106. Band Heft 1, März 2004

D 20577

Jahresbericht

der Deutschen Mathematiker-Vereinigung

1 – 2004

Herausgegeben von A. Krieg
unter Mitwirkung von
U. Gather, E. Heintze, B. Kawohl,
H. Lange, H. Triebel

Jahresbericht

der Deutschen Mathematiker-Vereinigung

Der „Jahresbericht" ist das offizielle Veröffentlichungsorgan der Deutschen Mathematiker-Vereinigung, für dessen inhaltliche Gestaltung im Auftrag des Präsidiums der jeweilige Herausgeber zuständig ist. Im „Jahresbericht" sollen vornehmlich Überblicksartikel über Teilgebiete der reinen und angewandten Mathematik, Nachrufe sowie historische Artikel und Buchbesprechungen veröffentlicht werden.

Manuskripte:

Alle für die Schriftleitung des Jahresberichts bestimmten Briefe und Manuskripte sind an Prof. Dr. A. Krieg zu richten. Für Buchbesprechungen ist Prof. Dr. H. Lange zuständig. Bücher, von denen eine Besprechung erfolgen soll, werden bei den Verlagen angefordert. Autoren von Buchbesprechungen und Artikeln werden gebeten, die vorhandenen LATEX-style-files für den Jahresbericht zu verwenden. Somit kann der Aufwand für die Satzarbeiten erheblich reduziert werden. Sollten Illustrationen in die Arbeiten integriert werden, können diese auch in das Satzsystem übernommen werden. Dazu ist es erforderlich, dass die Daten der Abbildungen nochmals in separaten Dateien einzeln abgespeichert werden. Die LATEX-style-files sind neben weiteren Informationen im Internet verfügbar unter

http://www.mathA.rwth-aachen.de/dmv/index.html

Auf Anfrage können die style-files auch auf Diskette zugeschickt werden.
Grundsätzlich sollen nur solche Manuskripte eingereicht werden, die nicht gleichzeitig an anderer Stelle zur Veröffentlichung eingereicht oder bereits veröffentlicht worden sind. Mit der Annahme zur Veröffentlichung erwirbt der Verlag das Recht zur Vervielfältigung und Verbreitung sowie das Recht der Übersetzung in andere Sprachen.

Verlag:

B. G. Teubner Verlag/GWV Fachverlage GmbH
Abraham-Lincoln-Straße 46
65189 Wiesbaden
http://www.teubner.de
http://www.gwv-fachverlage.de

Geschäftsführer: Dr. Hans-Dieter Haenel
Verlagsleitung: Dr. Heinz Weinheimer
Gesamtleitung Anzeigen: Thomas Werner
Gesamtleitung Produktion: Reinhard van den Hövel
Gesamtleitung Vertrieb: Gabriel Göttlinger

Abo-/Leserservice:

Tatjana Hellwig
Telefon: (06 11) 78 78-1 51
Fax: (06 11) 78 78-4 23
E-Mail: tatjana.hellwig@gwv-fachverlage.de

Marketing/Sonderdrucke:

Stefanie Hoffmann
Telefon: (06 11) 78 78-3 79
Fax: (06 11) 78 78-4 39
E-Mail: stefanie.hoffmann@gwv-fachverlage.de

Abonnentenverwaltung:

(Änderungen von Adressen und Bankverbindung, Rückfragen zu Rechnungen oder Mahnung)
VVA-Zeitschriftenservice, Abt. D6F6 / Jahresbericht der Deutschen Mathematiker-Vereinigung,
Postfach 7777, 33310 Gütersloh
Ursula Müller
Telefon: (0 52 41) 80-19 65
Fax: (0 52 41) 80-96 20
E-Mail: ursula.mueller@bertelsmann.de

Bezugsbedingungen:

Die Zeitschrift erscheint 4mal jährlich zum Jahresabonnementpreis von € 102,– (176,– sFr) inkl. Versandkosten. Der Bezug von Einzelheften ist nicht möglich. Schriftliche Kündigung des Abonnements spätestens sechs Wochen vor Ablauf des Bezugsjahres.
Für persönliche Mitglieder der DMV, die den Jahresbericht zu beziehen wünschen, ist der zwischen DMV und Verlag vereinbarte Bezugspreis maßgebend, der im Rahmen des Mitgliedsbeitrags erhoben wird.

Satz: Fotosatz Behrens, D-68723 Oftersheim
Druck: Thomas Müntzer, Bad Langensalza

ISSN 0012-0456

Inhalt

Jahresbericht der Deutschen Mathematiker-Vereinigung, 106. Bd. 2004, Nr. 1

Jahresbericht der Deutschen Mathematiker-Vereinigung, 106. Bd. 2004, Nr. 1

In den nächsten Heften erscheinende Arbeiten

D. Horstmann: From 1970 until present: The Keller-Segel model in chemotaxis and its consequences II
E. H. Lieb: Quantum Mechanics, The Stability of Matter and Quantum Electrodynamics
J. Matoušek, G. Ziegler: Topological lower bounds for the chromatic number: A hierarchy
K. Mohnke: Holomorphe Kurven, Hamiltonsche Dynamik und Symplektische Topologie
M. Väth: Einige topologische Methoden in der nichtlinearen Analysis

---

Anschriften der Herausgeber

Prof. Dr. Aloys Krieg, Lehrstuhl A für Mathematik, RWTH Aachen, 52056 Aachen
E-Mail: krieg@mathA.rwth-aachen.de

Prof. Dr. Ursula Gather, Lehrstuhl für Mathematische Statistik und industrielle Anwendungen, Universität Dortmund, 44221 Dortmund
E-Mail: gather@statistik.uni-dortmund.de

Prof. Dr. Ernst Heintze, Institut für Mathematik, Universität Augsburg, 86135 Augsburg
E-Mail: heintze@math.uni-augsburg.de

Prof. Dr. Bernhard Kawohl, Mathematisches Institut, Universität zu Köln, 50923 Köln
E-Mail: kawohl@mi.uni-koeln.de

Prof. Dr. Herbert Lange, Mathematisches Institut, Friedrich-Alexander-Universität Erlangen-Nürnberg, Bismarckstraße 1½, 91054 Erlangen
E-Mail: lange@mi.uni-erlangen.de

Prof. Dr. Hans Triebel, Mathematisches Institut, Friedrich-Schiller-Universität, Ernst-Abbe-Platz 1–4, 07740 Jena
E-Mail: triebel@minet.uni-jena.de

Bezugshinweis

Früher erschienene Bände (ab Band 68) des „Jahresberichts der Deutschen Mathematiker-Vereinigung" können durch den Buchhandel oder den Verlag bezogen werden.

Nachdruck der Bände 41 bis 67 liefert: Swets & Zeitlinger, Heereweg 347b, POB 810, NL-2160 SZ Lisse/Holland

Jahresbericht der Deutschen Mathematiker-Vereinigung, 106. Bd. 2004, Nr. 1

Vorwort

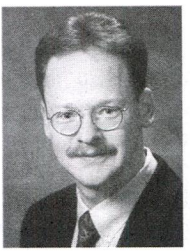
Die Konzeption des Jahresberichts sieht insbesondere Übersichtsartikel zu den verschiedenen Bereichen der reinen und angewandten Mathematik vor. In diesem Heft finden sie dazu zwei Beispiele.

Die angewandte Mathematik wird durch den Artikel der Arbeitsgruppe von Frau Pigeot (Bremen) abgedeckt. Ihr Beitrag beschreibt Anwendungen der Statistik in der Medizin und in den Biowissenschaften.

Der zweite Artikel stammt von Herrn Eschenburg (Augsburg) und ist der Geometrie zuzuordnen. Sein Beitrag beschreibt Verallgemeinerungen von Minimalflächen im euklidischen Raum und die zugehörigen Abbildungen, wobei eine Neuinterpretation von im Wesentlichen bekannten Fakten einen einheitlichen Blickwinkel auf diese Theorie ermöglicht.

A. Krieg





Iris Pigeot Karin Bammann Achim Reineke Nina Wawro Astrid Zierer

Statistical methods in genetics: From microarrays to genetic epidemiology – an overview

Abstract

- Keywords and Phrases: Association, genetic epidemiology, linkage, microarrays
- AMS subject classification: 92 D 30, 92 D 10

With the complete sequencing of the human genome in 2003 an immense step towards the biological understanding of human beings has been taken. This paper describes how the information at hand now can be used in the field of genetic epidemiology to analyze the role of genes in the etiology of diseases. We give a short introduction to genetics and review preprocessing and analytical techniques for microarray data before we cover traditional study designs. Finally we present in detail family and population based methods that enable to locate and identify candidate genes involved in the development of diseases. The outlook provides the reader with selected untackled problems and research areas.

Eingegangen: 25. 08. 2003

Iris Pigeot, Bremen Institute for Prevention Research and Social Medicine, Institute of Statistics, University of Bremen, pigeot@bips.uni-bremen.de
Karin Bammann, Bremen Institute for Prevention Research and Social Medicine, bammann@bips.uni-bremen.de
Achim Reineke, Bremen Institute for Prevention Research and Social Medicine, areineke@bips.uni-bremen.de
Nina Wawro, Institute of Statistics, University of Bremen, nina@math.uni-bremen.de
Astrid Zierer, Bremen Institute for Prevention Research and Social Medicine, Institute of Statistics, University of Bremen, astrid@math.uni-bremen.de

DMV
JAHRESBERICHT
DER DMV

1 Introduction

In recent years, the genetic information on individuals has become of growing interest with various consequences as for instance in forensics, medicine, and presumably for public health. Although the benefit of genetic research is quite obvious in certain respects, it may also lead to ethical problems, which have to be kept in mind when discussing its potential merits. We will focus here on genetic epidemiology which plays a major role when trying to identify those genes in humans, being at least partially responsible for particular diseases. The idea behind is to identify variants of the human DNA related to certain diseases with the aim to clarify the pathogenesis of the disease, to develop diagnostic tests, and finally to establish causal therapies. Despite the fact that many diseases are known to be at least partially genetically caused, the responsible genes typically have not yet been located. Furthermore, only a small fraction of the estimated ten thousands of human genes has been identified until now, although in course of the Human Genome Project the human genome has been completely mapped and sequenced in 2003. Thus, detecting genetic variants that are more frequently present in affected than in unaffected individuals still constitutes a challenging demand in research. It has, however, to be noted that only a few diseases are assumed to be monogenetic, which means that they are caused by a single gene as for instance hemophilia. Most diseases are suspected to be caused by several genes as for instance insulin-dependent diabetes mellitus where about 20 regions are assumed to contain related genes ([69]). Such diseases are called oligo- or polygenic. If the onset of a disease is additionally influenced by environmental factors, we speak of a complex disease. In polygenic diseases each single responsible gene only has a small effect on the disease and large sample sizes are required to detect a statistically significant influence.

Genetic epidemiology now comes into play when searching for candidate genes in populations where mainly two different approaches have to be distinguished. The first approach is to look for a more frequent occurrence of a specific allele in affected persons of a population compared to unaffected persons, i.e. for association between this allele and the disease phenotype. Another approach is applied within families and is based on the concept of linkage, i.e. it is investigated whether the disease is usually cosegregated with a specific marker allele.

This paper is intended to outline the complete path from collecting the data, over the most common study designs to the statistical methods for analyzing the data which typically call for special techniques being appropriate to cope on the one hand with the family structure being inherent in most data and on the other hand with the huge amount of genetic information and of additionally collected information on environmental factors. The statistical part focuses on methods for identifying candidate genes. The aspects treated in this paper are selected without the claim of completeness.

For this purpose, the paper organizes as follows. We first roughly repeat the most important terms of genetics. For obtaining genetic information on individuals, microarrays can be used. Different types and approaches for analyzing the data are presented in Section 3. Section 4 is devoted to the various study designs in genetic epidemiology. Family-based studies as well as population-based studies are dealt with from a statistical perspective in Section 5 where statistical methods to test for linkage and/or association

are discussed. Further techniques, problems, and research areas are addressed in the final section.

2 Genetic background

The entire genetic information of any living individual is coded in the DNA (desoxyribonucleic acid). The DNA forms a double strand of long sequences of four different nucleotide bases. Between these strands, the nucleotide bases adenine and thymine, guanine and cytosine, respectively, can build pairs (Watson-Crick base pairing) such that one strand of the DNA is a complement to the other and both strands fully determine each other. The human genome is organized in 23 pairs of chromosomes of which 22 are pairs of autosomes and one is a pair of gonosomes determining the sex of the individual. One set of chromosomes is inherited from the mother and the other from the father. A gene is a sequence of nucleotides of the DNA at a fixed locus which codes for a gene product (i.e. a protein), its different variants are called alleles. For each gene, the two alleles of the corresponding chromosome pair define the genotype of an individual. Every individual possesses either two different or two identical alleles at every locus, the individual is said to have a heterozygous or homozygous genotype, respectively.

The phenotype, in contrast, refers to a certain characteristic expressed in an individual. The relation of genotype and phenotype is defined by the so-called penetrance of a gene. Let A denote any one gene locus with alleles $A_1, \ldots, A_n$ and probabilities $Pr(A_i)$ for allele $A_i, i = 1, \ldots, n$, to occur. Let further X denote a corresponding phenotype. The penetrance f_{ij} of expressing X for the genotype (A_i, A_j) is given by the conditional probability

$$(1) \quad f_{ij} = Pr(X|A_i, A_j)$$

with $i, j = 1, \ldots, n$. A Mendelian segregation implies that penetrances are complete, that is 0 or 1. An allele A_i is called dominant if $f_{ij} = f_{ji} = f_{ii} = 1$ for $i \neq j$, it is called recessive if $f_{ij} = f_{ji} = 0$ for $i \neq j$ and $f_{ii} = 1$.

In the following, the phenotype affected vs. unaffected of certain diseases is of special interest. Monogenetic diseases are determined by just one gene locus and complete penetrances. However, most diseases of interest, e.g. cancer, are complex in that more than one gene locus is involved and penetrances are sometimes not symmetric due to genetic imprinting and usually incomplete, i.e. the probability of getting affected might be less than one or greater than zero although the genetic predisposition is given or not, respectively. Moreover, the incidence of a disease can also depend on certain environmental factors and so-called phenocopies might occur which means that an individual can become affected without having the particular disease gene.

2.1 Mendelian segregation

According to the Mendelian laws, each of the two alleles of a parent is transmitted independently and at random with probability 0.5 to an offspring. If the genotypes of the

parents are known, we can calculate probabilities for the offspring to have a certain genotype and – with known penetrances – as well to have a certain phenotype. In the situation of Figure 1 the probability for an offspring to be affected or unaffected according to the Mendelian laws is

$$Pr(\text{affected}) = \frac{1}{4} \quad \text{for a recessive mode of inheritance,}$$

$$Pr(\text{affected}) = \frac{3}{4} \quad \text{for a dominant mode of inheritance.}$$

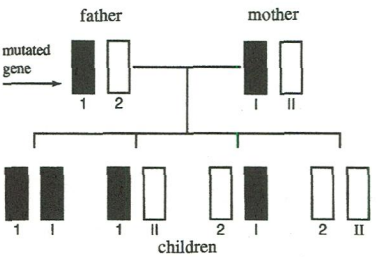


Figure 1: Pattern of inheritence

Another important fact resulting from the Mendelian laws is that the distribution of genotypes in a population is constant from the second generation on. This phenomenon, known as Hardy-Weinberg equilibrium, is valid only under certain conditions. Among others the population has to be large enough and mating has to be random. Of course, the probability of mating or the capability of reproducing living offspring may not depend on the genotype.

This approach can be extended to more than one gene locus, provided the inheritance of the different genes follows the Mendelian laws.

2.2 Linkage of gene loci

Deviations from the Mendelian laws can occur if two genes are not inherited independently. This happens, when they are located on the same chromosome such that they cannot always be separated during meiosis. Recombination is the exchange of parts of the parental chromosomes due to crossing-over of chromatides (the strands of the chromosomes). This results in the derivation of alleles from different parental chromosomes if an odd number of crossing-overs occurs (see Figure 2). The probability of this event is called recombination fraction θ.

The closer two genes are located to each other the more frequent they will be inherited together. In this case they are called 'linked'. If they are not linked joint inheritance is by chance only, i.e. with a probability of 0.5. In this case the recombination fraction θ equals 0.5, whereas $\theta = 0$ means tightest possible linkage as the probability of recombination is zero. The term linkage therefore refers to the location of two or more genes on the same chromosome and is related to gene loci and not to specific alleles. Following Donnelly [22], let S_{jk} be 0 or 1 denoting whether in meiosis j at locus k the maternal or

paternal gene of the parent is transmitted to the offspring. Then between two loci k and l, the recombination fraction is defined as

(2) $\theta_{kl} = Pr(S_{jk} \neq S_{jl})$,

assuming that the recombination probability does not depend on the meiosis j.

The basic idea of looking at two gene loci is to see whether the alleles of a marker gene co-segregate with the disease gene among related subjects which would imply that they are linked. This gives an idea of the location of the disease locus since the one of the marker locus is known. Thus, the identification of linkage to a marker locus is often the first stage in positional cloning of a disease gene.

In general, one has to distinguish between genetic distance and recombination fraction. The genetic distance d, measured in centi-Morgan (cM), refers to the expected number of crossing-over events between two loci. Only over short distances, e.g. less than 5 cM, a recombination fraction of 0.01 is about the same as a genetic distance of 1 cM. In contrast to the genetic distances of a series of loci, the corresponding recombination fractions are not additive. Mapping functions establish the relationship between these two measures of distance.

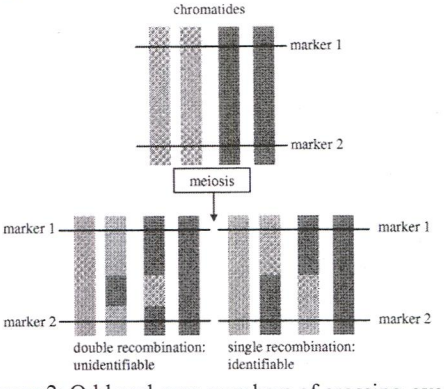


Figure 2: Odd and even numbers of crossing-overs

Assuming a homogenous Poisson process for the number of crossing-overs that occur, Haldane [40] relates the genetic distance and the recombination fraction θ by

(3) $d = -0.5 \ln(1 - 2\theta)$.

This model reflects the idea of 'no inference' between number and locations of crossing-over events. Another widely used mapping function was introduced by Kosambi [53]. If one assumes that the occurrence of one crossing-over prevents another one from taking place closely,

(4) $d = 0.25 \ln\left[\frac{(1 + 2\theta)}{(1 - 2\theta)}\right]$

is the appropriate relation of recombination fraction and genetic distance.

2.3 Expression of genes

Sometimes, it is not a gene itself, but rather the gene product that is of interest. Although each cell of an individual carries the same DNA, the cells work quite differently, e.g. liver cells, brain cells. The diversity arises from the activation of genes in a cell that means which proteins are actually built. The whole process of building proteins from genes is called gene expression and can be summarized basically by two steps. In a first step, genetic information of the double-stranded DNA is transcribed into the single-stranded messenger RNA (ribonucleic acid) or mRNA. This process is called transcription. In a second step, the information contained in the mRNA is translated into a protein. This process is called translation. Because of existing difficulties in measuring the amount of proteins built, the expression of a gene is rather evaluated by measuring the amount of mRNA in a cell.

3 Microarrays: A widespread data-sampling approach

Microarray technology opens a variety of new possibilities to study the human genome complexity ([13]). Microarrays derive their power and universality from a key property of DNA molecules, the complementary base-pairing. The number of complementary pairs needed to keep the two strands together depends, besides others, on environmental factors, mainly on the temperature. To get single stranded DNA it has to be denatured, this means to split the strands by rising the temperature. When two long complementary single strands of DNA meet, they tend to stick together even though the binding power of every single pair is only weak. The term hybridization refers to this annealing of nucleic acid strands from different sources according to the base pairing rules. Microarrays can be seen as an enhancement of the Southern or Northern Blots ([84]). These blots allow to analyze a fragment of the DNA or RNA, respectively. Compared to the blots (which are only able to analyze one sequence at a time) the microarray experiment facilitates the researcher to simultaneously consider tens of thousand of sequences or genes, respectively. Due to this huge amount of data and the typically only small number of repeated measures, the analysis of such arrays calls for a particular statistical approach.

Obviously, microarrays offer a broad range of applications. They are, for instance, used for genotyping, measuring the transcript abundance, identification of genome-phenotype relationships, or for determining identity by decent (see Section 5.1.2). They can also contribute to understand the networks of gene regulation. In medical practice, microarrays can be applied in cancer research to characterize the tumor on a genomic scale for a more reliable diagnosis and an improved treatment of cancer. Another application is in immunology, where the host genomic response to bacterial infection is studied.

Each microarray experiment consists of a number of steps: the determination of the queried DNA-sequence, the laboratory tasks, the image processing, the data normalization, and the pursuing analysis. Each step has an impact on the outcome of such an experiment.

3.1 Types of microarrays

There are two main types of microarrays, namely cDNA microarrays and high density oligonucleotide arrays. Using cDNA microarrays, DNA is spotted onto a solid support as for example a glass slide; with the latter, oligonucleotides consisting of up to 25 nucleotide bases are synthesized on to a silica slide by a process known as photolithography.

3.1.1 cDNA microarrays

cDNA microarrays consist of several DNA sequences spotted in a high density array on a glass slide using a robotic arrayer. For queries of mRNA, the samples, called targets, are reverse-transcribed into cDNA, labeled using a fluorescent dye or a radioactive marker and hybridized with the arrayed DNA sequences denoted as probes. The targets align to the corresponding probes, as described above (following the definition of probe and target given in [1]). After hybridization, the slides are scanned. The resulting image

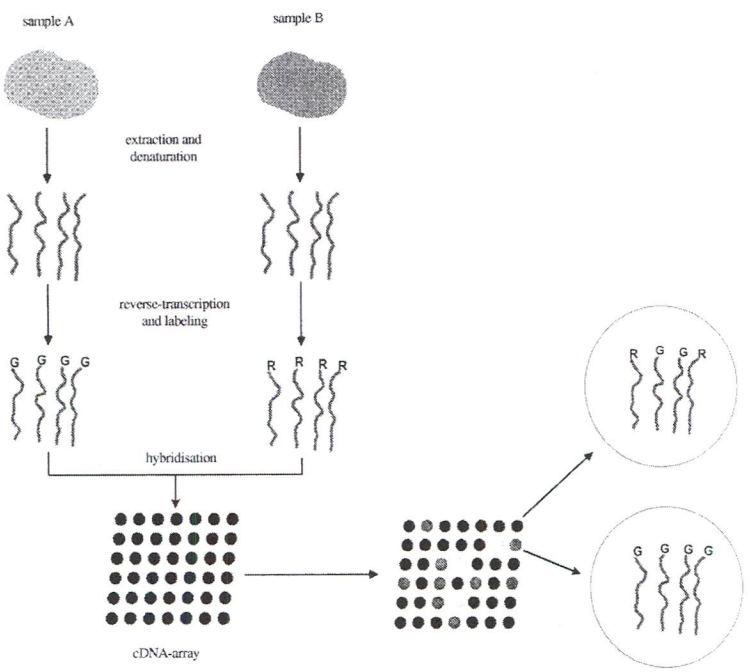


Figure 3: Two color cDNA microarray

can be used to determine the position and abundance of the aligned targets on the slide. Combined with the knowledge about the contents of the spotted array, the incidence of a queried sequence can be assessed. The resulting values can only be assessed relatively to each other, microarrays are not able to measure the absolute amount of occurrence because of hybridization kinetics. Comparing different slides therefore needs further

standardization and correction. One possibility to avoid some of the resulting problems is the parallel hybridization of two samples labeled with two different dyes on the same slide. These two color-settings allow a direct comparison of two samples. The relative abundance of the spotted DNA sequences may be assessed by monitoring the differential hybridization of the two samples. Figure 3 describes the main steps in a two color cDNA microarray experiment.

3.1.2 High density synthetic oligonucleotide arrays

Another approach to design microarrays is based on the usage of synthetic oligonucleotides. Synthetic linkers modified with photochemically removable protecting groups are attached to a glass substrate. Afterwards light is directed through a photolithographic mask to produce localized photodeprotection. Chemical building blocks are incubated with the surface, and chemical coupling occurs at those sites that have been illuminated in the preceding step. In the next step, a new mask is used. The light is directed to different regions of the substrate and the chemical cycle is repeated until the desired sequences are generated. Following the completed preparation, the array is ready for hybridization. This process is similar to the cDNA-array but solely one color can be used. The parallel hybridization of two samples is impossible. The design is based on sequence information alone, without the need for physical intermediates such as clones, polymerase chain reaction (PCR) products, or cDNAs, just to name some of them (see [59]).

To improve the abilities for error correction and to enhance the reliability the detector for a queried DNA-sequence is combined of different probes. Each so-called probe set consists of multiple oligonucleotides (probe cells) of different sequences designed to hybridize on different regions of the same RNA. The use of multiple independent detectors for the same gene improves the reliability of the results. An additional level of redundancy is achieved by the use of mismatch (MM) control probes. The mismatch probes are identical to their perfect match (PM) partners except for a single base difference in a central position. The probes are arranged as probe *pairs* PM and MM. The MM probes act as specificity controls that allow discrimination between 'real' signals and those due to non-specific or semi-specific hybridization (see Figure 4).

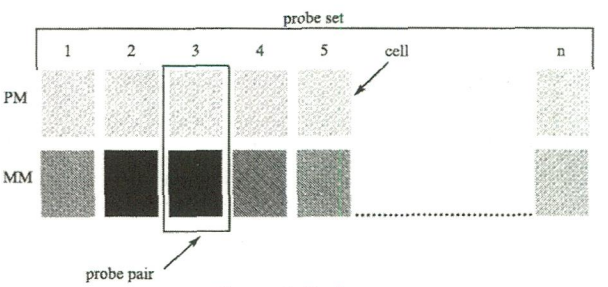


Figure 4: Probe set

3.2 Image analysis

Before any values can be computed and passed to the data analysis pipeline, appropriate image recognition has to be carried out. Currently, different methods are discussed in the literature. Simple methods use a threshold to locate the spot in the image and distinguish between background and foreground values. Alternative techniques search for connected groups of similarly colored pixels by defining a circle around the center of the located spot used to define an inner (spot) as well as an outer (background) region ([2]). More sophisticated algorithms are based on seeded region growing algorithms ([4]) or adaptive pixel selection (APS) ([76]).

3.3 Normalization

Microarrays simultaneously measure transcript abundances for thousands of genes in a cell population or tissue sample. Due to variations in sample treatment, labeling, dye efficiency and detection, the spatial position on the slide and differences between slides, the fluorescence intensities can in general not be compared directly ([41]). An appropriate standardization is required to deal with the systematic and random effects occurring as a consequence of the preceding process described. This standardization is often referred to as normalization. Normalization is meant to allow for comparability of the values within an array as well as between different arrays. Normalizing the intensities is elementary for the pursuing analysis. There are different methods available to normalize the data, depending on the experimental conditions, the type of array and further factors. In the following, we review some aspects of the preprocessing of the data.

The first decision to be made concerns the selection of genes (intensity values) to be used for normalization. Often the vast majority of intensity values is chosen, assuming only a few changed values between the arrays, which implies that the distribution of the majority of genes for each array is nearly the same ([47]). Another suggestion is to select so called housekeeping genes, a set of genes specifically chosen for normalization, existent in all targets. Housekeeping genes are typically highly expressed and do not allow to assess the dye biases for less expressed genes. To circumvent this disadvantage a specially designed microarray sample pool (MSP) may be selected as set of controls. The values of an MSP should span a wide intensity range such that they yield a 'correction factor', applicable to all classes of expression intensities facilitating the adjustment task ([106]). As control genes are used to correct for spatial and intensity based trends in the data, they should be present in a sufficient quantity and uniformly distributed across the whole slide.

An alternative method described for oligonucleotide arrays by Schadt et al. [76] and Tseng et al. [95] is based on an approximately invariant set of genes. A set of genes is said to be invariant if their intensity ranks are approximately the same for all considered experiments. In practice this method works well for normal settings with a small number of slides. In a complex setting where a large number of slides has to be normalized, the set of invariant genes is too small for comprehensive normalization. If there is a sufficiently large number of invariant genes, this method is similar to the proposals described above.

After having fixed the set of genes for normalization, different approaches for calculating or performing the adaption of the slides can be thought of. These may be roughly distinguished as scaling approaches, smoothing approaches, transformation methods such that the distribution of intensities is the same across all arrays, and model-based approaches. Of course, not all techniques discussed in the literature can be reviewed here. To give an idea of the variety of methods used for normalization let us describe a few as representatives of these major types.

For illustrative purposes, let us in the following define in a two color cDNA-setting y_{gj} as the intensity of gene g, $g = 1, \ldots, G$, and color j, $j = 1(\text{red}), 2(\text{green})$, $M_g = \log_2(y_{g1}/y_{g2})$ the difference in log expression values, and $A_g = \frac{1}{2}\log_2(y_{g1}y_{g2})$ the average of the log expression values. The use of a logarithmic scale, implicitly assuming multiplicative effects in the data, offers certain advantages. Typically the random variation of the original data roughly increases linearly with the average signal intensity. The use of an additive model, based on the logarithmic scale, simplifies the next steps of the analysis in so far as among others the variability tends to be constant.

A simple scaling approach is based on the calculation of the global median or mean of all selected values across the whole array. This constant correction term, briefly denoted as c, is then used to uniformly correct all intensity values M_g. The corrected values $\tilde{M}_g$ are thus calculated as

$$(5) \quad \tilde{M}_g = M_g - c, \; g = 1, \ldots, G.$$

This approach neglects differences caused e.g. by the spatial position on the slide, different intensity based biases or the specific character of the queried genes. These various sources of biases can be accounted for by non-linear methods using local regression to fit curves to noisy data by a multivariate smoothing procedure. Here, a linear or quadratic function of the predictor variables is fitted in a moving fashion, analogously to a moving average in time series analysis (see Cleveland [16]). Some examples applying robust locally weighted regression are given in Dudoit et al. [23], Åstrand [7] and Yang et al. [106]. The calibration factors are now estimated based on a suitable robust scatterplot smoother $c(A_g)$, which leads to

$$(6) \quad \tilde{M}_g = M_g/c(A_g), \; g = 1, \ldots, G.$$

This smoothing approach can be further generalized by fitting different curves $c_r(A_g), r = 1, \ldots, R$, to different regions of the array.

The next approach described tries to correct the probe intensity values for each array in a set of arrays by adjusting the different distributions. The quantile normalization technique introduced by Bolstad et al. [10] extends the underlying idea of a quantile-quantile plot, that the distributions of two data sets are the same if their quantiles lie on a diagonal, to I dimensions, I representing the number of all arrays to be compared. For this purpose, i.e. to achieve an adjustment of the different distributions, consider the vector of the s-th quantiles, $s = 1, \ldots, p$, $\boldsymbol{q}_s = (q_{s1}, \ldots, q_{sI})$ for all I arrays and calculate the mean quantiles across all arrays. These values then substitute the values of the data items in the original data set. This transformation implies that the quantiles all lie along the unit diagonal defined as $\boldsymbol{d} = (\frac{1}{\sqrt{I}}, \ldots, \frac{1}{\sqrt{I}})$ because of the above described projection of $\mathbf{q}$ onto $\mathbf{d}$ with

$$\text{proj}_{\boldsymbol{d}}\ \boldsymbol{q}_s = \left(\frac{1}{I}\sum_{i=1}^{I} q_{si}, \ldots, \frac{1}{I}\sum_{i=1}^{I} q_{si}\right).$$

Besides the more explorative methods presented above, model-based approaches can be found in the literature (cf. [15], [46], [48], [56], [65]). In case of two-color cDNA microarrays, Chen et al. [15] presumably present the first statistical approach which compares the expression levels of two different samples based on only one array with a huge number of genes. The samples are dyed with different colors and the ratio of the two intensities measured at one spot is inspected. Provided that an appropriate background correction has already been implemented in the preceding image analysis, the following multiplicative model for the intensities is introduced, where we follow the notation by Huber et al. [47]:

$$(7) \quad Y_{gk} = b_k\, \mu_{gk}(1 + \epsilon_{gk}), \qquad \epsilon_{gk} \sim N(0, v^2).$$

Here, Y_{gk} denotes the intensity for gene g, $g = 1, \ldots, G$, and sample k, $k = 1, 2$, μ_{gk} is the true intensity of the appropriate gene and sample, and b_k denotes the normalization factor for sample k, such that $E(Y_{gk}) = b_k \mu_{gk}$. The random variables Y_{g1} and Y_{g2} are supposed to be independent and the variance v^2 of the noise ϵ_{gk} is assumed to be constant for all genes and samples. This implies that the coefficient of variation of the Y_{gk} is constant and for any g and k the variance of Y_{gk} can be estimated which makes the model identifiable. The approximate distribution of the statistical test based on the ratio of the two intensities at a certain gene g can be derived under H_0 exploiting (7) to test the hypothesis

$$H_0 : \mu_{g1} = \mu_{g2} \quad \text{vs.} \quad H_1 : \mu_{g1} \neq \mu_{g2}.$$

In contrast to the normality assumption of Chen et al. [15], Newton et al. [65] for instance propose a model based on the assumption of independent and gamma-distributed intensities of the red and green samples.

As in many other fields, analysis of variance (ANOVA) are frequently applied in this context. Kerr et al. [48] introduce a model with four factors but without including all interaction terms, namely

$$(8) \quad \log Y_{ijkg} = \mu + A_i + D_j + V_k + G_g + (AG)_{ig} + (VG)_{kg} + \epsilon_{ijkg},$$

where μ denotes the overall average signal, A_i represents the effect of the i-th array, D_j the effect of the j-th dye, V_k the effect of the k-th variety, G_g the effect of the g-th gene, $(AG)_{ig}$ represents the interaction effect between array i and gene g, $(VG)_{kg}$ the interaction between variety k and gene g. The errors ϵ_{ijkg} are assumed to be independent and identically distributed with mean 0. Further effects can easily be accounted for by just enlarging the above model. However, this results in a loss of degrees of freedom required to estimate the error variance in the experiment. But note that the solely effects of interest are usually the interactions between varieties and genes, where the varieties refer to the categories of the factor under investigation such as tissue types. The model is fitted via least squares with usual parameter constraints to make it identifiable.

An interesting generalization of the approach of Chen et al. [15] is proposed by Huber et al. [46]. They model the error of non-differentially expressed genes according to

Rocke and Durbin [74] such that the variance of the measured intensities is quadratically dependent on the mean. Using the following variance-stabilizing transformation

$$h_i(y_{gi}) = \operatorname{arsinh}(a_i + b_i y_{gi})$$

with a_i and b_i appropriate parameters also accounting for a calibration, the variance of the transformed intensities becomes approximately independent of their mean. Both steps of data transformation, namely calibration and variance-stabilization, are incorporated in an appropriate statistical model as follows

$$(9) \quad h_i(Y_{gi}) = \mu_g + \epsilon_{gi}, \quad \epsilon_{gi} \sim N(0, v^2), \quad g \in \tilde{G},$$

where $\tilde{G}$ denotes the set of probes representing the vast majority of non-differentially expressed genes, $\mu_g = E(h(Y_{gi}))$. It allows to account for an arbitrary number of samples and can also be applied to oligonucleotide arrays. The model parameters are estimated maximizing the corresponding profile log-likelihood. For details we refer to Huber et al. [46].

There are also model-based approaches that account for the specific requirements of the oligonucleotide arrays. Since only one sample is hybridized on the array (at a time) the mismatch (MM) probes are used to validate the measured intensities at the perfect match (PM). Li and Wong [56] consider the following statistical model for one probe set, representing one particular gene g:

$$(10) \quad Y_{il} = PM_{il} - MM_{il} = \theta_i \phi_l + \epsilon_{il}, \ \sum_l \phi_l^2 = L, \ \epsilon_{il} \sim N(0, \sigma^2).$$

The difference between the perfect match and the mismatch probe is investigated for each individual probe l, $l = 1, \ldots, L$, of the probe set on the i-th array, $i = 1, \ldots, I$, where θ_i represents the so-called model-based expression index, ϕ_l the probe-specific rate of increase of corresponding PM response and ϵ_{il} the random error, assumed to be independent identically normally distributed according to equation (10). The above constraint w.r.t. ϕ_l is only to ensure the identifiability of the model. The parameters are estimated by an iterative application of the least-square method. As alternative, Li and Wong [57] also introduce a model which only accounts for the PM values.

Based on a suitable preprocessing of the data statistical analysis can be carried out which will be discussed in the next subsections.

3.4 Statistical analysis

The statistical analysis of data gained from microarrays is typically based on appropriately normalized data sets. The techniques applied range from statistical tests for possibly different gene expressions for instance under treatment versus control conditions to multivariate techniques primarily to detect structures in the data set which may then be used to reduce its dimensionality.

3.4.1 Differential expression

Investigating the expression of genes it is of special interest whether the levels of expression for two (or more) varieties, say e.g. due to two cell types or treatments, is different. First ideas to determine differential expression of genes in microarray analysis can be found in DeRisi et al. [20] and Schena et al. [80]. More sophisticated approaches based on various distributional assumptions are for example introduced by Chen et al. [15] and Newton et al. [65] to state differential expression for single-slide experiments.

For a comparison of multiple slides, based on the model proposed in Kerr et al. [48] the difference in the variety-gene interactions of two varieties $VG_{2g} - VG_{1g}$ can be estimated without any preceding normalization. Because typically neither is the assumption of normality fulfilled nor is the number of observations large enough to justify asymptotic arguments, the bootstrap percentile is suggested to calculate confidence intervals for the differences.

Another possibility to cope with the non-normality of the data is to perform a logarithmic transformation. Thus, let x_{kgm}, $m = 1, \ldots, n_k, k = 1, 2, g = 1, \ldots, G$, denote the log intensity ratios. The application of such a transformation beforehand is the rationale behind the application of the standard t-test statistic T_g to identify differentially expressed genes with

$$(11) \quad T_g = \frac{\bar{x}_{1g} - \bar{x}_{2g}}{\sqrt{\frac{s^2_{1g}}{n_1} + \frac{s^2_{2g}}{n_2}}}.$$

Here, $\bar{x}_{kg}$ denotes the mean expression level of gene g and variety k, s^2_{kg} the sample variance of the expression level of gene g, and n_k the number of replicates of variety k. The above t-test or the Welch's t-test in case of unequal variances are standard techniques to test for equal levels of expression for gene g. Nevertheless, the assumption of normality may even not hold after having applied a suitable normalization. Thus, Dudoit et al. [23] propose to use a permutation test instead, where the number of permutations is limited to 1000 if it is no longer feasible to perform all possible permutations. Another nonparametric alternative is given by the classical Wilcoxon rank sum test (see [25]).

A further undesired behavior of T_g (see (11)) can be observed when the estimated standard deviation is quite small, since then, even if the difference in the nominator is very small, T_g tends to large values leading to false positive results. Possible ways out are discussed in [27], [60], [96] who propose to add a suitable constant s_0 to the estimated standard deviation in the denominator of (11). This constant may be chosen such that the coefficient of variation is minimized or just as the 90th percentile of the standard errors of all genes.

In microarray experiments the number of genes G under investigation is usually high, up to several thousands, with a relatively low number of replicates. Since the aim of a statistical analysis is here to identify those of thousand candidate genes which have their expression levels changed and since this identification process is based on numerous statistical tests simultaneously performed for each single gene g, a multiple testing problem occurs. There are various approaches ([23], [24], [25], [26], [27], [60], [96]) to deal with this problem. Dudoit et al. [23] mention some methods for adjusting p-values

where the simplest is to consider the Bonferroni single step adjusted p-value for gene g given as

(12) $\tilde{p}_g = \min(Gp_g, 1)$.

Here, p_g denotes the unadjusted p-value, which is usually very conservative. A less conservative, stepwise procedure is proposed by Holm [43]. But since this procedure does not account for the correlation structure among the variables, it still can be improved. Therefore, Dudoit et al. [23] apply the procedure introduced by Westfall and Young [101] which exploits the correlations between the genes.

Besides the adjustment of the significance level α other criteria to account for multiplicity may be applied. In microarray settings the false discovery rate (FDR)

(13) FDR = E(proportion of rejected hypothesis that are actually true)

proposed by Benjamini and Hochberg [8] is frequently used instead (cf. Tusher et al. [96]).

Efron et al. [27] propose an empirical Bayes approach assuming there are two groups, the differentially and the non-differentially expressed genes, with prior probabilities p_1 and $p_0 = 1 - p_1$ and corresponding prior densities $f_1(z)$ and $f_0(z)$ for some kind of expression score Z. The mixture density is then given as $f(z) = p_0 f_0(z) + p_1 f_1(z)$. Application of Bayes' theorem yields the a posteriori probabilities of differential expression $p_1(z)$ and $p_0(z)$ as

(14) $p_1(z) = Pr\{\text{differently expressed}|Z = z\} = 1 - p_0 f_0(z)/f(z)$

(15) $p_0(z) = Pr\{\text{non-differently expressed}|Z = z\} = p_0 f_0(z)/f(z)$.

For Bayesian analysis, specification of the a priori probabilities and densities is required. For details, how to estimate the a posteriori probabilities in microarray settings see Efron et al. [27]. They also emphasize a close connection between the FDR and the empirical Bayes methodology.

3.4.2 Cluster analysis and classification

The high-dimensionality of microarray data calls for multivariate methods to analyze their structure and to reduce their dimensionality and complexity. Since a detailed presentation of such multivariate techniques would be beyond the scope of this review paper, only a brief overview over the methods typically applied in this field will be given. A standard tool to reveal the structure of a multidimensional data set is cluster analysis. On the one hand, it can be used to identify groups of genes which act similarly among the samples w.r.t. gene expression. This may be a sensible approach for instance to find groups of genes which are correlated to a specific phenotype. On the other hand, it is of interest to group e.g. samples of tumor cells to assign the type of tumor or to identify new subgroups. Besides hierarchical clustering techniques ([29], [42]), non-hierarchical methods as the k-means procedure and self-organizing maps ([90]) as well as model-based clustering approaches ([107]) are applied to microarray data.

Analogously to cluster analysis classification methods offer a valuable approach to tackle the problem of how to structure microarray data. There are several techniques,

for instance based on regression analysis ([28], [39]) including prediction analysis of microarray ([94]), discriminant analysis ([58]), and support vector machines ([12]).

4 Genetic epidemiology

Genetic epidemiology is 'the study of the role of genetic factors and their interaction with environmental factors in the occurrence of disease in human populations' ([55]). In contrast to common genetic studies the definition of genetic epidemiology refers explicitly to populations, underlining thus the epidemiologic component of the field (see [91]). Besides of general population studies, study frameworks can be e.g. population isolates, intermarriage between populations, and migration of populations into different environments. Much more emphasis than in traditional epidemiology, however, is on the inclusion of familial data. We would first like to address some study designs before entering in a more in depth discussion of the statistical methods applied in studies of genetic epidemiology.

Apart from traditional epidemiologic study designs, like case-control, cohort, or cross-sectional designs, which are potentially suitable to assess the influence of genetic factors on a disease, there are several newly proposed designs that are mainly developed directly from the traditional approaches and incorporate familial data. Following a short review of the cohort and the case-control design, the main study types in genetic epidemiology are presented according to their traditional sampling perspective, although studies on genetic factors inherently include a retrospective part, the genome of the study subjects.

4.1 Traditional study designs

Traditional study designs are especially suitable if one or more putative genes for a disease are known or if a limited number of hypotheses exists concerning the genetic factors. Genetic factors and even familial information can be included like any other exposure variable in the study, allowing thus the analysis of single factors, and of gene-gene and gene-environment interaction with statistical methods used in traditional epidemiology. The probably most intuitive study design in epidemiology is the *cohort study*, where a defined population under exposure or at risk of exposure is followed prospectively over a long period with regard to the exposure and possible sequels of the exposure. For different exposure groups, frequency measures like morbidity or mortality rates can then be calculated in a straightforward fashion and compared between exposure groups. If the disease of interest is rare and if the suspected risk factors are not uncommon in the population, the *case-control study* can be an economically attractive alternative to the time- and money-consuming cohort study. In case-control studies incident cases are compared to unaffected population or hospital controls, with regard to their exposure in the past. Odds ratios can then be calculated that estimate the relative effect of risk factors on the disease via comparisons of the exposure frequencies in the case and the control groups. Other approaches used in epidemiology are e.g. nested

case-control studies and case-cohort studies, where the cases and controls are sampled from an assembled cohort. However, cohort and case-control studies are, like most traditional approaches, subject to potential confounding because of undetected genetic heterogeneity within the study population (population stratification, cf. to Section 5.2). This problem is not unknown in traditional epidemiology, however, it can intensify severely when looking at genetic risk factors although the general impact of unaccounted genetic heterogeneity on study results is not clear since bias arises not necessarily in all situations (for a discussion see [92], [99], [14]). Most of the following non-traditional approaches try to circumvent confounding due to undetected genetic heterogeneity by incorporating family members of study subjects in the study design.

4.2 Retrospective designs: Derivatives of the case-control approach

Possible variants of the traditional case-control approach for studies in genetic epidemiology are family-based case-control designs, where (real or hypothetical) controls are derived from the family of the cases. In *case-sibling studies*, each case is matched to a sibling, who was unaffected at the age the case became affected (index age). Standard methods of matched case-control studies, like conditional logistic regression, can be used within the case-sibling design, analyzing both genetic factors and environmental data of cases and controls up to their index age, and gene-environment interaction. In *case-parent studies*, pseudo-siblings are constructed of the alleles of the cases' parents as hypothetical controls. For each case three pseudo-siblings are obtained by constructing the genotypes of those allelic combinations that were not transmitted to the respective case. The pseudo-siblings are presumed to be unaffected. For analyzing case-parent studies, standard methods of 1:3-matched case-control studies can be applied as well. However, because of the hypothetical control, main effects of environmental factors cannot be estimated within the case-parent design. While controlling securely for underlying population structure, both designs have other potential pitfalls, the case-sibling design requires dropping of cases without unaffected siblings, thus possibly introducing selection bias; the case-parent design requires the genotyping of both parents, and is therefore not suitable for late endpoints. Moreover, the dropping of incomplete family trios is here as well a potential source of selection bias. A special approach for analyzing gene-environment interaction, when genotyping of parents or siblings is not feasible, is the *case-only design*, which uses only information from affected subjects. The case-only design makes use from the fact, that for dichotomous genetic factors and dichotomous environmental factors the odds ratio for the genetic factor in relation to the environmental factor in cases is an estimate for the multiplicative interaction effect of genetic and environmental factor in the population, if the genetic and the environmental factor are independent in the unaffected part of the population ([68]). The gene-environment interaction effect can then be estimated via logistic regression with the gene as the dependent and the environmental factor as the independent variable. It should be noted though that the case-only design can of course be subject to confounding through population stratification and/or to bias from situations where the independence assumption does not hold. Moreover, the main effects of the genetic or environmental factors can-

not be estimated within the case-only design. An overview of retrospective study designs for assessing disease risks of genetic and environmental factors can be found in Weinberg and Umbach [100], comparisons of sample size requirements for the different retrospective approaches can be found in Gauderman [38].

4.3 Cross-sectional designs

From a historical perspective, there are two important cross-sectional study types that were used for assessing genetic risk factors: adoption studies and twin studies. In *adoption studies*, the variances in phenotypes in different familial constellations of families with at least one member not biologically related are compared to each other in order to separate genetic from environmental influential factors. These constellations comprise e.g. biological parents to biological children vs. adopted children, adopted vs. biological children, biological children that were raised together vs. biological children that were raised separately from each other. Similarly, in *twin studies*, the variance in phenotypes in monozygotic twins is compared to that of dizygotic twins. Both, adoption and twin studies can be subject to considerable sampling bias, especially if the study sample is constituted of volunteers. While adoption studies belong chiefly to the past not just because of methodological, but also because of legal and ethical problems; twin studies are still of interest. Since monozygotic twins have identical genomes, they are ideally suited to study environmental effects. Several countries have established twin registers, facilitating population-based twin studies and thus reducing sampling problems. For more information and recent developments, see [11].

A cross-sectional approach for estimating the age-specific penetrance of gene mutations is the *kin-cohort study* ([98]). Given the mode of inheritance, the penetrance is estimated using the genotype of a group of index persons and the phenotypes of the kin of these index persons. Although the design underlies strong assumptions and is therefore subject to several sources of bias (see [37]), the approach can theoretically be incorporated in other study designs, if information about the phenotypes of kin is included.

5 Statistical methods in genetic studies

5.1 Statistical methods in family-based studies

Family-based studies are carried out with the aim to determine the influence of a candidate gene on the disease status. The main idea is that if the potential gene is linked with the marker locus, relatives with similar respectively different disease status are more respectively less similar at the marker locus than expected under independence. We restrict ourselves to families with two generations.

5.1.1 Lod score method

Linkage may be investigated as already discussed in Section 2 by measuring the distance between marker and disease gene loci. Such a distance measure is given by the lod score roughly defined as the logarithm of a likelihood ratio. The likelihood in the nominator is calculated from the genotypic data of some observed pedigree. The likelihood in the denominator is derived assuming no linkage. More formally, following Thompson [93], let $\mathbf{Y}$ denote phenotypic data at trait and marker loci, where $Y_{i,\cdot}$ contains all data on individual i, and $Y_{\cdot,j}$ all data related to the genotype at locus j. Then, regarding the genotypes of individuals at multiple loci, denoted by $\mathbf{G}$, as latent variables leads to

$$(16)\quad Pr(\mathbf{Y}) = \sum_{\mathbf{G}} Pr(\mathbf{Y}|\mathbf{G})Pr(\mathbf{G}) = \sum_{\mathbf{G}} \left(\prod_{\text{observed } i} Pr(Y_{i,\cdot}|G_{i,\cdot}) \right) Pr(\mathbf{G}),$$

where

$$(17)\quad Pr(\mathbf{G}) = \prod_{\text{founders } i} Pr(G_{i,\cdot}) \prod_{\text{nonfounders } i} Pr(G_{i,\cdot}|G_{M_{i,\cdot}}, G_{F_{i,\cdot}}).$$

Formula (16) is based on several assumptions such that the individuals' phenotypes $Y_{i,\cdot}$ are conditionally independent given their genotypes $G_{i,\cdot}$. So-called founders are individuals without parental information in a certain pedigree. These observations are regarded as random draws from the population. The probabilities of the founder genotypes $Pr(G_{i,\cdot})$ are assumed to be determined by the allele frequencies and allelic associations in the population. For the nonfounders, the transmission probabilities $Pr(G_{i,\cdot}|G_{M_{i,\cdot}}, G_{F_{i,\cdot}})$, considered as conditionally independent given the genotypic information of their parents, can be calculated according to the Mendelian laws and the linkage relationships. If this assumption is satisfied, this reduces to investigating the recombination fractions θ which are then the only unknown parameters in the transmission probabilities. Thus, (16) gives the likelihood $L(\theta)$. Since θ is equal to 0.5 for unlinked loci the lod score is given as

$$(18)\quad Z(\theta) = \log_{10}(L(\theta)/L(0.5)).$$

Maximizing equation (18) with respect to θ, i.e. replacing θ with its maximum likelihood estimate, leads to a likelihood ratio test statistic suitable for investigating the null hypothesis of no linkage.

Formally, one has to test for the hypothesis of no linkage, i.e.

$$(19)\quad H_0 : \theta = 0.5 \quad \text{vs.} \quad H_1 : \theta < 0.5.$$

Under H_0, the lod score asymptotically follows a mixture of a $\chi^2_{\mathbf{k}}$-distribution and of 0.5 density at zero, corrected with a factor of $\frac{1}{2\ln(10)} = \frac{1}{4.6}$ (cf. [87]). It has become common sense to use 3 as bound to declare linkage and -2 to exclude linkage, derived by Morton [64] exploiting the theory of sequential probability ratio tests. This implicitly means that H_0 is rejected if

$$(20)\quad Z > \frac{1}{4.6}\chi^2_{2;0.001} = \frac{13.82}{4.6} \approx 3.$$

It should, however, be noted that the above statistical test is based on the assumption that the genetic model of segregation is completely specified by the recombination parameter. This holds for recessive and dominant traits but hardly for complex diseases which then require more complex algorithms for calculating the likelihood and more detailed information on the underlying trait model (see [93] for an overview).

Since the specification of the unknown trait model is accompanied by a high degree of uncertainty and because of the complexity of the likelihood computation as well as the enormous mass of marker data, alternative approaches have been proposed. These approaches, presented in the following subsection, do no longer assume a specific genetic trait model and have in this sense to be thought of as nonparametric. We will focus on sib-pair methods for dichotomous traits. For a review of extensions of the model-free methods presented below, for instance to quantitative traits, to larger pedigrees as well as for the analysis of covariates, multiple marker, and disease loci, we refer to Holmans [45].

5.1.2 Affected sib-pair methods

The general idea underlying the subsequent statistical tests for linkage is that, regardless of the trait model, related affected individuals will share genes identical by descent (IBD) not only at trait loci with higher probability but also at marker loci linked to those trait loci. Thus, using a highly polymorphic marker, having both unaffected parents and the affected sib-pair typed, it can be investigated whether the shared alleles at the marker locus of these two individuals are inherited from the same parent. Please note that only heterozygous parents are informative for sib-pair analysis (cf. Figure 5). If the marker is not linked to the disease gene, the probabilities of a sib-pair sharing 0, 1, or 2 alleles IBD are $\frac{1}{4}, \frac{1}{2}$, or $\frac{1}{4}$, respectively. If the marker is linked to the particular disease locus, the probability of sharing alleles IBD is expected to be increased. However, the probabilities of sharing alleles IBD depend on the mode of inheritance. As shown by Suarez et al. [88], the probability of 2 alleles IBD increases above 0.25 for a dominant trait whereas the probability of 1 allele IBD remains close to 0.5. Similarly, for a recessive trait it can be shown that again as expected the probability of 2 alleles IBD increases whereas the probability of 1 allele IBD drops below 0.5.

Instead of investigating the number of alleles shared IBD, one can look at the numbers of alleles identity by state (IBS). With this concept, it is of interest whether two individuals have a copy of a particular allele in common, regardless of whom it was inherited. Thus, IBS $\geq$ IBD (cf. Figure 5). Although IBS yields less information it might be a way out of problems caused by incomplete polymorphic marker loci and/or untyped parents ([54]), but obviously methods based on IBS will be in general less powerful. Let us finally illustrate IBD and IBS for a sib-pair in Figure 5.

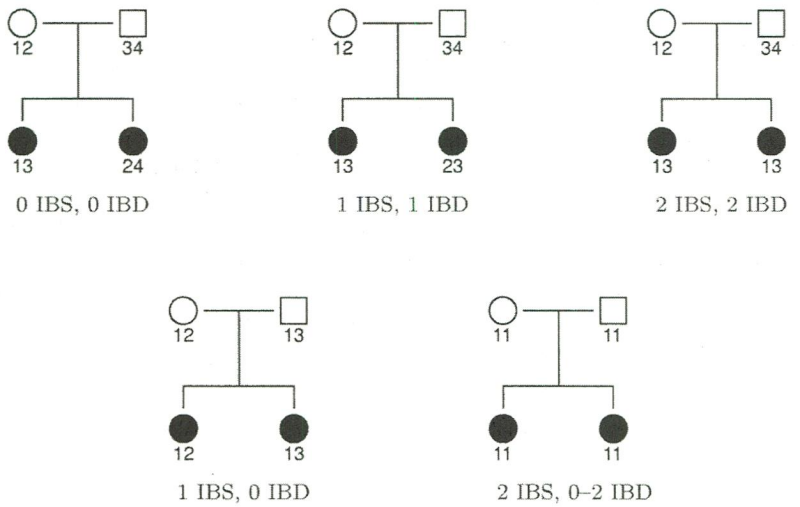


Figure 5: Identity by descent and identity by state

Exploiting the IBD approach leads to various statistical methods for testing for linkage. Some of the classical procedures in this field will be reviewed in the following without the claim of completeness and without dwelling into details.

The maximum lod score test

The MLS (maximum lod score) test by Risch [73] is based on the maximum likelihood method and the lod score. Let us briefly review the basic idea of the MLS statistic. Let α_i denote the prior probability that two siblings share i marker alleles IBD, i.e. $\alpha_i = Pr(\text{IBD} = i)$, and let z_i be the posterior probability that two siblings share i marker alleles IBD given they are both affected, i.e.

$$z_i = Pr(\text{IBD} = i | D_1 = 1, D_2 = 1),\ i = 0, 1, 2,$$

where $D_j, j = 1, 2$, refers to the disease phenotype of the two individuals with 1 being affected and 0 else. The posterior probabilities z_i have to be estimated from the observed marker information. For an affected pair, let finally w_i be the probability of the observed marker phenotypes, denoted as $\mathbf{Y}_{\text{marker}}$, of the sib-pair given that they share i marker alleles IBD, i.e.

$$w_i = Pr(\mathbf{Y}_{\text{marker}} | \text{IBD} = i),\ i = 0, 1, 2.$$

With w_{ik} denoting the corresponding probability for the kth pair, the likelihood of the observed data for the kth affected pair hence results in

$$(21) \quad L_k = Pr(\mathbf{Y}_{\text{marker,k}}) = \sum_{i=0}^{2} z_i w_{ik},$$

which gives the joint likelihood for all N pairs as $L = \prod_{k=1}^{N} L_k$.

With the above probabilities the hypothesis of no linkage means that the prior and the posterior probabilities of sharing i marker alleles IBD should be equal, i.e. $H_0 : \mathbf{z} = \boldsymbol{\alpha}$ with $\mathbf{z} = (z_0, z_1, z_2)$ and $\boldsymbol{\alpha} = (\alpha_0, \alpha_1, \alpha_2)$, where the latter is equal to $(\frac{1}{4}, \frac{1}{2}, \frac{1}{4})$. This is to be tested against $H_1 : \mathbf{z} \neq \boldsymbol{\alpha}$. The appropriate test statistic is based on the likelihood ratio which reads as

$$(22) \quad \Lambda = \prod_{k=1}^{N} \left(\sum_{i=0}^{2} z_i w_{ik} \right) \Big/ \left(\sum_{i=0}^{2} \alpha_i w_{ik} \right).$$

Thus, we obtain the MLS statistic as $T_{\mathrm{MLS}} = \max_{\mathbf{z}} \log_{10} \Lambda$. Taking into account that under H_0 T_{MLS} is asymptotically χ^2_2-distributed and using again a small level of significance of 0.001 to prevent from false positive results, linkage can be declared if $T_{\mathrm{MLS}} > 3$ which coincides with the bound proposed by Morton [64].

The posterior probabilities $\mathbf{z}$ can be estimated by maximum likelihood which yields the maximum of the lod score since its denominator is constant with respect to $\mathbf{z}$. For this purpose, the EM-algorithm by Dempster et al. [19] can be applied. This leads to the following recursive formula for $z_l, l = 0, 1, 2$:

$$(23) \quad z_l' = \frac{1}{N} \sum_{k=1}^{N} \frac{z_l w_{lk}}{\sum_{i=0}^{2} z_i w_{ik}}.$$

In addition

$$(24) \quad \begin{aligned} w_i &= Pr(\mathbf{Y}_{\mathrm{marker}} | \mathrm{IBD} = i) \\ &= \frac{Pr(\mathrm{IBD} = i | \mathbf{Y}_{\mathrm{marker}}) Pr(\mathbf{Y}_{\mathrm{marker}})}{Pr(\mathrm{IBD} = i)} \\ &\propto Pr(\mathrm{IBD} = i | \mathbf{Y}_{\mathrm{marker}}). \end{aligned}$$

As it is sufficient to estimate the w_i ratios instead of estimating the exact values, the test statistic (22) can be derived using the proportional relationship (24). Examples of how to calculate the lod score and possible extensions of this approach can be found in Risch [73].

The above test statistic does not take into account that not all possible vectors $\mathbf{z}$ correspond to a genetic trait model. As shown by Holmans [44] the set of possible sharing probabilities for sib-pairs at each disease-susceptibility locus falls within a triangle bounded by the three lines $z_0 = 0$, $z_1 = 0.5$, and $z_1 = 2z_0$. Regardless of the recombination fraction θ this also holds for the sharing probabilities at the marker locus. Maximizing the lod score under this restriction leads to a more powerful test since it can be shown that the asymptotic distribution of the T_{MLS} under the above triangle condition is a mixture of a χ^2_1 and a χ^2_2 distribution (for details we refer to [44]). As demonstrated by Holmans [44] rejection of the null hypothesis of no linkage if

$$(25) \quad T_{\mathrm{MLS}} > \frac{\chi^2_{1;0.001}}{4.6} \approx 2.3$$

gives a statistical test with a significance level very close to 0.001 in all situations considered by the author. An alternative to the MLS test is given by the minimax test introduced by Whittemore and Tu [103].

The 'goodness-of-fit' test

Blackwelder and Elston [9] consider different types of test statistics that are mainly based on the conditional distribution of the number of marker alleles shared IBD by sib-pairs conditional on the sibs' disease status.

For deriving the χ^2-statistic of goodness-of-fit we compare the observed numbers of sibs with i marker alleles shared IBD and j affected members, $i,j = 0,1,2$, with those expected under the null hypothesis of no linkage. Regardless of the disease status of the sibs, the expected proportions of sib-pairs sharing 0,1, or 2 alleles IBD are, as already mentioned above, $\frac{1}{4}$, $\frac{1}{2}$, and $\frac{1}{4}$, respectively. Let r_{ji} denote the observed numbers of sib-pairs with i marker alleles shared IBD and j affected members and $n_j = \sum_{i=0}^{2} r_{ji}$ the number of sib-pairs with j affected members. The corresponding expected numbers, in the following denoted as $e_{ji} = n_j p_{ji}$, can be calculated using the expected proportions as $e_{j2} = e_{j0} = n_j/4$ and $e_{j1} = n_j/2$. Since one would not usually sample unaffected sib-pairs we will restrict ourselves to the case $n_0 = 0$. The corresponding null hypothesis of no linkage then reads as

$$(26)\quad H_0 : p_{22} = p_{12} = p_{20} = p_{10} = \frac{1}{4};\ p_{21} = p_{11} = \frac{1}{2},$$

and the according χ^2-statistic can be calculated as

$$(27)\quad \chi^2 = \sum_{j=1}^{2}\sum_{i=0}^{2} \frac{(r_{ji} - e_{ji})^2}{e_{ji}}$$

which is approximately χ^2-distributed with four degrees of freedom under the null hypothesis.

The genotypic test

The genotypic test considers those sib-pairs only, where both have status affected. Thus, the null hypothesis can be formulated as $H_0 : p_{22} = p_{20} = \frac{1}{4};\ p_{21} = \frac{1}{2}$ and the corresponding test statistic (27) reduces to

$$(28)\quad \chi^2_{genotypic} = \frac{(r_{22} - n_2/4)^2}{n_2/4} + \frac{(r_{21} - n_2/2)^2}{n_2/2} + \frac{(r_{20} - n_2/4)^2}{n_2/4},$$

which is approximately χ^2-distributed with two degrees of freedom.

The two-allele test

The two-allele test, originally proposed by Day and Simons [18] and Suarez et al. [88], only accounts for those affected sib-pairs who share two alleles IBD which gives the hypothesis of no linkage as $H_0 : p_{22} = p_{12} = \frac{1}{4}$. Depending on whether we further restrict ourselves to those sib-pairs where both are affected ($n_1 = 0$) or not ($n_1 \neq 0$) this test statistic results in

$$(29)\quad \chi^2_{allele,1} = \frac{(r_{22} - n_2/4)^2}{n_2/4} + \frac{(r_{12} - n_1/4)^2}{n_1/4} \quad \text{for} \quad n_1 \neq 0$$

$$(30)\quad \chi^2_{allele,2} = \frac{(r_{22} - n_2/4)^2}{n_2/4} \quad \text{for} \quad n_1 = 0.$$

Under H_0, both statistics are asymptotically χ^2_1-distributed.

Instead of (29) and (30) we may directly use approximately normally distributed test statistics for the one-sided testing problems given by the null hypothesis from above and the following alternatives of linkage $H_1 : p_{22} > p_{12}$ if $n_1 \neq 0$ and $H_1 : p_{22} > \frac{1}{4}$ for $n_1 = 0$, respectively. The test statistics can be calculated as

$$(31) \quad Z_1 = \frac{\hat{p}_{22} - \hat{p}_{12}}{s_1} \quad \text{with} \quad s_1^2 = 3\left(\frac{1}{n_1} + \frac{1}{n_2}\right)/16 \quad \text{for} \quad n_1 \neq 0$$

$$(32) \quad Z_2 = \frac{\hat{p}_{22} - \frac{1}{4}}{s_2} \quad \text{with} \quad s_2^2 = \frac{3}{16n_2} \quad \text{for} \quad n_1 = 0,$$

where the variances of the nominator, s_1^2 and s_2^2, respectively, are derived under the null hypothesis of no linkage.

The mean test

The last statistical test we want to discuss is known as mean test. It compares the observed mean number of alleles shared IBD given the disease status with their expected mean number under the null hypothesis where the latter equals 1. Thus, the null hypothesis of no linkage can be formulated as $H_0 : p_{21} + 2p_{22} = p_{11} + 2p_{12} = 1$. For calculating the observed mean number given j individuals of one sib-pair are affected ($\overline{\text{IBD}}_j$) we have to estimate the unknown conditional probabilities $p_{ji} = Pr(\text{IBD} = i | D_1 + D_2 = j)$ which yields $\hat{p}_{ji} = r_{ji}/n_j$ and thus $\overline{\text{IBD}}_j = 0 \cdot \hat{p}_{j0} + 1 \cdot \hat{p}_{j1} + 2 \cdot \hat{p}_{j2}$.

For $n_1 \neq 0$ the alternative hypothesis of linkage given as $H_1 : p_{21} + 2p_{22} > p_{11} + 2p_{12}$ can be tested using the following test statistic

$$(33) \quad Z_3 = \frac{\hat{p}_{21} + 2\hat{p}_{22} - \hat{p}_{11} - 2\hat{p}_{12}}{s_3} \quad \text{with} \quad s_3^2 = \left(\frac{1}{n_1} + \frac{1}{n_2}\right)/2.$$

Restricting again our analysis to sib-pairs with two affected members reduces the alternative to $H_1 : p_{21} + 2p_{22} > 1$ and

$$(34) \quad Z_4 = \frac{\hat{p}_{21} + 2\hat{p}_{22} - 1}{s_4} \quad \text{with} \quad s_4^2 = \frac{1}{2n_2}.$$

Under H_0, (33) and (34) are approximately standard normally distributed, where the variances are again derived under the assumption of no linkage.

Various of the above test statistics have been compared by Blackwelder and Elston [9] analytically, rather than by simulations, assuming a specific genetic trait model. Their results show that the mean test and the goodness-of-fit test come up with significance levels very close to the nominal ones, whereas the two-allele test tends to be liberal or conservative depending on the underlying sampling scheme. Furthermore, the mean test is demonstrated to be more powerful than the other two tests. Regarding the sampling scheme, the authors recommend choosing sib-pairs with both members affected which is shown to be the most powerful strategy. Thus, there are mainly two reasons to base linkage analysis on affected pairs only, namely the potential misclassification of sibs regarding their disease status and the greater power obtained from affected sib-pairs. But as the authors stress upon, these results heavily depend on the assumption concerning the IBD distribution in the absence of linkage.

Further discussions of statistical tests for linkage in sib-pairs are found in [79], [49] and [35], where among others additional genetic restrictions are assumed to derive more powerful tests.

5.2 Statistical methods in population-based studies

In contrast to linkage, which concerns the physical distance of alleles at two loci, namely the marker and a hypothetical trait locus, it might be looked for association between the trait and a specific marker allele in population-based studies as for instance in case-control studies for dichotomous traits. This statistical concept investigates whether a particular allele occurs more often in the affected members of a population than expected assuming independence between trait and marker locus. Let us for instance consider insulin-dependent diabetes mellitus (IDDM), denoted as D_1, and HLA-D3, denoted as M_1. If the independence assumption holds the probability for their joint occurrence is given by the product of the marginal probabilities, i.e.

$$Pr(M_1, D_1) = Pr(M_1)Pr(D_1).$$

However, this antigen clearly occurs more frequently in affected persons, indicating a positive association. In general, association occurs if

$$(35) \quad Pr(M_i, D_j) = Pr(M_i)Pr(D_j) + \delta_{ij},\ \delta_{ij} \neq 0,$$

where M denotes a marker gene with alleles $M_i, i = 1, \ldots, m$, and D a disease gene with alleles $D_j, j = 1, \ldots, n$, respectively. In case that $\delta_{ij} = Pr(M_i, D_j) - Pr(M_i)Pr(D_j) \neq 0$ we speak of linkage disequilibrium. A positive value of δ_{ij} indicates that the marker and the disease allele jointly occur more often than expected under independence. Please note that this term is not to be mixed up with the physical concept of linkage introduced in the previous section. Linkage is not a necessary condition for linkage disequilibrium (to be discussed in more detail below) although for tight linkage, linkage disequilibrium retains over many generations because of the following close relation between these two concepts of association and linkage. Let us denote with δ_r linkage disequilibrium in the rth generation and with δ_0 in the initial population then it holds

$$\delta_r = (1 - \theta)^r \delta_0.$$

This relation implies that for tight linkage, i.e. for $\theta \approx 0$ the initial linkage disequilibrium δ_0 remains nearly constant whereas in case of no linkage, i.e. $\theta = 0.5$, linkage disequilibrium measured by δ_r decreases rather fast being e.g. only 30% of the initial value in the 5th generation.

Of course, it has to be taken into account that the disease gene is in general not known and can therefore not directly be observed. The only information we have is the individuals' phenotype that probably indicates a potential disease gene. Thus, at the end it has to be searched for disease genes that are associated with the disease phenotype.

Although such population-based studies are advantageous compared to family-based studies in some respect, especially for diseases with late age of onset, where obtaining parental genotypes can be difficult, they also suffer from certain drawbacks.

Typically a large number of cases is required since often only a small gene effect can be observed. Moreover, the p-value of the corresponding statistical test for association has to be small to claim a positive effect due to the problem of multiple testing. Another point is, that the observed association between marker locus and disease phenotype can have several reasons. It remains unclear, whether there is indeed linkage between the marker locus and the disease gene which helps to locate the position of the disease gene. Another possibility might be that the marker allele itself is presumably the disease allele and a candidate gene might be identified. Obviously, this is the ideal case since such an association cannot be eliminated by recombination. The most crucial explanation for an observed association is confounding for instance due to ethnicity resulting from population stratification.

The latter problem will be illustrated by the following fictitious example from the literature. Let us consider two dichotomous loci with alleles A, a and B, b, respectively, and two populations where within each population Hardy-Weinberg equilibrium holds. Let us further assume that in population 1 the probability for the occurrence of A equals 1, whereas $Pr(B) = Pr(b) = 0.5$. Thus, the probabilities of the joint occurrence of these two alleles are as follows: $Pr(A, B) = Pr(A, b) = 0.5, Pr(a, B) = Pr(a, b) = 0$. In population 2, it is, in contrast, assumed that $Pr(A) = Pr(a) = 0.5$ and $Pr(B) = 1$ which gives $Pr(A, B) = Pr(a, B) = 0.5, Pr(A, b) = Pr(a, b) = 0$. Supposed that in a population-based study we observe an 50:50 mixture of both populations then $Pr(A, B) = 0.5, Pr(A, b) = Pr(a, B) = 0.25, Pr(a, b) = 0$, and $Pr(A) = Pr(B) = 0.75$ which yields

$$\delta_{A,B} = Pr(A, B) - Pr(A)Pr(B) = 0.5 - 0.75^2 = -0.0625.$$

Hence, we observe linkage disequilibrium in the mixed population although association of the two genes is absent in each single population. Population structure might thus lead to a false positive result, although as for instance Whittaker and Morris [102] point out that the question concerning the possible contribution of this phenomenon to the type I error rate of well designed case-control studies is still open where in their opinion 'the impact of population structure is likely to be small in most cases'.

Currently, two major approaches are discussed in the literature to cope with the problem of population stratification. Pritchard et al. [71], [70] propose a two-step procedure. In the first step a model-based cluster analysis is performed on several unlinked marker to identify a potential population structure and to assign each individual to an unstructured subpopulation. In the second step, a statistical test for independence of a candidate gene and the disease phenotype is carried out, where the null hypothesis is extended to independence within subpopulations. The other approach to adjust for population structure, is known as genomic control ([21]). Based on a set of unlinked markers across the genome the effect of population structure is estimated. This estimate is then used to adjust the critical value for the test statistic at the candidate locus.

To circumvent the above problems, especially to assure that the controls are drawn from the same genetic population as the cases, it might be advisable to perform family-based association studies instead of case-control association studies. The related statistical methods are based on comparing the frequencies that a particular allele has been inherited or not, respectively. We will focus here on two main methods, namely the hap-

lotype relative risk (HRR) and the transmission/disequilibrium test (TDT) as well as on modifications of the latter.

The haplotype relative risk

The haplotype relative risk (HRR) is introduced by Rubinstein et al. [75] and again picked up by Falk and Rubinstein [33]. The HRR uses virtual controls constructed of the parental alleles or haplotypes not present in the affected child. Hence, the HRR has on the one hand the advantage that cases and controls belong to the same genetic population and that the quality of controls is assured among others due to simultaneous typing. On the other hand it is more restrictive with respect to the necessary data as complete parental genotype information is needed. This information may be, as already mentioned, difficult or even impossible to obtain for diseases with late age of onset. In addition, three persons have to be typed for one case-control pair.

Provided that our sample consists of N complete family trios with heterozygous parents and one affected child the HRR can be derived as follows. Let us denote with M_1 the marker allele of interest and with M_2 any other allele than M_1 at this locus. Then, HRR is defined as

$$(36) \quad \text{HRR} = \frac{p_1}{1-p_1} \Big/ \frac{p_2}{1-p_2}$$

with p_1 being the probability that the child carries M_1 which means that at least one parent has transmitted the particular allele to the child and p_2 the probability that at least one of the non-transmitted parental alleles is M_1. The HRR can be estimated by

$$(37) \quad \widehat{\text{HRR}} = \frac{N_{11}N_{22}}{N_{12}N_{21}},$$

where N_{11} to N_{22} denote the corresponding observed frequencies summarized in a 2×2-table as depicted in Table 1.

	M_1	M_2	Σ
Affected child	N_{11}	N_{12}	$2N$
Fictitious control	N_{21}	N_{22}	$2N$

Table 1: Observed frequencies of transmitted and non-transmitted parental marker alleles to affected child and fictitious control for N completely typed family trios; M_1 marker allele of interest.

Investigating the statistical properties of the HRR the recombination fraction as well as the linkage disequilibrium have to be taken into account. Following Ott [66] and Knapp et al. [51] we consider two loci, the disease locus with D_1 denoting the allele of interest and D_2 the normal allele and the marker locus with alleles M_1 and M_2. Let us denote with p the population frequency of D_1 and with q the population frequency of M_1. Then, in case of allelic independence between M_1 and D_1 the probability of the haplotype M_1D_1 is $Pr(D_1, M_1) = pq$. Accounting for potential linkage disequilibrium we get

$$\begin{aligned}
Pr(D_1, M_1) &= pq + \delta, \\
Pr(D_2, M_1) &= (1-p)q - \delta, \\
Pr(D_1, M_2) &= p(1-q) - \delta, \\
Pr(D_2, M_2) &= (1-p)(1-q) + \delta.
\end{aligned}$$

Ott [66] derives the joint distribution of transmitted and non-transmitted marker alleles for one parent as summarized in Table 2, where also the recombination fraction θ is accounted for. The joint distribution of transmitted and non-transmitted marker genotypes in families with one affected child can be found in Knapp et al. [51] where it is assumed that no selection and/or mutation occurs in the production of gametes, that the Hardy-Weinberg equilibrium holds for the marker-disease genotypes, and that there is random mating regarding marker and disease genotypes.

	non–transmitted		
transmitted	M_1	M_2	Σ
M_1	$q^2 + \frac{\delta q}{p}$	$q(1-q) + \frac{\delta(1-\theta-q)}{p}$	$q + \frac{\delta(1-\theta)}{p}$
M_2	$q(1-q) + \frac{\delta(\theta-q)}{p}$	$(1-q)^2 - \frac{\delta(1-q)}{p}$	$1 - q - \frac{\delta(1-\theta)}{p}$
Σ	$q + \frac{\delta\theta}{p}$	$1 - q - \frac{\delta\theta}{p}$	1

Table 2: Joint distribution of transmitted and non-transmitted marker alleles for one parent.

Thus, statistical independence of transmitted and non-transmitted alleles only hold if $\theta\delta = 0$, i.e. if $\theta = 0$ and/or $\delta = 0$. Based on the above joint distribution Ott [66] demonstrates that on the one hand HRR $= 1$ if $\delta = 0$ or $\theta = 0.5$. On the other hand HRR $\neq 1$ only if $\delta \neq 0$ and $\theta \neq 0.5$. Thus, equivalence of the relative risk (RR) and the HRR for a recessive trait only holds if $\theta = 0$ ([66]) which seems to be implicitly assumed by Falk and Rubinstein [33] in their proof of equivalence of the HRR with the RR. Knapp et al. [51] extend this result to arbitrary modes of inheritance. They show that

(38) $|\text{HRR} - 1| \leq |\text{RR} - 1|$ if $\theta > 0$.

The above inequality (38) means that the strength of association measured by HRR never exceeds the one measured by RR. That is in case of no association (RR=1) the HRR also equals 1. In addition, for $\theta = 0.5$ the HRR always equals 1, irrespective of δ. Looking at the estimates of RR and HRR, Knapp et al. [51] point out that there is even in case of a positive recombination fraction no tendency for the HRR to favor a false-positive association compared to the classical RR.

The transmission/disequilibrium test

Spielman et al. [86] propose the so-called transmission/disequilibrium test (TDT) as test for linkage between a marker locus and a potential disease locus which is not affected by the problems due to population stratification. It is based on the idea, as already exploited by Ott [66] and discussed above, to look whether the marker allele of interest is transmitted or not to the affected child. But in contrast to the HRR, the paired information is used, which results in a situation as demonstrated in Table 3.

	non–transmitted		
transmitted	M_1	M_2	Σ
M_1	N_{11}	N_{12}	$N_{1\cdot}$
M_2	N_{21}	N_{22}	$N_{2\cdot}$
			$2N$

Table 3: Observed frequencies of transmitted and non-transmitted parental marker allele pairs to affected child for N completely typed family trios; M_1 marker allele of interest.

If the parent is homozygous at the marker locus the transmitted and non-transmitted allele coincide and will lead to an entry to the cells denoted as N_{11} or N_{22}, respectively, depending on the parental allele. If the parent is heterozygous at the marker locus we get an entry to the N_{12} cell if M_1 is transmitted and M_2 not and vice versa to the N_{21} cell if M_2 is transmitted and M_1 not. The joint distribution of transmitted and non-transmitted marker alleles for one parent can be found in Table 2 from which it becomes obvious that the TDT might be used as test for linkage with a genetic marker when population association has been found as introduced by Spielman et al. [86] but also as test for association if linkage is present as pointed out by Elston [30]. To be more precise if there is no difference in the probabilities of transmitting the M_1 or the M_2 allele to the affected child the corresponding probabilities, denoted as p_{12} and p_{21}, should be identical with

$$p_{12} = q(1-q) + \frac{\delta(1-\theta-q)}{p}$$

$$p_{21} = q(1-q) + \frac{\delta(\theta-q)}{p}.$$

Closer inspection of p_{12} and p_{21} gives that they coincide if $\delta(1-\theta-q)/p$ and $\delta(\theta-q)/p$ equal zero. This is fulfilled in absence of association ($\delta = 0$) and/or of linkage ($\theta = 0.5$). Thus, we get $H_0 : p_{12} = p_{21}$ is equivalent to $H_0 : \theta = 0.5$ and/or $H_0 : \delta = 0$. Under H_0 and with τ denoting the probability that a heterozygous parent transmits M_1, it holds that $N_{12} \sim \text{Bin}(N_{12} + N_{21}; \tau = 0.5)$ which leads to an exact binomial test or to the approximate McNemar-test with test statistic

$$(39) \quad T_{TDT} = \frac{(N_{12} - N_{21})^2}{N_{12} + N_{21}}.$$

Under H_0 (39) is approximately χ_1^2-distributed. The resulting test is known as TDT. Alternatively, it may also be derived as a score test by exploiting the binomial distribution of N_{12} which gives the following log-likelihood $\ln L(\tau)$

$$\ln L(\tau) = N_{12}\ln(\tau) + N_{21}\ln(1-\tau)$$

and thus the score-test as

$$T_{TDT} = \frac{Z^2}{V}$$

with

$$Z = \frac{\partial \log L(\tau)}{\partial \tau} \quad \text{and} \quad V = \frac{\partial^2 \log L(\tau)}{\partial \tau^2}$$

which finally leads to the above TDT statistic when being evaluated at $\tau = 0.5$ ([102]).

Let us briefly discuss its statistical properties as test for linkage in the presence of association ([32]). The TDT can be applied without being restricted to certain modes of inheritance whenever the data are collected from families with one or more affected offspring or from pedigrees, where monozygotic twin pairs have to be excluded. It keeps the level for the type I error even in case of population stratification. It has, however, to be noted that it has no power if there is no association between the genes at disease and

marker locus since for $\delta = 0$ the probabilities p_{12} and p_{21} always coincide, irrespectively of θ. In addition, the more association there is, i.e. the greater δ, the higher the power of the TDT. Furthermore, it should be mentioned that the TDT is also appropriate for testing whether the HRR differs from one.

The TDT has been introduced for binary marker data and two parents with one affected child typed. For extensions to families with more than one affected child we refer to the original paper by Spielman et al. [86], and for dealing with quantitative traits see Allison [5]. Further extensions can be found in the literature, some of them to be reviewed below, where Whittaker and Morris [102] point out that their approach for deriving the TDT as score-test is advantageous with respect to possible generalizations as for instance to multi-allelic marker loci.

Regarding the latter let us mention Sham and Curtis [81] who provide a log-linear test for a multi-allelic marker locus with alleles $M_1, \ldots, M_m$ which is based on the comparison of the marginal frequencies of the transmitted and non-transmitted alleles, i.e. $N_{1\cdot}, \ldots, N_{m\cdot}$ and $N_{\cdot 1}, \ldots, N_{\cdot m}$, and Schaid [77] who proposes the so-called generalized TDT (GDTD) and maxTDT. The former considers the differences between $N_{i\cdot}$ and $N_{\cdot i}$ which gives an $m-1$ dimensional vector of differences to be weighted with the inverse of the estimated covariance matrix to yield the GDTD. For the maxTDT the usual TDT is calculated for each allele i lumping all other alleles as 'non-i'. The maximum of the resulting m TDT statistics is chosen as the maxTDT test statistic.

In addition, situations may arise where the parental genotypes are missing such that the TDT can no longer be calculated. One way out would be the sib-TDT proposed by Spielman and Ewens [85]. Under the null hypothesis of no linkage and no association the probability of transmitting a parental marker allele to a child is independent of the disease status of the child which means that the distribution of genotypes in the offspring should be independent of the disease status within families. This gives the justification to derive a statistical test in case that the parents are not typed by comparing the numbers of M_1 alleles in affected children and in their unaffected sibs. Taking now the number of M_1 alleles in affected sibs, i.e. conditional on the sibs being affected, leads to a test statistic which can be shown to be, appropriately standardized, asymptotically normally distributed. Analogously to the TDT, this test is only valid to test for association if the minimal family configuration is used which means here one affected and one unaffected sib.

Another possibility to tackle the problem of missing genotype data of the parents is to try to infer it from the genotype data of the children. Treating these inferred genotypes as if they were really observed leads to possible bias as for instance to an inflation of the type I error rate as shown by Curtis [17]. To cope with this problem Knapp [50] derives a statistical test, known as reconstruction combined-TDT (RC-TDT), where the parental genotype information is reconstructed and then it is corrected for the bias pointed out by Curtis [17]. For this purpose, a multi-allelic marker and affected as well as unaffected sibs are needed. In addition, Knapp [50] gives a detailed list of those specific data situations where such a reconstruction is feasible.

If only information is available on a biallelic marker for one affected child and one parent Sun et al. [89] propose the so-called 1-TDT which is based on constructing a non-iterative odds ratio estimator derived from internal case-control studies.

Of course, the list of possible modifications of the TDT to cope with various situations such as missing data, quantitative traits, multi-allelic marker loci, or pedigrees cannot be exhaustively treated within this paper. More information on a variety of possible approaches can be found in Whittaker and Morris [102].

Finally, let us again come back to the fact that the TDT can also be regarded as a test for association. For this purpose, however, only families with one affected child are to be used since deriving the distribution of the TDT statistic under the null hypothesis it has to be assumed that the transmissions from the parent to the affected child are independent which does no longer hold if a heterozygous parent transmits alleles to two affected children in the presence of linkage of disease and marker gene. It has additionally to be noted that in this situation the TDT is the same testing procedure as before but it can only test simultaneously for linkage and association and not solely for one or the other. That is, as Ewens and Spielman [32] stress upon, the distribution of the TDT statistic in the absence of linkage regardless of association is the same as in the absence of association regardless of linkage.

6 Outlook

Although the techniques presented above have been, at least partly, especially developed to cope with the particular demands of the study designs and the data structures in genetic epidemiology as for instance the affected sib-pair method, the haplotype relative risk, or the transmission/disequilibrium test, more sophisticated methods are still required to handle the further increasing complexity of the data. Due to new technologies and developments in genetics, genomewide scans have become feasible to map complex-disease genes which makes it is necessary to derive statistical approaches of high-resolution haplotype or multiple-marker linkage disequilibrium of complex diseases. In a recent paper, Fan and Knapp [34] investigate high-resolution linkage disequilibrium mapping methods based on haplotype maps or microsatellite maps. Statistical tests are derived to test for association between a disease locus and two haplotype blocks or two markers, where these tests are further discussed regarding their appropriateness and their statistical properties.

Besides these new developments that are only briefly addressed in this paper, other important fields could not be covered in full detail. One of the most obvious statistical problems in this field is due to possibly false positive associations resulting from the enormous number of statistical tests which are applied to check for potential candidate genes (cf. Section 3). As a typical example consider, say, five presumed disease genes and 20.000 candidate genes. If now for each single test of association statistical significance is concluded at a p-value smaller than 0.05 then the probability to detect false positive associations is about 99.5%. Decreasing the individual level to values of about 10^{-8} seems to be appropriate to adjust for multiplicity. Of course, other multiple adjustment techniques might be thought of being less restrictive as adaptive designs ([52]), but nevertheless the required adjustment leads to rather small significance levels for the individual tests. An alternative way out is for instance provided by the method of sample-splitting as suggested in [67].

Another research topic relates to pedigrees consisting of more than two generations which leads to even more complicated data structures. Just to mention two examples let us come back to the lod-score, a likelihood-based method which can be exploited to test for linkage, and the TDT. Elston and Stewart [31] propose an algorithm for calculating the likelihood recursively for pedigrees. Regarding the TDT, Martin et al. [62] derive an extension to account for pedigrees where, roughly spoken, the information of allelic transmissions on a pedigree level is summed up.

In our presentation of the statistical methods to test for linkage or association, the focus was on qualitative traits. Quantitative traits may be treated by formulating adequate regression models where it has to be checked whether linear models are sufficient (cf. [5], [72], [6], [36] and [3]) which would be the case if normality of the response variable can be assumed. If, however, other exponential family distributions seem to be more adequate as underlying distributions, generalized linear models as introduced by McCullagh and Nelder [63] may be adapted. Higher flexibility in modelling such traits can be achieved by exploiting neural networks as introduced in genetic epidemiology by Lucek et al. [61]. Regression models feature in any case the possibility to be rather easily extended to account for a polygenic component, shared family environment, multi-allelic markers, further covariates such as environmental factors, and interaction terms. Especially the latter are of importance when investigating complex diseases since they allow for modelling gene-gene and gene-environment interactions. First approaches to account for such interactions can be found in [78], [105], [97], [83], [104], and especially based on graphical models in [82]. In our view, particularly the problem of adequately accounting for gene-gene and gene-environment interactions is of high practical relevance and deserves further research.

Acknowledgement

Financial support by 'Senator für Bildung und Wissenschaft der Freien Hansestadt Bremen' is gratefully acknowledged.

References

[1] The Chipping Forecast. *Nature Genetics*, 21 suppl, 1999.
[2] *GenePix Pro 4.1 Microarray acquisition and analysis software for the GenePix 4000 and 4100 series scanners – Users guide and tutorial*. Axon Instruments Inc., 2002.
[3] G. R. Abecasis, L. R. Cardon, and W. O. C. Cookson. A general test of association for quantitative traits in nuclear families. *American Journal of Human Genetics*, 66(1):279–292, 2000.
[4] R. Adams and L. Bischof. Seeded region growing. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(6):641–647, 1994.
[5] D. B. Allison. Transmission-disequilibrium tests for quantitative traits. *American Journal of Human Genetics*, 60(3):676–690, 1997.
[6] D. B. Allison, M. Heo, N. L. Kaplan, and E. R. Martin. Sibling-based tests of linkage and association for quantitative traits. *American Journal of Human Genetics*, 64(6):1754–1764, 1999.
[7] M. Åstrand. Contrast normalization of oligonucleotide arrays. *Journal of Computational Biology*, 10(1):95–102, 2003.

[8] Y. Benjamini and Y. Hochberg. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series B*, 57(1):289–300, 1995.

[9] W. C. Blackwelder and R. C. Elston. A comparison of sib-pair linkage tests for disease susceptibility loci. *Genetic Epidemiology*, 2(1):85–97, 1985.

[10] B. Bolstad, R. A. Irizarry, M. Åstrand, and T. P. Speed. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics*, 19(2):185–193, 2003.

[11] D. Boomsma, A. Busjahn, and L. Peltonen. Classical twin studies and beyond. *Nature Reviews Genetics*, 3(11):872–882, 2002.

[12] M. P. S. Brown, W. N. Grundy, D. Lin, N. Cristianini, C. W. Sugnet, T. S. Furey, M. J. Ares, and D. Haussler. Knowledge-based analysis of microarray gene expression data using support vector machines. *Proceedings of the National Academy of Sciences of the United States of America*, 97(1):262–267, 2000.

[13] P. O. Brown and D. Botstein. Exploring the new world of the genome with DNA microarrays. *Nature Genetics*, 21 suppl:33–37, 1999.

[14] L. R. Cardon and L. J. Palmer. Population stratification and spurious allelic association. *The Lancet*, 361(9357):598–604, 2003.

[15] Y. Chen, E. R. Dougherty, and M. L. Bittner. Ratio-based decisions and the quantitative analysis of cDNA microarray images. *Journal of Biomedical Optics*, 2(4):364–374, 1997.

[16] W. S. Cleveland. Robust locally weighted regression and smoothing scatterplots. *Journal of the American Statistical Association*, 74(368):829–836, 1979.

[17] D. Curtis. Use of siblings as controls in case-control association studies. *Annals of Human Genetics*, 61(4):319–333, 1997.

[18] N. E. Day and M. J. Simons. Disease susceptibility genes – their identification by multiple case family studies. *Tissue Antigens*, 8(2):109–119, 1976.

[19] A. P. Dempster, N. M. Laird, and D. B. Rubin. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Series B*, 39:1–38, 1977.

[20] J. DeRisi, L. Penland, P. O. Brown, M. L. Bittner, P. S. Meltzer, M. Ray, Y. Chen, Y. A. Su, and J. M. Trent. Use of a cDNA microarray to analyse gene expression patterns in human cancer. *Nature Genetics*, 14(4):457–460, 1996.

[21] B. Devlin and K. Roeder. Genomic control for association studies. *Biometrics*, 55(4):997–1004, 1999.

[22] K. P. Donnelly. The probability that related individuals share some section of genome identical by descent. *Theoretical Population Biology*, 23(1):34–63, 1983.

[23] S. Dudoit, Y. H. Yang, M. C. Callow, and T. P. Speed. Statistical methods for identifying differentially expressed genes in replicated cDNA microarray experiments. *Statistica Sinica*, 12(1):111–139, 2002.

[24] B. Efron. Robbins, empirical bayes, and microarrays. Technical Report 2001-30B/219, Dept. of Statistics Stanford University, 2001.

[25] B. Efron, J. D. Storey, and R. J. Tibshirani. Microarrays, empirical Bayes methods, and false discovery rates. Technical Report 2001-23B/217, Dept. of Statistics Stanford University, 2001.

[26] B. Efron and R. J. Tibshirani. Empirical Bayes methods and false discovery rates for microarrays. *Genetic Epidemiology*, 23(1):70–86, 2002.

[27] B. Efron, R. J. Tibshirani, J. D. Storey, and V. Tusher. Empirical Bayes analysis of a microarray experiment. *Journal of the American Statistical Association*, 96(456):1151–1160, 2001.

[28] P. H. C. Eilers, J. M. Boer, G.-J. van Ommen, and H. C. van Houwelingen. Classification of microarray data with penalized logistic regression. In M. L. Bittner, Y. Chen, A. N. Dorsel, and E. R. Dougherty, editors, *Microarrays: Optical technologies and informatics*, 4266, pages 187–198. SPIE, 2001.

[29] M. B. Eisen, P. T. Spellman, P. O. Brown, and D. Botstein. Cluster analysis and display of genome-wide expression patterns. *Proceedings of the National Academy of Sciences of the United States of America*, 95(25):14863–14868, 1998.

[30] R. C. Elston. Linkage and association. *Genetic Epidemiology*, 15(6):565–576, 1998.
[31] R. C. Elston and J. Stewart. A general model for the genetic analysis of pedigree data. *Human Heredity*, 21(6):523–542, 1971.
[32] W. J. Ewens and R. S. Spielman. The transmission/disequilibrium test: History, subdivision, and admixture. *American Journal of Human Genetics*, 57(2):455–464, 1995.
[33] C. T. Falk and P. Rubinstein. Haplotype relative risks: An easy reliable way to construct a proper control sample for risk calculations. *Annals of Human Genetics*, 51(3):227–233, 1987.
[34] R. Fan and M. Knapp. Genome association studies of complex diseases by case-control designs. *American Journal of Human Genetics*, 72(4):850–868, 2003.
[35] J. J. Faraway. Improved sib-pair linkage test for disease susceptibility loci. *Genetic Epidemiology*, 10(4):225–233, 1993.
[36] D. W. Fulker, S. S. Cherny, P. C. Sham, and J. K. Hewitt. Combined linkage and association sib-pair analysis for quantitative traits. *American Journal of Human Genetics*, 64(1):259–267, 1999.
[37] M. H. Gail, D. Pee, and R. Carroll. Kin-cohort designs for gene characterization. *Journal of the National Cancer Institute Monographs*, 1999(26):55–60, 1999.
[38] W. J. Gauderman. Sample size requirements for matched case-control studies of gene-environment interaction. *Statistics in Medicine*, 21(1):35–50, 2002.
[39] D. Ghosh. Singular value decomposition regression models for classication of tumors from microarray experiments. In R. B. Altman, A. K. Dunker, L. Hunter, K. Lauderdale, and T. E. Klein, editors, *Pacific Symposium on Biocomputing 2002*, volume 7, pages 18–29. World Scientific, 2002.
[40] J. B. S. Haldane. The combination of linkage values, and the calculation of distances between the loci of linked factors. *Journal of Genetics*, 8:299–309, 1919.
[41] A. J. Hartemink, D. K. Gifford, T. S. Jaakkola, and R. A. Young. Maximum likelihood estimation of optimal scaling factors for expression array normalization. In M. L. Bittner, Y. Chen, A. N. Dorsel, and E. R. Dougherty, editors, *Microarrays: Optical technologies and informatics*, 4266, pages 132–140. SPIE, 2001.
[42] J. Herrero, A. Valencia, and J. Dopazo. A hierarchical unsupervised growing neural network for clustering gene expression patterns. *Bioinformatics*, 17(2):126–136, 2001.
[43] S. Holm. A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 6:65–70, 1979.
[44] P. Holmans. Asymptotic properties of affected-sib-pair linkage analysis. *American Journal of Human Genetics*, 52(2):362–374, 1993.
[45] P. Holmans. Nonparametric linkage. In D. J. Balding, M. Bishop, and C. Cannings, editors, *Handbook of Statistical Genetics*, pages 487–505. Wiley, 2001.
[46] W. Huber, A. von Heydebreck, H. Sültmann, A. Poustka, and M. Vingron. Variance stabilization applied to microarray data calibration and to the quantification of differential expression. *Bioinformatics*, 18 suppl(1):96–104, 2002.
[47] W. Huber, A. von Heydebreck, and M. Vingron. Analysis of microarray gene expression data. In D. J. Balding, M. Bishop, and C. Cannings, editors, *Handbook of Statistical Genetics*, preprint. Wiley, 2nd edition, 2003.
[48] M. K. Kerr, M. Martin, and G. A. Churchill. Analysis of variance for gene expression microarray data. *Journal of Computational Biology*, 7(6):819–837, 2000.
[49] M. Knapp. A powerful test of sib-pair linkage for disease susceptibility. *Genetic Epidemiology*, 8(2):141–143, 1991.
[50] M. Knapp. The transmission/disequilibrium test and parental-genotype reconstruction: The reconstruction-combined transmission/disequilibrium test. *American Journal of Human Genetics*, 64(3):861–870, 1999.
[51] M. Knapp, S. A. Seuchter, and M. P. Baur. The haplotype-relative-risk (HRR) method for analysis of association in nuclear families. *American Journal of Human Genetics*, 52(6):1085–1093, 1993.

[52] I. R. König, H. Schäfer, A. Ziegler, and H.-H. Müller. Reducing sample sizes in genome scans: Group sequential study designs with futility stops. *Genetic Epidemiology, accepted*, 2003.
[53] D. D. Kosambi. The estimation of map distances from recombination values. *Annals of Eugenics*, 12:172–175, 1944.
[54] K. Lange. The affected sib-pair method using identity by state relations. *American Journal of Human Genetics*, 39(1):148–150, 1986.
[55] J. M. Last, editor. *A Dictionary of Epidemiology*. Oxford University Press, 4th edition, 2001.
[56] C. Li and W. H. Wong. Model-based analysis of oligonucleotide arrays: Expression index computation and outlier detection. *Proceedings of the National Academy of Sciences of the United States of America*, 98(1):31–36, 2001.
[57] C. Li and W. H. Wong. Model-based analysis of oligonucleotide arrays: Model validation, design issues and standard error application. *Genome Biology*, 2(8):research 0032.1–0032.11, 2001.
[58] W. Li and Y. Yang. How many genes are needed for a discriminant microarray data analysis? In S. M. Lin and K. F. Johnson, editors, *Methods of microarray data analysis*, pages 137–150. Kluwer Academic Publishers, 2001.
[59] R. J. Lipshutz, S. P. A. Fodor, T. R. Gingeras, and D. J. Lockhart. High density synthetic oligonucleotide arrays. *Nature Genetics*, 21 suppl:20–24, 1999.
[60] I. Lönnstedt and T. P. Speed. Replicated microarray data. *Statistica Sinica*, 12(1):31–46, 2002.
[61] P. R. Lucek, J. Hanke, J. Reich, S. A. Solla, and J. Ott. Multi-locus nonparametric linkage analysis of complex trait loci with neural networks. *Human Heredity*, 48(5):275–284, 1998.
[62] E. R. Martin, M. P. Bass, and N. L. Kaplan. Correcting for a potential bias in the pedigree disequilibrium test. *American Journal of Human Genetics*, 68(4):1065–1067, 2001.
[63] P. McCullagh and J. A. Nelder. *Generalized Linear Models*. Chapman and Hall, 2nd edition, 1989.
[64] N. E. Morton. Sequential tests for the detection of linkage. *American Journal of Human Genetics*, 7:277–318, 1955.
[65] M. A. Newton, C. M. Kendziorski, C. S. Richmond, F. R. Blattner, and K. W. Tsui. On differential variability of expression ratios: Improving statistical inference about gene expression changes from microarray data. *Journal of Computational Biology*, 8(1):37–52, 2001.
[66] J. Ott. Statistical properties of the haplotype relative risk. *Genetic Epidemiology*, 6(1):127–130, 1989.
[67] M. J. Owen, P. Holmans, and P. McGuffin. Association studies in psychiatric genetics. *Molecular Psychiatry*, 2(4):270–273, 1997.
[68] W. W. Piegorsch, C. R. Weinberg, and J. A. Taylor. Non-hierarchical logistic models and case-only designs for assessing susceptibility in population-based case-control studies. *Statistics in Medicine*, 13(2):153–162, 1994.
[69] F. Pociot and M. F. McDermott. Genetics of type 1 diabetes mellitus. *Genes and Immunity*, 3(5):235–249, 2002.
[70] J. K. Pritchard, M. Stephens, and P. Donnelly. Inference of population structure using multilocus genotype data. *Genetics*, 155(2):945–959, 2000.
[71] J. K. Pritchard, M. Stephens, N. A. Rosenberg, and P. Donnelly. Association mapping in structured populations. *American Journal of Human Genetics*, 67(1):170–181, 2000.
[72] D. Rabinowitz. A transmission disequilibrium test for quantitative trait loci. *Human Heredity*, 47(6):342–350, 1997.
[73] N. Risch. Linkage strategies for genetically complex traits. III. The effect of marker polymorphism on analysis of affected relative pairs. *American Journal of Human Genetics*, 46(2):242–253, 1990.
[74] D. M. Rocke and B. Durbin. A model for measurement error for gene expression arrays. *Journal of Computational Biology*, 8(6):557–569, 2001.
[75] P. Rubinstein, M. Walker, C. Carpenter, C. Carrier, J. Krassner, C. T. Falk, and F. Ginsberg. Genetics of HLA disease associations: The use of the haplotype relative risk (HRR) and the

'haplo-delta' (Dh) estimates in juvenile diabetes from three racial groups. *Human Immunology*, 3(4):384, 1981.

[76] E. E. Schadt, C. Li, B. Ellis, and W. H. Wong. Feature extraction and normalization algorithms for high-density oligonucleotide gene expression array data. *Journal of Cellular Biochemistry*, 84 suppl(37):120–125, 2001.

[77] D. J. Schaid. General score tests for associations of genetic markers with disease using cases and their parents. *Genetic Epidemiology*, 13(5):423–449, 1996.

[78] D. J. Schaid. Case-parents design for gene-environment interaction. *Genetic Epidemiology*, 16(3):261–273, 1999.

[79] D. J. Schaid and T. G. Nick. Sib-pair linkage tests for disease susceptibility loci: Common tests vs. the asymptotically most powerful test. *Genetic Epidemiology*, 7(5):359–370, 1990.

[80] M. Schena, D. Shalon, R. W. Davis, and P. O. Brown. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science*, 270(5235):467–470, 1995.

[81] P. C. Sham and D. Curtis. An extended transmission/disequilibrium test (TDT) for multi-allele marker loci. *Annals of Human Genetics*, 59(3):323–336, 1995.

[82] N. A. Sheehan, B. Gulbrandtsen, M. S. Lund, and D. A. Sorensen. Bayesian MCMC mapping of quantitative trait loci in a half-sib design: A graphical model perspective. *International Statistical Review*, 70(2):241–267, 2002.

[83] S. L. Slager, D. J. Schaid, L. Wang, and S. N. Thibodeau. Candidate-gene association studies with pedigree data: Controlling for environmental covariates. *Genetic Epidemiology*, 24(4):273–283, 2003.

[84] E. Southern, K. Mir, and M. Shchepinov. Molecular interactions on microarrays. *Nature Genetics*, 21 suppl:5–9, 1999.

[85] R. S. Spielman and W. J. Ewens. A sibship test for linkage in the presence of association: The sib transmission/disequilibrium test. *American Journal of Human Genetics*, 62(2):450–458, 1998.

[86] R. S. Spielman, R. E. McGinnis, and W. J. Ewens. Transmission test for linkage disequilibrium: The insulin gene region and insulin-dependent diabetes mellitus (IDDM). *American Journal of Human Genetics*, 52(3):506–516, 1993.

[87] K. Strauch. *Kopplungsanalyse bei genetisch komplexen Erkrankungen mit genomischem Imprinting und Zwei-Genort-Krankheitsmodellen*. Urban & Vogel, 2002.

[88] B. K. Suarez, J. Rice, and T. Reich. The generalized sib pair IBD distributions: Its use in the detection of linkage. *Annals of Human Genetics*, 42:87–94, 1978.

[89] F. Sun, W. D. Flanders, Q. Yang, and M. J. Khoury. A new method for estimating the risk ratio in studies using case-parental control design. *American Journal of Epidemiology*, 148(9):902–909, 1998.

[90] P. Tamayo, D. K. Slonim, J. P. Mesirov, Q. Zhu, S. Kitareewan, E. Dmitrovsky, E. S. Lander, and T. R. Golub. Interpreting patterns of gene expression with self-organizing maps: Methods and application to hematopoietic differentiation. *Proceedings of the National Academy of Sciences of the United States of America*, 96(6):2907–2912, 1999.

[91] D. C. Thomas. Genetic epidemiology with a capital "E". *Genetic Epidemiology*, 19(4):289–300, 2000.

[92] D. C. Thomas and J. S. Witte. Point: Population stratification: A problem for case-control studies of candidate-gene associations? *Cancer Epidemiology Biomarkers and Prevention*, 11(6):505–512, 2002.

[93] E. A. Thompson. Linkage analysis. In D. J. Balding, M. Bishop, and C. Cannings, editors, *Handbook of Statistical Genetics*, pages 541–563. Wiley, 2001.

[94] R. J. Tibshirani, T. Hastie, B. Narasimhan, and G. Chu. Diagnosis of multiple cancer types by shrunken centroids of gene expression. *Proceedings of the National Academy of Sciences of the United States of America*, 99(10):6567–6572, 2002.

[95] G. C. Tseng, M.-K. Oh, L. Rohlin, J. C. Liao, and W. H. Wong. Issues in cDNA microarray analysis: Quality filtering, channel normalization, models of variations and assessment of gene effects. *Nucleic Acids Research*, 29(12):2549–2557, 2001.

[96] V. G. Tusher, R. J. Tibshirani, and G. Chu. Significance analysis of microarrays applied to the ionizing radiation response. *Proceedings of the National Academy of Sciences of the United States of America*, 98(9):5116–5121, 2001.

[97] D. M. Umbach and C. R. Weinberg. The use of case-parent triads to study joint effects of genotype and exposure. *American Journal of Human Genetics*, 66(1):251–261, 2000.

[98] S. Wacholder, P. Hartge, J. P. Struewing, D. Pee, M. McAdams, L. Brody, and M. Tucker. The kin-cohort study for estimating penetrance. *American Journal of Epidemiology*, 148(7):623–630, 1998.

[99] S. Wacholder, N. Rothman, and N. Caporaso. Counterpoint: Bias from population stratification is not a major threat to the validity of conclusions from epidemiological studies of common polymorphisms and cancer. *Cancer Epidemiology Biomarkers and Prevention*, 11(6):513–520, 2002.

[100] C. R. Weinberg and D. M. Umbach. Choosing a retrospective design to assess joint genetic and environmental contributions to risk. *American Journal of Epidemiology*, 152(3):197–203, 2000.

[101] P. H. Westfall and S. S. Young, editors. *Resampling-based multiple testing: Examples and methods for p-value adjustment*. Wiley, 1993.

[102] J. C. Whittaker and A. P. Morris. Family-based tests of association and/or linkage. *Annals of Human Genetics*, 65(5):407–419, 2001.

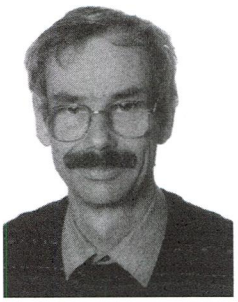
[103] A. S. Whittemore and I.-P. Tu. Simple, robust linkage tests for affected sibs. *American Journal of Human Genetics*, 62(5):1228–1242, 1998.

[104] Q. Yang, M. J. Khoury, J. M. Friedman, and W. D. Flanders. On the use of population attributable fraction to determine sample size for case-control studies of gene-environment interaction. *Epidemiology*, 14(2):161–167, 2003.

[105] Q. Yang, M. J. Khoury, F. Sun, and W. D. Flanders. Case-only design to measure gene-gene interaction. *Epidemiology*, 10(2):167–170, 1999.

[106] Y. H. Yang, S. Dudoit, P. Luu, D. M. Lin, V. Peng, J. Ngai, and T. P. Speed. Normalization for cDNA microarray data: A robust composite method addressing single and multiple slide systematic variation. *Nucleic Acids Research*, 30(4):e15, 2002.

[107] K. Y. Yeung, C. Fraley, A. Murua, A. E. Raftery, and W. L. Ruzzo. Model-based clustering and data transformations for gene expression data. *Bioinformatics*, 17(10):977–987, 2001.



Pluriharmonic maps, twisted loops and twistors

J.-H. Eschenburg

Abstract

- Keywords and Phrases: Associated families, twistors, loop groups
- Mathematics subject classification: 53C35, 53C43, 53C55; 22E67

A minimal surface in euclidean space has two very special properties: (A) It allows a twisted circle of isometric deformations preserving the tangent plane (Associated family), and (B) it is just the real part of a holomorphic map (Weierstrass representation). In fact, these two properties hold more generally for a pluriharmonic map f of a simply connected complex manifold into euclidean space. If instead the target space is a Riemannian symmetric space P, Property (A) essentially remains true, however by lack of global parallel displacements a parallel isomorphism between the tangent spaces of the associated family is needed. Consequently Property (B) gets more complicated: f arises by projecting a "superhorizontal" holomorphic map $\hat{f}$ into a certain infinite dimensional flag manifold (adjoint orbit) fibering over P. the "universal twistor space". The map $\hat{f}$ takes values in a finite dimensional sub-twistor space iff the associated family is trivial.

Eingegangen: 24.11.2003

Institut für Mathematik, Universität Augsburg, D-86135 Augsburg,
eschenburg@math.uni-augsburg.de

DMV
JAHRESBERICHT
DER DMV

Introduction

Among the most beautiful objects in Geometry are the minimal surfaces in 3-space. One of their spectacular properties is the existence of a (so called *associated*) family of deformations preserving the interior distances and the surface normal while rotating the principal directions. The best known example is the deformation of the catenoid into the helicoid, cf. `http://www.ag.jku.at/verbieg_en.html`. It starts and ends with the catenoid which however is turned inside-out during the deformation. This is an example of a *twisted loop* of surfaces: It comes back to its original shape, but only after applying a *point reflection* on the ambient space. The same phenomenon occurs when euclidean 3-space is replaced with a *symmetric space* P, a Riemannian manifold with isometric point reflections at every point. Moreover, the minimal surface can be replaced by a harmonic map of a surface into P. In fact, a surface is only the easiest example of a complex manifold, and we may equally well consider a *pluriharmonic* map of an arbitrary complex manifold. Also these maps can be deformed by twisted loops, described as mappings into some space of loops. These are holomorphic mappings of a certain kind, and therefore pluriharmonic maps can be obtained from holomorphic data; the Weierstrass representation for minimal surfaces is the best known example.

Sometimes the situation is rigid and the twisted loop of deformations arises only by isometries of the ambient space; such pluriharmonic maps are called *isotropic*. Of course they can exist only if the point reflections in P can be deformed to the identity; symmetric spaces with this property are called *inner*. This case is much simpler and has been studied for a long time, starting with the work of Calabi [4]. It is closely connected with the concept of a *twistor* (cf. [3]) which was first investigated by R. Penrose [13] in connection to Relativity. Roughly speaking, a twistor is a complex structure on a tangent space of an inner symmetric space P, and the set of all twistors forms a fibre bundle over P. The original example studied by Penrose was the "classical" twistor fibration $\mathbb{C}P^3 \to S^4$; any point of complex projective 3-space can be viewed as a complex structure on some tangent space of S^4. Each isotropic pluriharmonic map into a symmetric space can be lifted to some twistor space over P. The most classical twistor fibration was also the most successfull: With its help R. Bryant [1] gave an explicit description of all minimal spheres in S^4.

In the present survey article we want to explain, following work of K. Uhlenbeck, J. Rawnsley and others, how the general (non-isotropic) case can also be understood in terms of a generalized twistor theory. This more general twistor space is infinite dimensional, a space of twisted loops, and since it contains all other twistor spaces, we would like to call it "universal". This is mainly a re-interpretation of well known facts which however could give a unified view point to the theory. Most of the details missing in this survey can be found in [7]. We thank J. Dorfmeister for many useful hints and discussion.

0 Harmonic maps

A smooth manifold M is called *Riemannian* if there is an inner product on each of its tangent spaces, depending smoothly on the base point. For any smooth curve $c : [a, b] \to M$ the *length* $\mathcal{L}(c) = \int_a^b |c'(t)|\, dt$ is defined, giving M the structure of a metric space which locally is approximated by euclidean space. All euclidean notions are applicable but have different properties. The main difference arises for the *parallel displacement* of tangent vectors and the corresponding differentiation of tangent vector fields, the so called *Levi-Civita derivative*: The parallel displacement becomes path dependent, and the Levi-Civita derivatives with respect to two coordinate directions do not commute; in fact their commutator is the basic invariant of Riemannian geometry, the *curvature tensor*.

Classical euclidean geometry is investigated by using substructures: lines, planes etc. In Riemannian geometry, the rôle of lines is taken by *geodesic lines* which have parallel tangent vectors and which locally minimize the length and also the *energy* $\mathcal{E}(c) = \int_a^b |c'(t)|^2\, dt$ among all curves connecting two given point. If metric completeness is assumed, any two points are joined by a geodesic line, like in euclidean geometry. But what are the substitutes for planes and higher dimensional subspaces? A plane in space contains the line passing through any two of its points. In Riemannian geometry, a submanifold with this property is called *totally geodesic*. However such submanifolds are very rare unless we restrict attention to spaces of constant curvature which are locally just spherical, euclidean or hyperbolic spaces.

Therefore we consider another generalization of geodesics to higher dimensions, using the energy minimizing property. If M and P are Riemannian manifolds and $f : M \to P$ a smooth map, the derivative of f at a point $x \in M$ is a linear map $df_x : T_xM \to T_{f(x)}P$ between the corresponding tangent spaces of M and P. The vector space $\mathrm{Hom}(T_xM, T_{f(x)}P)$ inherits an inner product and hence a norm from the inner products on T_xM and $T_{f(x)}P$ given by the Riemannian metrics; in fact

$$(1) \quad |df_x|^2 = |df_x.e_1|^2 + \ldots + |df_x.e_m|^2$$

for an orthonormal basis $(e_1, \ldots, e_m)$ of T_xM. For any compact subset $M_o \subset M$ the *energy* or *Dirichlet integral* of $f|M_o$ is

$$(2) \quad \mathcal{E}(f|M_o) = \int_{M_o} |df_x|^2\, dv_x$$

where dv_x denotes the volume element of M at x determined by the Riemannian metric. A map $f : M \to P$ is called *harmonic* if the variation of its energy vanishes,

$$(3) \quad \delta\mathcal{E}(f|M_o) := \frac{d}{ds}\mathcal{E}(f_s|M_o)|_{s=0} = 0$$

for any compact subset $M_o \subset M$, where $f_s : M \to P$ is any smooth variation of f with $f = f_s$ outside M_o. As always, this variational principle is equivalent to its Euler differential equation:

$$(4) \quad \Delta f := \mathrm{trace}\, Ddf = 0$$

involving the Levi-Civita derivative D on $\mathrm{Hom}(TM, TP)$ which is induced by the Riemannian metrics on M and P.

Harmonic maps exist for all dimensions of M. A harmonic map of the real line $M = \mathbb{R}$ is just a geodesic. The next case when M is a surface is most interesting since then the energy (2) is invariant under conformal changes of the metric g on M. In fact, if g is replaced with g/μ^2 for some smooth positive scaling function μ on M, then $|df_x|^2$ takes up a factor $\mu(x)^2$ while the 2-dimensional volume element dv_x is divided by $\mu(x)^2$, hence the energy remains unchanged. But an oriented surface with a conformal class of metrics is nothing else than a 1-dimensional complex manifold where the complex structure J on the tangent space is the rotation by the angle $\pi/2$. Hence harmonic maps $f : M \to P$ are already defined when P is Riemannian but M is only a complex 1-dimensional manifold (a *Riemann surface*) without specified metric. We will see that locally all harmonic maps of Riemann surfaces into symmetric spaces can be obtained in terms of meromorphic functions on M as has been shown in [8]. These maps became interesting to physicists under the name σ-models (cf. [6]).

If $\dim M > 2$, harmonic maps in general do not have such nice properties. However there is an interesting special case where the methods of complex analysis still apply. Let M be a *complex* manifold of any dimension. A map $f : M \to P$ is called *pluriharmonic* if $f|C$ is harmonic for any complex 1-dimensional submanifold (*complex curve*) $C \subset M$. If we compare harmonic maps of surfaces to geodesics, then pluriharmonic maps play the rôle of totally geodesic submanifolds, and they do not always exist as we shall see. However there are many interesting examples. Under certain conditions, a harmonic map $f : M \to P$ of a Kähler manifold M is automatically pluriharmonic, in particular this holds if P has nonpositive curvature operator (cf. [15], [11]).

1 Associated families and symmetric spaces

Another peculiarity for $\dim M = 2$ is the existence of so called *associated families* of harmonic maps. Consider the case $P = \mathbb{R}^n$. Then (4) becomes $f_{22} = -f_{11}$ (where the indices mean partial derivatives). This is the integrability condition for the differential form $df \circ J = f_2 dx_1 - f_1 dx_2$; in other words, $df \circ J$ is a closed iff f is harmonic. Hence locally $df \circ J = d\tilde{f}$ for another (so called *conjugate*) harmonic function $\tilde{f}$, and taking linear combinations, we obtain a circle of such maps $f_\theta = f\cos\theta + \tilde{f}\sin\theta$ with

$$(5) \quad df_\theta = df \circ R_\theta$$

where R_θ denotes the rotation by the angle θ on the tangent space of M. This is called the *associated family* of f. This result can be extended in two ways. We may replace the surface by an arbitrary complex manifold M and put $R_\theta = I\cos\theta + J\sin\theta$ where I is the identity and J the complex structure on TM. Moreover we replace $\mathbb{R}^n$ by any *symmetric space* P.

A Riemannian manifold P is called *symmetric* if for every $p \in P$ there is an isometry $s_p : P \to P$ fixing p with derivative $(ds_p)_p = -I$; this is called *point reflection* or *symmetry at* p. Thus s_p reverses any geodesic line passing through p. As a consequence, the group of isometries $I(P)$ acts transitively on P since any two points $o, q \in P$ can be

joined by a geodesic segment $c : [0, 1] \to P$, and we can map o to q by the point reflection s_p where $p = c(\frac{1}{2})$ is the midpoint of c. Locally, symmetric spaces are characterized by the fact that the curvature tensor

$$R^P(X, Y)Z = [D_X, D_Y]Z - D_{[X,Y]}Z$$

(acting on tangent vector fields X, Y, Z on P) is *parallel* on P, i.e. it commutes with the parallel displacements on P. On any tangent space T_pP it defines a trilinear map (a so called *Lie triple product*) R^P which completely encodes the local structure of the symmetric space P. In particular, any isometric linear map $\phi : T_pP \to T_qP$ preserving R^P (i.e. $R^P(\phi X, \phi Y)\phi Z = \phi R^P(X, Y)Z$) "extends" to an isometry $g \in I(P)$ with $g(p) = q$ and $dg_p = \phi$. For $p = q$ any such ϕ will be called an *automorphism* of T_pP; it extends to an isometry of P fixing the point p.

Equation (5) as it stands can hold only for $P = \mathbb{R}^n$ where all tangent spaces are identified by global parallel displacement. But otherwise df_x and $d(f_\theta)_x$ take values in different tangent spaces, $T_{f(x)}P$ and $T_{f_\theta(x)}P$. Thus we replace (5) by

(6) $\quad df_\theta = \Phi_\theta \circ df \circ R_\theta$

where $\Phi_\theta(x)$ is an isomorphism between $T_{f(x)}P$ and $T_{f_\theta(x)}P$ for any $x \in M$ which is as nice as possible:

- $\Phi_\theta(x)$ is a linear isometry preserving the curvature tensor R^P,
- $\Phi_\theta(x)$ is parallel with respect to x.

A family of smooth maps $f_\theta : M \to P$ satisfying (6) will be called an *associated family* of $f = f_0$. A main result of [10] characterizes pluriharmonic maps by associated families:

Theorem 1. *Let M be a simply connected complex manifold and P a symmetric space of nonpositive or nonnegative curvature. Then a smooth map $f : M \to P$ is pluriharmonic if and only if it has an associated family f_θ. This is uniquely determined up to isometries of P.*

From the uniqueness we can derive more properties of the associated family. Since $R_{\theta+\pi} = -R_\theta$, we obtain a solution $(f_{\theta+\pi}, \Phi_{\theta+\pi})$ of (6) for the rotation angle $\theta + \pi$ from a solution (f_θ, Φ_θ) for θ, namely

(7) $\quad f_{\theta+\pi} = f_\theta, \quad \Phi_{\theta+\pi} = -\Phi_\theta.$

In the second equation we may replace $-\Phi_\theta(x)$ by $\Phi_\theta(x)s_{f(x)}$ obtaining

(8) $\quad \Phi_{\theta+\pi}(x) = \Phi_\theta(x)s_{f(x)},$

since $\Phi_\theta(x)$ can be considered as an isometry of P sending $f(x)$ onto $f_\theta(x)$ and $s_{f(x)}$ acts as $-I$ on $T_{f(x)}P$.

A class of particular interest is formed by the so called *isotropic* pluriharmonic maps whose associated family is constant: $f_\theta = f$ for all θ (up to isometries of P). Then (6) becomes

(9) $\quad df = \Phi_\theta \circ df \circ R_\theta,$

and this time each $\Phi_\theta(x)$ is an isometry preserving R^P and mapping $T_{f(x)}P$ onto itself.

We can choose $\theta \mapsto \Phi_\theta(x)$ to be a one-parameter group, a homomorphism of the unit circle: If (f, Φ_θ) and $(f, \Phi_{\theta'})$ are solutions of (9) for θ and θ', then $(f, \Phi_\theta \Phi_{\theta'})$ is a solution for $\theta + \theta'$ (since $R_{\theta+\theta'} = R_\theta R_{\theta'}$) and hence we can assume

(10) $\quad \Phi_{\theta+\theta'} = \Phi_\theta \Phi_{\theta'}, \quad \Phi_\pi = -I$

2 Twistor lifts

Twistors (cf. [13], [3]) have been introduced in order to apply complex analysis to non-complex symmetric spaces. On a complex (so called *hermitian*) symmetric space there is a *complex structure* on any tangent space T_pP which by definition is a Lie triple automorphism j with $j^2 = -I$. In complex coordinates, j is just the multiplication by $i = \sqrt{-1}$, and it is invariant under parallel displacements. But if no such complex structure is given, the idea is to consider the set of *all possible* complex structures on T_pP. Any of these belongs to a one-parameter group of automorphims of T_pP called *twistors at p*. More precisely, if $G = I(P)^o$ denotes the identity component of the isometry group $I(P)$, a *twistor at p* is by definition a smooth homomorphism $\tau : S^1 = \mathbb{R}/(2\pi\mathbb{Z}) \to G$ fixing p and passing through the point reflection $s_p = \tau_\pi$; the corresponding complex structure on T_pP is given by $j = \tau_{\pi/2}$. Of course this is possible only if the point reflections belong to the identity component of the isometry group; symmetric spaces with this property are called *inner*. E.g. the even dimensional spheres are inner, but not so the odd dimensional ones.

A *twistor space* Z over P is the conjugacy class of some twistor τ_o within G. If τ is a twistor at p, then $g\tau g^{-1}$ is a twistor at gp, and since G acts transitively, Z contains twistors at all points of P. Thus Z fibres over P with fibre Z_p being the set of all twistors at p; let $\mathrm{p} : Z \to P$ be the projection. Further Z is a complex manifold: Every $\tau \in Z$ defines a complex structure $\tau_{\pi/2}$ on $T_{\mathrm{p}(\tau)}P$ which extends canonically to a complex structure on $T_\tau Z$. In fact Z can be viewed as an adjoint orbit (an orbit of the adjoint representation of G on its Lie algebra $\mathfrak{g}$) since the one-parameter group is determined by its infinitesimal generator in $\mathfrak{g}$, and it is well known that all adjoint orbits are complex manifolds (quotients of complex Lie groups by closed complex subgroups).

The easiest example is the 4-dimensional sphere $P = S^4 \subset \mathbb{R}^5$ where the twistors $\tau \in Z_p$ are one-parameter groups of orthogonal matrices fixing p and acting by oriented planar rotations on two orthogonal planes in $p^\perp$. Any such τ is conjugate under $SO(5)$ to $(\tau_o)_\theta = \mathrm{diag}\,(1, \rho_\theta, \rho_\theta)$ with $\rho_\theta = \left(\begin{smallmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{smallmatrix}\right)$; more precisely, Z is the conjugacy class of τ_o in $SO(5)$. The stabilizer of τ_o is the subgroup $U(2) \subset SO(4) \subset SO(5)$ and hence $Z \cong SO(5)/U(2)$ is complex projective 3-space $\mathbb{C}P^3$ (recall that $SO(5) = PSp(2) \subset PU(4)$ acts transitively on $\mathbb{C}P^3 = PU(4)/U(3)$ and $PSp(2) \cap U(3) = U(2)$).

Also in the general case twistors are composed by planar rotations which however may have different velocities: The tangent space $\mathfrak{p} = T_pP$ is an orthogonal sum of subspaces $\mathfrak{p}_k$ on which τ acts by planar rotations $\rho_{k\theta}$. But since $\tau_\pi = -I$, the weights k must be odd integers. Following [3], we call a twistor τ *canonical* if the lowest weight space $\mathfrak{p}_1$ generates $\mathfrak{p}$ as a Lie triple algebra, i.e. $\mathfrak{p} = \mathfrak{p}_1 + R^P(\mathfrak{p}_1, \mathfrak{p}_1, \mathfrak{p}_1) + \ldots$ The space $\mathfrak{p} = T_pP$

can be embedded naturally into $T_\tau Z$ as the *horizontal subspace* for the fibration $\mathrm{p} : Z \to P$, and the subspace $\mathfrak{p}_1 \subset \mathfrak{p} \subset T_\tau Z$ will be called *superhorizontal*.

Now let $f : M \to P$ be an isotropic pluriharmonic map which is *full*, i.e. $f(M)$ does not belong to a proper totally geodesic subspace of P. Then by the results of the previous chapter $\Phi(x) = (\theta \mapsto \Phi_\theta(x))$ is a twistor at $f(x)$ for any $x \in M$ and thus defines a map $\Phi : M \to Z$ with $\mathrm{p} \circ \Phi = f$, the so called *twistor lift*. From the parallelity of $x \mapsto \Phi(x)$ we see that $d\Phi$ takes values in the horizontal bundle of the fibration $\mathrm{p} : Z \to P$ ("horizontal" and "parallel" are just the same notions for the principal bundle $G \to P$ and its associated bundles.). More precisely, using (9) we see that $d\Phi$ takes values in the superhorizontal subbundle on which the twistor has weight one and hence agrees with the complex structure j (more precisely, with the rotation group generated by j); this shows that Φ is also holomorphic. Vice versa it is easy to see that the projection of a superhorizontal holomorphic map is an isotropic pluriharmonic map. Thus we obtain:

Theorem 2. *Isotropic pluriharmonic maps $f : M \to P$ are precisely the projections of holomorphic superhorizontal maps into twistor spaces over P.*

3 Loop space lifts

Now let us consider an arbitrary pluriharmonic map $f : M \to P$. Fixing a base point $p \in P$ we have a canonical projection $\pi : G \to P$, $\pi(g) = gp$ where G is the identity component of the isometry group of P. If we also fix a suitable basis B of the vector space T_pP, then $g(B)$ is a basis for $T_{gp}P$ and hence G can be considered as a certain set of bases (*frames*) of the tangent spaces of P. On a contractible open subset $M' \subset M$, the map f can be lifted to G, yielding a smooth map $F : M' \to G$ which projects onto $f|M'$, i.e. $f(x) = F(x)p$. This is called a *local framing* of f since it provides each tangent space $T_{f(x)}P$ with a frame $F(x)$. Obviously, two such framings $F, \tilde{F}$ differ by a map into the isotropy group $K = \{g \in G;\ gp = p\}$, more precisely, $\tilde{F} = FF_K$ for some smooth map $F_K : M' \to K$.

We have already seen that pluriharmonic maps come in associated one-parameter families f_θ satisfying (6). A framing $F : M' \to G$ of f defines also a framing $F_\theta = \Phi_\theta F$ for f_θ. More generally we may put

$$F_\theta = g_\theta \Phi_\theta F \tag{11}$$

for an arbitrary isometry $g_\theta \in G$, replacing f_θ with $g_\theta f_\theta$. We will use this freedom as follows: We fix base points $x_o \in M$ and $p = f(x_o) \in P$. We may assume $F(x_o) = I$, and we choose $g_\theta = \Phi_\theta(x_o)^{-1}$ whence $F_\theta(x_o) = I$ for all θ. Recall that $F(x) \in G$ maps p to $f(x)$ and hence conjugates the point reflections s_p and $s_{f(x)}$. From (8) we obtain $g_{\theta+\pi} = s_p g_\theta$ and hence $F_{\theta+\pi}(x) = g_{\theta+\pi}\Phi_{\theta+\pi}(x)F(x) = s_p g_\theta \Phi_\theta(x) s_{f(x)} F(x) = s_p F_\theta(x) s_p$, thus

$$F_{\theta+\pi}(x) = \sigma(F_\theta(x)) \tag{12}$$

where $\sigma \in \mathrm{Aut}(G)$ denotes the conjugation by s_p. Therefore each map $\theta \mapsto F_\theta(x)$ belongs to the *twisted loop group*

$$\hat{G} = \Lambda_\sigma(G) = \{\gamma : S^1 \to G;\ \gamma_{\theta+\pi} = \sigma(\gamma_\theta)\} \tag{13}$$

where the loops γ are sufficiently regular (e.g. of class H^1). Thus we have obtained a map $\mathcal{F} : M' \to \hat{G}$ with $\mathcal{F}(x)_\theta := F_\theta(x)$. If we had chosen another framing $\tilde{F} = FF_K$, we would have got a corresponding map $\tilde{\mathcal{F}} = \mathcal{F}F_K$. By projecting $\mathcal{F}$ to the coset space $\hat{Z} = \hat{G}/K$ (where $K \subset \hat{G}$ denotes the subgroup of constant loops in the isotropy group $K \subset G$) we get a map $\bar{\mathcal{F}} = \mathcal{F}K$ which is independent of the choice of F and hence is globally defined on M. This space $\hat{Z}$ fibres over P where the projection $\hat{\mathrm{p}} : \hat{Z} \to P$ is the evaluation of the loop at the initial point: $\hat{\mathrm{p}}(\gamma K) = \gamma_0 p$. Now we have constructed a smooth map $\bar{\mathcal{F}} : M \to \hat{Z}$ which is a lift of $f : M \to P$, i.e. $\hat{\mathrm{p}} \circ \bar{\mathcal{F}} = f$.

This loop space $\hat{Z}$ is again a complex manifold, a quotient of two complex loop groups: $\hat{Z} = \hat{G}^c/\hat{G}^+$ where $\hat{G}^c$ is the set of loops into the complexified group[1] G^c and $\hat{G}^+$ is the subgroup of those $\gamma : S^1 \to G^c$ such that both γ, γ^{-1} extend to analytic maps on the unit disk in $\mathbb{C}$. If G^c is a matrix group, we may write each $\gamma \in \hat{G}^c$ as a matrix Fourier series $\gamma_\theta = \sum_{k \in \mathbb{Z}} A_k e^{ik\theta}$, and $\gamma \in \hat{G}^+$ iff $A_k = 0$ for $k < 0$ and the same is true for γ^{-1}. It can be shown that $\hat{\mathcal{F}} : M \to \hat{Z}$ is holomorphic; in fact one constructs a holomorphic lift into $\hat{G}^c$ using the parallelity of Φ_θ. Moreover it follows from (6) that the differential of $\hat{\mathcal{F}}$ takes values in a finite dimensional homogeneous subbundle of the tangent bundle of $\hat{Z}$, which at the base point consists of the simplest possible (finite) Fourier series $Ae^{i\theta} + \bar{A}e^{-i\theta}$; this will be called the *superhorizontal subbundle*. Vice versa, we can characterize pluriharmonic maps by this property. Thus we arrive at a theorem which looks quite similar to the one in the isotropic case:

Theorem 3. *General pluriharmonic maps $f : M \to P$ are precisely the projections of holomorphic superhorizontal maps into the loop space $\hat{Z}$ over P.*

Remarks. 1. In fact, the differential $d\hat{\mathcal{F}}$ can be described in terms of a $\mathfrak{p}^c$-valued holomorphic differential form on M, called *normalized potential*. If M is a surface, this may be an arbitrary meromorphic 1-form, but in higher dimensions an integrability condition is needed (*curved flat condition*). In [8] it was shown how to obtain f back from the potential. This formula allows to compute explicit examples.

2. The loop group $\hat{G}^c$ acting on $\hat{Z}$ preserves the set of holomorphic superhorizontal maps $\hat{\mathcal{F}} : M \to \hat{Z}$. Hence it induces an action on the set of pluriharmonic maps which is a special case of the so called *dressing action*.

4 The "universal twistor"

Can the previous construction also be viewed as a sort of twistor lift? To answer this question we first try to understand the twistor construction as a special case of the loop space lift. This is not difficult since both times we have used Φ_θ. In the isotropic case, $\Phi_\theta(x)$ is a one-parameter group conjugate to a fixed twistor τ for each $x \in M$, and from $F_\theta = (\Phi_\theta(x_o))^{-1}\Phi_\theta F$ we obtain (after a slight modification of our frame F)

(14) $\quad \mathcal{F} = \tau F \tau^{-1}$

Recall that $\hat{Z} = \hat{G}/K$ and $Z = Ad(G)\tau = G/H$ where H is the centralizer of τ. Motivated by (14) we consider the group homomorphism

(15) $\quad \rho_\tau : G \to \hat{G}, \quad g \mapsto \tau g \tau^{-1}$.

The loop $\rho_\tau(g)$ is constant if and only if g commutes with τ, i.e. $g \in H$. Thus ρ_τ induces an equivariant embedding $\bar{\rho}_\tau : Z = G/H \to \hat{Z} = \hat{G}/K$. This is holomorphic and preserves superhorizontality. In fact, $d\rho_\tau$ maps a (complex) eigenvector X of $Ad(\tau)$ corresponding to an eigenvalue $e^{ik\theta}$ into the corresponding Fourier monomial: $(d\rho_\tau(X))_\theta = Ad(\tau_\theta)X = e^{ik\theta}X$.

The link between Z and $\hat{Z}$ becomes even more apparent if we consider $\hat{Z}$ like Z as an adjoint orbit. In fact, denoting the Lie algebra of $\hat{G}$ by $\hat{\mathfrak{g}}$ (consisting of the loops in $\mathfrak{g}$), we get an embedding $\hat{Z} \to \hat{\mathfrak{g}}$, $\gamma K \mapsto \gamma'\gamma^{-1}$ where $\gamma' = \frac{d}{d\theta}\gamma_\theta$. This can be considered as an adjoint orbit if we enlarge the Lie algebra $\hat{\mathfrak{g}}$ by an element δ with $ad(\delta)\xi := \xi'$ for any (sufficiently regular) $\xi \in \hat{\mathfrak{g}}$. Assuming G to be a matrix group we can represent each $\gamma \in \hat{G}$ as a multiplication operator and δ as a differential operator on matrix valued loops. Thus

$$(16) \quad Ad(\gamma)\delta = \gamma\delta\gamma^{-1} = \delta - \gamma'\gamma^{-1},$$

and the mapping $\gamma K \mapsto \delta - \gamma'\gamma^{-1}$ is an embedding of $\hat{Z}$ as the adjoint orbit of δ in the enlarged Lie algebra.[2] Comparing with $Z = Ad(G)\tau$ we conclude that the one-parameter group generated by δ should be a "universal twistor" $\hat{\tau}$. This does not belong to $\hat{G}$ itself but to the automorphism group $\mathrm{Aut}(\hat{\mathfrak{g}})$; it is the shift of the loop parameter: For any $\xi \in \hat{\mathfrak{g}}$ we have

$$(17) \quad \hat{\tau}_\theta(\xi)_{\tilde{\theta}} = \xi_{\theta+\tilde{\theta}}.$$

Lemma. *All twistors τ acting on $\mathfrak{g}$ by the adjoint representation are restrictions of the "universal twistor" $\hat{\tau}$ on $\hat{\mathfrak{g}}$, more precisely,*

$$(18) \quad d\rho_\tau \circ Ad(\tau) = \hat{\tau} \circ d\rho_\tau$$

Proof. We have $\mathfrak{g}^c = \sum_k \mathfrak{g}_k$ with $Ad(\tau_\theta)X = e^{ik\theta}X$ for any $X \in \mathfrak{g}_k$. On the other hand, for any $X \in \mathfrak{g}_k$ we have $\hat{\tau}_\theta(d\rho_\tau(X)) = \hat{\tau}_\theta(\tilde{\theta} \mapsto e^{ik\tilde{\theta}}X) = e^{ik\theta}d\rho_\tau(X)$ which proves the claim. □

We sum up our discussion by the following

Theorem 4. *Any pluriharmonic map $f : M \to P$ is the projection of a holomorphic superhorizontal map $\bar{\mathcal{F}}$ into the universal twistor space $\hat{Z}$. The map f is isotropic iff $\bar{\mathcal{F}}(M)$ is contained in one of the finite dimensional twistor spaces $Z_\tau \subset \hat{Z}$.*

Remark. There is an important difference between the finite dimenisonal twistors τ and the universal one $\hat{\tau}$: The "universal twistor" does not act on P but on loop spaces, and therefore it is not a twistor in the sense of Section 2. But we may pass to the infinite dimensional symmetric space $\hat{P} = \hat{G}/\hat{K}$ which consists of the loops in P, and clearly $\hat{\tau}$ is a twistor on $\hat{P}$. This space also fibres over P via the evaluation at the initial point, and the loop space fibration factorizes over $\hat{P}$ as $\hat{Z} \to \hat{P} \to P$.

Notes

1 We can think of G as being a real matrix group defined by algebraic equations which may be complexified.

2 This is the corresponding twisted affine Kac-Moody Lie algebra without the central extension which is not essential for the adjoint representation, cf. [14].

References

[1] R. Bryant: *Conformal and minimal immersions of compact surfaces into the 4-sphere*, J. Diff. Geom. **17** (1982), 455–473.

[2] F.E. Burstall and M.A. Guest: *Harmonic two-spheres in compact symmetric spaces, Math. Ann.* **309** (1997) 541–57.

[3] F.E. Burstall and J.H. Rawnsley: *Twistor Theory for Riemannian Symmetric Spaces*, Springer L.N. in Math. **1424** (1990).

[4] E. Calabi: *Minimal immersions of surfaces in euclidean spheres*, J. Diff. Geom. **1** (1967), 111–125.

[5] S.S. Chern, J.Wolfson: *Minimal surfaces by moving frames*, Am. J. Math. **105** (1983), 59–83.

[6] A.M. Din, W.J. Zakrzewski: *General classical solutions of the $\mathbb{C}P^{n-1}$ model*, Nucl. Phys. B 174 (1980), 397–406.

[7] J. Dorfmeister, J.-H. Eschenburg: *Pluriharmonic maps, loop groups and twistor theory*, to appear in J. Glob. An. and Geom.

[8] J. Dorfmeister, F. Pedit, and H. Wu: *Weierstrass type representations of harmonic maps into symmetric spaces*, Comm. Anal. Geom. **6** (1998), 633–668.

[9] J. Eells, J.C. Wood: *Harmonic maps from surfaces to complex projective spaces*, Adv. Math. **49** (1983), 217–263.

[10] J.-H. Eschenburg and R. Tribuzy: *Associated families of pluriharmonic maps and isotropy*, manuscripta math. **95** (1998), 295–310.

[11] N, Ferreira, M. Rigoli, R. Tribuzy: Isometric immersions of Kähler manifolds, Rend. Sem. Mat. Univ. Padova **90** (1993).

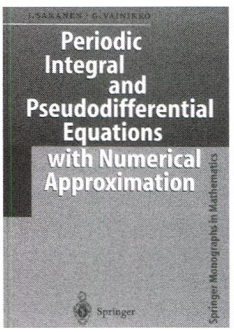
[12] Y. Ohnita: *Harmonic maps into symmetric spaces and integrable system theory*, in: Sophia Kokyuroku in Mathematics 45 (2002), The proceedings of the conference "Theory of Lie Groups and Manifolds" (Sophia University, 2001.3.5-3.7), Dep. of Math., Sophia Univ., 2002, 81-98.

[13] R. Penrose: *Twistor Algebra*, J. Math. Phys. **8** (1967), 345–366.

[14] A. Pressley and G. Segal: *Loop Groups*, Oxford Science Publications, Clarendon Press, 1986.

[15] J.H. Sampson: *Application of harmonic maps to Kähler geometry*, Contemp. Math. **39** (1986), 125–134.

[16] K. Uhlenbeck: *Harmonic maps into Lie groups (classical solutions of the chiral model)*, J. Diff. Geom. **30** (1989), 1–50.



J. Saranen, G. Vainikko
Periodic Integral and Pseudodifferential Equations with Numeric Approximation

Berlin u. a., Springer, 2002, 452 S., EUR 79,95

Die Analysis von Integralgleichungen in allen ihren Variationen und die dazu gehörende numerische Analysis ist ein weitgefächertes und seit jeher aktuelles und stimulierendes Gebiet. Daher ist es verständlich, dass ständig Bücher erscheinen, die sich diesen oder jenen Aspekten der Theorie und ihren Anwendungen annehmen.

Die Autoren des zu besprechenden Buches erklären in der Einleitung, dass dieses als Lehrbuch für graduierte oder postgraduierte Studenten, aber auch als wissenschaftliche Monographie angesehen werden kann. Der Inhalt gruppiert sich um die klassischen Integraloperatoren, wie sie bei der Behandlung von ebenen Randwertproblemen in Gebieten mit glattem Rand mittels der Randintegralgleichungsmethode entstehen (Laplace- und Helmholtzgleichung, biharmonische Gleichung, usw.). Das Buch ist in sich geschlossen konzipiert und enthält ausschließlich bewiesene Resultate. Die ersten fünf Kapitel bieten klassisches Material und dienen als Einführung in die Randintegralgleichungsmethode für die oben genannten Randwertprobleme. Hier werden auch die Hilfsmittel bereitgestellt (einschließlich der Theorie der Fredholmoperatoren), auf die immer wieder zurückgegriffen wird. In den Kapiteln 6 und 7 werden schließlich entsprechend eine Klasse allgemeiner periodischer Integraloperatoren behandelt (deren Kerne auch periodische Distributionen sein können) und periodische Pseudodifferentialoperatoren untersucht, wobei diese beiden Kapitel eng miteinander verzahnt sind und den Rahmen bilden, in dem die eingangs erwähnten Randintegraloperatoren untersucht werden. Diese Analysis wird fast ausschließlich in der Hilbertraumskala der periodischen Sobolevräume $H^\mu(\mu \in \mathbb{R})$ vorgenommen. Ein Grund dafür dürfte sein, dass die numerische Analysis, der fast die gesamte zweite Hälfte des Buches gewidmet ist, in der genannten Skala noch relativ einfach entwickelt werden kann. Eine Ausnahme bildet das 4. Kapitel, in dem der wichtige Cauchysche singuläre Integraloperator in den periodischen Hölderräumen $C^\alpha(0 < \alpha < 1)$ studiert wird. Diese Resultate werden nur herangezogen, um diesen Operator in den Räumen H^μ zu erklären. Dieser Umweg kann vermieden werden, wenn man eine elegante Idee benutzt, wie sie im 1. Kapitel des bekannten Buches von I. Gohberg und N. Krupnik [2] zu finden ist. Überdies unterläuft den Autoren im Beweis des (korrekt formulierten) Lemmas 4.1.5 ein Lapsus: Der Beweis ist bemerkenswert kurz – und nicht schlüssig. Sie benutzen die falsche Behauptung, dass $C^1(\Gamma)$ in $C^\alpha(\Gamma)$ $(0 < \alpha < 1)$ dicht ist. Ansonsten ist dieser Teil des Buches, der der Analysis gewidmet ist, durchaus lesbar. Anfänger oder Einsteiger werden schätzen, dass dieser in sich geschlossen ist und alle Behauptungen bewiesen werden. Kritisch verbleibt anzumerken, dass jegliche Hinweise auf den möglichen Ausbau der gebotenen Theorie fehlen. Beispielsweise wird nicht erwähnt, dass periodische Pseudodifferentialoperatoren auch in der wichtigen Skala der Hölder-Zygmundräume studiert werden können.

Ab dem 8. Kapitel wird auf die numerische Analysis periodischer Integralgleichungen eingegangen. Zunächst stehen polynomiale Verfahren im Vordergrund (Galerkin- und Kollokationsverfahren, volldiskretisierte Versionen). Auch hier wird die gebotene Theorie mit vollständigen Beweisen versehen. Breiten Raum nehmen Fragen

ein, die sich mit schnellen und effizienten Algorithmen befassen. In diesem Punkt leistet das vorliegende Buch für die betrachteten Gleichungen mehr als andere einschlägige Monographien. Das 13. Kapitel ist ausschließlich Splineapproximationsverfahren gewidmet, wobei sich die Autoren hier auf die Untersuchung von Stabilitäts- und Konvergenzuntersuchungen einschränken. Als Splineräume fungieren Räume glattester Splines auf äquidistanten Gittern, und die Untersuchungen stützen sich maßgeblich auf gewisse (wohlbekannte) Rekurrenzbeziehungen für die Fourierkoeffizienten dieser Splinefunktionen.

Diesen der numerischen Analysis gewidmeten Teil sehe ich weitaus kritischer. Es werden hier Themen behandelt, die den Autoren nahe stehen. Dies ist zweifelsohne zu akzeptieren. Nicht zu akzeptieren ist aber, dass der Eindruck vermittelt wird, dass es neben diesen Themen keine weiteren gibt bzw. damit die Thematik im Wesentlichen abgehandelt ist. In Wirklichkeit stellen die Ausführungen der Autoren nur ein kleines Segment in der heute bekannten numerischen Analysis für die betrachteten Operatoren und ihrer Modifikationen dar. Selbst für dieses Segment sind die Literaturhinweise höchst eigenwillig gestaltet. Das Gleiche gilt für die geschichtliche Einordnung. Damit nehmen die Autoren, ob bewusst oder unbewusst, eine Wertung vor. Einige exemplarische Beispiele seien dazu angeführt:

- Im 11. Kapitel werden gewisse Integralgleichungen auf nichtgeschlossenen Kurvenstücken betrachtet und über eine Periodizierungstechnik auf den periodischen Fall zurückgeführt. Diese Methode ist ziemlich restriktiv und spielt für Integralgleichungen auf nichtgeschlossenen Kurven eine Außenseiterrolle. Für solche Gleichungen gibt es andere und sehr anpassungsfähige Methoden, die das Randverhalten der Lösungen (und dies kann ziemlich kompliziert sein) in geeigneter Weise berücksichtigen. Den Autoren ist dies sicher bekannt; warum teilen sie dann dies ihren Lesern nicht mit, und warum fehlt jeder Hinweis auf das Werk von P. Junghanns und seiner Koautoren (siehe MathSciNet)?
  Einen knappen geschichtlichen Abriß (der bis etwa 1990 reicht) kann man zu der im Kapitel 11 behandelten Problematik in den „Notes and comments" zu den Kapiteln 9 und 12 in [5] finden.
- Im Punkt 12.8 wird auf Lokalisierungstechniken eingegangen und als Quelle auf eine 1985 erschienene Arbeit von Arnold und Wendland verwiesen. Zu dieser Zeit waren lokale Prinzipien in der numerischen Analysis bereits wohlbekannt; die Ausarbeitung dieser Technik ist mit anderen Namen verknüpft. Man vergleiche etwa die „Notes and comments" zu Chapter 3 in [1], Chapter 7 in [2] und „Notes and references" zu Chapter 4 in [4]. Ferner mußte ich mit Befremden feststellen, dass die frühen Arbeiten von S. Prößdorf und seinen Mitarbeitern zu Splineapproximationsverfahren in den Literaturangaben zum Kapitel 13 völlig ausgeblendet wurden. So galt die Arbeit S. Prößdorf, G. Schmidt „A finite element collocation method for singular integral equations", Math. Nachr. 100 (1981), 33–60, als Durchbruch.
- Im Kapitel 13 findet sich nirgends ein Hinweis, dass es zu dem dort bevorzugten Zugang interessante und weitreichende Alternativen gibt, wie sie etwa in [5], Kapitel 10 und 13, dargelegt wurden.
- In Anwendungen treten häufig Systeme von Integralgleichungen auf. So können die in der Mechanik wichtigen singulären Integralgleichungen mit Konjugation auf solche Systeme zurückgeführt werden. Dieser Umstand wird mit keinem Wort erwähnt. Hinzuzufügen ist, dass in diesem Fall häufig neue und nichttriviale Überlegungen angestellt werden müssen.

Abschließend möchte ich vermerken, dass dieses Buch durchaus als Einführung in die im Titel genannten Gegenstände dienen kann, aber keineswegs den gegenwärtigen Forschungsstand einigermaßen ädequat wie-

dergibt und auch nicht als Orientierungshilfe dienen kann.

Literatur

[1] A. Böttcher, B. Silbermann: *Invertibility and Asymptotics of Toeplitz Matrices*. Akademie-Verlag, Berlin, 1983.

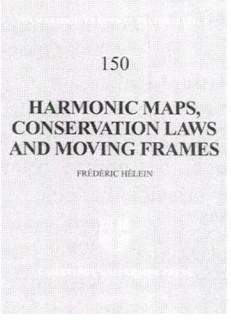
[2] A. Böttcher, B. Silbermann: *Analysis of Toeplitz operators*. Akademie-Verlag, Berlin, 1989 und Springer-Verlag Berlin, 1990.

[3] I. Gohberg, N. Krupnik: *Einführung in die Theorie der eindimensionalen singulären Integraloperatoren*. Birkhäuser Verlag, Basel, 1979.

[4] R. Hagen, S. Roch, B. Silbermann: *C^*-algebras and Numerical Analysis*. Marcel Dekker, Inc., New-York, Basel, 2001.

[5] S. Prößdorf, B. Silbermann: *Numerical Analysis for Integral and related Operator Equations*. Akademie-Verlag, Berlin, 1991 und Birkhäuser-Verlag, Basel, 1991.

Chemnitz B. Silbermann



F. Hélein
Harmonic Maps, Conservation Laws and Moving Frames (Cambridge Tracts in Math. 150)

Cambridge University Press, 2002, 264 S., £ 47,50

James Eells beginnt sein Vorwort zu Héleins Buch „Harmonic maps, conservation laws and moving frames" (fortan HCM) mit den folgenden Zeilen: *„Harmonic maps between Riemannian manifolds provide a rich display of both differential geometric and analytical phenomena. These aspects are inextricably intertwined – a source of undiminished fascination."*

Eine große Anzahl von Vorträgen, Büchern und Arbeiten über harmonische Abbildungen fangen mit ähnlichen – wenn auch auf eine etwas nüchternere Art formulierten – Äußerungen an. In den meisten Fällen werden jedoch sehr schnell die geometrischen und analytischen Aspekte entkoppelt, und man konzentriert sich dann auf einen der beiden Aspekte. Bemerkenswert an HCM ist, dass hier beide Aspekte ausführlich betrachtet werden.

Nach einer gründlichen Einführung in das Fachgebiet präsentiert der Autor ein Anzahl neuer Resultate über harmonische Abbildungen und aus verwandten Fachbereichen, deren Beweise teilweise raffinierte Techniken aus der Analysis und der Differentialgeometrie verwenden. In allen Fällen werden die notwendigen Techniken im Buch ausführlich erklärt, motiviert und entwickelt; es werden nicht nur im Wesentlichen eigenständige Beweise einer Reihe von neuen, tiefgreifenden Ergebnissen auf diesem Gebiet präsentiert, sondern man gewinnt auch eine Wertschätzung für eine Anzahl von Techniken, die nicht unbedingt zum Standardrepertoire gehören. Darüber hinaus werden auch weitere mögliche Anwendungen dieser Techniken vorgestellt.

Das Buch befasst sich mit harmonischen Abbildungen. Betrachtet man zwei Riemann'sche Mannigfaltigkeiten $(\mathcal{M}, g)$ und $(\mathcal{N}, h)$, so ist eine harmonische Abbildung $u : \mathcal{M} \to \mathcal{N}$ ein kritischer Punkt der Dirichlet-Energie

$$E(u) := \tfrac{1}{2} \int_{\mathcal{M}} g^{\alpha\beta}(x) h_{ij}(u(x)) \frac{\partial u^i}{\partial x^\alpha} \frac{\partial u^j}{\partial x^\beta} dvol_g .$$

Genauer gesagt: Unter relativ milden Voraussetzungen an $\mathcal{M}$ und $\mathcal{N}$ lässt sich auf natürliche Weise der Sobolevraum $H^{1,2}(\mathcal{M}, \mathcal{N})$ definieren. Eine *schwach harmonische Abbildung* ist ein kritischer Punkt von E im Raum $H^{1,2}(\mathcal{M}, \mathcal{N})$. Ist ein solcher

kritischer Punkt darüber hinaus zweimal stetig differenzierbar, so nennt man ihn eine *(klassische) harmonische Abbildung*. Eine derartige Abbildung u löst die zu E assoziierte Euler-Gleichung, welche in lokalen Koordinaten um x_0 in $\mathcal{M}$ und $u(x_0)$ in $\mathcal{N}$ gegeben ist durch

$$\Delta_g u^i + g^{\alpha\beta}(x)\Gamma^i_{jk}(u(x))\frac{\partial u^j}{\partial x^\alpha}\frac{\partial u^k}{\partial x^\beta} = 0\,.$$

Hier ist Δ_g der Laplace-Beltrami Operator auf $\mathcal{M}$ und die Γ^i_{jk} sind die Christoffel-Symbole zweiter Art auf der Koordinatenumgebung von $u(x_0)$. Harmonische Abbildungen können somit als Verallgemeinerungen sowohl von harmonischen Funktionen als auch von Geodätischen interpretiert werden. In verschiedenen Teilgebieten der Physik sind harmonische Abbildungen von Interesse; einige von diesen physikalischen Bezügen werden in der Einleitung von HCM kurz besprochen.

Die Standardtheorie der elliptischen Systeme zeigt, dass stetige schwach harmonische Abbildungen glatt und deshalb (klassische) harmonische Abbildungen sind. Ein Schwerpunkt der Forschung über harmonische Abbildungen ist die Frage, ob schwach harmonische Abbildungen ipso facto glatt sind. Im Allgemeinen gilt dies nicht: So ist zum Beispiel die radiale Projektion von der n-dimensionalen euklidischen Kugel auf ihren Rand für $n \geq 3$ eine schwach harmonische Abbildung, die im Nullpunkt unstetig ist. Unter geeigneten Einschränkungen sind schwach harmonische Abbildungen jedoch glatt, zum Beispiel wenn die vorgebene Bildmannigfaltigkeit $\mathcal{N}$ eine nicht-positive Schnittkrümmung aufweist, oder wenn das Bild der schwach harmonischen Abbildung in einer geodätisch konvexen Teilmenge von $\mathcal{N}$ enthalten ist.

Ein weiterer Forschungsschwerpunkt befasst sich mit der partiellen Regularität von schwach harmonischen Abbildungen. Hier sind Eigenschaften der singulären Menge, d. h. der Menge der Unstetigkeitsstellen einer schwach harmonischen Abbildung, von Interesse. Zu diesen Eigenschaften zählt etwa eine obere Schranke der Hausdorff-Dimension der singulären Menge. Wie ein Ergebnis von Rivière aus dem Jahr 1995 zeigt, sind Einschränkungen notwendig, um überhaupt eine sinnvolle partielle Regularitätstheorie erhalten zu können. Rivière produzierte nirgends stetige schwach harmonische Abbildungen. Zu den typischen Klassen von Abbildungen, die auf partielle Regularität untersucht werden, gehören *energieminimierende* und *stationäre* schwach harmonische Abbildungen. Erstere minimieren die Dirichlet-Energie E bezüglich gegebener fester Randdaten oder in einer gegebenen Homotopieklasse. Zweitere sind kritische Punkte von E bezüglich geeigneter Variationen des Definitionsbereiches.

Die beiden Schwerpunkte – die volle und die partielle Regularität – bilden den Rahmen für die in HCM dargestellten Ergebnisse.

Das erste Kapitel enthält eine Einführung in die Theorie der harmonischen Abbildungen. Der Autor konzentriert sich dabei auf die Definition verschiedener Klassen schwach harmonischer Abbildungen; er erklärt und motiviert Fragen nach der partiellen und vollen Regularität, die in den nachfolgenden Kapiteln behandelt werden. Desweiteren wird das *Erhaltungsgesetz* als einer der zentralen Gedanken des Buches eingeführt; grob gesagt liefert ein solches Gesetz das Verschwinden gewisser Größen auf ganz $\mathcal{M}$, z. B. der Divergenz eines bestimmten Vektorfeldes. Ermöglicht werden solche Erhaltungsaussagen durch bestimmte Symmetrien im Definitions- oder Zielbereich.

Das zweite Kapitel befasst sich mit symmetrischen Mannigfaltigkeiten, wobei harmonischen Abbildungen von Riemann'schen Flächen in Sphären besondere Aufmerksamkeit gewidmet wird. Speziell geht der Autor darauf ein, wie gewisse Symmetrieeigenschaften der Zielmannigfaltigkeit es ermöglichen, Kompaktheitsaussagen für Folgen schwach harmonischer Abbildungen in der schwachen Topologie auf $H^{1,2}$ zu gewinnen. Der Höhepunkt des Kapitels ist ein bahnbrechender Beweis des Autors

aus dem Jahre 1990, welcher die Glattheit von schwach harmonischen Abbildungen von Riemann'schen Flächen in euklidische Sphären liefert. Der Beweis stützt sich in entscheidender Weise auf ein Ergebnis von Henry Wente aus dem Jahr 1969. Wente betrachtete das Dirichlet-Problem zum System

$$-\Delta\phi = \{a,b\} := \frac{\partial a}{\partial x}\frac{\partial b}{\partial y} - \frac{\partial a}{\partial y}\frac{\partial b}{\partial x}$$

auf einem beschränkten 2-dimensionalen Definitionsbereich Ω, mit $a,b \in H^{1,2}(\Omega)$. (Wente war vor allem an Spezialfällen dieses Systems interessiert, nämlich an Systemen, die von konform parametriesierten Flächen konstanter mittlerer Krümmung gelöst werden.) Die Standardtheorie der elliptischen Gleichungen ergibt, dass eine schwache Lösung ϕ zu L^q gehört für jedes $q \in [1,\infty)$ und darüber hinaus die Ableitung $D\phi$ in L^p liegt für jedes $p \in [1,2)$. Wente zeigte, dass die besondere Struktur der Inhomogenität es ermöglicht, eine geringe (aber entscheidende) Verbesserung zu gewinnen, nämlich $\phi \in L^\infty \cap H^{1,2}$. Hiermit konnte er die Glattheit schwacher Lösungen ableiten.

Im dritten Kapitel mit dem Titel *Compensations and exotic function spaces* befasst sich der Autor zunächst mit dem Beweis von Wentes Ergebnis. Er fährt dann mit der Beschreibung verwandter Kompensationsphänomena nach Müller (1990) und Coifman-Lions-Myers-Semmes (1993) fort, was zu einer Einführung in die Theorie der Hardy- und Lorentz-Räume führt. Das Kapitel endet mit der Darstellung eines Ergebnisses von Craig Evans aus dem Jahr 1991: Stationäre harmonische Abbildungen aus einer m-dimensionalen Mannigfaltigkeit in eine euklidische Sphäre haben verschwindendes $m-2$-dimensionales Hausdorff-Maß.

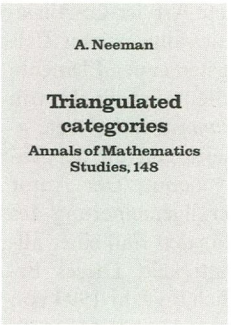
Im Mittelpunkt des vierten Kapitels steht die Erweiterung des Regularitätssatzes des Autors aus Kapitel 2. Hier betrachtet der Autor allgemeine Zielmannigfaltigkeiten. Der Grundgedanke (vom Autor bescheiden als „naiv" bezeichnet, obwohl die Ausführung ziemlich subtil ist) besteht darin, die Kompensationsargumente des vorigen Kapitels durch eine geeignete Art der Lokalisierung an die Situation eines allgemeinen Ziels anzupassen. Die wichtigsten neuen Elemente – und dritter Bestandteil des Titels von HCM – sind *moving frames* (mitbewegte n-Beine), insbesondere diejenigen, die der Autor „Coulomb frames" nennt. Der Autor fährt fort mit einer Verallgemeinerung des Resultates von Evans für den Fall einer allgemeinen Zielmannigfaltigkeit. Dieses Ergebnis wurde ursprünglich im Jahr 1993 von Fabrice Bethuel bewiesen. Der Autor erörtert ebenfalls eine Erweiterung der Kompaktheitsresultate aus Kapitel 2.

Das letzte – und relativ kurze – Kapitel beschreibt weitere Anwendungen von Kompensationsphänomenen, um höhere Regularität für Flächen, die gewisse Integralbedingungen erfüllen, zu gewinnen.

Durch die ausführliche Präsentation ist HCM als Lehrbuch für eine Vorlesung im Hauptstudium oder für ein Seminar bestens geeignet. Da es sich außerdem mit relativ neuen Ergebnissen auf diesem Fachgebiet auseinandersetzt, ist es auch für Mathematiker interessant, die auf ähnlichen Gebieten Forschungsarbeit leisten, zumal der Autor sich immer wieder auf offene Probleme bezieht.

Es gibt nur sehr wenige und auch nur geringfügige Kritikpunkte: Im ersten Kapitel gibt es ein paar Ungenauigkeiten in der Darstellung von existierenden Ergebnissen; manchmal ist es schwer festzustellen, ob der Autor gerade eine Standardnotation oder -terminologie verwendet, oder ob sie vom Autor selbst geprägt wurde; das Englisch ist manchmal etwas exzentrisch (HCM ist die zweite Ausgabe eines Buches, das ursprünglich 1996 auf Französisch erschien). Diesen letzten Punkt kann man eher als Zeichen des Enthusiasmus des Autors für dieses Thema bewerten. Er möchte seiner Leserschaft nicht nur die Thematik, sondern auch seine Begeisterung dafür näher bringen. Mit HCM ist ihm dies vollständig gelungen.

Erlangen J. Grotowski



A. Neeman
Triangulated categories
Annals of Mathematics Studies, 148

Princeton Univ. Press, Princeton, 2001, 449 S., $ 46,–

Warum ein Buch über triangulierte Kategorien? Sie gehören heute zu den wichtigen Hilfsmitteln in Algebra, Geometrie und Topologie, so daß man vielleicht den Untertitel „Triangulated categories for the working mathematician" erwartet. Tatsächlich hat Neeman allerdings mehr zu bieten. Neben einer elementaren Einführung in die Theorie der triangulierten Kategorien werden zwei Konzepte ausführlich behandelt, nämlich *Brownsche Darstellbarkeit* und *Bousfield Lokalisierung*. Beide Konzepte haben ihren Ursprung in der algebraischen Topologie der 60er und 70er Jahre, wurden jedoch in den 90er Jahren in einer Allgemeinheit weiterentwickelt, die unter anderem zu schönen Anwendungen in der algebraischen Geometrie und in der Darstellungstheorie endlicher Gruppen führte.

Um die Thematik des vorliegenden Bandes zu verstehen, vergleichen wir trianguliert mit abelschen Kategorien. Eine abelsche Kategorie kann man sich als additive Kategorie (d. h. die Morphismenmengen bilden abelsche Gruppen und die Komposition ist bilinear) vorstellen, in der gewisse Sequenzen von Morphismen (die sogenannten *kurzen exakten Sequenzen*) ausgezeichnet sind und gewisse Axiome erfüllen. Eine triangulierte Kategorie ist ebenso eine additive Kategorie, in der gewisse Sequenzen von Morphismen (die sogenannten *exakten Dreiecke*) ausgezeichnet sind und gewissen Axiome erfüllen.

Das klassische Beispiel einer triangulierten Kategorie ist die *derivierte Kategorie* $\mathbf{D}(\mathcal{A})$ einer abelschen Kategorie $\mathcal{A}$. In diesem Fall erhält man eine Einbettung

$$\mathcal{A} \longrightarrow \mathbf{D}(\mathcal{A}), \quad X \longmapsto \bar{X},$$

und jede kurze exakte Sequenz $0 \to X \to Y \to Z \to 0$ in $\mathcal{A}$ induziert ein exaktes Dreieck $\bar{X} \to \bar{Y} \to \bar{Z} \to \bar{X}[1]$ in $\mathbf{D}(\mathcal{A})$. Insbesondere liefert ein Element in $\mathrm{Ext}^1_{\mathcal{A}}(Z, X)$ einen Morphismus $\bar{Z} \to \bar{X}[1]$ in $\mathbf{D}(\mathcal{A})$, und allgemeiner gilt

$$\mathrm{Ext}^n_{\mathcal{A}}(Z, X) \cong \mathrm{Hom}_{\mathbf{D}(\mathcal{A})}(\bar{Z}, \bar{X}[n])$$

für jedes $n \geq 0$.

Die derivierte Kategorie ermöglicht es also derivierte Funktoren zu berechnen. Verdier hat diese Idee in seiner Thèse erstmals formalisiert und in diesem Zusammenhang triangulierte Kategorien eingeführt [5]. Parallel dazu hat seinerzeit Puppe in der algebraischen Topologie dieselben Eigenschaften für die stabile Homotopiekategorie herausgearbeitet [4].

Betrachtet man abelsche Kategorien, so spielen häufig zusätzliche Eigenschaften eine wichtige Rolle. Klassisch sind sogenannte *Grothendieck Kategorien*, also abelsche Kategorien mit exakten induktiven Limites und einem Generator. Beispielsweise ist jede Modulkategorie eine Grothendieck Kategorie. In der Welt der triangulierten Kategorien sucht man zu Recht nach einem vergleichbaren Konzept. Bewährt haben sich die *kompakt erzeugten* triangulierten Kategorien, denn Brownsche Darstellbarkeit und Bousfield Lokalisierung lassen sich auf diese Klasse von triangulierten Kategorien übertragen. Die derivierte Kategorie einer Modulkategorie oder die stabile Homotopiekategorie der algebraischen Topologie sind klassische Beispiele für kompakt erzeugte triangulierte Kategorien. Allerdings treten in der Praxis triangulierter Kategorien auf, die nicht kompakt erzeugt sind. Hier setzt Neeman mit seiner Verallgemeinerung an und entwickelt das Konzept der *wohlerzeugten* (englisch: *well generated*) triangulierten Kategorie.

Die Diskussion der wohlerzeugten Kategorien bildet das Herzstück dieser Monographie. Die Definition ist zugegebenermaßen kompliziert, läßt sich jedoch glücklicherweise erheblich vereinfachen [1]. Die entscheidende Idee geht zurück auf Freyd und wird in diesem Buch ausführlich behandelt. Es handelt sich quasi um eine Umkehrung der Konstruktion von Verdier, d. h. eine triangulierte Kategorie wird universell in eine abelsche Kategorie eingebettet, so daß exakte Dreiecke in exakte Sequenzen überführt werden. Diese universelle abelsche Kategorie ist im allgemeinen zu groß, aber lässt sich im Fall der wohlerzeugten Kategorien geschickt verkleinern ohne daß zu viel Information verloren geht.

Wohlerzeugte Kategorien verallgemeinern also die kompakt erzeugten Kategorien, und das erklärte Ziel dieser Verallgemeinerung sind die Übertragung von Brownscher Darstellbarkeit und Bousfield Lokalisierung auf wohlerzeugte triangulierte Kategorien. Der *Brownsche Darstellbarkeitssatz* besagt für eine triangulierte Kategorie $\mathcal{T}$ mit beliebigen Koprodukten, daß jeder kontravariante Funktor $F: \mathcal{T} \to Ab$ in die Kategorie der abelschen Gruppen genau dann *darstellbar* ist (d. h. von der Form $\mathrm{Hom}_{\mathcal{T}}(-, X)$ für ein Objekt X) wenn F exakte Dreiecke in exakte Sequenzen und Koprodukte in Produkte überführt. Dieser Satz ist äußerst nützlich. Beispielsweise erhält man einen eleganten Beweis für die Grothendieck Dualität indem man beobachtet, daß der Brownsche Darstellbarkeitssatz im Fall der derivierten Kategorie eines Schemas anwendbar ist [2].

Von einer *Bousfield Lokalisierung* spricht man sofern eine triangulierte Unterkategorie $\mathcal{S} \subseteq \mathcal{T}$ vorliegt, und der Quotientenfunktor $\mathcal{T} \to \mathcal{T}/\mathcal{S}$ im Sinne Verdiers einen Rechtsadjungierten besitzt. Im Fall der Bousfield Lokalisierung ist die triangulierte Unterkategorie $\mathcal{S}$ *lokalisierend*, d. h. abgeschlossen unter Koprodukten. Diese Bedingung an $\mathcal{S}$ ist im allgemeinen jedoch nicht hinreichend. Beispielsweise wird gezeigt, daß die Bousfield Lokalisierung existiert falls $\mathcal{S}$ von einer Menge von Objekten erzeugt wird. In diesem Fall sind übrigens auch $\mathcal{S}$ und der Quotient $\mathcal{T}/\mathcal{S}$ wohlerzeugt.

Insgesamt beruht ein Großteil des Materials in diesem Buch auf dem Wechselspiel zwischen triangulierter und abelscher Struktur. Dieses wird ergänzt durch eine Reihe von Anhängen, in denen vorwiegend das notwendige Material über abelsche Kategorien bereitgestellt wird. Damit ist die Darstellung im wesentlichen in sich abgeschlossen. Lediglich ein paar Grundkenntnisse über Kategorien und Funktoren, oder aus der homologischen Algebra werden vorausgesetzt. Dagegen verzichtet der Autor weitgehend auf Beispiele. Ein gewisses Interesse des Lesers an allgemeiner Theorie wird also vorausgesetzt. Hilfreich sind die historischen Bemerkungen am Schluß eines jeden Abschnitts. Sie vermitteln den Standpunkt des Autors und bieten eine gute Orientierung.

Kehren wir zur Ausgangsfrage nach dem Zweck eines solchen Buches über triangulierte Kategorien zurück, so ist die Antwort des Rezensenten zwiespältig. Sucht man eine elementare Einführung in die Theorie der triangulierten Kategorien, so ist man im Vergleich mit Verdiers Thèse mindestens ebenso gut bedient, Französischkenntnisse vorausgesetzt. Positiv erwähnt sei jedoch Neemans elegante Darstellung des sogenannten Oktaederaxioms. Moderne Aspekte, beispielsweise t-Strukturen, werden leider nicht berücksichtigt, wobei allerdings das Versprechen „We will see them [the six gluing functors] again in the discussion of t-structures." auf Seite 319 vom guten Willen des Autors zeugt.

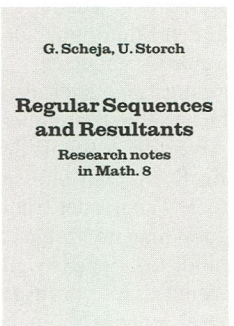
Das Buch wird zu einer wertvollen Lektüre sobald man sich für *große* triangulierte Kategorien interessiert, also Kategorien mit beliebigen Koprodukten. Hier leistet Neeman echte Pionierarbeit, und diese verdient uneingeschränkte Anerkennung.

Literatur

[1] H. Krause: On Neeman's well generated triangulated categories. Documenta Math. **6** (2001), 119–125.

[2] A. Neeman: The Grothendieck duality theorem via Bousfield's techniques and Brown representability. J. Amer. Math. Soc. **9** (1996), 205–236.
[3] A. Neeman: A survey of well generated triangulated categories. Preprint (2002).
[4] D. Puppe: On the structure of stable homotopy theory. Colloquium on algebraic topology, Aarhus Universitet Matematisk Institut (1962), 65–71.
[5] J. L. Verdier: Des catégories dérivées des catégories abéliennes. Astérisque **239** (1996).

Paderborn H. Krause



G. Scheja, U. Storch
Regular Sequences and Resultants
Research notes in Math. 8

Natick, Mass., AK Peters, 2001, 142 S., $ 30,–

In dem Buch von Scheja und Storch wird die Eliminationstheorie in gewichteten projektiven Räumen behandelt, die über beliebigen kommutativen, noetherschen Ringen definiert sind. Dabei spielen reguläre Sequenzen und vollständige Durchschnitte eine wichtige Rolle. Demgemäß werden im zweiten der insgesamt vier Kapitel des Buches reguläre Sequenzen und vollständige Durchschnitte ausführlich behandelt, und zwar für beliebige noethersche Ringe in der lokalen als auch der globalen Version. Besonderes Augenmerk wird den graduierten vollständigen Durchschnitten gewidmet. Diese sind Restklassenringe von homogenen regulären Folgen in einem Polynomring über einem noetherschen Ring, dessen Variablen positive Gewichte besitzen.

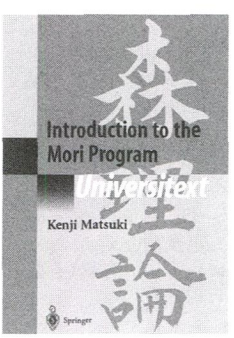
Für die Definition von Resultanten sind generische reguläre Sequenzen von Bedeutung. Hierbei ist der Grundring des gewichteten Polynomrings selbst ein Polynomring über den ganzen Zahlen, und die Polynome der Sequenz besitzen als Koeffizienten die Unbestimmten des Grundrings. Das schwierige Problem, wann Sequenzen von homogenen generischen Polynomen eine reguläre Folge bilden, wird in dem Buch umfassend diskutiert. Dies führt zu dem Begriff der streng zulässigen Folgen. Insbesondere werden generische Binome charakterisiert, die streng zulässig sind.

Im weiteren Verlauf wird der Hauptsatz der Eliminationstheorie für projektive Räume vorgestellt und es wird gezeigt, dass das generische Eliminationsideal bezüglich streng zulässiger Folgen ein Haupt- und Primideal ist. Danach wird das Eliminationsideal bezüglich beliebiger homogener regulärer Folgen bestimmt. Insbesondere wird ausgeführt, dass das Eliminationsideal ein divisorielles Ideal ist, falls der Grundring ein ganz abgeschlossener Integritätsbereich ist.

Das letzte Kapitel des Buches schließlich behandelt Resultanten. Diese sind insofern von Bedeutung, als sie dieselbe Nullstellenmenge wie das Eliminationsideal haben, darüberhinaus aber bessere funktorielle Eigenschaften besitzen. Die Autoren führen zunächst das Resultantenideal einer regulären Folge homogener Polynome eines Polynomrings über einem ganz abgeschlossenen noetherschen Integritätsbereich ein. Dabei stimmt die Länge der Folge mit der Anzahl der Variablen des Polynomrings überein. Es wird gezeigt, dass dieses Resultantenideal wieder divisoriell ist. Mit Hilfe des Koszulkomplexes, welcher der regulären Folge zugeordnet ist, wird ein kanonischer Erzeuger des Resultantenideals konstruiert. Dies führt zu expliziten Resultantenformeln, und verallgemeinert die von Hurwitz bekannte klassische Resultante, wo alle Variablen das Gewicht 1 haben.

Das Buch ist eine außerordentlich sorgfältig ausgearbeitete moderne Einführung in die Theorie der Resultanten. Es erfordert vom Leser solide Kenntnisse der kommutativen und homologischen Algebra, sowie einige Grundkenntnisse der algebraischen Geometrie. Die Autoren haben sich bemüht, die speziell hier erforderlichen Techniken und Sachverhalte, etwa Eigenschaften numerischer Halbgruppen, die sonst in der Literatur gar nicht oder nur schwer auffindbar sind, zusammenzustellen und im Detail zu erklären. Dadurch ist das Buch weitgehend in sich geschlossen und kommt in seinen Beweisen ohne weitere Literaturverweise aus. Darüberhinaus wurde den Abschnitten Übungsaufgaben, historische Anmerkungen und Ergänzungen beigefügt, die aber für das Verständnis des Haupttextes nicht unbedingt notwendig sind. Somit ist das Buch als Textbuch einer Spezialvorlesung, oder als Grundlage für ein Seminar bestens geeignet.

Essen J. Herzog



K. Matsuki
Introduction to the Mori Program

Berlin u. a., Springer, 2002, 478 S., EUR 74,95

Das Mori-Programm und die birationale Geometrie von projektiven Varietäten der Dimension 3 waren zentrale Themen in der algebraischen Geometrie der 1980er Jahre. Heute ist die Mori-Theorie zu einem unverzichtbaren Werkzeug geworden. Vereinfacht gesagt, geht es darum

1. unter allen projektiven Varietäten, die zu einer gegebenen Mannigfaltigkeit birational sind, besonders einfache Varietäten, die „minimalen Modelle", zu finden und
2. die Eigenschaften der minimalen Modelle zu studieren und birationale Abbildungen zwischen den minimalen Modellen zu klassifizieren.

Im „Minimalen Modell Programm", oder „Mori-Programm", wird Aufgabe (1) gelöst, indem zu jeder projektiven Mannigfaltigkeit eine Folge von birationalen Abbildungen konstruiert wird, an deren Ende entweder ein sehr spezieller Faserraum oder ein „minimales Modell" steht, also eine Varietät, deren kanonisches Bündel nef ist. Das Programm entstand aus den Arbeiten [Mor79] und [Mor82], in denen S. Mori die Existenz von rationalen Kurven auf nicht-minimalen Mannigfaltigkeiten zeigte. Mit Hilfe dieser Kurven konstruierte er Abbildungen, die ein direktes Analogon zur klassischen Kontraktion von (-1)-Kurven auf Flächen sind. Motiviert durch diese Ergebnisse entstand in dem folgenden Jahrzehnt die moderne Mori-Theorie durch die Arbeit vieler Mathematiker, unter anderem Y. Kawamata, J. Kollár, Y. Miyaoka, S. Mori, M. Reid, V. Sarkisov und V. Shokurov. Zumindest für 3-dimensionale Varietäten wurde ein einigermaßen befriedigendes Bild der birationalen Geometrie gewonnen.

Kenji Matsuki, der die Entwicklung von Anfang an begleitet hat und einer der Autoren der grundlegenden Arbeit [KMM87] war, bietet in dem vorliegenden Buch eine engagiert geschriebene und gut lesbare Darstellung der Theorie. Matsuki legt großen Wert darauf, alle auftretenden Konzepte und Begriffe ausführlich zu motivieren und die Nützlichkeit jedes neuen Konzeptes zuerst an vertrauten Beispielen zu beleuchten. Viele kleine in den Text eingestreute Übungsaufgaben helfen sehr beim Verstehen, der Stil ist direkt und angenehm zu lesen. Das Einzige, was mich persönlich oft gestört hat, ist die Zerrissenheit des Layouts in den technischen Abschnitten.

Das Buch gliedert sich grob in drei Teile:

Geometrie von Flächen In den ersten beiden Kapiteln, die zusammen etwa ein Viertel des Buches ausmachen, wird die klassische birationale Geometrie von Flächen wiederholt. Viele Begriffe der höherdimensionalen Theorie werden hier bereits eingeführt und durch Beispiele erklärt. Mir hat besonders gut gefallen, wie klar Matsuki zeigt, dass eine logarithmische Variante der Mori-Theorie notwendig ist, auch wenn man zunächst nur an dem „normalen" Mori-Programm interessiert ist.

Moris Programm Nach der aufwändigen Einführung werden im dritten Kapitel die Hauptresultate der Mori-Theorie zunächst ohne Beweis zusammengestellt. Matsuki zeigt an Beispielen, warum in der höherdimensionalen Theorie Singularitäten auftreten, und warum die Einführung einer neuen Klasse von birationalen Transformationen, den „Flips", notwendig ist.

In den Kapiteln 4–8 werden der Verschwindungssatz von Kawamata-Viehweg, der Satz über die Basispunktfreiheit und der Kegelsatz bewiesen. Die Beweise orientieren sich an im Wesentlichen an den Originalarbeiten, etwa [KMM87], und bieten dem Fachmann nichts Neues. Der Autor beschränkt sich bei der Klassifikation der auftretenden Singularitäten auf den Flächenfall. Im Kapitel 9 behandelt Matsuki den noch verbliebenen Teil des Mori-Programms, die Sätze über die Existenz und die Terminierung von Flips. Der äußerst lange und schwierige Existenzbeweis wird nicht gegeben.

Abbildungen zwischen minimalen Modellen Die verbleibenden Kapitel befassen sich mit der Geometrie der rationalen Kurven und mit der Geometrie von birationalen Abbildungen zwischen minimalen Modellen und zwischen Mori-Faserräumen. Als besonderes Bonbon wird im abschließenden Kapitel die ganze Theorie noch einmal für torische Varietäten dargelegt.

Fazit: Kenji Matsuki hat meiner Meinung nach ein durchweg gelungenes Buch vorgelegt. Das Werk unterscheidet sich von den bekannten Darstellungen [CKM88], [MP97] und [KM98] und dem Originalartikel [KMM87] im Wesentlichen durch die Ausführlichkeit und durch die Fülle an konkreten Beispielen. Es werden lediglich Grundkenntnisse in algebraischer Geometrie und der Geometrie von Flächen vorausgesetzt, so dass sich das Buch für ein Oberseminar gut eignet.

Literatur

[CKM88] H. Clemens, J. Kollár, and S. Mori. Higher dimensional complex geometry. *Astérisque*, 16, 1988.

[KM98] J. Kollár and S. Mori. *Birational geometry of algebraic varieties*, volume 134 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 1998. With the collaboration of C. H. Clemens and A. Corti, Translated from the 1998 Japanese original.

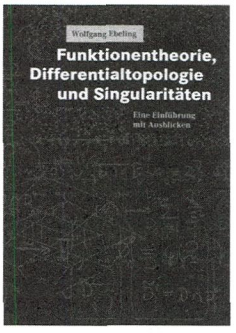
[KMM87] Y. Kawamata, K. Matsuda, and K. Matsuki. Introduction to the minimal model problem. *Algebraic Geometry*, Sendai, 1987.

[Mor79] S. Mori. Projective manifolds with ample tangent bundles. *Ann. of Math.*, 110:593–606, 1979.

[Mor82] S. Mori. Threefolds whose canonical bundles are not numerically effective. *Ann. of Math., II. Ser.*, 116:133–176, 1982.

[MP97] Y. Miyaoka and Th. Peternell. *Geometry of higher-dimensional algebraic varieties*, volume 26 of *DMV Seminar*. Birkhäuser Verlag, Basel, 1997.

Bayreuth S. Kebekus



W. Ebeling
Funktionentheorie, Differentialtopologie und Singularitäten
Eine Einführung mit Ausblicken

Braunschweig/Wiesbaden, Vieweg, 2001, 303 S., EUR 29,90

Das vorliegende Buch gibt eine Einführung in Teile der höheren Funktionentheorie (Theorie der Riemannschen Flächen, Funktionentheorie mehrerer komplexer Veränderlicher) sowie Grundlagen und einige neuere Ergebnisse der Theorie der Singularitäten.

Bei der Breite der angesprochenen Gebiete ist von vornherein zu erwarten, dass in keinem von ihnen auch nur entfernt der Anspruch auf Vollständigkeit erhoben wird. Schon die Theorie der Riemannschen Flächen ist für sich genommen so umfangreich, dass sie sich kaum in ein einziges Lehrbuch komprimieren lässt. Aus diesem Grund ist von vornherein eine restriktive Auswahl erforderlich.

Diese orientiert sich im vorliegenden Buch vor allem an den Bedürfnissen der Theorie der Singularitäten. Dieses Gebiet hat sich in der zweiten Hälfte des zwanzigsten Jahrhunderts etabliert. Auch wenn es hier Fragen gibt, die sich einfach erklären lassen und die geometrisch transparent sind, gibt es eine Reihe von Hilfsmitteln aus verschiedenen Gebieten, die für die Arbeit unerlässlich sind: aus der Funktionentheorie, der Differentialtopologie und der algebraischen Topologie.

Zu den Begriffen, die aus der algebraischen Topologie benötigt werden, gehören abgesehen von den singulären Homologiegruppen die Überlagerungen, die Fundamentalgruppe und die Homotopiegruppen. Dementsprechend konzentriert sich das Buch im Bereich der Riemannschen Flächen auf Themen, die mit der Überlagerungstheorie zusammenhängen.

Ähnliches gilt im Bereich der Funktionentheorie mehrerer Veränderlicher: Da es sich bei der Singularitätentheorie um eine lokale Theorie handelt, wird hier primär die lokale Theorie analytischer Mengen behandelt. Jeder, der eine Vorlesung über dieses Thema gehalten hat, weiß, dass hierbei die kommutative Algebra wesentlich zu Hilfe genommen werden muss – der elementar-geometrischen Argumentation sind Grenzen gesetzt. Dies gilt zum Beispiel für die saubere Einführung des Dimensionsbegriffs.

Im Rahmen der Differentialtopologie wird aus ähnlichen Gründen die Theorie der differenzierbaren Faserbündel in den Vordergrund gestellt.

Vom Standpunkt der Singularitätentheorie erfolgt in diesem Buch eine Konzentration auf den klassischen Fall der isolierten Hyperflächensingularitäten. Die Behandlung der Milnorfaserung darf hier natürlich nicht fehlen. Für eine detaillierte Untersuchung der entsprechenden Monodromie sind jedoch weitere Hilfsmittel angebracht, die in diesem Buch entwickelt werden, insbesondere die Morsifikation, die ein Heranziehen der Picard-Lefschetz-Formeln erlaubt. Aber auch weitere Themen wie universelle Entfaltung und Klassifikation von Singularitäten werden darüber hinaus angesprochen.

Nun zum Aufbau: Das erste Kapitel ist der Theorie der Riemannschen Flächen gewidmet. Sie wird vom Standpunkt der Überlagerungstheorie aus dargestellt. Bekanntlich verwendet man in der Theorie der Riemannschen Flächen im Gegensatz zur algebraischen Topologie einen abgeschwächten Überlagerungsbegriff, deswegen setzt man die beteiligten Räume zweckmässig als hausdorffsch voraus (auch wenn dies nicht immer wiederholt wird). Nicht nur historisch, sondern auch systematisch wichtig ist die Einführung der Riemannschen Fläche einer me-

romorphen, insbesondere einer algebraischen Funktion – damit ergibt sich der Zusammenhang zwischen algebraischen Kurven und Riemannschen Flächen.

Im zweiten Kapitel wird die lokale Theorie der analytischen Mengen behandelt. Der algebraische Apparat wird dabei so entwickelt, dass eine Begründung der geometrischen Sachverhalte ermöglicht wird. Zu den behandelten Begriffen zählen Regularität, Dimension, Bildmenge.

Im dritten Kapitel geht es primär um die analytische Untersuchung isolierter Singularitäten holomorpher Funktionen, am Ende steht dabei die Klassifikation der einfachen Singularitäten. Ein nützliches Hilfsmittel ist dabei der Begriff der (universellen) Entfaltung. Darüber hinaus werden vorab grundlegende Begriffe aus der Differentialtopologie wie Vektorfelder, Transversalität und Liegruppen angesprochen.

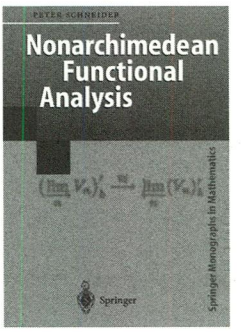
Im vierten Kapitel wird ein wichtiges Gebiet aus der Differentialtopologie behandelt: die differenzierbaren Faserungen. Zu den in diesem Zusammenhang angesprochenen Themen zählen die Holonomiegruppe, Schnitt- und Verschlingungszahlen und die Zopfgruppe.

Diese Theorie ist wichtig für die topologische Untersuchung von Singularitäten im letzten Kapitel. Im Zusammenhang mit der Milnorfaserung gelangt man zu verschiedenen Invarianten wie Monodromie, Variation und Seifertform, der Übergang zu einer Morsifikation führt dabei zu einem tieferen Verständnis der geometrischen Situation. Mit dem „Milnorgitter" wird ein Begriff angesprochen, der zum Forschungsgebiet des Autors gehört; am Schluss wird ein Ausblick auf neuere Entwicklungen gegeben.

Das Buch wendet sich nicht nur an Studenten, sondern auch an Hochschullehrer, die eine einschlägige Vorlesung oder ein entsprechendes Seminar planen. Hierzu wird im Vorwort eine Reihe praktikabler Vorschläge gemacht.

Das Buch ist sorgfältig verfasst, die Voraussetzungen werden deutlich gemacht. Es bietet die Möglichkeit zu verschiedenartigem Einsatz in der Lehre wie zum Selbststudium (etwa zur Spezialisierung für Diplomanden, zur Einarbeitung für Doktoranden). Insgesamt ist das Buch daher sehr empfehlenswert.

Münster H. Hamm



P. Schneider
Nonarchimedean Functional Analysis
Springer Monographs in Mathem

Berlin u. a., Springer, 2002, 156 S., € 39,95

Nonarchimedean Functional Analysis is the study of Functional Analysis where the usual real or complex scalar field is replaced by a nonarchimedean valued complete field K, for example the field of the p-adic numbers. The Founding Father of this discipline was the Dutch mathematician A.F. Monna with a series of papers, beginning in 1943, and the area was further developed in the sixties and seventies, culminating in the book of A.C.M. van Rooij, Non-archimedean Functional Analysis (Marcel Dekker, 1978) in which the then 'state of the art' was established. The book mainly treated Banach space theory whereas in the thesis of J. van Tiel (1965) the foundations of locally convex spaces were introduced. After that, numerous papers in the area appeared, but no systematic books until 2001, when the present work under review was published. It is the first textbook seriously covering locally convex theory over K, so, for that reason alone, it is most welcome.

The main motivation for writing the book, according to the author, was the growing interest from other areas, for example number

theory (p-adic modular forms and deformation of Galois representations) and representation theory of p-adic reductive groups, in methods from Functional Analysis. Yet, the book is self-contained, complete with all proofs, and therefore attractive also to those who are not acquainted with the above areas. In fact, the reader is assumed to have only a basic knowledge of linear algebra and point set topology. However, to enjoy the book one should be familiar with the basics of classical Functional Analysis.

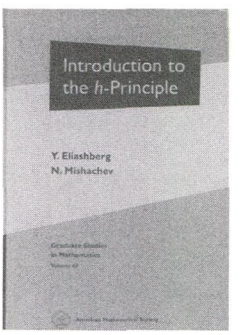
The book has four chapters. In the first two the foundations on Banach and locally convex spaces are established. Most of this theory can already be found in the work cited above. The real story starts in Chapter 3 that treats duality theory. With an eye on the possible users from other disciplines the author restricts himself to the case where K is spherically (= maximally) complete, the reason being that in this case the Hahn Banach Theorem holds without restriction, and the fact that in the envisaged applications K is discretely valued, hence spherically complete. On one hand this is certainly a valid point of view, on the other hand it is a pity that the generally interested reader does not get easy access to the interesting case where K is not spherically complete (e.g. the completion of the algebraic closure of the p-adic number field). Be that as it may, by this restriction the theory runs more smoothly than in the general case.

In Chapter 3 the basic notions of c-compactness and compactoidity are introduced. They can be viewed as a 'convexification' of compactness, precompactness respectively, needed since K may be not locally compact. With the help of these notions the Mackey topology is shown to be the finest admissible topology. After studying reflexivity attention is given to compact inductive limits (a linear map is called compact if there is a neighbourhood of 0 whose image has c-compact closure). In Chapter 4 the usual injective and projective tensor product topologies are introduced, and they turn out to be equal (!). This has as a consequence that the nonarchimedean notions of Schwarz spaces and nuclear spaces are the same. Completely continuous, compact, nuclear maps are studied. A trace function is introduced on the set of nuclear maps, and finally, a theory of Fredholm operators (kernel and cokernel are finite-dimensional) is developed.

The theory of Chapters 3 and 4 is rather recent, interesting and deep.

The book is well-written, with care for details. Recommended.

Nijmegen W. H. Schikhof



Y. Eliashberg, N. Mishachev
Introduction to the h-Principle (Grad Studies in Math. 48)

Providence, Am. Math. Soc., 2002, 206 S., $ 30,–

The homotopy principle, or *h*-principle, refers to the "flexibility" of certain differential (in)equalities, in the sense that their solution can be reduced to a purely homotopy theoretic problem. It was discovered by J. Nash and S. Smale in the 1950ies and formalized (under the name w.h.e.-principle) by Y. Eliashberg and M. Gromov in the early 1970ies. The development culminated in Gromov's 1986 book *Partial Differential Relations* (where he also introduced the term *h*-principle).

Instances of the *h*-principle appear in all branches of differential geometry, e.g.: the Nash-Kuiper theorem on C^1-isometric immersions; Smale's theorem on sphere immersions; Oka's principle for Stein manifolds; Lagrangian immersions in symplectic mani-

folds; and Legendrian immersions in contact manifolds.

Despite its importance, the state of the literature on the h-principle was quite unsatisfactory. Gromov's book was renowned for its generality and comprehensiveness, less for its readibility. The situation improved with D. Spring's 1998 monograph *Convex Integration Theory*. The present book takes a different, and more direct, approach to the h-principle via "holonomic approximation". To describe this, we need some notation.

For a smooth fibration $X \to V$ denote by $X^{(r)}$ the space of r-jets of (local) sections in X. The jet space fibers naturally over V. The r-jet of any (global) section $f : V \to X$ yields a section $J^r_f : V \to X^{(r)}$ of this fibration. Such a section is called *holonomic*. A *(partial) differential relation* is simply a subset $\mathcal{R}$ of $X^{(r)}$. One says that $\mathcal{R}$ satisfies the *h-principle* if every section $F : V \to \mathcal{R} \subset X^{(r)}$ is homotopic, through sections in $\mathcal{R}$, to a holonomic section.

In Chapters 1–3 of the book, Eliashberg and Mishachev set up and prove a general *Holonomic Approximation Theorem*: If F is a section of $X^{(r)} \to V$ near a submanifold $A \subset V$ of positive codimension, then there exists a C^0-small diffeotopy $h^t : V \to V$ and a holonomic section $\tilde{F}$ near $h^1(A)$ which is C^0-close to F near $h^1(A)$. This result immediately implies the h-principle near submanifolds of positive codimension for open differential relations that are invariant under diffeomorphisms of V.

The proof of the Holonomic Approximation Theorem, albeit elementary, is by no means easy to digest. But the authors spell it out in considerable detail, illustrating it with helpful figures, so that with enough time and energy one can follow the arguments.

The reader's efforts are rewarded by numerous applications of the Holonomic Approximation Theorem in differential topology, which occupy the following 5 chapters. Classical examples include Smale's sphere eversion (the two-sphere in $\mathbb{R}^3$ can be turned inside out through immersions), and the Smale-Hirsch h-principle for immersions of positive codimension.

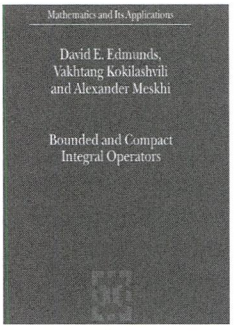
Some of the biggest successes, and the most spectacular failures, of the h-principle have occured in symplectic and contact geometry. Given Eliashberg's research in these two areas over the last two decades, it comes as no surprise that Chapters 9–16 are devoted to applications in symplectic and contact geometry. Chapter 9 provides the necessary background material. In the ensuing 7 chapters, the following results are proved (among others): existence of symplectic and contact structures on open manifolds; C^0-approximation by isosymplectic and isocontact embeddings of codimension at least 4; h-principle for isosymplectic and isocontact immersions of positive codimension; h-principle for Legendrian and Lagrangian immersions.

Chapter 11 mentions some of the failures of the h-principle in symplectic and contact geometry. However, this discussion is very sketchy, and the interested reader should consult some of the numerous textbooks and survey articles on symplectic topology, e.g. McDuff and Salamon's *Introduction to Symplectic Topology*.

The remaining chapters are concerned with convex integration. From the outset, the authors restrict to the case of 1-jets. The main theorem is the h-principle for "ample" differential relations in the 1-jet space (Chapter 18). Using this, the authors prove in Chapter 21 one of the most surprising results in differential geometry: the Nash-Kuiper theorem that any "short" embedding of positive codimension of Riemannian manifolds can be C^0-approximated by a C^1-isometric embedding. For example, the standard two-sphere of radius 1 admits a C^1-isometric embedding into an arbitrarily small ball in $\mathbb{R}^3$!

In my opinion, this is an excellent book which makes an important theory accessible to graduate students in differential geometry – and to senior researchers who never had the stamina to work through Gromov's book.

München K. Cieliebak



D. E. Edmunds, V. Kokilashvili, A. Meskhi
Bounded and Compact Integral Operators
Math. and its Applications

Dordrecht u. a., Kluver Acad. Publ. 2002, 643 S., € 199,–

This book deals with mapping properties of classical integral operators and their generalisations preferably in weighted L_p-spaces and corresponding weighted Lorentz-spaces $L_{p,q}$ defined on intervals, the half-line $\mathbb{R}_+$, the line $\mathbb{R}$, sets in $\mathbb{R}^n$ or, more generally, measure spaces (X, μ). Here (X, μ) might be a homogeneous space where the measure μ satisfies the doubling condition or a non-homogeneous space where μ does not necessarily has the doubling property. The main aim is to find criteria (this means necessary and sufficient conditions) under which the operators considered are bounded, or compact. This includes some qualitative assertions about the degree of compactness expressed in terms of singular numbers. There are also a few applications to Fourier multipliers in some function spaces defined on $\mathbb{R}^n$. The classical integral operators considered include:

$$Hf(x) = \int_0^x f(y)\,dy, \quad x > 0,$$

(Hardy type),

$$R_\alpha f(x) = \int_0^x \frac{f(t)}{(x-t)^{1-\alpha}}\,dt,$$

(fractional integrals),

$$\mathcal{R}_\alpha f(x) = \int_{-\infty}^x (x-y)^{\alpha-1} f(y)\,dy,$$

(Riemann -Liouville),

$$\mathcal{W}_\alpha f(x) = \int_x^\infty (y-x)^{\alpha-1} f(y)\,dy, \quad \text{(Weyl)},$$

$$T_{\gamma ab} f(x) = \int_a^b \frac{f(y)}{|x-y|^{1-\gamma}}\,dy,$$

(potential type),

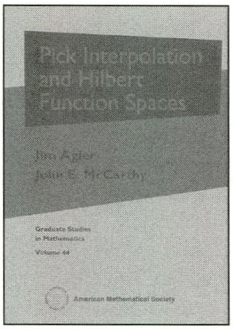
but also

$$B_+^\alpha f(x) = \int_{|y|<|x|} \frac{(|x|^2 - |y|^2)^\alpha}{|x-y|^n} f(y)\,dy,$$

(in $\mathbb{R}^n$).

This book is mainly based on the authors' own results obtained within the last years. But it covers in many respects the classical theory (since the assertions are often criteria). Some chapters have almost encyclopedic character and may serve as a source both for researchers and consumers of inequalities for integral operators. The book has 9 chapters and a list of open problems. The references cover the state-of-art. There are 310 items. The 9 chapters are: 1. Hardy-type operators, 2. Fractional integrals on the real line, 3. One-sided maximal functions, 4. Ball fractional integrals, 5. Potentials on $\mathbb{R}^N$, 6. Fractional integrals on measure spaces, 7. Singular numbers, 8. Singular integrals, 9. Multipliers of Fourier transforms.

Jena H. Triebel



J. Agler, J. E. Mc Carthy
Pick Interpolation and Hilbert Function Spaces
Grad Studies in Math. 44

Providence, American Math. Soc., 2002, 308 S., $ 49,–

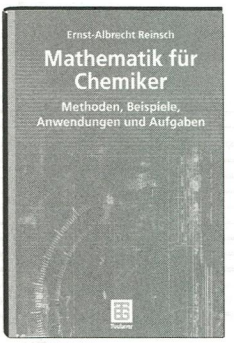
Let $\lambda_1, \ldots, \lambda_N$ be N points in the unit disc $\mathbb{D}$ in the complex plane and let $w_1, \ldots, w_N$ be complex numbers. The classical Pick problem (Pick, 1916) deals with the question of whether there exists a holomorphic function ϕ,

$$\phi : \quad \mathbb{D} \mapsto \mathbb{D} \quad \text{with} \quad \phi(\lambda_j) = w_j$$
where $j = 1, \ldots, N$.

The space $H^\infty(\mathbb{D})$ of all bounded analytic functions in $\mathbb{D}$ can be indentified with the space of all multipliers of the Hardy space H^2. Then the classical Pick problem can be reformulated in terms of this multiplier space $H^\infty(\mathbb{D})$. The book deals with the following generalisation. A Hilbert function space is a Hilbert space $\mathcal{H}$ of functions on some set X (preferably, but not necessarily, holomorphic functions in subsets of $\mathbb{C}^n$) such that the evaluation at each point of X is a nonzero continuous functional on $\mathcal{H}$. Let $\lambda_1, \ldots, \lambda_N$ be points of X and let $w_1, \ldots, w_N$ be complex numbers. The Pick problem is the question when does there exist a multiplier ϕ of $\mathcal{H}$ of norm at most one such that $\phi(\lambda_j) = w_j$. The main aim of the book is to analyse what spaces have the Pick property. This is done in 16 chapters: 0. Prerequisites and notation, 1. Introduction, 2. Kernels and function spaces, 3. Hardy spaces, 4. $P^2(\mu)$, 5. Pick redux, 6. Qualitative properties of the solution of the Pick problem in $H^\infty(\mathbb{D})$. 7. Characterizing kernels with complete Pick property, 8. The universal Pick kernel, 9. Interpolating sequences, 10. The model theory I: Isometries, 11. The bidisk, 12. The extremal three point problem on $\mathbb{D}^2$, 13. Collections of kernels, 14. Model theory II: Function spaces, 15. Localization.

Jena H. Triebel

Mathematik in Beispiel, Theorie und Anwendung



Ernst-Albrecht Reinsch
Mathematik für Chemiker
Methoden, Beispiele, Anwendungen und Aufgaben
2004. 536 S. Br. EUR 34,90
ISBN 3-519-00443-7

Inhalt

Zahlen - Lineare Algebra - Funktionen - Differentialrechnung - Integralrechnung - Reihenentwicklung - Entwicklung nach Funktionensystemen, Fouriertransformation - Differentialgleichungen - Gruppentheorie - Fehler- und Ausgleichsrechnung

Das Buch

Wie kann man chemische Probleme mathematisch beschreiben und lösen? In Theorie, Anwendung und Beispiel zeigt dieses Buch dem Studierenden die erforderlichen mathematischen Methoden praxisbezogen und verständlich. Viele Übungsaufgaben machen fit für Prüfung und Praxis.

Teubner

Abraham-Lincoln-Str. 46
65189 Wiesbaden
Fax 0611.7878-420
www.teubner.de

ISSN 0012-0456