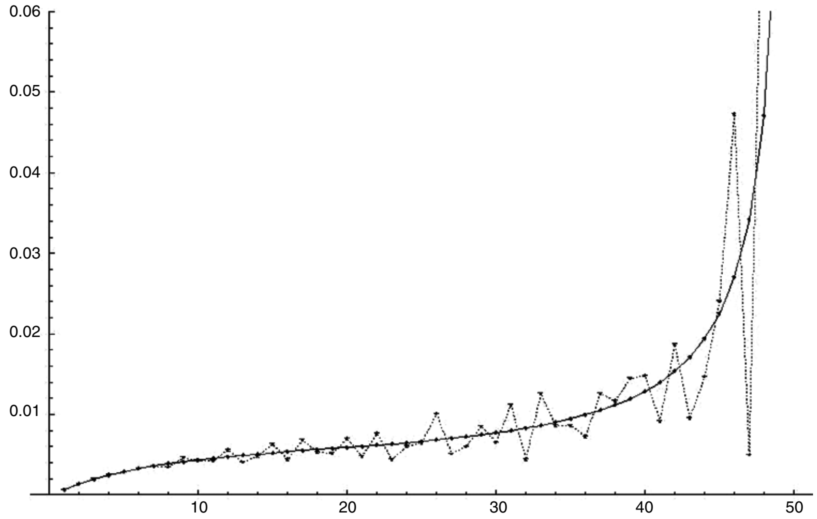
Let us assume that the risk free zero rate (with continuous compounding) is given as $R(0,T) = 0.06–0.03 \exp(-T/4)$ (zero rate R per year, T in years).

If α is assumed to be piecewise constant, then determining it from market data leads to a linear triangular system and could be solved, in principle, recursively.

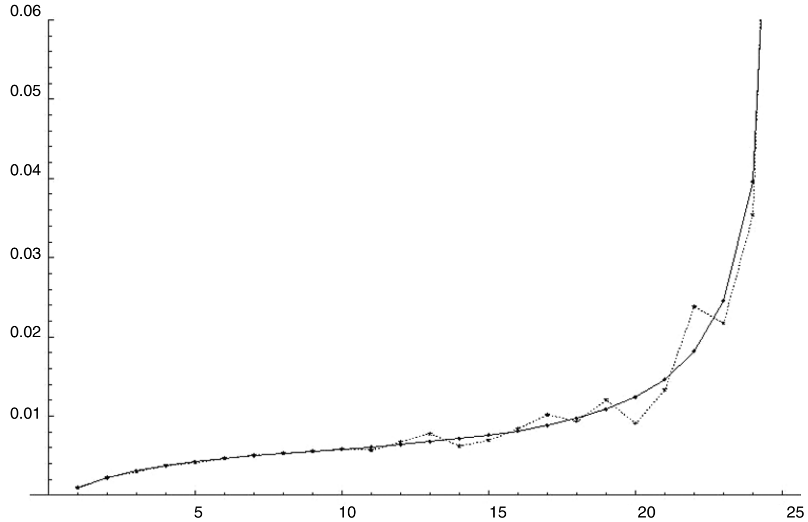
In practice, the right hand side of the linear system is not known exactly, but contains noise (bid-ask-spreads of swap rates, business-day-shifting due to holidays, to name a few).

If we use this naïve approach and have a (random) relative noise level of at most 1 percent, then we obtain

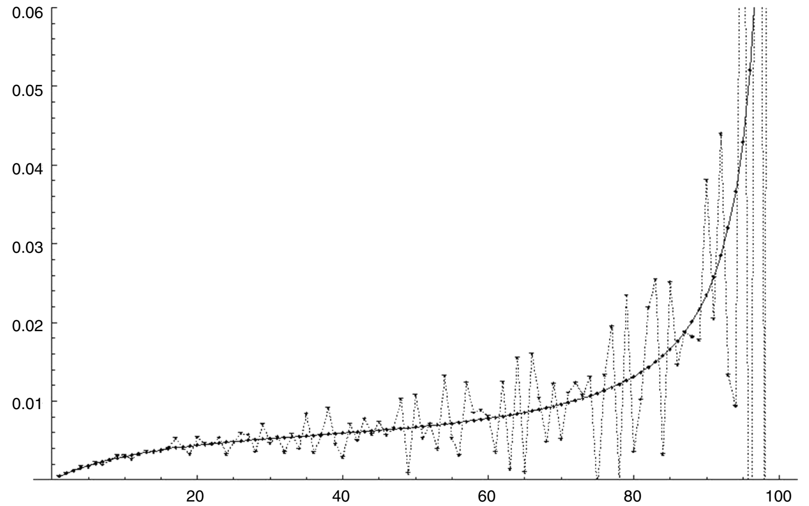




Semi-annual



Drift term in Hull-White. Data used for 25 years, annual time steps. Smooth (synthetic) solution, nonsmooth solution obtained from noisy data. Reversion speed = 0.08 (per year), volatility = 0.7%



Quarter-annual

Don't try to calculate the parameters too exactly

These results show some numerical error, which is not too severe. If one wants to get more accurate results, the natural idea would be to make the calculation grid finer. So we do the same calculations also for semi-annual, quarter-annual and monthly time-steps:

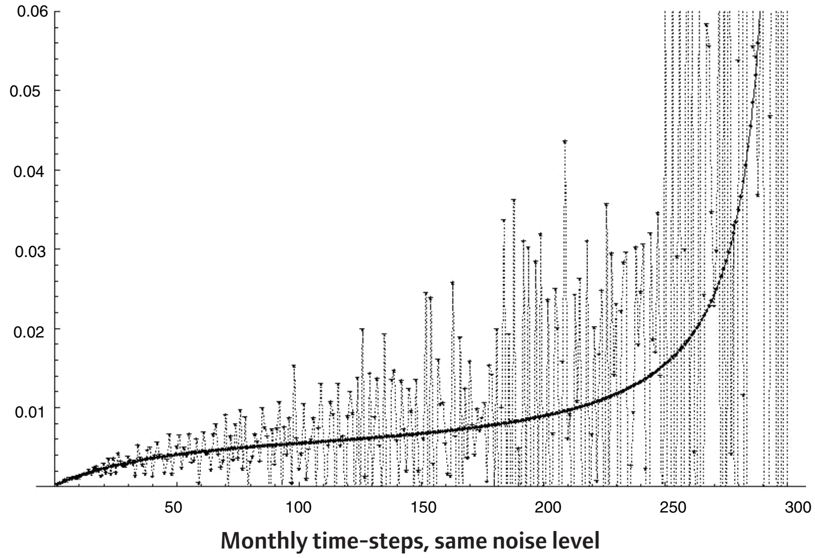
What is happening here?

A functional analytic setting

We recapitulate:

We had to solve a linear integral equation "of the first kind" (meaning that the function we are looking for occurs only inside the integral) with a noisy right hand side. When the error level remained at 1 percent relative error, the solutions got worse and worse, the finer the grid was.

Why did the oscillations in the above pictures arise?



The simplest example of a first-kind integral equation, its solution being x = f', is

$$\int_0^s x(t)dt = f(s).$$

If we perturb the right hand side by low-amplitude/high-frequency noise and thus use noisy data of the form

$$f_{\delta,n}(s) := f(s) + \delta \cdot \sin(ns/\delta),$$

then

$$\left\| f_{\delta,n} - f \right\|_\infty \leq \delta \quad \text{and} \quad \left\| D(f_{\delta,n}) - D(f) \right\|_\infty = n.$$

Thus, arbitrarily small perturbations in the right hand side f may lead to arbitrarily large perturbations in the solution if these perturbations are of high frequency. If we solve this simple integral equation numerically by discretisation, the same effect as above appears, namely that the error becomes larges as the discretization becomes finer.

The proper abstract setting in the framework of "functional analysis" of a first-kind integral equation is a linear operator equation

$$Tx = y$$

between Hilbert spaces X,Y. In both considered cases, the operator T is "compact", which is an abstract way of saying that it is smoothing. If we define a generalised "solution" as least-squares, minimum norm solution

$$(\text{minimise } \|x\| \text{ among all solutions of } \|Tx - y\| \to \text{ min})$$

Then, if T is compact between function spaces, the (generalised) inverse of T is an unbounded operator, which means that the (generalised) solution depends discontinuously on the date; this in turn implies that even finite dimensional approximations are unstable, and the instability increased with dimension. Now we have seen the reasons why these oscillations appeared and became worse with finer discretisation in the example above.

Problems whose solutions depend discontinuously on the data are called "ill posed" and typically appear when modelling "inverse problems" like parameter identification problems. This "ill posedness" can also be quantified and is, for integral equations, more severe the smoother the kernel of the integral operator is.

Regularisation: How to treat ill-posed problems properly

If we think of the (linear) operator equation Tx=y from above, one (by now classical) way to stabilise it is to solve (Tikhonov regularisation)

$$x_\alpha^\delta = \arg\min(\left\| Tx - y^\delta \right\|^2 + \alpha \|x\|^2)$$

instead of $Tx = y^\delta$.

For a linear equation, this is equivalent to solving

$$(T^*T + \alpha I)x_\alpha^\delta = T^*y^\delta$$

with T^* being the adjoint operator of T.

It can be shown (Engl-Hanke-Neubauer) that if the "regularisation parameter" α is chosen appropriately (depending on the noise level δ), then the regularised solution converges to the true solution, when δ tends to zero.

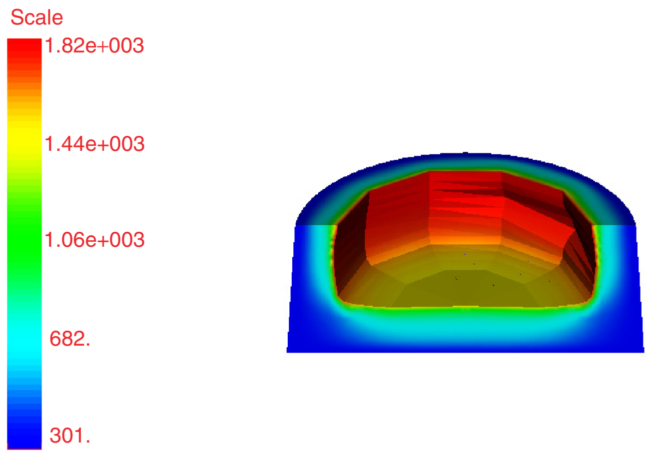
In the 1980-s, we developed a theory including implementable optimal parameter choice strategies for various regularisation techniques for linear problems (Tikhonov, Landweber iteration, maximum entropy), which we extended to nonlinear problems in the 90's. This is essential because most parameter identification problems are intrinsically non-linear. In finance, for example, the price of an option depends in a non-linear way on the volatility surface.

A blast furnace example

Obviously, parameter identification problems do not only arise in computational finance but also in science and engineering applications (for a survey, see, e.g. [Engl, Kügler]).

For example, in ironmaking by the usage of blast furnaces, liquid slag is chemically quite aggressive and therefore erodes the so-called hearth at the bottom of the furnace. This erosion makes it necessary to rebuild a blast furnace after 10 years or so, which may mean 8 years or 12 years. Shutting down a blast furnace which produces several million tons of metallic iron per year for several months, costs huge sums of money, and therefore it is essential to know if the wall of the furnace is thick enough for a continued safe operation.

Mathematically, this leads to a non-linear 3D anisotropic heat equation, where the domains of iron/slag in the interior and the domain of the brickwork is unknown and has to be determined by parameter identification techniques. The data are continuous temperature measurements of thermocouples within the brickwork. Due to the extremely rough conditions in the hearth, it happens quite frequently that thermocouples are destroyed or that they flow away into the slag, and noise is quite high. Additionally, this is a so-called sidewards heat equation (all reliable measurements are outside the brick-slag-interface) which is "severely ill posed" (in the sense of the quantification of ill-posedness mentioned above) and hence notoriously unstable. A heuristic reason for this are the strong smoothing properties of heat conduction and, more general, of diffusion "forward in time".



Calculated 3D hearth of a blast furnace. Temperatures in Kelvin.

Nevertheless, the following picture shows that by proper regularisation techniques, reasonable results can be achieved. This project is carried out in cooperation with Siemens VAI.

Adjoint problems

In the above example, one starts with a guess for the geometry of the interior boundary of the wall, calculates the temperature distribution by appropriate 3D finite elements, and then tries to minimise the Tikhonov functional consisting of the sum of squares of the distance to the measurements plus a penalty function which penalises oscillating boundaries.

To solve this optimisation problem, one needs to calculate the derivative of the solution of a 3D-heat-equation (the forward problem) with respect to the boundary of the wall which is described by hundreds of unknowns. If you tried to do this by forward or central differences, you would need to solve hundreds of 3D-nonlinear equations for one gradient calculation. When solving one forward problem takes 5 minutes (which is quite fast), it would take you 6 to 12 hours to calculate one gradient, and several days or even weeks to come to the solution which might have changed in the meantime.

The technique of adjoint problems allows to calculate the gradient with roughly the same effort as ONE forward problem, and therefore to find the actual estimate for the wall thickness ("estimate" because there is noise which cannot be removed completely) within a few hours.

Alternatively to solving the minimisation problem to which Tikhonov regularisation amounts, one could use iterative regularisation methods like regularised Gauss-Newton or Landweber iteration, where the crucial regularising effect comes from stopping the iteration at the right time: iteration too far leads to error amplification. See e.g. [Engl, Scherzer].

In all these iterative methods, adjoint problems have to be solved.

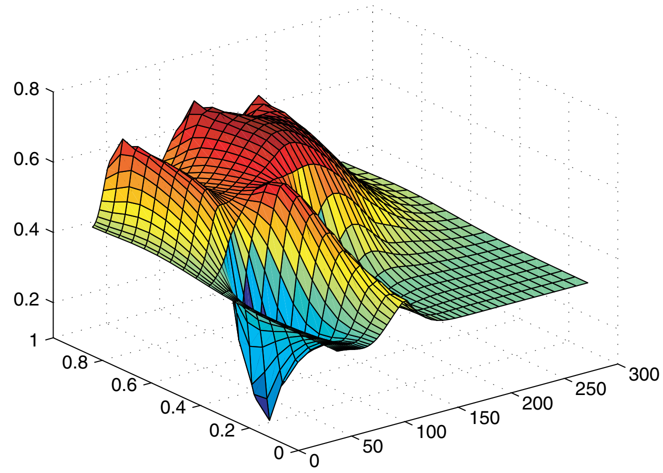
Back to finance: Nonlinear inverse problems

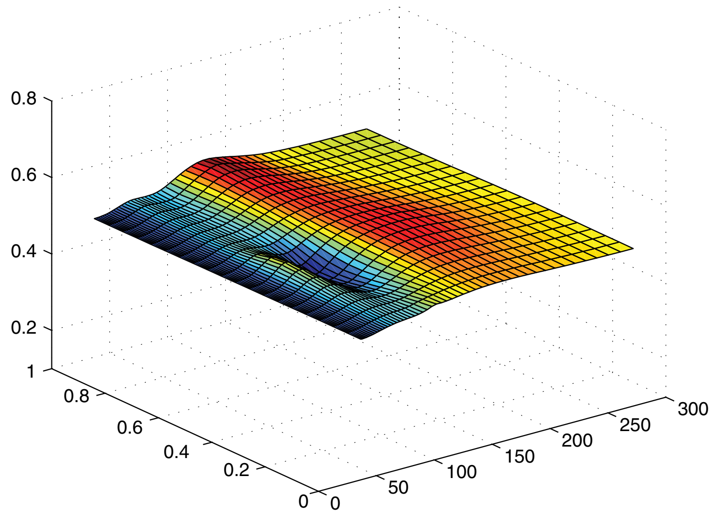
Similar techniques should be – and it turns out that they are – applicable to parameter identification problems in partial differential equations in computational finance.

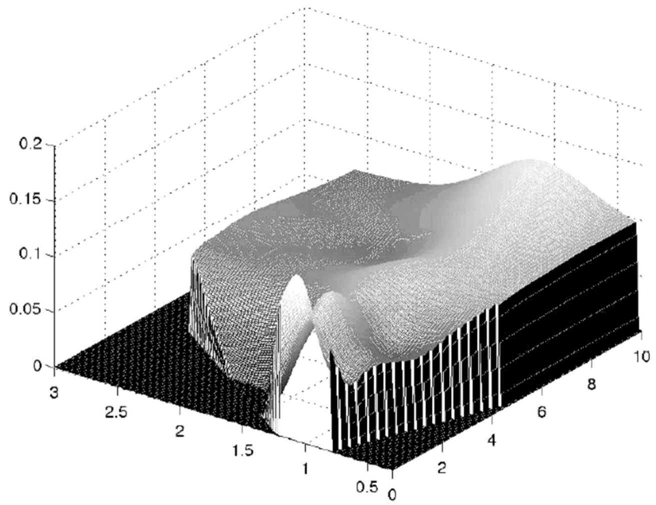
The identification of local volatility [Dupire] by regularisation techniques has been analyzed in [Egger, Engl] and extended (for example to incomplete input data) in the UnRisk PRICING ENGINE. The following figures show a local volatility surface without and with Tikhonov regularisation.

Of course, the forward equation to be solved here is much easier to solve than the 3D blast furnace equation and can be done within fractions of a second. Nevertheless, time requirements for mark-to-market prices are so tight that again adjoint techniques have to be applied. For Lévy models, leading to fatter tails, the analysis including convergence rates has been done in [Kindermann et al].

Our first example at the very beginning was a Hull-White model, which has the advantage of closed form solutions for bonds, caps and swaptions. If, we consider Black-Karasinski at the other extreme of widely used short-rate models, the calibration process needs the solution of a family of partial differential equations (for the bonds, caps, swaptions) in each iteration of minimizing a Tikhonov functional. For more information, see [UnRisk, Chapter 4.19]







Local speed function for the Merton jump-diffusion model of [Andersen – Andreasen]

Conclusions

Identification and calibration are inverse problems, which have the property of "ill posedness" which in turn leads to surprising effects: noise is amplified, this noise amplification is the more severe the smoother the problem data are, and finer discretisation/ more iterations make these problems even more severe. Based on the well-developed mathematical theory of regularisation, which is about optimal compromises between approximation and stability, one can efficiently and stably solve such inverse problems.

Acknowledgements

I want to thank Andreas Binder and his UnRisk team for valuable discussions in computational finance and for providing detailed examples.

Links

If you are interested in more Industrial Mathematics, linking basic research to applications in science, engineering and finance, visit http://www.ricam.oeaw.ac.at/media

REFERENCES

Andersen, Andreasen: Jump-diffusion processes: Volatility smile fitting and numerical methods for option pricing. Review of Derivatives Research 4, 231–262, 2000.
Binder, Engl, Grever, Mörwald: Optimal cooling strategies in continuous casting of steel with variable casting speed, Inverse Problems in Engineering 2, 289–300, 1996.
Binder, Engl, Vessella: Some inverse problems for a nonlinear parabolic equation connected with continuous casting of steel: Stability estimates and regularization, Numerical Funct. Analysis and Optim. 11, 643–671, 1990.
Binder, Hanke, Scherzer: On the Landweber Iteration for Nonlinear Ill-Posed Problems, J. Inv. Ill-Posed Problems 4, 381–389, 1996.
Brigo, Mercurio: Interest Rate Models. Theory and Practice: With Smile, Inflation and Credit, Springer Finance 2006.
Burger, Eisenberg, Engl: Inverse problems related to ion channel selectivity, SIAM Journal on Applied Mathematics 67, 960–989, 2007.
Burger, Engl, Markowich, Pietra: Identification of doping profiles in semiconductor devices, Inverse Problems 17, 1765–1795, 2001.
Chen, Engl, Landl, Zeman: Optimal strategies for the cooling of steel strips in hot strip mills, Inverse Problems in Engineering 2, 103–118, 1995.
Dupire: Pricing with a smile. Risk 7, 18–20, 1994.
Egger, Engl: Tikhonov regularization applied to the inverse problem of option pricing: convergence analysis and rates, Inverse Problems 21, 1027–1045, 2005.
Engl, Hanke, Neubauer: Regularisation of Inverse Problems, Kluwer, Dordrecht, 2nd ed., 2000.
Engl, Kügler: Nonlinear Inverse Problems: Theoretical Aspects and Some Industrial Applications, in: Capasso and Periaux (eds.), Multidisciplinary Methods for Analysis, Optimization and Control of Complex Systems, Springer Heidelberg, Series Mathematics in Industry, 3–48, 2005.
Engl, Neubauer: On an inverse problem from magnetostatics, in: Computational Techniques and Applications: CTAC-89 (eds: Hogarth and Noye), Hemisphere Publ. Comp., Washington, D.C., 1990.
Engl, Scherzer: Convergence rates results for iterative methods for solving nonlinear ill-posed problems, in: Colton, Engl, Louis, McLaughlin, Rundell (eds.), Surveys on Solution Methods for Inverse Problems, Springer, Wien/New York, 7-34, 2000.
Kindermann, Mayer, Albrecher, Engl: Identification of the local speed function in a Levy model for option pricing, to appear in: Journal of Integral Equations and Applications
Shreve: Stochastic Calculus for Finance Vol.II, Springer, 2004.
Wilmott: Derivatives: The Theory and Practice of Financial Engineering, Wiley 1988.
UnRisk Manual: Available from www.unriskderivatives.com

Author

Heinz W. Engl is Full Professor for Industrial Mathematics at the Johannes Kepler University in Linz, Director of the Radon Institute for Computational and Applied Mathematics of the Austrian Academy of Sciences (RICAM), where he also leads the inverse problems group, and scientific director of the Industrial Mathematics Competence Center (IMCC).

He is full member of the Austrian Academy of Sciences, of the European Academy of Sciences and Arts, and Fellow of the Institute of Physics (London).

At the International Congress for Industrial and Applied Mathematics (Zurich, July 2007) he has been awarded the Pioneer Prize 2007, established by SIAM (Philadelphia) for "pioneering work introducing applied mathematical methods and scientific computing techniques to an industrial problem area or a new scientific field of applications".

During the last years, Heinz has been Visiting Fellow at St. Catherine's College, Oxford, and has organised a special semester on inverse problem and its application at UCLA.

Recently, he has been elected Vice Rector for Research at the University of Vienna. He will assume this position in October 2007.