Appendix A
Auxiliary Material

A.1 Orlicz Norms

We frequently use Orlicz norms $\|\cdot\|_\psi$ of random variables. Given a convex nondecreasing function $\psi : \mathbb{R}_+ \mapsto \mathbb{R}_+$ with $\psi(0) = 0$ and a random variable η on a probability space $(\Omega, \Sigma, \mathbb{P})$, define

$$\|\eta\|_\psi := \inf\left\{C > 0 : \mathbb{E}\psi\left(\frac{|\eta|}{C}\right) \le 1\right\}$$

(see Ledoux and Talagrand [101], van der Vaart and Wellner [148], de la Pena and Giné [50]). If we want to emphasize the dependence of the Orlicz norms on the probability measure, we write $\|\cdot\|_{L_\psi(\mathbb{P})}$ (similarly, $\|\cdot\|_{L_\psi(P)}$, $\|\cdot\|_{L_\psi(\Pi)}$, etc).

If $\psi(u) = u^p$ for some $p \ge 1$, then the ψ-norm coincides with the usual L_p-norm. Some other useful choices of function ψ correspond to Orlicz norms in spaces of random variables with subgaussian or subexponential tails. For $\alpha > 0$, define

$$\psi_\alpha(u) := e^{u^\alpha} - 1,\ u \ge 0.$$

Most often, ψ_2- and ψ_1-norms are used (the first one being the "subgaussian norm" and the second one being the "subexponential norm"). Note that, for $\alpha < 1$, the function ψ_α is not convex and, as a result, $\|\cdot\|_{\psi_\alpha}$ is not a norm. However, to overcome this difficulty, it is enough to modify ψ_α in a neighborhood of 0. As it is common in the literature, we ignore this minor inconvenience and use $\|\cdot\|_{\psi_\alpha}$ as a norm even for $\alpha < 1$. Moreover, usually, we need the ψ_α-norms for $\alpha \ge 1$. The following bounds are well known (see [148], p. 95):

$$\|\eta\|_{\psi_{\alpha_1}} \le (\log 2)^{\alpha_1/\alpha_2}\|\eta\|_{\psi_{\alpha_2}},\ 1 \le \alpha_1 \le \alpha_2$$

and, for all $p \in (m-1, m]$, $m = 2, 3, \ldots$ $\|\eta\|_{L_p} \le m!\|\eta\|_{\psi_1}$.

It easily follows from the definition of ψ_α-norms that

V. Koltchinskii, *Oracle Inequalities in Empirical Risk Minimization and Sparse Recovery Problems*, Lecture Notes in Mathematics 2033, DOI 10.1007/978-3-642-22147-7,

$$\mathbb{P}\{|\eta| \geq t\} \leq 2\exp\left\{-\left(\frac{t}{\|\eta\|_{\psi_\alpha}}\right)^\alpha\right\}.$$

Another well known fact is that, for many convex nondecreasing functions ψ, including $\psi(u) = u^p$ with $p \geq 1$ and ψ_α with $\alpha \geq 1$, for all $N \geq 1$ and for all random variables $\eta_1, \ldots, \eta_N$

$$\left\|\max_{1\leq k\leq N} \eta_k\right\|_\psi \leq K \max_{1\leq k\leq N} \|\eta_k\|_\psi \psi^{-1}(N),$$

where K is a constant depending on ψ (see, e.g., [148], Lemma 2.2.2).

A.2 Classical Exponential Inequalities

Let $X_1, \ldots, X_n$ be independent random variables with $\mathbb{E}X_j = 0,\ j = 1, \ldots, n$. We state below several classical exponential bounds for the sum

$$S_n := X_1 + \cdots + X_n.$$

Denote $B_n^2 := \mathbb{E}X_1^2 + \cdots + \mathbb{E}X_n^2$.

- *Bernstein's inequality*. Suppose $|X_j| \leq U$, $j = 1, \ldots, n$. Then,

$$\mathbb{P}\{S_n \geq t\} \leq \exp\left\{-\frac{t^2}{2B_n^2\left(1 + \frac{tU}{3B_n^2}\right)}\right\}.$$

- *Bennett's inequality*. Suppose $|X_j| \leq U$, $j = 1, \ldots, n$. Then,

$$\mathbb{P}\{S_n \geq t\} \leq \exp\left\{-\frac{B_n^2}{U^2} h\left(\frac{tU}{B_n^2}\right)\right\},$$

  where $h(u) := (1 + u)\log(1 + u) - u$.
- *Hoeffding's inequality*. Suppose $a_j < b_j, j = 1, \ldots, n, \quad X_j \in [a_j, b_j]$, $\mathbb{E}X_j = 0$, $j = 1, \ldots, n$. Then,

$$\mathbb{P}\left\{S_n \geq t\right\} \leq \exp\left\{-\frac{2t^2}{\sum_{j=1}^n (b_j - a_j)^2}\right\},\ t \geq 0.$$

- *Bernstein's type inequality for ψ_1-random variables*. Suppose $\|X_j\|_{\psi_1} \leq V$. Then,

$$\mathbb{P}\{S_n \geq t\} \leq \exp\left\{-c\left(\frac{t^2}{nV^2} \bigwedge \frac{t}{V}\right)\right\}$$

with some universal constant $c > 0$.

Bernstein's inequality easily implies that, for all $t > 0$, with probability at least $1 - e^{-t}$

$$|S_n| \le C(B_n \sqrt{t} \vee Ut),$$

where C is a numerical constant. We frequently use this form of Bernstein's inequality and other inequalities of similar type.

A.3 Properties of ♯- and ♭-Transforms

Here we provide some properties of ♯- and ♭-transforms introduced in Sect. 4.1 and used in the construction of excess risk bounds. The proofs of these properties are rather elementary. We are mainly interested in ♯-transform.

1. If $\psi(u) = o(u)$ as $u \to \infty$, then the function $\psi^\sharp$ is defined on $(0, +\infty)$ and is a nonincreasing function on this interval.
2. If $\psi_1 \le \psi_2$, then $\psi_1^\sharp \le \psi_2^\sharp$. Moreover, if $\psi_1(\delta) \le \psi_2(\delta)$ either for all $\delta \ge \psi_2^\sharp(\varepsilon)$, or for all $\delta \ge \psi_1^\sharp(\varepsilon) - \tau$ with an arbitrary $\tau > 0$, then also $\psi_1^\sharp(\varepsilon) \le \psi_2^\sharp(\varepsilon)$.
3. For all $a > 0$,
$$(a\psi)^\sharp(\varepsilon) = \psi^\sharp(\varepsilon/a).$$
4. If $\varepsilon = \varepsilon_1 + \cdots + \varepsilon_m$, then
$$\psi_1^\sharp(\varepsilon) \bigvee \cdots \bigvee \psi_m^\sharp(\varepsilon) \le (\psi_1 + \cdots + \psi_m)^\sharp(\varepsilon) \le \psi_1^\sharp(\varepsilon_1) \bigvee \cdots \bigvee \psi_m^\sharp(\varepsilon_m).$$
5. If $\psi(u) \equiv c$, then
$$\psi^\sharp(\varepsilon) = c/\varepsilon.$$
6. If $\psi(u) := u^\alpha$ with $\alpha \le 1$, then
$$\psi^\sharp(\varepsilon) := \varepsilon^{-1/(1-\alpha)}.$$
7. For $c > 0$, denote $\psi_c(\delta) := \psi(c\delta)$. Then
$$\psi_c^\sharp(\varepsilon) = \frac{1}{c}\psi^\sharp(\varepsilon/c).$$
If ψ is nondecreasing and $c \ge 1$, then
$$c\psi^\sharp(u) \le \psi^\sharp(u/c).$$
8. For $c > 0$, denote $\psi_c(\delta) := \psi(\delta + c)$. Then for all $u > 0, \varepsilon \in (0, 1]$
$$\psi_c^\sharp(u) \le (\psi^\sharp(\varepsilon u/2) - c) \vee c\varepsilon.$$

Recall the definitions of functions of concave type and strictly concave type from Sect. 4.1.

9. If ψ is of concave type, then $\psi^\sharp$ is the inverse of the function

$$\delta \mapsto \frac{\psi(\delta)}{\delta}.$$

In this case,

$$\psi^\sharp(cu) \geq \psi^\sharp(u)/c$$

for $c \leq 1$ and

$$\psi^\sharp(cu) \leq \psi^\sharp(u)/c$$

for $c \geq 1$.

10. If ψ is of strictly concave type with exponent γ, then for $c \leq 1$

$$\psi^\sharp(cu) \leq \psi^\sharp(u)c^{-\frac{1}{1-\gamma}}.$$

A.4 Some Notations and Facts in Linear Algebra

Let L be a linear space. The following notations are frequently used: l.s.(B) for a linear span of a subset $B \subset L$,

$$\text{l.s.}(B) := \left\{ \sum_{j=1}^{n} \lambda_j x_j : n \geq 1, \lambda_j \in \mathbb{R}, x_j \in B \right\};$$

conv(B) for its convex hull,

$$\text{conv}(B) := \left\{ \sum_{j=1}^{n} \lambda_j x_j : n \geq 1, \lambda_j \geq 0, \sum_{j=1}^{n} \lambda_j = 1, x_j \in B \right\};$$

and conv$_s(B)$ for its symmetric convex hull,

$$\text{conv}_s(B) := \left\{ \sum_{j=1}^{n} \lambda_j x_j : n \geq 1, \lambda_j \in \mathbb{R}, \sum_{j=1}^{n} |\lambda_j| \leq 1, x_j \in B \right\}.$$

For vectors $u, v \in \mathbb{C}^m$ or $u, v \in \mathbb{R}^m$, $\langle u, v \rangle$ denotes the standard Euclidean inner product of u and v; $|u|$ denotes the corresponding norm of u. Notations $\|u\|_{\ell_2}$ or $\|u\|_{\ell_2^m}$ are also used for the same purpose.

For vectors u, v in $\mathbb{C}^m$ (or other real and complex Euclidean spaces), $u \otimes v$ denotes their tensor product, that is, the linear transformation defined by

$$(u \otimes v)x = \langle v, x\rangle u.$$

Given a subspace $L \subset \mathbb{C}^m$ (more generally, a subspace of any Euclidean space), P_L denotes the orthogonal projection onto L and $L^{\perp}$ denotes the orthogonal complement of L.

We use the notations $\mathbb{M}_{m_1,m_2}(\mathbb{R})$ and $\mathbb{M}_{m_1,m_2}(\mathbb{C})$ for the spaces of all $m_1 \times m_2$ matrices with real or complex entries, respectively. In the case when $m_1 = m_2 = m$, we use the notations $\mathbb{M}_m(\mathbb{R})$ and $\mathbb{M}_m(\mathbb{C})$. The space of all Hermitian $m \times m$ matrices is denoted by $\mathbb{H}_m(\mathbb{C})$. For $A, B \in \mathbb{H}_m(\mathbb{C})$, the notation $A \leq B$ means that $B - A$ is nonnegatively definite.

We denote by $\mathrm{rank}(A)$ the rank of a matrix A and by $\mathrm{tr}(A)$ the trace of a square matrix A. Given $A \in \mathbb{M}_{m_1,m_2}(\mathbb{C})$, A^* denotes its adjoint matrix. We use the notations $\langle \cdot, \cdot \rangle$ for the Hilbert–Schmidt inner product of two matrices of the same size,

$$\langle A, B\rangle = \mathrm{tr}(AB^*),$$

$\|\cdot\|$ for the operator norm of matrices and $\|\cdot\|_p,\ p \geq 1$ for their *Schatten p-norm*:

$$\|A\|_p := \Big(\sum_k \sigma_k^p(A)\Big)^{1/p},$$

where $\{\sigma_k(A)\}$ denote the singular values of the matrix A (usually, arranged in a nonincreasing order). In particular, $\|\cdot\|_2$ is *the Hilbert–Schmidt* or *Frobenius norm* and $\|\cdot\|_1$ is *the nuclear norm*. The notation $\|\cdot\|$ is reserved for the operator norm. Given a probability distribution Π in $\mathbb{H}_m(\mathbb{C})$, we also associate with a matrix $B \in \mathbb{H}_m(\mathbb{C})$ the linear functional $\langle B, \cdot\rangle$ and define the $L_2(\Pi)$ norm of B as the $L_2(\Pi)$-norm of this functional:

$$\|B\|^2_{L_2(\Pi)} := \int_{\mathbb{H}_m(\mathbb{C})} \langle B, x\rangle^2 \Pi(dx).$$

We use the corresponding inner product $\langle \cdot, \cdot \rangle_{L_2(\Pi)}$ in the space of matrices.

For a matrix $S \in \mathbb{H}_m(\mathbb{C})$ of rank r with spectral decomposition

$$S = \sum_{j=1}^r \lambda_j (e_j \otimes e_j),$$

where $e_1, \dots, e_r$ are the eigenvectors corresponding to the non-zero eigenvalues $\lambda_1, \dots, \lambda_r$, define *the support* of S as $\mathrm{supp}(S) := \mathrm{l.s.}(\{e_1, \dots, e_r\})$. Also, define

$$|S| := \sqrt{S^2} = \sum_{j=1}^r |\lambda_j|(e_j \otimes e_j)$$

and

$$\mathrm{sign}(S) := \sum_{j=1}^{r} \mathrm{sign}(\lambda_j)(e_j \otimes e_j).$$

It is well known that the subdifferential of the convex function $\mathbb{H}_m(\mathbb{C}) \ni S \mapsto \|S\|_1$ has the following representation (see, e.g., [151]):

$$\partial \|S\|_1 = \Big\{ \mathrm{sign}(S) + P_{L^\perp} W P_{L^\perp} : \|W\| \le 1 \Big\},$$

where $L := \mathrm{supp}(S)$.

Some other facts of linear algebra used in Chap. 9 can be found in [21].

References

1. Adamczak, R. (2008) A tail inequality for suprema of unbounded empirical processes with applications to Markov chains. *Electronic Journal of Probability*, 13, 34, 1000-1034.
2. Adamczak, R., Litvak, A.E., Pajor, A. and Tomczak-Jaegermann, N. (2011) Restricted isometry property of matrices with independent columns and neighborly polytopes by random sampling. *Constructive Approximation*, to appear.
3. Affentranger, F. and Schneider, R. (1992) Random Projections of Regular Simplices. *Discrete Comput. Geom.*, 7(3), 219–226.
4. Ahlswede, R. and Winter, A. (2002) Strong converse for identification via quantum channels. *IEEE Transactions on Information Theory*, 48, 3, pp. 569–679.
5. Alexander, K.S. (1987) Rates of growth and sample moduli for weighted empirical processes indexed by sets. *Probability Theory and Related Fields*, 75, 379–423.
6. Anthony, M. and Bartlett, P. (1999) Neural Network Learning: Theoretical Foundations. Cambridge University Press.
7. Arlot, S. and Bartlett, P. (2011) Margin adaptive model selection in statistical learning. *Bernoulli*, 17, 2, 687–713.
8. Arlot, S. and Massart, P. (2009) Data-driven calibration of penalties for least-squares regression. *J. Machine Learning Research*, 10, 245–279.
9. Aubin, J.-P. and Ekeland, I. (1984) Applied Nonlinear Analysis. J. Wiley & Sons, New York.
10. Audibert, J.-Y. (2004) Une approche PAC-bayésienne de la théorie statistique de l'apprentissage. PhD Thesis, University of Paris 6.
11. Baraud, Y. (2002) Model selection for regression on a random design. *ESAIM: Probability and Statistics*, 6, 127–146.
12. Barron, A., Birgé, L. and Massart, P. (1999) Risk Bounds for Model Selection via Penalization. *Probability Theory and Related Fields*, 113, 301–413.
13. Bartlett, P. (2008) Fast rates for estimation error and oracle inequalities for model selection. *Econometric Theory*, 24, 2, 545–552.
14. Bartlett, P., Boucheron, S. and Lugosi, G. (2002) Model selection and error estimation. *Machine Learning*, 48, 85–113.
15. Bartlett, P., Bousquet, O. and Mendelson, S. (2005) Local Rademacher Complexities, *Annals of Statistics*, 33, 4, 1497-1537.
16. Bartlett, P., Jordan, M. and McAuliffe, J. (2006) Convexity, Classification and Risk Bounds. *Journal of American Statistical Association*, 101(473), 138–156.
17. Bartlett, P. and Mendelson, S. (2006) Empirical Minimization. *Probability Theory and Related Fields*, 135(3), 311–334.
18. Bartlett, P., Mendelson, S. and Neeman, J. (2010) ℓ_1-regularized linear regression: persistence and oracle inequalities. *Probability Theory and Related Fields*, to appear.

V. Koltchinskii, *Oracle Inequalities in Empirical Risk Minimization and Sparse Recovery Problems*, Lecture Notes in Mathematics 2033, DOI 10.1007/978-3-642-22147-7,

19. Bartlett, P. and Williamson, R.C. (1996) Efficient agnostic learning of neural networks with bounded fan-in. *IEEE Transactions on Information Theory*, 42(6), 2118–2132.
20. Ben-Tal, A. and Nemirovski, A. (2001) Lectures on Modern Convex Optimization. Analysis, Algorithms and Engineering Applications. MPS/SIAM Series on Optimization, Philadelphia.
21. Bhatia, R. (1997) Matrix Analysis. Springer, New York.
22. Bickel, P., Ritov, Y. and Tsybakov, A. (2009) Simultaneous Analysis of LASSO and Dantzig Selector. *Annals of Statistics*, 37, 4, 1705–1732.
23. Birgé, L. and Massart, P. (1997) From Model Selection to Adaptive Estimation. In: Festschrift for L. Le Cam. Research Papers in Probability and Statistics. D. Pollard, E. Torgersen and G. Yang (Eds.), 55-87. Springer, New York.
24. Birgé, L. and Massart, P. (2007) Minimal Penalties for Gaussian Model Selection. *Probability Theory and Related Fields*.138, 33–73.
25. Blanchard, G., Bousquet, O. and Massart, P. (2008) Statistical performance of support vector machines. *Annals of Statistics*, 36, 2, 489–531.
26. Blanchard, G., Lugosi, G. and Vayatis, N. (2003) On the rate of convergence of regularized boosting classifiers. *Journal of Machine Learning Research*, 4, 861–894.
27. Bobkov, S. and Houdré, C. (1997) Isoperimetric constants for product probability measures. *Annals of Probability*, v. 25, no. 1, 184–205.
28. Borell, C. (1974) Convex measures on locally convex spaces. *Arkiv för Matematik*, 12, 239–252.
29. Boucheron, S., Bousquet, O. and Lugosi, G. (2005) Theory of Classification. A Survey of Some Recent Advances. *ESAIM Probability & Statistics*, 9, 323–371.
30. Boucheron, S., Bousquet, O., Lugosi, G. and Massart, P. (2005) Moment inequalities for functions of independent random variables. *Annals of Probability*, 33, 2, 514–560.
31. Boucheron, S., Lugosi, G. and Massart, P. (2000) A sharp concentration inequality with applications. *Random Structures and Algorithms*, 16, 277–292.
32. Boucheron, S. and Massart, P. (2010) A high dimensional Wilks phenomenon. *Probability Theory and Related Fields*, to appear.
33. Bousquet, O. (2002) A Bennett concentration inequality and its application to suprema of empirical processes. *C.R. Acad. Sci. Paris*, 334, 495–500.
34. Bousquet, O., Koltchinskii, V. and Panchenko, D. (2002) Some local measures of complexity of convex hulls and generalization bounds. In: COLT2002, Lecture Notes in Artificial Intelligence, 2375, Springer, 59 - 73.
35. Bunea, F., Tsybakov, A. and Wegkamp, M. (2007) Aggregation for Gaussian regression, *Annals of Statistics*, 35,4, 1674–1697.
36. Bunea, F., Tsybakov, A. and Wegkamp, M. (2007) Sparsity oracle inequalities for the LASSO. *Electronic Journal of Statistics*, 1, 169–194.
37. Bunea, F., Tsybakov, A. and Wegkamp, M. (2007) Sparse density estimation with ℓ_1 penalties. In: *Proc. 20th Annual Conference on Learning Theory (COLT 2007)*, Lecture Notes in Artificial Intelligence, Springer, v. 4539, pp. 530–543.
38. Candes, E. (2008) The restricted isometry property and its implications for compressed sensing, *C. R. Acad. Sci. Paris, Ser. I*, 346, 589-592.
39. Candes, E. and Plan, Y. (2009) Near-ideal model selection by ℓ_1-minimization. *Annals of Statistics*, 37, 5A, 2145-2177.
40. Candes, E. and Plan, Y. (2009) Tight Oracle Bounds for Low-Rank Matrix Recovery from a Minimal Number of Random Measurements. *IEEE Transactions on Information Theory*, to appear.
41. Candes, E. and Recht, B. (2009) Exact matrix completion via convex optimization. *Foundations of Computational Mathematics*, 9(6), 717–772.
42. Candes, E., Romberg, J. and Tao, T. (2006) Stable Signal Recovery from Incomplete and Inaccurate Measurements. *Communications on Pure and Applied Mathematics*, 59, 1207–1223.
43. Candes, E., Rudelson, M., Tao, T. and Vershynin, R. (2005) Error Correction via Linear Programming. *Proc. 46th Annual IEEE Symposium on Foundations of Computer Science (FOCS05)*, IEEE, 295–308.

44. Candes, E. and Tao, T. (2007) The Dantzig selector: statistical estimation when p is much larger than n. *Annals of Statistics*, 35,6, 2313–2351.
45. Candes, E. and Tao, T. (2010) The power of convex relaxation: Near-optimal matrix completion. *IEEE Transactions on Information Theory*, 56, 2053–2080.
46. Catoni, O. (2004) Statistical Learning Theory and Stochastic Optimization. *Ecole d'Eté de Probabilités de Saint-Flour XXXI -2001*, Lecture Notes in Mathematics, **1851**, Springer, New York.
47. Cucker, F. and Smale, S. (2001) On the mathematical foundations of learning. *Bull. Amer. Math. Soc.*, 39, 1–49.
48. Dalalyan, A. and Tsybakov, A. (2008) Aggregation by exponential weighting, sharp PAC-Bayesian bounds and sparsity. *Machine Learning*, 72, 39–61.
49. DeVore, R., Kerkyacharian, G., Picard, D. and Temlyakov, V. (2006) Approximating Methods for Supervised Learning. *Foundations of Computational Mathematics*, 6, 3–58.
50. de la Pena, V. and Giné, E. (1998) Decoupling: From Dependence to Independence, Springer, New York.
51. Devroye, L., Györfi, G. and Lugosi, G. (1996) A Probabilistic Theory of Pattern Recognition. Springer, New York.
52. Donoho, D.L. (2004) Neighborly Polytopes and Sparse Solution of Underdetermined Linear Equations. Technical Report, Stanford.
53. Donoho, D.L. (2006) For Most Large Underdetermined Systems of Equations the Minimal ℓ^1-norm Near-Solution Approximates the Sparsest Near-Solution. *Communications on Pure and Applied Mathematics*, 59, 7, 907–934.
54. Donoho, D.L. (2006) For Most Large Underdetermined Systems of Linear Equations the Minimal ℓ^1-norm Solution is also the Sparsest Solution. *Communications on Pure and Applied Mathematics*, 59, 797–829.
55. Donoho, D.L. (2006) Compressed Sensing. *IEEE Transactions on Information Theory*, 52, 4, 1289–1306.
56. Donoho, D.L., Elad, M. and Temlyakov, V. (2006) Stable Recovery of Sparse Overcomplete Representations in the Presence of Noise. *IEEE Transactions on Information Theory*, 52, 1, 6–18.
57. Donoho, D.L. and Tanner, J. (2005) Neighborliness of Randomly-Projected Simplices in High Dimensions. *Proc. National Academy of Sciences*, 102(27), 9446–9451.
58. Dudley, R.M. (1978) Central Limit Theorems for Empirical Measures, *Annals of Probability*, 6,6, 899–929.
59. Dudley, R.M. (1999) Uniform Central Limit Theorems. Cambridge University Press.
60. Einmahl, U. and Mason, D. (2000) An empirical processes approach to the uniform consistency of kernel type function estimators. *J. Theoretical Probability*, 13, 1–37.
61. Fromont, M. (2007) Model selection by bootstrap penalization for classification. *Machine Learning*, 66, 2-3, 165–207.
62. van de Geer, S. (1999) Empirical Processes in M-estimation. Cambridge University Press. Cambridge.
63. van de Geer, S. (2008) High-dimensional generalized linear models and the Lasso, *Annals of Statistics*, 36, 2, 614–645.
64. Giné, E. and Guillou, A. (2001) A law of the iterated logarithm for kernel density estimators in the precense of censoring. *Annales Inst. H. Poincaré (B), Probabilités et Statistiques*, 37, 503–522.
65. Giné, E., Koltchinskii, V. and Wellner, J. (2003) Ratio Limit Theorems for Empirical Processes. In: *Stochastic Inequalities*, Birkhäuser, pp. 249–278.
66. Giné, E. and Koltchinskii, V. (2006) Concentration Inequalities and Asymptotic Results for Ratio Type Empirical Processes. *Annals of Probability*, 34, 3, 1143-1216.
67. Giné, E. and Nickl, R. (2010) Adaptive estimation of a distribution function and its density in sup-norm loss by wavelet and spline projections. *Bernoulli*, 16, 4, 1137–1163.
68. Giné, E. and Mason, D. (2007) On local U-statistic processes and the estimation of densities of functions of several variables. *Annals of Statistics*, 35, 1105-1145.

69. Giné, E. and Zinn, J. (1984) Some limit theorems for empirical processes. *Annals of Probability* 12, 929-989.
70. Gross, D. (2011) Recovering Low-Rank Matrices From Few Coefficients in Any Basis. *IEEE Transactions on Information Theory*, 57, 3, 1548–1566.
71. Gross, D., Liu, Y.-K., Flammia, S.T., Becker, S. and Eisert, J. (2010) Quantum state tomography via compressed sensing. *Phys. Rev. Lett.*, 105(15):150401, October 2010.
72. Györfi, L., Kohler, M., Krzyzak, A. and Walk, H. (2002) A Distribution-Free Theory of Nonparametric Regression. Springer.
73. Hanneke, S. (2011) Rates of convergence in active learning, *Annals of Statistics*, 39, 1, 333–361.
74. Johnstone, I.M. (1998) Oracle Inequalities and Nonparametric Function Estimation. In: Documenta Mathematica, *Journal der Deutschen Mathematiker Vereinigung*, Proc. of the International Congress of Mathematicians, Berlin, 1998, v.III, 267–278.
75. Keshavan, R.H., Montanari, A. and Oh, S. (2010) Matrix completion from a few entries. *IEEE Transactions on Information Theory*, 56, 2980–2998.
76. Klartag, B. and Mendelson, S. (2005) Empirical Processes and Random Projections. *Journal of Functional Analysis*, 225(1), 229–245.
77. Klein, T. (2002) Une inégalité de concentration à gauche pour les processus empiriques. *C.R. Acad. Sci. Paris*, Ser I, 334, 500–505.
78. Klein, T. and Rio, E. (2005) Concentration around the mean for maxima of empirical processes. *Annals of Probability*, 33,3, 1060–1077.
79. Kohler, M. (2000) Inequalities for uniform deviations of averages from expectations with applications to nonparametric regression. *J. of Statistical Planning and Inference*, 89, 1–23.
80. Koltchinskii, V. (1981) On the central limit theorem for empirical measures. *Theory of Probability and Mathematical Statistics*, 24, 71–82.
81. Koltchinskii, V. (2001) Rademacher penalties and structural risk minimization. *IEEE Transactions on Information Theory*, 47(5), 1902–1914.
82. Koltchinskii, V. (2005) Model selection and aggregation in sparse classification problems. *Oberwolfach Reports: Meeting on Statistical and Probabilistic Methods of Model Selection, October 2005*.
83. Koltchinskii, V. (2006) Local Rademacher Complexities and Oracle Inequalities in Risk Minimization. *Annals of Statistics*, 34, 6, 2593–2656.
84. Koltchinskii, V. (2009) Sparsity in penalized empirical risk minimization, *Annales Inst. H. Poincaré (B), Probabilités et Statistiques*, 45, 1, 7–57.
85. Koltchinskii, V. (2009) The Dantzig Selector and Sparsity Oracle Inequalities. *Bernoulli*, 15,3, 799-828.
86. Koltchinskii, V. (2009) Sparse Recovery in Convex Hulls via Entropy Penalization, *Annals of Statistics*, 37, 3, 1332-1359.
87. Koltchinskii, V. (2010) Rademacher Complexities and Bounding the Excess Risk in Active Learning, *Journal of Machine Learning Research*, 2010, 11, 2457–2485.
88. Koltchinskii, V. (2010) Von Neumann entropy penalization and low rank matrix estimation. Preprint.
89. Koltchinskii, V. (2010) A remark on low rank matrix recovery and noncommutative Bernstein type inequalities. Preprint.
90. Koltchinskii, V., Lounici, K. and Tsybakov, A. (2010) Nuclear norm penalization and optimal rates for noisy low rank matrix completion. *Annals of Statistics*, to appear.
91. Koltchinskii, V. and Minsker, S. (2010) Sparse recovery in convex hulls of infinite dictionaries. In: *COLT 2010, 23rd Conference on Learning Theory*, A. Kalai and M. Mohri (Eds), Haifa Israel, pp. 420–432.
92. Koltchinskii, V. and Panchenko, D. (2000) Rademacher processes and bounding the risk of function learning. In: Giné, E., Mason, D. and Wellner, J. *High Dimensional Probability II*, 443–459.
93. Koltchinskii, V. and Panchenko, D. (2002) Empirical margin distributions and bounding the generalization error of combined classifiers. *Annals of Statistics*, 30, 1, 1–50.

94. Koltchinskii, V., Panchenko, D. and Lozano, F. (2003) Bounding the generalization error of convex combinations of classifiers: balancing the dimensionality and the margins. *Ann. Appl. Probab.,* **13**, 1, 213–252.
95. Koltchinskii, V. and Panchenko, D. (2005) Complexities of convex combinations and bounding the generalization error in classification. *Annals of Statistics*, 33, 4, 1455–1496.
96. Koltchinskii, V. and Yuan, M. (2008) Sparse Recovery in Large Ensembles of Kernel Machines. In: *21st Annual Conference on Learning Theory COLT-2008*, Helsinki, 229–238, Omnipress.
97. Koltchinskii, V. and Yuan, M. (2010) Sparsity in Multiple Kernel Learning, *Annals of Statistics*, 38, 3660–3695.
98. Lecué, G. and Gaiffas, S. (2010) Sharp oracle inequalities for high-dimensional matrix prediction. *IEEE Transactions on Information Theory*, to appear.
99. Lecué, G. and Mendelson, S. (2009) Aggregation via empirical risk minimization. *Probability Theory and Related Fields*, 145(3–4), 591–613.
100. Ledoux, M. (2001) The Concentration of Measure Phenomenon. Mathematical Surveys and Monographs, Vol. 89, American Mathematical Society.
101. Ledoux, M. and Talagrand, M. (1991) Probability in Banach Spaces. Springer-Verlag, New York.
102. Lieb, E.H. (1973) Convex trace functions and the Wigner-Yanase-Dyson conjecture. *Adv. Math.*, 11, 267–288.
103. Lugosi, G. and Vayatis, N. (2004) On the Bayes-risk consistency of regularized boosting methods. *Annals of Statistics*, 32, 1, 30–55.
104. Lugosi, G. and Wegkamp, M. (2004) Complexity regularization via localized random penalties. *Annals of Statistics*, 32, 4, 1679–1697.
105. Mammen, E. and Tsybakov, A. (1999) Smooth discrimination analysis. *Annals of Statistics* 27, 1808–1829.
106. Massart, P. (2000) Some applications of concentration inequalities to statistics. *Ann. Fac. Sci. Toulouse (IX)*, 245–303.
107. Massart, P. (2007) Concentration Inequalities and Model Selection. *Ecole d'ete de Probabilités de Saint-Flour 2003*, Lecture Notes in Mathematics, Springer.
108. Massart, P. and Nedelec, E. (2006) Risk bounds for statistical learning. *Annals of Statistics*, 34, 5, 2326–2366.
109. Massart, P. and Meynet, C. (2010) An ℓ_1-Oracle Inequality for the Lasso. *Electronic Journal of Statistics*, to appear.
110. McAllester, D.A. (1998) Some PAC-Bayesian theorems. In *Proc. 11th Annual Conference on Learning Theory*, pp. 230–234, ACM Press.
111. Meier, L., van de Geer, S. and Bühlmann, P. (2009) High-dimensional additive modeling, *Annals of Statistics*, 37, 3779–3821.
112. Mendelson, S. (2002) Improving the sample complexity using global data. *IEEE Transactions on Information Theory*, 48, 1977–1991.
113. Mendelson, S. (2002) Geometric parameters of kernel machines. In: COLT 2002, Lecture Notes in Artificial Intelligence, 2375, Springer, 29–43.
114. Mendelson, S. (2008) Lower bounds for the empirical minimization algorithm. *IEEE Transactions on Information Theory*, 54, 8, 3797–3803.
115. Mendelson, S. (2010) Empirical processes with a bounded ψ_1 diameter. *Geometric and Functional Analysis*, 20,4, 988–1027.
116. Mendelson, S. and Neeman, J. (2010) Regularization in kernel learning. *Annals of Statistics*, 38, 1, 526–565.
117. Mendelson, S., Pajor, A. and Tomczak-Jaegermann, N. (2007) Reconstruction and subgaussian operators in Asymptotic Geometric Analysis. *Geometric and Functional Analysis*, 17(4), 1248–1282.
118. Negahban, S. and Wainwright, M.J. (2010) Restricted strong convexity and weighted matrix completion: optimal bounds with noise. Preprint.

119. Nemirovski, A. (2000) Topics in non-parametric statistics, In: P. Bernard, editor, *Ecole d'Et'e de Probabilités de Saint-Flour XXVIII*, 1998, Lecture Notes in Mathematics, Springer, New York.
120. Nielsen, M.A. and Chuang, I.L. (2000) Quantum Computation and Quantum Information, Cambridge University Press.
121. Pajor, A. and Tomczak-Jaegermann, N. (1985) Remarques sur les nombres d'entropie d'un opérateur et de son transposé. *C.R. Acad. Sci. Paris Sér. I Math.*, 301(15), 743–746.
122. Pollard, D. (1982) A central limit theorem for empirical processes. *Journal of the Australian Mathematical Society*, A, 33, 235–248.
123. Pollard, D.(1984) Convergence of Stochastic Processes, Springer, New York.
124. Recht, B. (2009) A Simpler Approach to Matrix Completion. *Journal of Machine Learning Research*, to appear.
125. Recht, B., Fazel, M. and Parrilo, P. (2010) Guaranteed minimum rank solutions of matrix equations via nuclear norm minimization. *SIAM Review*, 52, 3, 471–501.
126. Rigollet, P. and Tsybakov, A. (2011) Exponential screening and optimal rates of sparse estimation. *Annals of Statistics*, 39, 731–771.
127. Rohde, A. and Tsybakov, A. (2011) Estimation of high-dimensional low rank matrices. *Annals of Statistics*, 39, 2, 887–930.
128. Rudelson, M. (1999) Random vectors in the isotropic position. *J. Funct. Analysis*, 164(1), 60–72.
129. Rudelson, M. and Vershynin, R. (2005) Geometric Approach to Error Correcting Codes and Reconstruction of Signals. *Int. Math. Res. Not.*, no 64, 4019-4041.
130. Rudelson, M. and Vershynin, R. (2010) Non-asymptotic theory of random matrices: extreme singular values. *Proceedings of the International Congress of Mathematicians*, Hyderabad, India.
131. Saumard, A. (2010) Estimation par Minimum de Contraste Régulier et Heuristique de Pente en Sélection de Modéles. PhD Thesis, Université de Rennes 1.
132. Shen, X. and Wong, W.H. (1994) Convergence rate of sieve estimates. *Annals of Statistics*, 22, 580-615.
133. Simon, B. (1979) Trace Ideals and their Applications. Cambridge University Press.
134. Steinwart, I. (2005) Consistency of support vector machines and other regularized kernel machines. *IEEE Transactions on Information Theory*, 51, 128–142.
135. Steinwart, I. and Christmann, A. (2008) Support Vector Machines. Springer.
136. Steinwart, I. and Scovel, C. (2007) Fast rates for support vector machines using Gaussian kernels. *Annals of Statistics*, 35, 575–607.
137. Talagrand, M. (1994) Sharper bounds for Gaussian and empirical processes. *Annals of Probability,* 22, 28-76.
138. Talagrand, M. (1996) A new look at independence. *Annals of Probability,* 24, 1-34.
139. Talagrand, M. (1996) New concentration inequalities in product spaces. *Invent. Math.,* 126, 505-563.
140. Talagrand, M. (2005) The Generic Chaining. Springer.
141. Tibshirani, R. (1996) Regression shrinkage and selection via Lasso, *J. Royal Statist. Soc., Ser B*, **58**, 267–288.
142. Tropp, J.A. (2010) User-friendly tail bounds for sums of random matrices. *Foundations of Computational Mathematics*, to appear.
143. Tsybakov, A. (2003) Optimal rates of aggregation. In: *Proc. 16th Annual Conference on Learning Theory (COLT) and 7th Annual Workshop on Kernel Machines*, Lecture Notes in Artificial Intelligence, **2777**, Springer, New York, 303–313.
144. Tsybakov, A. (2004) Optimal aggregation of classifiers in statistical learning. *Annals of Statistics*, 32, 135–166.
145. Tsybakov, A. and van de Geer, S. (2005) Square root penalty: adaptation to the margin in classification and in the edge estimation. *Annals of Statistics*, 33, 3, 1203–1224.
146. Vapnik, V. (1998) Statistical Learning Theory. John Wiley & Sons, New York.

147. Vapnik, V. and Chervonenkis, A. (1974) Theory of Pattern Recognition. Nauka, Moscow (in Russian).
148. van der Vaart, A.W. and Wellner, J.A. (1996) Weak Convergence and Empirical Processes. With Applications to Statistics. Springer-Verlag, New York.
149. van der Vaart, A.W. and Wellner, J.A. (2011) A Local Maximal Inequality under Uniform Entropy. *Electronic Journal of Statistics*, 5, 192–203.
150. Vershik, A.M. and Sporyshev, P.V. (1992) Asymptotic behavior of the number of faces of random polyhedra and the neighborliness problem. *Selecta Math. Soviet.*, 11(2), 181–201.
151. Watson, G. A. (1992) Characterization of the subdifferential of some matrix norms. *Linear Algebra Appl.*, 170, 33-45.
152. Yang, Y. (2000) Mixing strategies for density estimation. *Annals of Statistics*, **28**, 75–87.
153. Yang, Y. (2004) Aggregating regression procedures for a better performance. *Bernoulli*, **10**, 25–47.
154. Zhang, T. (2001) Regularized Winnow Method. In: *Advances in Neural Information Processing Systems 13 (NIPS2000)*, T.K. Leen, T.G. Dietrich and V. Tresp (Eds), MIT Press, pp. 703–709.
155. Zhang, T. (2004) Statistical behavior and consistency of classification methods based on convex risk minimization. *Annals of Statistics*, 32, 1, 56–134.
156. Zhang, T. (2006) From epsilon-entropy to KL-complexity: analysis of minimum information complexity density estimation. *Annals of Statistics*, 34, 2180–2210.
157. Zhang, T. (2006) Information Theoretical Upper and Lower Bounds for Statistical Estimation. *IEEE Transactions on Information Theory*, 52, 1307–1321.

Index

V. Koltchinskii, *Oracle Inequalities in Empirical Risk Minimization and Sparse Recovery Problems*, Lecture Notes in Mathematics 2033, DOI 10.1007/978-3-642-22147-7,

Programme of the school

Main lectures

Richard Kenyon	Lectures on dimers
Vladimir Koltchinskii	Oracle inequalities in empirical risk minimization and sparse recovery problems
Yves Le Jan	Markov paths, loops and fields

Short lectures

Michael Allman	Breaking the chain
Pierre Alquier	Lasso, iterative feature selection and other regression methods satisfying the “Dantzig constraint”
Jürgen Angst	Brownian motion and Lorentzian manifolds, the case of Robertson-Walker space-times
Witold Bednorz	Some comments on the Bernoulli conjecture
Charles Bordenave	Spectrum of large random graphs
Cédric Boutillier	The critical Ising model on isoradial graphs via dimers
Robert Cope	Modelling in birth-death and quasi-birth-death processes
Irmina Czarna	Two-dimensional dividend problems
Michel Émery	Geometric structure of Azéma martingales
Christophe Gomez	Time reversal of waves in random waveguides: a super resolution effect
Nastasiya Grinberg	Semimartingale decomposition of convex functions of continuous semimartingales
François d’Hautefeuille	Entropy and finance
Erwan Hillion	Ricci curvature bounds on graphs
Wilfried Huss	Internal diffusion limited aggregation and related growth models
Jean-Paul Ibrahim	Large deviations for directed percolation on a thin rectangle
Liza Jones	Infinite systems of non-colliding processes
Andreas Lagerås	General branching processes conditioned on extinction

V. Koltchinskii, *Oracle Inequalities in Empirical Risk Minimization and Sparse Recovery Problems*, Lecture Notes in Mathematics 2033, DOI 10.1007/978-3-642-22147-7,

Benjamin Laquerrière	Conditioning of Markov processes and applications
Krzysztof Łatuszynski	"If and only if" conditions (in terms of regeneration) for $\sqrt{n}$-CLTs for ergodic Markov chains
Thierry Lévy	The Poisson process indexed by loops
Wei Liu	Spectral gap and convex concentration inequalities for birth-death processes
Gregorio Moreno	Directed polymers on a disordered hierarchical lattice
Jonathan Novak	Deforming the increasing subsequence problem
Ecaterina Sava	The Poisson boundary of lamplighter random walks
Bruno Schapira	Windings of the SRW on $\mathbb{Z}^2$ or triangular lattice and averaged Dehn function
Laurent Tournier	Random walks in Dirichlet environment
Nicolas Verzelen	Kullback oracle inequalities for covariance estimation
Vincent Vigon	LU-factorization and probability
Guillaume Voisin	Pruning a Lévy continuum random tree
Peter Windridge	Blocking and pushing interactions for discrete interlaced processes
Olivier Wintenberger	Some Bernstein's type inequalities for dependent data

List of participants

38th Probability Summer School, Saint-Flour, France

July 6–19, 2008

Lecturers

Richard KENYON	Brown University, Providence, USA
Vladimir KOLTCHINSKII	Georgia Inst. Technology, Atlanta, USA
Yves LE JAN	Université Paris-Sud, France

Participants

Michael ALLMAN	Warwick Univ., UK
Pierre ALQUIER	Univ. Paris Diderot, F
Jurgen ANGST	Univ. Strasbourg, F
Jean-Yves AUDIBERT	ENPC, Marne-la-Vallée, F
Witold BEDNORZ	Warsaw Univ., F
Charles BORDENAVE	Univ. Toulouse, F
Cédric BOUTILLIER	Univ. Pierre et Marie Curie, Paris, F
Pierre CONNAULT	Univ. Paris-Sud, Orsay, F
Robert COPE	Univ. Queensland, St Lucia, Australia
Irmina CZARNA	Univ. Wroclaw, Poland
Bassirou DIAGNE	Univ. Orléans, F
Roland DIEL	Univ. Orléans, F
Michel EMERY	Univ. Strasbourg, F
Mikael FALCONNET	Univ. Joseph Fourier, Grenoble, F
Benjamin FAVETTO	Univ. Paris Descartes, F
Jacques FRANCHI	Univ. Strasbourg, F
Laurent GOERGEN	ENS, Paris, F
Christophe GOMEZ	Univ. Paris Diderot, F
Nastasiya GRINBERG	Warwick Univ., UK
François HAUTEFEUILLE	London, UK
Erwan HILLION	Univ. Paul Sabatier, Toulouse, F
Wilfried HUSS	Graz Univ. Technology, Austria
Jean-Paul IBRAHIM	Univ. Paul Sabatier, Toulouse, F

V. Koltchinskii, *Oracle Inequalities in Empirical Risk Minimization and Sparse Recovery Problems*, Lecture Notes in Mathematics 2033, DOI 10.1007/978-3-642-22147-7,

Liza JONES	Univ. Oxford, UK
Adrien KASSEL	ENS, Paris, F
Andreas LAGERA S	Univ. Gothenburg & Chalmers Univ. Technology, Sweden
Benjamin LAQUERRIERE	Univ. La Rochelle, F
Krzysztof LATUSZYNSKI	Warsaw School Economics, Poland
Thierry LÉVY	ENS, Paris, F
Biao LI	Univ. Blaise Pascal, Clermont-Ferrand, F
Wei LIU	Univ. Blaise Pascal, Clermont-Ferrand, F
Karim LOUINICI	Univ. Paris Diderot, F
Sybille MICAN	Univ. München, Germany
Stanislav MINSKER	Georgia Inst. Technology, Atlanta, USA
Gregorio MORENO FLORES	Univ. Paris Diderot, F
Jonathan NOVAK	Queen's Univ., Kingston, Canada
Jean PICARD	Univ. Blaise Pascal, Clermont-Ferrand, F
Erwan SAINT LOUBERT BIÉ	Univ. Blaise Pascal, Clermont-Ferrand, F
Ecaterina SAVA	Technical Univ. Graz, Austria
Catherine SAVONA	Univ. Blaise Pascal, Clermont-Ferrand, F
Bruno SCHAPIRA	Univ. Paris-Sud, Orsay, F
Eckhard SCHLEMM	Univ. Toronto, Canada & Freie Univ. Berlin, Germany
Arseni SEREGIN	Univ. Washington, USA
Laurent SERLET	Univ. Blaise Pascal, Clermont-Ferrand, F
Frédéric SIMON	Univ. Claude Bernard, Lyon, F
Laurent TOURNIER	Univ. Claude Bernard, Lyon, F
Kiamars VAFAYI	Univ. Leiden, NL
Nicolas VERZELEN	Univ. Paris-Sud, Orsay, F
Vincent VIGON	Univ. Strasbourg, F
Guillaume VOISIN	Univ. Orléans, F
Peter WINDRIDGE	Univ. Warwick, UK
Olivier WINTERBERGER	Univ. Paris 1, F
Lorenzo ZAMBOTTI	Univ. Pierre et Marie Curie, Paris, F
Jean-Claude ZAMBRINI	GFMUL, Lisbon, Portugal

LECTURE NOTES IN MATHEMATICS

Springer

Edited by J.-M. Morel, B. Teissier; P.K. Maini

Editorial Policy (for the publication of monographs)

1. Lecture Notes aim to report new developments in all areas of mathematics and their applications - quickly, informally and at a high level. Mathematical texts analysing new developments in modelling and numerical simulation are welcome.
   Monograph manuscripts should be reasonably self-contained and rounded off. Thus they may, and often will, present not only results of the author but also related work by other people. They may be based on specialised lecture courses. Furthermore, the manuscripts should provide sufficient motivation, examples and applications. This clearly distinguishes Lecture Notes from journal articles or technical reports which normally are very concise. Articles intended for a journal but too long to be accepted by most journals, usually do not have this "lecture notes" character. For similar reasons it is unusual for doctoral theses to be accepted for the Lecture Notes series, though habilitation theses may be appropriate.

2. Manuscripts should be submitted either online at www.editorialmanager.com/lnm to Springer's mathematics editorial in Heidelberg, or to one of the series editors. In general, manuscripts will be sent out to 2 external referees for evaluation. If a decision cannot yet be reached on the basis of the first 2 reports, further referees may be contacted: The author will be informed of this. A final decision to publish can be made only on the basis of the complete manuscript, however a refereeing process leading to a preliminary decision can be based on a pre-final or incomplete manuscript. The strict minimum amount of material that will be considered should include a detailed outline describing the planned contents of each chapter, a bibliography and several sample chapters.
   Authors should be aware that incomplete or insufficiently close to final manuscripts almost always result in longer refereeing times and nevertheless unclear referees' recommendations, making further refereeing of a final draft necessary.
   Authors should also be aware that parallel submission of their manuscript to another publisher while under consideration for LNM will in general lead to immediate rejection.

3. Manuscripts should in general be submitted in English. Final manuscripts should contain at least 100 pages of mathematical text and should always include

   – a table of contents;
   – an informative introduction, with adequate motivation and perhaps some historical remarks: it should be accessible to a reader not intimately familiar with the topic treated;
   – a subject index: as a rule this is genuinely helpful for the reader.

   For evaluation purposes, manuscripts may be submitted in print or electronic form (print form is still preferred by most referees), in the latter case preferably as pdf- or zipped psfiles. Lecture Notes volumes are, as a rule, printed digitally from the authors' files. To ensure best results, authors are asked to use the LaTeX2e style files available from Springer's web-server at:

   ftp://ftp.springer.de/pub/tex/latex/svmonot1/ (for monographs) and
   ftp://ftp.springer.de/pub/tex/latex/svmultt1/ (for summer schools/tutorials).

Additional technical instructions, if necessary, are available on request from lnm@springer.com.

4. Careful preparation of the manuscripts will help keep production time short besides ensuring satisfactory appearance of the finished book in print and online. After acceptance of the manuscript authors will be asked to prepare the final LaTeX source files and also the corresponding dvi-, pdf- or zipped ps-file. The LaTeX source files are essential for producing the full-text online version of the book (see http://www.springerlink.com/openurl.asp?genre=journal&issn=0075-8434 for the existing online volumes of LNM). The actual production of a Lecture Notes volume takes approximately 12 weeks.

5. Authors receive a total of 50 free copies of their volume, but no royalties. They are entitled to a discount of 33.3 % on the price of Springer books purchased for their personal use, if ordering directly from Springer.

6. Commitment to publish is made by letter of intent rather than by signing a formal contract. Springer-Verlag secures the copyright for each volume. Authors are free to reuse material contained in their LNM volumes in later publications: a brief written (or e-mail) request for formal permission is sufficient.

Addresses:
Professor J.-M. Morel, CMLA,
École Normale Supérieure de Cachan,
61 Avenue du Président Wilson, 94235 Cachan Cedex, France
E-mail: morel@cmla.ens-cachan.fr

Professor B. Teissier, Institut Mathématique de Jussieu,
UMR 7586 du CNRS, Équipe "Géométrie et Dynamique",
175 rue du Chevaleret
75013 Paris, France
E-mail: teissier@math.jussieu.fr

For the "Mathematical Biosciences Subseries" of LNM:

Professor P. K. Maini, Center for Mathematical Biology,
Mathematical Institute, 24-29 St Giles,
Oxford OX1 3LP, UK
E-mail : maini@maths.ox.ac.uk

Springer, Mathematics Editorial, Tiergartenstr. 17,
69121 Heidelberg, Germany,
Tel.: +49 (6221) 487-8259

Fax: +49 (6221) 4876-8259
E-mail: lnm@springer.com