Dynamical approaches to linearly constrained convex minimization

Von der Fakultät Konstruktions-, Produktions- und Fahrzeugtechnik
der Universität Stuttgart zur Erlangung der Würde eines
Doktor-Ingenieurs (Dr.-Ing.) genehmigte Abhandlung

Vorgelegt von

Simon K. Niederländer

aus Waldkirchen

Hauptberichter:	Prof. Dr.-Ing. Frank Allgöwer
Mitberichter:	Prof. Dr. rer. nat. habil. Sergey Dashkovskiy
	Prof. Dr. rer. nat. Fabian Wirth

Tag der mündlichen Prüfung: 20. Dezember 2022

Institut für Systemtheorie und Regelungstechnik

Universität Stuttgart

2022

Contents

Abstract

This thesis contributes towards the dynamical viewpoint of solving linearly constrained convex minimization problems.

In a first approach, we revisit the classical Arrow–Hurwicz (AH) differential system in view of its mini-maximizing properties with respect to the Lagrangian associated with the convex minimization problem. Using tools from monotone operator theory, we investigate the asymptotic properties of the (AH) differential system and provide conditions for which its solutions converge towards a saddle point of the Lagrangian. Our convergence analysis thereby mainly relies on a "Lagrangian identity" which naturally extends on the well-known descent property associated with the continuous steepest descent method. We further present various new asymptotic decay rate estimates on the solutions of the (AH) differential system. In particular, we show that its solutions decay at an exponential rate according to the worst-case estimates known for the classical damped harmonic oscillator. We conclude our discussion by pointing out that the results on the (AH) differential system directly translate to the more general case of solving structured convex minimization problems.

Our second approach concerns the study of a Liénard-type evolution system relative to the convex minimization problem. As a decisive feature, the inertial dynamics (ID) are governed by a Hessian-driven "damping term" associated with the convex function to be minimized and "potential effects" induced by the linear constraints. Drawing on energy-like arguments that capture the dissipative nature of the inertial dynamics (ID) by means of a Bregman distance, we investigate the limiting behavior of its solutions and characterize their limit by following a hierarchical minimization principle. More precisely, we show that the limit of a solution of the (ID) differential system is characterized in terms of a Bregman-like projection of its initial data onto the set of feasible points of the convex minimization problem. In addition, we provide fast asymptotic decay rate estimates on the solutions of the (ID) differential system based on a quadratic curvature condition on the objective function of the convex minimization problem. We complement our study with the fact that the second-order Liénard-type inertial dynamics (ID) admit a first-order representation in terms of the Arrow–Hurwicz differential system (AH).

Bibliography

F. Álvarez. On the minimizing property of a second order dissipative system in Hilbert spaces. *SIAM J. Control Optim.*, 38:1102–1119, 2000.

F. Álvarez and A. Cabot. Asymptotic selection of viscosity equilibria of semilinear evolution equations by the introduction of a slowly vanishing term. *Discrete Contin. Dyn. Syst.*, 15:921–938, 2006.

F. Álvarez, H. Attouch, J. Bolte, and P. Redont. A second-order gradient-like dissipative dynamical system with Hessian-driven damping. Application to optimization and mechanics. *J. Math. Pures Appl.*, 81:747–779, 2002.

F. Álvarez, J. Bolte, and O. Brahic. Hessian Riemannian gradient flows in convex programming. *SIAM J. Control Optim.*, 43:477–501, 2004.

A. S. Antipin. Minimization of convex functions on convex sets by means of differential equations. *Differ. Equ.*, 30:1365–1375, 1994.

V. Apidopoulos, N. Ginatta, and S. Villa. Convergence rates for the heavy-ball continuous dynamics for non-convex optimization, under Polyak–Łojasiewicz condition. *J. Global Optim.*, 84:563–589, 2022.

K. J. Arrow and L. Hurwicz. A gradient method for approximating saddle points and constrained maxima. *RAND Corp., Santa Monica, CA*, pages P–223, 1951.

K. J. Arrow, L. Hurwicz, and H. Uzawa. *Studies in linear and non-linear programming.* Stanford University Press, Stanford, CA, 1958.

H. Attouch and A. Cabot. Asymptotic stabilization of inertial gradient dynamics with time-dependent viscosity. *J. Differ. Equ.*, 263:5412–5458, 2017.

H. Attouch and A. Cabot. Convergence of damped inertial dynamics governed by regularized maximally monotone operators. *J. Differ. Equ.*, 264:7138–7182, 2018.

H. Attouch and P.-É. Maingé. Asymptotic behavior of second-order dissipative evolution equations combining potential with non-potential effects. *ESAIM Control Optim. Calc. Var.*, 17:836–857, 2011.

H. Attouch and J. Peypouquet. Convergence of inertial dynamics and proximal algorithms governed by maximally monotone operators. *Math. Program.*, 174:391–432, 2019.

H. Attouch, X. Goudou, and P. Redont. The heavy ball with friction method, I. The continuous dynamical system: Global exploration of the local minima of a real-valued function by asymptotic analysis of a dissipative dynamical system. *Commun. Contemp. Math.*, 2:1–34, 2000.

H. Attouch, J. Bolte, P. Redont, and M. Teboulle. Singular Riemannian barrier methods and gradient-projection dynamical systems for constrained optimization. *Optimization*, 53:435–454, 2004.

H. Attouch, G. Buttazzo, and G. Michaille. *Variational analysis in Sobolev and BV spaces*. MOS-SIAM Series on Optimization. Society for Industrial and Applied Mathematics, Philadelphia, 2014.

H. Attouch, A. Cabot, and M.-O. Czarnecki. Asymptotic behavior of nonautonomous monotone and subgradient evolution equations. *Trans. Amer. Math. Soc.*, 370:755–790, 2018.

H. Attouch, Z. Chbani, J. Fadili, and H. Riahi. Fast convergence of dynamical ADMM via time scaling of damped inertial dynamics. *J. Optim. Theory Appl.*, 193:704–736, 2022a.

H. Attouch, Z. Chbani, J. Fadili, and H. Riahi. First-order optimization algorithms via inertial systems with Hessian driven damping. *Math. Program.*, 193:113–155, 2022b.

H. Attouch, R. I. Boţ, and E. R. Csetnek. Fast optimization via inertial dynamics with closed-loop damping. *J. Eur. Math. Soc.*, 25:1985–2056, 2023.

J. P. Aubin and A. Cellina. *Differential inclusions*. Grundlehren der mathematischen Wissenschaften 264. Springer, New York, 1984.

J.-F. Aujol, C. Dossal, and A. Rondepierre. Convergence rates of the heavy-ball method for quasi-strongly convex optimization. *SIAM J. Optim.*, 32:1817–1842, 2022.

J.-F. Aujol, C. Dossal, and A. Rondepierre. Convergence rates of the heavy-ball method under the Łojasiewicz property. *Math. Program.*, 198:195–254, 2023.

J.-B. Baillon. Un exemple concernant le comportement asymptotique de la solution du problème $\mathrm{d}u/\mathrm{d}t + \partial\phi(u) \ni 0$. *J. Funct. Anal.*, 28: 369–376, 1978.

J.-B. Baillon and H. Brézis. Une remarque sur le comportement asymptotique des semigroupes non linéaires. *Houston J. Math.*, 2:5–7, 1976.

J.-B. Baillon and G. Haddad. Quelques propriétés des opérateurs angle-bornés et n-cycliquement monotones. *Israel J. Math.*, 26:137–150, 1977.

I. Barbălat. Systèmes d'équations différentielles d'oscillations non linéaires. *Rev. Roumaine Math. Pures Appl.*, 4:267–270, 1959.

V. Barbu. *Nonlinear semigroups and differential equations in Banach spaces.* Noordhoff International Publishing, Leyden, 1976.

H. H. Bauschke and P. L. Combettes. *Convex analysis and monotone operator theory in Hilbert spaces.* CMS Books in Mathematics. Springer, New York, 2017.

H. H. Bauschke, J. M. Borwein, and P. L. Combettes. Bregman monotone optimization algorithms. *SIAM J. Control Optim.*, 42:596–636, 2003.

P. Bégout, J. Bolte, and M. A. Jendoubi. On damped second-order gradient systems. *J. Differ. Equ.*, 259:3115–3143, 2015.

R. Bellman. The stability of solutions of linear differential equations. *Duke Math. J.*, 10:643–647, 1943.

M. Benzi, G. H. Golub, and J. Liesen. Numerical solution of saddle point problems. *Acta Numer.*, 14:1–137, 2005.

J. Bolte and M. Teboulle. Barrier operators and associated gradient-like dynamical systems for constrained minimization problems. *SIAM J. Control Optim.*, 42:1266–1292, 2003.

J. Bolte, T. P. Nguyen, J. Peypouquet, and B. W. Suter. From error bounds to the complexity of first-order descent methods for convex functions. *Math. Program.*, 165:471–507, 2017.

R. I. Boţ and E. R. Csetnek. Second order forward-backward dynamical systems for monotone inclusion problems. *SIAM J. Control Optim.*, 54:1423–1443, 2016.

R. I. Boţ and D.-K. Nguyen. Improved convergence rates and trajectory convergence for primal-dual dynamical systems with vanishing damping. *J. Differ. Equ.*, 303:369–406, 2021.

R. I. Boţ, E. R. Csetnek, and S. C. László. Tikhonov regularization of a second order dynamical system with Hessian driven damping. *Math. Program.*, 189:151–186, 2021.

L. M. Bregman. The relaxation method of finding the common point of convex sets and its application to the solution of problems in convex programming. *U.S.S.R. Comput. Math. and Math. Phys.*, 7:200–217, 1967.

H. Brézis. Monotonicity methods in Hilbert spaces and some applications to nonlinear partial differential equations. *In Contributions to Nonlinear Funct. Analysis, Acad. Press*, pages 101–156, 1971.

H. Brézis. *Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert.* Mathematics Studies 5. North-Holland, Amsterdam, 1973.

H. Brézis. Asymptotic behavior of some evolution systems. *In Nonlinear Evolution Equations, Madison, 1977, Acad. Press*, pages 141–154, 1978.

H. Brézis. *Functional analysis, Sobolev spaces and partial differential equations.* Springer, New York, 2011.

H. Brézis and F. E. Browder. Nonlinear ergodic theorems. *Bull. Amer. Math. Soc.*, 82:959–961, 1976.

F. E. Browder. Nonlinear operators and nonlinear equations of evolution in Banach spaces. *In Nonlinear Functional Analysis, Proceedings of Symposia in Pure Math., Amer. Math. Soc.*, 1976.

R. E. Bruck. Asymptotic convergence of nonlinear contraction semigroups in Hilbert spaces. *J. Funct. Anal.*, 18:15–26, 1975.

A. Cabot and P. Frankel. Asymptotics for some semilinear hyperbolic equations with non-autonomous damping. *J. Differ. Equ.*, 252:294–322, 2012.

A. Cabot, H. Engler, and S. Gadat. On the long time behavior of second order differential equations with asymptotically small dissipation. *Trans. Amer. Math. Soc.*, 361:5983–6017, 2009.

H. Cartan. *Calcul différentiel.* Collection Méthodes. Hermann, Paris, 1967.

J. Céa. *Optimisation, théorie et algorithmes.* Dunod, Paris, 1971.

Z. Chbani and H. Riahi. Existence and asymptotic behaviour for solutions of dynamical equilibrium systems. *Evol. Equ. Control Theory*, 3:1–14, 2014.

G. Chen and M. Teboulle. Convergence analysis of a proximal-like minimization algorithm using Bregman functions. *SIAM J. Optim.*, 3: 538–543, 1993.

R. Chill and E. Fašangová. Gradient systems. *Proceedings of the 13th International Internet Seminar*, 2010.

E. A. Coddington and N. Levinson. *Theory of ordinary differential equations.* International Series in Pure and Applied Mathematics. McGraw Hill Co., New York, 1955.

P. L. Combettes. Monotone operator theory in convex optimization. *Math. Program.*, 170:177–206, 2018.

R. Cominetti, J. Peypouquet, and S. Sorin. Strong asymptotic convergence of evolution equations governed by maximal monotone operators with Tikhonov regularization. *J. Differ. Equ.*, 245:3753–3763, 2008.

M. G. Crandall and A. Pazy. Semi-groups of nonlinear contractions and dissipative sets. *J. Funct. Anal.*, 3:376–418, 1969.

C. M. Dafermos. Asymptotic behavior of solutions of evolution equations. *In Nonlinear Evolution Equations, Madison, 1977, Acad. Press*, pages 103–123, 1978.

K. Deimling. *Ordinary differential equations in Banach spaces.* Lecture Notes in Mathematics. Springer, New York, 1977.

M. Edelstein. The construction of an asymptotic center with a fixed-point property. *Bull. Amer. Math. Soc.*, 78:206–208, 1972.

I. Ekeland and R. Témam. *Convex analysis and variational problems.* Classics in applied mathematics. Society for Industrial and Applied Mathematics, Philadelphia, 1999.

S. D. Flåm and A. Ben-Israel. Approximating saddle points as equilibria of differential inclusions. *J. Math. Anal. Appl.*, 141:264–277, 1989.

J. Franců. Weakly continuous operators. Application to differential equations. *Appl. Math.*, 39:45–56, 1994.

T. H. Gronwall. Note on the derivatives with respect to a parameter of the solutions of a system of differential equations. *Ann. of Math.*, 20: 292–296, 1919.

O. Güler. Convergence rate estimates for the gradient differential inclusion. *Optim. Methods Softw.*, 20:729–735, 2005.

J. K. Hale. *Asymptotic behavior of dissipative systems.* Mathematical Surveys and Monographs 25. American Mathematical Society, Providence, Rhode Island, 1988.

A. Haraux. *Systèmes dynamiques dissipatifs et applications.* Recherches en Mathématiques Appliquées 17, Masson, Paris, 1991.

A. Haraux and M. A. Jendoubi. *The convergence problem for dissipative autonomous systems.* Springer Briefs in Mathematics. Springer, New York, 2015.

A. Haraux and E. Zuazua. Decay estimates for some semilinear damped hyperbolic problems. *Arch. Ration. Mech. Anal.*, 100:191–206, 1988.

G. H. Hardy, J. E. Littlewood, and G. Pólya. *Inequalities.* Cambridge University Press, New York, 1952.

J.-B. Hiriart-Urruty and C. Lemaréchal. *Convex analysis and minimization algorithms I.* Grundlehren der mathematischen Wissenschaften 305. Springer, New York, 1993.

J. Jahn. *Introduction to the theory of nonlinear optimization.* Springer, New York, 2007.

T. Kato. Nonlinear semigroups and evolution equations. *J. Math. Soc. Japan*, 19:508–520, 1967.

Y. Kōmura. Nonlinear semi-groups in Hilbert space. *J. Math. Soc. Japan*, 19:493–507, 1967.

T. Kose. Solutions of saddle value problems by differential equations. *Econometrica*, 24:59–70, 1956.

K. Kurdyka. On gradients of functions definable in o-minimal structures. *Ann. Inst. Fourier*, 48:769–783, 1998.

E. Landau. Einige Ungleichungen für zweimal differentiierbare Funktionen. *Proc. Lond. Math. Soc.*, 11:43–49, 1913.

B. Lemaire. An asymptotical variational principle associated with the steepest descent method for a convex function. *J. Convex Anal.*, 3: 63–70, 1996.

A. Liénard. Étude des oscillations entretenues. *Rev. gén. d'électr.*, 23: 901–912 and 946–954, 1928.

T. Lin and M. I. Jordan. A control-theoretic perspective on optimal high-order optimization. *Math. Program.*, 195:929–975, 2022.

S. Łojasiewicz. Une propriété topologique des sous-ensembles analytiques réels. *Les Équations aux Dérivées Partielles, Éditions du centre National de la Recherche Scientifique, Paris*, 117:87–89, 1963.

S. Łojasiewicz. Sur les trajectoires du gradient d'une fonction analytique. *Seminari di Geometria, Bologna (1982/83), Università degli Studi di Bologna, Bologna*, pages 115–117, 1984.

R. H. Martin, Jr. *Nonlinear operators and differential equations in Banach spaces.* John Wiley & Sons, New York, 1976.

G. J. Minty. Monotone (nonlinear) operators in Hilbert space. *Duke Math. J.*, 29:341–346, 1962.

G. Moroşanu. *Nonlinear evolution equations and applications.* Mathematics and its Applications. Kluwer Academic Publishers Group, Dordrecht, 1988.

A. S. Nemirovski and D. B. Yudin. Cesari convergence of the gradient method of approximating saddle points of convex-concave functions. *Dokl. Akad. Nauk SSSR*, 239:1056–1059, 1978.

Y. Nesterov. A method of solving a convex programming problem with convergence rate $\mathcal{O}(1/k^2)$. *Sov. Math. Dokl.*, 27:372–376, 1983.

Y. Nesterov. *Lectures on convex optimization.* Springer Optimization and its Applications. Springer, New York, 2018.

S. K. Niederländer. Second-order dynamics with Hessian-driven damping for linearly constrained convex minimization. *SIAM J. Control Optim.*, 59:3708–3736, 2021.

S. K. Niederländer. On the Arrow–Hurwicz differential system for linearly constrained convex minimization. *Optimization*, 2023. URL `https://doi.org/10.1080/02331934.2023.2215799`.

H. Okochi. A note on asymptotic strong convergence of nonlinear contraction semigroups. *Proc. Japan Acad. Math. Sci.*, 56:83–84, 1980.

Z. Opial. Weak convergence of the sequence of successive approximations for nonexpansive mappings. *Bull. Amer. Math. Soc.*, 73:591–597, 1967.

J. Palis and W. de Melo. *Geometric theory of dynamical systems.* Springer, New York, 1982.

G. B. Passty. Ergodic convergence to a zero of the sum of monotone operators in Hilbert space. *J. Math. Anal. Appl.*, 72:383–390, 1979.

A. Pazy. Strong convergence of semigroups of nonlinear contractions in Hilbert space. *J. Anal. Math.*, 34:1–35, 1978a.

A. Pazy. On the asymptotic behavior of semigroups of nonlinear contractions in Hilbert space. *J. Funct. Anal.*, 27:292–307, 1978b.

A. Pazy. Semi-groups of nonlinear contractions and their asymptotic behavior. *In Nonlinear Analysis and Mechanics: Heriot-Watt Symposium III, Pitman, London*, pages 36–134, 1979.

J. Peypouquet. *Convex optimization in normed spaces.* Springer Briefs in Optimization. Springer, New York, 2015.

J. Peypouquet and S. Sorin. Evolution equations for maximal monotone operators: Asymptotic analysis in continuous and discrete time. *J. Convex Anal.*, 17:1113–1163, 2010.

B. T. Polyak. Gradient methods for the minimisation of functionals. *U.S.S.R. Comput. Math. and Math. Phys.*, 3:864–878, 1963.

B. T. Polyak. Some methods of speeding up the convergence of iteration methods. *U.S.S.R. Comput. Math. and Math. Phys.*, 4:1–17, 1964.

B. T. Polyak. Iterative methods using Lagrange multipliers for solving extremal problems with constraints of the equation type. *U.S.S.R. Comput. Math. and Math. Phys.*, 10:42–52, 1970.

B. T. Polyak. *Introduction to optimization.* Optimization Software, New York, 1987.

B. T. Polyak and N. V. Tret'yakov. The method of penalty estimates for conditional extremum problems. *U.S.S.R. Comput. Math. and Math. Phys.*, 13:42–58, 1973.

S. Reich. Nonlinear evolution equations and nonlinear ergodic theorems. *Nonlinear Anal.*, 1:319–330, 1977.

R. Reissig, G. Sansone, and R. Conti. *Qualitative Theorie nichtlinearer Differentialgleichungen.* Pubblicazioni dell'Istituto Nazionale di Alta Matematica. Edizioni Cremonese, Roma, 1963.

R. T. Rockafellar. Monotone operators associated with saddle-functions and minimax problems. *In Nonlinear Functional Analysis, Proceedings of Symposia in Pure Math., Amer. Math. Soc.*, pages 241–250, 1969a.

R. T. Rockafellar. Convex functions, monotone operators and variational inequalities. *In Theory and Applications of Monotone Operators, Proceedings of a NATO Advanced Study Institute*, pages 34–65, 1969b.

R. T. Rockafellar. *Convex analysis.* Princeton University Press, Princeton, 1970.

R. T. Rockafellar. Saddle-points and convex analysis. *In Differential Games and Related Topics, North-Holland*, pages 109–127, 1971.

O. Sebbouh, C. Dossal, and A. Rondepierre. Convergence rates of damped inertial dynamics under geometric conditions and perturbations. *SIAM J. Optim.*, 30:1850–1877, 2020.

G. R. Sell and Y. You. *Dynamics of evolutionary equations.* Applied Mathematical Sciences. Springer, New York, 2002.

L. Simon. Asymptotics for a class of non-linear evolution equations, with applications to geometric problems. *Ann. of Math.*, 118:525–571, 1983.

W. Su, S. Boyd, and E. Candès. A differential equation for modeling Nesterov's accelerated gradient method: Theory and insights. *J.*

Mach. Learn. Res., 17:1–43, 2016.

A. Tikhonov and V. Arsenine. *Méthodes de résolution de problèmes mal posés*. Mir, Moscow, 1974.

V. I. Venets. Continuous algorithms for solution of convex optimization problems and finding saddle points of convex-concave functions with the use of projection operators. *Optimization*, 16:519–533, 1985.

J. A. Walker. *Dynamical systems and evolution equations*. Mathematical Concepts and Methods in Science and Engineering. Plenum Press, New York, 1980.

K. Yosida. *Functional analysis*. Classics in Mathematics. Springer, New York, 1995.

E. Zeidler. *Nonlinear functional analysis and its applications I: Fixed-point theorems*. Springer, New York, 1986.

E. Zeidler. *Nonlinear functional analysis and its applications II/A: Linear monotone operators*. Springer, New York, 1990a.

E. Zeidler. *Nonlinear functional analysis and its applications II/B: Nonlinear monotone operators*. Springer, New York, 1990b.