Doctoral Thesis

Non-holonomy, critical manifolds and stability in constrained Hamiltonian systems

Author(s):
Chen, Thomas

Publication Date:
1999

Permanent Link:
https://doi.org/10.3929/ethz-a-002054454

Rights / License:
In Copyright - Non-Commercial Use Permitted
NON-HOLONOMY, CRITICAL MANIFOLDS AND STABILITY IN CONSTRAINED HAMILTONIAN SYSTEMS

A dissertation submitted to the
SWISS FEDERAL INSTITUTE OF TECHNOLOGY ZURICH
(ETH Zürich)

for the degree of
Doctor of Technical Sciences

presented by
Thomas Chen, Dipl. Ing. ETH, Dipl. Phys. ETH
born October 6, 1968
Swiss citizen

accepted on the recommendation of
Prof. Dr. Hans Brauchli, examiner
Prof. Dr. Eduard Zehnder, co-examiner

1999
ACKNOWLEDGEMENTS

I warmly thank my adviser Prof. Hans Brauchli for the possibility to carry out this work, for original and inspiring ideas, and a great amount of freedom in the choice of topics and methods. I am profoundly grateful for his great generosity to support my second study in theoretical physics which went along with this thesis. I also thank him for many highly interesting conference trips, for numerous occasions at which I enjoyed his and Julia's hospitality, and for the special events initiated by him, such as the annual seminar week in the alps.

I am profoundly grateful to Prof. E. Zehnder for very generously dedicating his precious time to a great number of discussions that were quintessential for my proper understanding of the topics presented in the sequel, and for accepting the coadvisorship for this thesis. I have learned immensely from his insights and clarity of reasoning, and his uplifting humour and friendliness were much appreciated.

I am deeply grateful to Prof. J. Fröhlich for his advice and support, and a great number of discussions during the last years which have taught me very important lessons, from which this work has also substantially benefitted.

I thank the institute of mechanics, and the head of the institute Prof. M. B. Sayir, for an excellent infrastructure, and for supporting my activities outside the department of mechanical engineering.

Many inspirations leading to results presented in this thesis stem from discussions with Mario Clerici, whom I thank for sharing his insights. I thank Marc von Wattenwyl for his interest and many discussions. Many thanks for everything to the secretary of the institute Traude Junker, who has the patience of a saint. I also thank Prof. Miguel Sofer for his help at early stages of this work. I am much indebted to Prof. Hiroaki Yoshimura for highly interesting and inspiring discussions, and to Prof. Oliver O'Reilly for his friendliness and hospitality during a conference trip to California in fall 1997. I thank Prof. J. E. Marsden for presenting the department of control and dynamical systems at Caltech to Prof. Yoshimura and me in fall 1997, and for pointing out very relevant references.

I heartily thank my parents and sisters, especially Susanne, for their unwavering human and moral support, and for their example of perseverance. I am profoundly grateful to all my friends for their endurance, and for being around when I badly needed them. I thank all my relatives in the United States and Taiwan for their encouragement and support.
SUMMARY

This thesis in the field of theoretical mechanics discusses a particular class of constrained Hamiltonian systems. It addresses questions concerned with geometrical and global topological issues about the critical manifolds, the stability of equilibria, aspects about numerical stability, and technical applications.

In the engineering sciences, constrained Hamiltonian systems are typically encountered in the numerical simulation of multibody systems. We will study a number of properties exhibited by such systems using methods of symplectic geometry and geometrical mechanics. We believe that this will make many qualitative issues particularly transparent. The mathematical tools employed in many parts of this thesis are, unfortunately, not standard for an engineering text, but are unavoidable for the topics in discussion. In the last chapter, we have summarized examples, illustrations, and practical implications of our results in an easily accessible way.

In chapter one, we introduce the general setting of the constrained Hamiltonian systems that will interest us. We are given a 2n-dimensional symplectic manifold \((M, \omega)\) together with a Hamiltonian function \(H : M \to \mathbb{R}\), and a smooth, almost symplectic, rank 2k distribution \(V\). We explicitly construct the \(\omega\)-skew selfadjoint bundle projector \(P : TM \to V\), and an almost complex structure \(J\), together with the associated Kähler metric \(g\). They are related to \(\omega\) by way of \(\omega(X, Y) = g(JX, Y)\), where \(X\) and \(Y\) are arbitrary vector fields. We show that \(g\) can be picked in a manner that \(P\) is selfadjoint, and that it commutes with \(J\).

The flow of the dynamical system of interest is generated by the vector field \(PX_H\), where \(X_H\) is the Hamiltonian vector field of the system at the absence of constraints. The corresponding dynamical equations can be written as \(\dot{x} = PJ\text{grad}H|_x = JP\text{grad}H|_x\), where the gradient is defined by \(g\). Due to the fact that \(P\) is explicitly given, one can straightforwardly write down the explicit expression for the Dirac bracket of the constrained system.

Chapter two addresses the geometry and global topology of the critical set \(\mathcal{C}\) of the constrained Hamiltonian system. The main technical tool used for this purpose is the gradient-like flow \(\phi_t\) generated by the vector field \(P\text{grad}H\), whose critical set is also \(\mathcal{C}\). Throughout
this chapter, it is assumed that $H : M \to \mathbb{R}$ is a Morse function, and it is proved that gener-
cically, $C$ is a normal hyperbolic, $2(n - k)$-dimensional fixed point submanifold of $M$ with
respect to $\phi_t$. The generalization of Morse theory developed by C. Conley and E. Zehnder
is used to prove a topological formula for closed, compact $C$, that is strongly reminiscent of
the Morse-Bott inequalities. An alternative proof, based on the use of the Witten complex
in Morse theory, is also presented.

In chapter three, we address constrained Hamiltonian mechanical systems. The phase
space is the cotangent bundle $T^*Q$ of an orientable Riemannian manifold $Q$, where the
Riemannian metric is defined by the kinetic energy. We consider Pfaffian constraints, which
are, by definition, linear conditions imposed on the generalized velocities. They represent
the case of highest technical relevance. It is shown that the physical orbits of such a system
are confined to a $n + k$-dimensional submanifold $M_{\text{phys}} \subset T^*Q$, which we refer to as the
"physical sheet".

We study the constrained physical system by use of an auxiliary constrained Hamilto-
nian system of the type introduced in chapter one, which is defined on the whole embedding
space $T^*Q$ of $M_{\text{phys}}$. This auxiliary system exhibits $M_{\text{phys}}$ as an invariant manifold, on
which it generates the physical flow. A number of questions related to numerical simula-
tions, symmetries, etc. are addressed.

In chapter four, we study the global topology of the critical manifold of the physical
constrained system, and derive Conley-Zehnder inequalities as in chapter two. We also
analyze the stability of equilibria of such systems, for the purpose of which the projector
formalism developed in the first three chapters proves to be highly efficient. The case of
asymptotic behaviour is treated with an application of center manifold theory, while the
case of critical stability is only analyzed with an application of averaging theory.

In chapter five, some technical applications of our results are presented. We list a
number of characteristic properties of the critical manifold of a constrained Hamiltonian me-
chanical system. We also propose a method to numerically construct the generic connectivity
components of the critical manifold in certain cases. The topological results of the second
chapter are illustrated with a two dimensional example, we present examples of mechanical
systems to illustrate some of the typical phenomena that occur in non-holonomic systems.
ZUSAMMENFASSUNG


Gebundene Hamilton'sche Systeme treten in den Ingenieurwissenschaften typischerweise im Zusammenhang mit der Simulation von Vielkörpersystemen auf. Wir werden Methoden der symplektischen Geometrie und der geometrischen Mechanik zu ihrem Studium benützen, wodurch sich viele qualitative Aspekte in besonders transparenter Art darstellen lassen. Der mathematische Apparat, der in einem Grossteil dieser Arbeit verwendet wird, ist leider nicht herkömmlich für einen Ingenieurtext, jedoch lässt er sich aufgrund der Themmatik nicht vermeiden. Im letzten Kapitel haben wir Beispiele, Illustrationen und technische Anwendungen unserer Resultate in einfach zugänglicher Weise dargelegt.

Die Grundlagen der Klasse gebundener Hamilton'scher Systeme, die für unsere Arbeit interessant sind, werden im ersten Kapitel eingeführt. Sie bestehen aus einer $2n$-dimensionalen symplektischen Mannigfaltigkeit $(M, \omega)$ und einer Hamiltonfunktion $H : M \to \mathbb{R}$, sowie einer fast symplektischen Distribution $V$ vom Faserrang $2k$. Wir konstruieren den $\omega$-schießadjungierten Bündelprojektor $P : TM \to V$ und führen eine fast komplexe Struktur $J$ mit einer zugehörigen Kählermetrik $g$ ein. Für alle Vektorfelder $X, Y$ in $TM$ gilt demnach $\omega(X, Y) = g(JX, Y)$. Insbesondere kann $g$ in einer Weise gewählt werden, dass $P$ symmetrisch ist und daher mit $J$ kommutiert.

Das für uns wesentliche dynamische System wird durch \( \dot{x} = PX_H\big|_x \) beschrieben, wobei $X_H$ das Hamilton'sche Vektorfeld des ungebundenen Systems darstellt. Die Bewegungsgleichungen lassen sich in der Form \( \dot{x} = PJ\text{grad}H\big|_x = JP\text{grad}H\big|_x \) schreiben, wobei der Gradient durch die Metrik $g$ definiert wird. Der Bündelprojektor $P$ erlaubt es uns, die Dirac-Klammer des gebundenen Systems in direkter Weise explizit hinzuschreiben.

Im zweiten Kapitel beschäftigen wir uns mit der Geometrie und Topologie der kritischen Menge $\mathcal{C}$ eines solchen dynamischen Systems und beweisen, dass es sich bei ihr generischerweise um eine $2(n-k)$-dimensionale Untermannigfaltigkeit von $M$ handelt. Unser wichtigstes
Werkzeug zum Studium der Topologie von $\mathcal{C}$ besteht aus dem gradientenartigen Fluss, der durch das Vektorfeld $P\text{grad}H$ erzeugt wird. Offensichtlich stimmt seine kritische Menge mit $\mathcal{C}$ überein. In diesem Kapitel gehen wir von der Annahme aus, $H : M \to \mathbb{R}$ sei eine Morse-Funktion und wir beweisen, dass $\mathcal{C}$ generischerweise bezüglich des gradientenartigen Flusses normal hyperbolisch ist. Für eine randfreie, kompakte symplektische Mannigfaltigkeit $M$ beweisen wir, mit Hilfe der Verallgemeinerung der Morse Theorie durch C. Conley und E. Zehnder, eine topologische Formel für $\mathcal{C}$, welche den zentralen Ungleichungen der Morse-Bott Theorie stark ähnelt. Darüberhinaus geben wir einen zweiten, alternativen Beweis desselben Resultats, der auf der Existenz des Witten Komplexes in der Morse Theorie basiert.

Im dritten Kapitel studieren wir gebundene Hamilton'sche Systeme in der Mechanik. In diesem Fall besteht der Phasenraum aus dem Kotangentialbündel einer Riemann'schen Mannigfaltigkeit $Q$, deren Metrik durch die kinetische Energie bestimmt wird. Wir betrachten ausschließlich den Fall Pfaff'scher Bindungen, welche durch lineare Bedingungen an die verallgemeinerten Geschwindigkeiten ausgedrückt werden. Hierdurch lassen sich bereits die meisten technisch relevanten Situationen abdecken. Die Lösungsorbits dieses Systems befinden sich in einer $n + k$-dimensionalen Untermannigfaltigkeit $\mathcal{M}_{\text{phys}}$ von $T^*Q$, welche wir als 'physikalisches Blatt' bezeichnen.

Zur Untersuchung dieses Systems führen wir ein dynamisches Hilfssystem auf $T^*Q$ ein, welches ein gebundenes Hamilton'sches System derjenigen Art darstellt, welches wir im ersten Kapitel eingeführt haben. Es weist $\mathcal{M}_{\text{phys}}$ als invariante Mannigfaltigkeit auf und erzeugt auf $\mathcal{M}_{\text{phys}}$ die physikalische Dynamik. Eine Anzahl von Eigenschaften dieser Systeme werden näher untersucht und in ihrer Eignung hinsichtlich numerischer Implementation beurteilt. Wir stellen ein besonderes Hilfssystem vor, welches $\mathcal{M}_{\text{phys}}$ als grenzstabile invariante Mannigfaltigkeit aufweist und Symmetrien des physikalischen gebundenen Systems in geeigneter Weise fortsetzt.

Im fünften Kapitel verweisen wir auf mögliche technische Anwendungen der Resultate aus den ersten vier Kapiteln. Qualitative Eigenschaften der kritischen Mannigfaltigkeit gebundener mechanischer Systeme werden aufgezählt und eine Methode wird vorgestellt, die es erlaubt, die generischen Zusammenhangskomponenten in besonderen Fällen vollständig numerisch zu rekonstruieren. Die topologischen Resultate des zweiten und vierten Kapitels werden anhand eines zweidimensionalen Beispiel anschaulich erklärt und mechanische Beispielssysteme werden diskutiert, um physikalische Phänomene in nichtholonom gebundenen Systemen zu veranschaulichen.
Seite Leer /
Blank leaf
CHAPTER 1

CONSTRAINED HAMILTONIAN SYSTEMS ON SYMPLECTIC MANIFOLDS

In this thesis, various aspects about constrained Hamiltonian systems will be addressed within the framework of geometrical mechanics. This chapter is dedicated to the definition of the class of dynamical systems that will be of central interest.

We will first present a short introductory section that summarizes fundamental facts about Hamiltonian dynamics on symplectic manifolds, but we will exclude all aspects that are not directly relevant for the later chapters.

In a next step, we will impose a particular class of linear phase space constraints upon Hamiltonian systems on symplectic manifolds. We will construct certain efficient technical tools for their analysis, comprising a specific bundle projector that is skew symmetric relative to the symplectic structure, an almost complex structure that commutes with the projection, and a Kähler metric compatible with the almost complex and symplectic structures, relative to which the projection is orthogonal.

Our approach has been inspired by a particular description of constrained multibody systems [10] that has been studied for many years at the institute of mechanics. We also remark that an earlier dissertation on a very similar subject matter has been authored by R. Weber [40] at this institute.

1.1 A minisurvey of Hamiltonian dynamics

We will here give a minisurvey of fundamental facts about Hamiltonian dynamics on symplectic manifolds.

Most proofs will be omitted, but can, in case of interest, easily be looked up in the literature on the topic. For a detailed introduction to geometrical mechanics, we refer the reader for instance to [1, 28]. In addition, all detailed differentiability questions will be
omitted; we will from here on, without further mention, assume that all geometrical objects
and structures shall be smooth.

A symplectic manifold is a pair \((M, \omega)\) that consists of a smooth, orientable manifold
\(M\) of even dimension \(2n\), together with a closed, non-degenerate two form \(\omega\), which is called
the symplectic structure. Non-degeneracy of the symplectic structure, together with the even
dimension of \(M\), imply that the \(2n\) form \(\omega^n\) defines a volume form on \(M\).

The Hamiltonian

\[ H : M \to \mathbb{R} \]

is assumed to be a smooth function. It defines the Hamiltonian vector field \(X_H\) by way of

\[ i_{X_H} \omega = -dH; \]

\(i\) and \(d\) denote the operations of interior multiplication and exterior derivation (we assume
that the reader is familiar with basic differential and symplectic geometry). The value of the
Hamiltonian \(H\) at a point \(x\) is called the energy of the point \(x\).

The Hamiltonian flow on \((M, \omega)\) is the one-parameter group of diffeomorphisms

\[ \Phi_t : \mathbb{R} \times M \to M \]

\[ (t, x) \mapsto \Phi_t(x) \]

that is generated by the Hamiltonian vector field; we will always assume that \(X_H\) is complete.
The variable \(t\) parametrizes the time evolution of this dynamical system, and clearly,

\[ \frac{d}{dt} \Phi_t(x) = X_H |_{\Phi_t(x)} \]

for all \(t\), and all \(x \in M\). \(M\) is referred to as the phase space of the system.

If the two form \(\omega\) is non-degenerate, but fails to be closed, \(\omega\) is said to be an almost
symplectic structure, and accordingly, the pair \((M, \omega)\) is an almost symplectic manifold. On
the other hand, it is clear that for a symplectic manifold, \(\omega\) represents an element of the
second de Rham cohomology group of \(M\), \(H^2(M)\). If this element is the trivial cohomology
class, there is a one form \(\theta_0\) such that \(\omega = -d\theta_0\). \(\theta_0\) is referred to as the symplectic one form,
or the 'one form potential' of \(\omega\). However, for a closed compact manifold, it is not possible
that \(\omega\) belongs to the trivial cohomology class, since in this case, the volume of \(M\) with
respect to the measure \(\omega^n\), which is then also an exact form, would be zero. This result is in
contradiction to the condition of non-degeneracy. Symplectic structures on closed compact
symplectic manifolds are never exact.
1.1.1 The Darboux theorem

The Darboux theorem implies that the local geometry of a symplectic manifold is essentially unique. It states that every sufficiently small neighbourhood in $M$ admits a coordinate chart $(x_1, \ldots, x_n, y^1, \ldots, y^n)$, relative to which the symplectic structure can locally be written as

$$\omega = dx_1 \wedge dy^1 + \ldots + dx_n \wedge dy^n.$$ 

Among other things, this implies that locally, $\omega$ always admits the one form potential

$$\theta_0 = \sum_k y^k dx_k.$$ 

Furthermore, this also implies that the symplectic structure can locally always be represented in terms of the **symplectic standard matrix**

$$J = \begin{pmatrix} 0 & 1_n \\ -1_n & 0 \end{pmatrix}.$$  \hfill (1.1)

In non-geometrical analytical dynamics, this representation of the symplectic structure is usually the starting point of the discussion of Hamiltonian dynamics. The Darboux theorem implies that all symplectic manifolds of a given dimension are locally equivalent to each other.

1.1.2 Symplectomorphisms

A diffeomorphism $f : M \to M$ is said to preserve the symplectic structure if its pullback maps $\omega$ to itself, that is, $f^* \omega = \omega$. It is then called a **symplectic diffeomorphism**, or a **symplectomorphism**. The set of symplectomorphisms of $M$ defines the **symplectomorphism group** $Sp(M)$, where the group operation is the concatenation of maps. To learn more on the structure of $Sp(M)$, one may for instance consult the beautiful books [27, 24] on symplectic topology.

The Hamiltonian flow $\Phi_t$ preserves the symplectic structure, which means that it consists of a one parameter group of symplectomorphisms. This is deduced from

$$\frac{d}{dt} \Phi_t^* \omega = \Phi_t^* L_{X_H} \omega$$

$$= \Phi_t^* ((d \circ i_{X_H} + i_{X_H} \circ d) \omega)$$

$$= \Phi_t^* (-d^2 H + i_{X_H} d \omega)$$

$$= 0; \hfill (1.2)$$
the symbol 'L' denotes the Lie derivative. The action of the Lie derivative on vector fields is given by $L_Y(X) = [X, Y]$, where $[\cdot, \cdot]$ denotes the Lie bracket, defined by $[X, Y](f) = X(Y(f)) - Y(X(f))$; $X$ and $Y$ denote vector fields, and $f$ is a smooth function.

We stress that the property of closedness of the symplectic structure has been used in the proof of the symplecticity of the flow map. The latter implies that the symplectic volume form $\Omega_\omega = \omega^n/n!$ is preserved, which is expressed by

$$\Phi_t^*\Omega_\omega = \Omega_\omega,$$

a result known as Liouville's theorem. The symplectic volume form is also referred to as the phase volume, or Liouville measure.

1.1.3 Poisson structures and energy conservation

It suffices to have a Poisson manifold in order to define a Hamiltonian system, which belongs to a much broader class of manifolds than the class of symplectic manifolds. For instance, the dimension of a Poisson manifold need not be an even number.

A Poisson structure is a skew symmetric bilinear form $\{\cdot, \cdot\}$ on the algebra $C^\infty(M)$ of smooth functions on $M$, which is not necessarily non-degenerate. It is a derivative in both arguments, and satisfies the Jacobi identity

$$\{f, \{g, h\}\} + \{h, \{f, g\}\} + \{g, \{h, f\}\} = 0.$$ 

The statement that it is a derivative in both arguments means that $\{f, g\}$ acts on $f$ and $g$ as a bilinear form on the one forms $df$ and $dg$. Therefore, the operator $\{f, \cdot\}$ defines a vector field, and the Poisson bracket itself defines a rank two tensor, which is a symplectic structure if and only if it is non-degenerate. The second part of this statement is the content of the Pauli-Jost theorem, cf. for instance [26]. The Hamiltonian vector field in the picture of Poisson geometry is defined as $X_H = \{\cdot, H\}$.

Every symplectic manifold carries a natural Poisson structure that is defined in terms of

$$\{f, g\} = \omega(X_f, X_g),$$

but the converse is, as we have said, not true. The skew symmetry of the pairing is evident; the crucial ingredient that relates Poisson and symplectic structures, if possible, is the
closedness of $\omega$, which one needs in order to prove the Jacobicity of the Poisson bracket. Consequently, almost symplectic structures do not define Poisson structures. One can write the time derivative of every smooth function $f$ along the orbits of the Hamiltonian flow as

$$\dot{f} = \{f, H\};$$

here the dot in $\dot{f}$ denotes the time derivative.

An important consequence of the antisymmetry of the Poisson bracket is that the value of the Hamiltonian as a function on $M$ is constant along all orbits, since

$$\dot{H} = \{H, H\} = 0.$$ 

Therefore, the manifold $M$ is foliated into invariant hypersurfaces which correspond to the energy level surfaces

$$\Sigma_E \equiv \{x \in M \mid H(x) = E\}, \quad (1.3)$$

with $M = \cup_E \Sigma_E$.

This concludes our truly small minisurvey of Hamiltonian dynamics.

### 1.2 Induced Hamiltonian systems on symplectic submanifolds

Submanifolds of a symplectic manifold are more subtle objects than submanifolds of, say, a Riemannian manifold, the reason being that the pullback of the symplectic structure to a submanifold defines a two form which does not necessarily exhibit any properties which are consistently satisfied everywhere on the submanifold. Submanifolds for which the pullback of the symplectic structure has consistent properties everywhere are suitably 'positioned' in $M$, such that their tangent bundles respect the symplectic structure in a specific manner to be made precise. Here, we recall the most important families of submanifolds that are 'adapted' to the symplectic structure, see also [1, 27].

Let us first consider linear subspaces $F \subset E$ of a linear symplectic vector space $(E, \omega_E)$. To this end, we let

$$F^\perp \equiv \{Y \in E \mid \omega_E(Y, Y') = 0 \forall Y' \in F\}$$

denote the symplectic complement of $F$ in $E$. Then, the following important cases can be listed for $F$: 

11
(1) The linear subspace $F$ is called *isotropic* if it is a subset of its own symplectic complement $F^\perp$, which means that the pairing $\omega_E(Y, Y')$ vanishes for all vectors $Y, Y'$ in $F$.

(2) The linear subspace $F$ is called *co-isotropic* if it contains its symplectic complement, that is, any vector $Y$ necessarily is an element of $F$ if $\omega(Y, Y')$ is zero for all $Y'$ in $F$.

(3) Denoting the vector space complement of $F$ by $F'$, such that $E = F \oplus F'$, $F$ is a *Lagrangian* subspace if both $F$ and $F'$ are isotropic. Every Lagrangian subspace $F$ has half the dimension of $E$, and coincides with its symplectic complement, that is, $F = F^\perp$.

(4) $F$ is called a *symplectic* subspace if $\omega_E$ restricted to $F \times F$ is non-degenerate. A symplectic subspace always has an even dimension.

If the properties of the linear subspaces presently discussed are satisfied on the whole tangent bundle of a submanifold of $M$, the following characteristic cases can be summarized.

A submanifold $A$ of the symplectic manifold $(M, \omega)$ is called an *isotropic* (co-isotropic, symplectic) submanifold if for every $x \in A$, the tangent space $T_xA$ is mapped to an isotropic (co-isotropic, symplectic) linear subspace of $T_{j(x)}M$ by the pushforward $j_*$ of the inclusion map.

Some alternative definitions are as follows. It can be deduced from the definition given here that the submanifold $A$ is isotropic if and only if the symplectic structure is pulled back to zero by the inclusion map, $j^*\omega = 0$. Furthermore, a different way to say that the submanifold $A$ is Lagrangian is to require it to be isotropic, and to assume the existence of an isotropic subbundle $K$ of the vector bundle $TM |_A$ in a manner that one obtains the Whitney sum $TM |_A = TA \oplus K$. If $A$ is a Lagrangian submanifold, its dimension is given by half the dimension of $M$, and each fibre of its tangent bundle satisfies $T_xA = (T_xA)^\perp$, where the symplectic complement is fibrewise defined in the obvious manner. The proofs of these statements and a more detailed discussion can for instance be found in [1, 27].

For every symplectic submanifold $A$ of $M$, there is a naturally induced Hamiltonian system on $A$, which is obtained in the following way. Let us assume that $A$ is smoothly embedded in $M$ via $j : A \to M$. Then, the symplectic structure on $M$ is pulled back to a two form $\omega_A = j^*\omega$ on $A$. This two form is non-degenerate because $A$ is symplectic, and it is closed since

$$d\omega_A = d(j^*\omega) = j^*(d\omega) = 0$$

follows from the commutativity of the exterior derivative with the pullback map.
Therefore, the symplectic structure on $M$ is pulled back to an induced symplectic structure on $A$; likewise, the Hamiltonian $H$ on $M$ is pulled back to a Hamiltonian $H_A$ on $A$ via $H_A = H \circ j$, and the one form $dH$ is pulled back to the section $j^*dH = dH_A$ of $T^*A$. This gives us the triple $(A, \omega_A, H_A)$, which is a Hamiltonian system on the symplectic submanifold $A$. This is the natural induced Hamiltonian system obtained from 'pulling back' the Hamiltonian system $(M, \omega, H)$ to $A$.

Accordingly, the induced Poisson bracket on $A$ is obtained from

$$\{f_A, g_A\}_A = \{f \circ j, g \circ j\}.$$  

Here the functions $f, g$ are arbitrary elements of $C^\infty(M)$ which are pulled back to functions $f_A = f \circ j$ and $g_A = g \circ j$ in $C^\infty(A)$. Finally, the induced Hamiltonian vector field $X_{H_A}$ on $A$ is defined by

$$\iota_{X_{H_A}}^\omega_A = -dH_A.$$  

Let $j_* : TA \to TM$ denote the pushforward of the inclusion map. It follows that $j_*X_{H_A}$ is the restriction of the section $X_H |_{j(A)}$ of $TM |_{j(A)}$ to $j_*TA$.

1.3 Almost symplectic distributions and associated natural structures

The constrained Hamiltonian systems of interest require a particular type of vector subbundles $V \subset TM$ to be given. The dynamics of such a system is generated by the component of $X_H$ that lies in $V$.

To present a natural example of such a system, let us assume that $M$ is completely foliated into a disjoint union of symplectic submanifolds $M = \bigcup \lambda A_\lambda$, with $\lambda$ in some index set. The union of the tangent bundles of all leaves

$$V = \bigcup \lambda TA_\lambda$$

is a vector space subbundle of $TM$, or a distribution over the base manifold $M$.

Because every leaf $A_\lambda$ is symplectic, $\omega_V \equiv \omega|_{V \times V}$ is non-degenerate (that is, for every linearly independent family of vector fields $\{X_i\}$ which locally spans $V$, the $2k \times 2k$-matrix $[\omega(X_i, X_j)]$ is invertible). A distribution $V$ for which $\omega_V$ is non-degenerate will be referred to as an almost symplectic distribution (note that every almost symplectic distribution must necessarily have an even rank).
1.3.1 Integrability and non-integrability

The distribution $V$ obtained in this manner is not only almost symplectic, but also integrable. A distribution $V$ over a base manifold $M$ is, by definition, called integrable or, in the language of analytical dynamics, holonomic, if $M$ is completely foliated into submanifolds in a manner that $V$ is the union of the tangent bundles of all leaves. Notice that the notion of 'holonomy' has a completely different meaning in differential geometry. Otherwise, $V$ is called non-integrable, non-holonomic or bracket generating. The expression 'bracket generating' is due to the following reason.

Every differentiable manifold $M$ admits an intrinsic, skew symmetric, bilinear mapping $[,] : \Gamma(TM) \times \Gamma(TM) \rightarrow \Gamma(TM)$, known as the Lie bracket. $\Gamma(TM)$ denotes the space of sections of $TM$. The Lie bracket maps a pair of vector fields $X$, $Y$ to a vector field $[X,Y]$, whose action on $f \in C^\infty(M, \mathbb{R})$ is given by

$$[X,Y](f) = X(Y(f)) - Y(X(f)).$$

The $j$-th component of $[X,Y]$ in a local chart is given by $X^iY^j_a - Y^iX^j_a$, and the pair $(\Gamma(TM), [\cdot, \cdot])$ has the structure of a Lie algebra. Here, we have introduced the abbreviated notation $(\cdot)_r := \frac{\partial}{\partial x^r}(\cdot)$.

Consider now a vector subbundle $V$ of $TM$ to be given, together with the associated space of sections $\Gamma(V)$. The question whether $V$ is locally integrable or not can be determined by use of a criterion due to Frobenius. The Frobenius theorem states that $V$ is locally integrable if and only if the pair $(\Gamma(V), [\cdot, \cdot])$ is a Lie subalgebra of $(\Gamma(TM), [\cdot, \cdot])$. In other words, a distribution is locally integrable if and only if its associated space of sections is closed under the Lie bracket. If $V$ is non-integrable, there are pairs of sections of $V$ for which the Lie bracket has a non-vanishing component in the complement $\bar{V}$ of $V$ in $TM$. This is the reason for the expression 'bracket generating'.

Although a distribution may be locally integrable, there are topological obstructions to global integrability. A classical example of a necessary condition for global integrability is the requirement that all Pontryagin classes $P_{r(V)}$ of the complementary distribution $V^\perp$ of $V$ have to vanish for $r > 2 \text{rank}(V^\perp)$, cf. for instance [9].

An alternative formulation of Frobenius' theorem can be stated as follows. Denote by $P : TM \rightarrow V$ the projection map, which we pick to be orthogonal with respect to some
auxiliary Riemannian metric $g$, and let $\bar{P}$ be the complementary projection. Then, $V$ is locally integrable if and only if

$$\bar{P}[PX, PY] = 0$$

(1.4)

is satisfied for all sections $X$ and $Y$ of $TM$. So the Frobenius criterion can be understood as a condition that is imposed on the projection operator. In a local chart, one can straightforwardly verify that the condition

$$P_i^r P_j^s (P^k_{r,s} - P^k_{s,r}) = 0$$

(1.5)

is equivalent to (1.4).

The repeated application of the Lie bracket to the spanning vector fields of a distribution defines the flag of a distribution, which is a sequence of subspaces $V_i$ of $TM$, with $V_1 = V$, that is defined by the recursion relation

$$V_i = [V, V_{i-1}];$$

so $V_i$ denotes the subbundle of $TM$ that is spanned by the sections of $V$, and by the Lie bracket commutators of all sections of $V$ and $V_{i-1}$. If all elements $V_i$ of the flag are distributions (that is, if the fibre ranks are all everywhere constant), the flag is called regular.

Because $TM$ is finite dimensional, the flag stabilizes, that is, there is some number $r$ such that $V_i = V_{i-1}$ holds for all $i$ larger than $r$. The smallest number $r$ which satisfies this condition is called the degree of non-holonomy of the distribution $V$. If the case $V_r = TM$ holds, $V$ is called totally non-holonomic or strongly bracket generating.

1.3.2 Existence of a $\omega$-skew selfadjoint bundle projector

In this subsection, we will prove that the bundle projector $P : TM \to V$ can always be picked to be skew-selfadjoint with respect to the symplectic structure. Our proof is inspired by a similar construction presented in [10] for a distribution over a Riemannian manifold. This projector, together with further auxiliary structures that will be defined in the next section, will prove to be a very efficient tool in the analysis of a constrained Hamiltonian system.
Let $V$ be an almost symplectic, rank $2k$ distribution over the symplectic manifold $(M, \omega)$. We claim that there exists a smooth, $g$-orthogonal bundle projector

$$P : TM \to V,$$

$P^2 = P,$ which satisfies

$$\omega(PY, Y') = \omega(Y, PY')$$

for all sections $Y$ and $Y'$ of $TM$. This property will be referred to as $\omega$-skew selfadjointness.

In order to prove this statement, let us pick an arbitrary linearly independent family of vector fields $\{X_1, \ldots, X_{2k}\}$ which locally spans $V$. Since $V$ is almost symplectic, the functions $C_{ij} = \omega(X_i, X_j)$ are the matrix components of a $GL(2k, \mathbb{R})$ valued function on $M$, which has an inverse $C^{ij}$. We claim that $P$ is locally given by the tensor

$$P = C^{ij}X_i \otimes \theta_j.$$  \hfill (1.7)

The one forms $\theta_j$ are defined by $\omega(X_j, \cdot)$.

First of all, it is obvious that the right hand side of the asserted formula (1.7) indeed is a tensor.

The fact that it is also a projector follows from

$$(C^{ij}X_i \otimes \theta_j)^2 = C^{ij}C_{jk}C^{kl}X_i \otimes \theta_l,$$

which uses that $\langle \theta_j, X_k \rangle = C_{jk}$ is given, as one verifies from the definition of $C_{ij}$. Here, $\langle \cdot, \cdot \rangle$ denotes the canonical pairing of one forms and vector fields. From $C^{ij}C_{jk} = \delta^i_k$, one understands that the right hand side of (1.7) is a projector as claimed, whose range, as one easily verifies, is given by $V$.

In order to prove $\omega$-skew self adjointness, we pick two arbitrary vector fields $X$ and $Y$, and insert the asserted formula (1.7) for $P$. One gets

$$\omega(PX, Y) = \omega(C^{kl}X_i \langle \theta_l, X \rangle, Y)$$

$$= C^{kl}\langle \theta_l, X \rangle \omega(X_k, Y)$$

$$= C^{kl}\langle \theta_l, X \rangle \langle \theta_k, Y \rangle$$

$$= -C^{lk}\langle \theta_l, X \rangle \langle \theta_k, Y \rangle$$

$$= -\omega(X_l, X)C^{lk}\langle \theta_k, Y \rangle$$

$$= \omega(X, C^{lk}X_i \langle \theta_k, Y \rangle)$$

$$= \omega(X, PY),$$

which shows that (1.7) is indeed correct. This construction only assumes the invertibility of $C_{ij}$, so it can be carried out for every almost symplectic distribution $V$.  

16
1.3.3 Further auxiliary structures

We will next prove that there exist a Riemannian metric $g$ and an almost complex structure $J$, that is, a smooth bundle isomorphism $J : TM \to TM$ with $J^2 = -1$, which are compatible with the symplectic structure. This precisely means that

$$\omega(X,Y) = g(JX,Y)$$

(1.8)

is satisfied for all sections $X, Y$ of $TM$. In addition, $g$ is hermitean, which means that $g(JX, JY) = g(X, Y)$, so that it is even a Kähler metric (a hermitean metric is Kähler if the associated two form, which is here given by $\omega$, is closed).

Furthermore, we will prove that $g$ can be picked in a manner that $P$ is symmetric with respect to the Riemannian structure, and that it commutes with $J$. The quadruple of compatible data consisting of $\omega$, $P$, $g$, and $J$ will be a very useful technical tool in all that follows.

Let us first address the construction of the almost complex structure $J$ and the Kähler metric $g$. We start by picking an arbitrary smooth Riemannian metric $\tilde{g}$ on $M$, relative to which $P$ shall be symmetric. Such a metric always exists; for instance, one may choose an arbitrary Riemannian metric $g'$ on $M$, and use it to define $\tilde{g}(X, Y) = g'(PX, PY) + g'(PX, PY)$, where $P$ denotes the projection complementary to $P$.

Since the symplectic structure is non-degenerate and smooth, there is a non-degenerate, smooth bundle mapping $K$ which is defined by

$$\omega(X,Y) = \tilde{g}(KX,Y).$$

(1.9)

It is skew symmetric with respect to $\tilde{g}$, because the symplectic structure is antisymmetric, so the adjoint operator of $K$ with respect to $\tilde{g}$ is $K^* = -K$.

The bundle map $K$ can be used to construct an almost symplectic structure. Since $K^*K = -K^2$ is smooth, positive definite, non-degenerate and symmetric, there is a unique smooth, positive definite, symmetric bundle map $A$ defined by $A^2 = -K^2$, which commutes with $K$. This immediately implies that the bundle map $J = KA^{-1}$ satisfies $J^2 = -1$, so that it defines an almost complex structure. This is the standard proof for the fact that every almost symplectic manifold admits an almost complex structure, see for instance [13].

Because $A$ is positive definite and symmetric, one can define a new metric by $g(X,Y) \equiv \tilde{g}(AX,Y)$, which obviously satisfies

$$\omega(X,Y) = g(JX,Y).$$

(1.10)
Moreover, this metric is hermitean, since
\[ g(JX, JY) = \bar{g}(KX, A^{-1}KY) \]
\[ = -\bar{g}(X, K^2A^{-1}Y) \]
\[ = \bar{g}(X, AY) \]
\[ = g(X, Y). \]

Because \( \omega \) is closed, \( g \) is not only hermitean, but even Kähler.

The projector \( P \) is symmetric with respect to \( g \), as one concludes from the following consideration. The fact that \( P \) commutes with \( K \) follows from
\[ \bar{g}(KPX, Y) = \omega(PX, Y) = \omega(X, PY) = \bar{g}(KX, PY) = \bar{g}(PKX, Y) \]
for all \( X \) and \( Y \), since \( P \) is symmetric with respect to \( \bar{g} \). \( P \) commutes with \( K \), so also commutes with \( K^2 \); thus it commutes with \( A^2 \). The linear operator \( A^2 \) is positive definite and symmetric, therefore commutativity with the symmetric operator \( P \) implies that \( P \) also commutes with \( A \). It immediately follows that \( P \) is symmetric with respect to \( g \).

Finally, let us prove that \( P \) commutes with \( J \). To this end, we observe that
\[ g(JPX, Y) = \omega(PX, Y) = \omega(X, PY) = g(JX, PY) = g(PJX, Y) \]
holds for all \( X \) and \( Y \), as a result of which one concludes that \( PJ = JP \) is satisfied, as claimed. This shows, among other things, that the almost complex structure restricts to a bundle map \( J : V \to V \).

If the almost complex structure is covariantly constant with respect to the Levi-Civita connection associated to \( g \), it is a complex structure. If \( \omega \) admits a compatible Kähler metric \( g \) together with a complex structure \( J \), \( M \) is a Kähler manifold.

### 1.4 Definition of the constrained Hamiltonian system

We are finally prepared to define the constrained Hamiltonian systems that will be of interest in this thesis. No other types of constrained Hamiltonian systems will be considered. Assume that a Hamiltonian system \( (M, \omega, H) \) is given, together with a rank \( 2k \) almost symplectic distribution \( V \).
Furthermore, let $X_H$ denote the Hamiltonian vector field associated to $(M, \omega, H)$, and let $P$ be the $\omega$-skew selfadjoint bundle projector $P : TM \to V$. The dynamical system of interest is defined by the vector field that corresponds to the component of $X_H$ in $V$,

$$X^c_H \equiv PX_H. \quad (1.11)$$

The notion of a constrained system is due to the fact that the components of the Hamiltonian vector field in the symplectic complement $V^\perp$ of $V$ are constrained to be zero. The dynamical system determined by $X^c_H$ will henceforth be referred to as the constrained Hamiltonian system associated to $(M, \omega, H, V)$. The flow map generated by this vector field is denoted by

$$\Phi^c_t : M \to M.$$  

The orbits associated to this flow obey the equations of motion

$$\dot{x}(t) = X^c_H |_{x(t)}. \quad (1.12)$$

Evidently, they are tangent to the distribution $V$. Using the compatible quadruple $(J, g, \omega, P)$ constructed above, (1.12) can be written as

$$\dot{x} = PJ \text{grad} H = JP \text{grad} H, \quad (1.13)$$

where the gradient is defined relative to the Kähler metric $g$. We recall that the gradient of a smooth function $f$ is defined by $df = g(\text{grad} f, \cdot)$. 

Due to the $\omega$-skew selfadjointness of $P$, one can straightforwardly write down a generalization of the Poisson bracket in terms of

$$\{f, g\}_D \equiv \omega(PX_f, PX_g) \quad (1.14)$$

$$\equiv \omega_V(X_f, X_g),$$

where $X_f$ and $X_g$ are the respective Hamiltonian vector fields of the smooth functions $f$ and $g$. This antisymmetric pairing of functions is again a first order derivative in both arguments, and is referred to as the Dirac bracket.

Time derivatives along orbits of the constrained system can, in direct analogy to Hamiltonian dynamics in the Poisson picture, be written in the form

$$\dot{f} = \{f, H\}_D.$$
The pair \((C^\infty(M), \{\cdot, \cdot\}_D)\) in general fails to be a Lie algebra. This is because Jacobicity does not hold for the Dirac bracket if \(V\) is non-integrable (one observes that Jacobicity holds for the Dirac bracket if and only if the Frobenius condition on \(P\) is satisfied).

The antisymmetry of the Dirac bracket implies that the Hamiltonian function \(H\) of the unconstrained system remains an integral of motion, since

\[
\dot{H} = \{H, H\}_D = 0.
\]

Therefore, the energy level surfaces \(\Sigma_E\) of the \(H\) are invariant manifolds with respect to the flow of the constrained Hamiltonian system presently described.

### 1.4.1 Integrable distributions and symplecticity

The Hamiltonian flow is a one parameter group of symplectomorphisms. In case of a constrained Hamiltonian system, this result can be generalized if the almost symplectic distribution \(V\) is integrable.

In fact, the one parameter group of two forms \(\Phi_t^*\omega\) obtained from pulling back the symplectic structure via the flow of the constrained system obeys

\[
\frac{d}{dt}\Phi_t^*\omega = \Phi_t^*L_{X^t}\omega = \Phi_t^*(di_{X^t}\omega + i_{X^t}d\omega) = -\Phi_t^*d(P^i_kH_i dx^k) = -\Phi_t^*(P^i_kH_i dx^l \wedge dx^k) = -\frac{1}{2}\Phi_t^*((P^i_k - P^i_l)H_i dx^l \wedge dx^k).
\]

Hence, the restriction of \(\frac{d}{dt}\Phi_t^*\omega\) to \(V\) yields that

\[
\frac{d}{dt}\Phi_t^*\omega \big|_{V\times V} (X, Y) = -\frac{1}{2}\Phi_t^*((H_i(P^i_k - P^i_l)P^k_rP^l_sX^rY^s))
\]

is satisfied for all sections \(X\) and \(Y\) of \(V\). The Frobenius condition states that the expression on the right hand side vanishes everywhere if and only if \(V\) is holonomic.

This implies that for integrable \(V\), the restriction of \(\Phi_t^*\omega\) to \(V\) is time independent, as a result of which it necessarily coincides with the value at the initial time \(t = 0\), which is \(\omega \big|_{V\times V} = \omega_V\). So integrability of \(V\) implies that

\[
\Phi_t^*\omega \big|_{V\times V} = \omega_V
\]
holds for all times. This relationship generalizes the property of symplecticity of the Hamiltonian flow.

A direct corollary that generalizes the Liouville theorem in the context of constrained Hamiltonian systems that are defined in terms of integrable, almost symplectic distributions, is expressed by

$$\Phi^*_t \omega^k |_{V^*} = \omega^k_V.$$  

The interpretation hereof is that the restriction of $\omega^k_V$ to every integral manifold of $V$ precisely corresponds to the symplectic volume form that is induced by the Liouville measure on the embedding space $M$ (up to a factor $\frac{1}{k}$).

The present considerations conclude this chapter, whose main emphasis has been dedicated to the formulation of the basic elements of constrained Hamiltonian systems. In the following chapter, the structure of the set of equilibria will be investigated in some detail. It possesses a wealth of beautiful structure, of which we will only be able to describe a small fraction, due to the limited scope of this thesis.
CHAPTER 2

THE GEOMETRY AND TOPOLOGY OF THE CRITICAL MANIFOLD

The set of equilibrium solutions of a dynamical system is generally referred to as its critical set. The main purpose of this chapter is to study geometrical and global topological properties of the critical sets \( C \) exhibited by the type of constrained Hamiltonian systems introduced in the first chapter. We will mainly be interested in structural questions about \( C \), with an emphasis on the generic case. For instance, an application of Sard's theorem will demonstrate that generically, \( C \) is a smooth \( 2(n - k) \)-dimensional submanifold \( C_{\text{gen}} \subset M \), which we will refer to as the critical manifold.

Furthermore, we will show that there is a fairly rigid relationship between \( C_{\text{gen}} \) and the critical set of the unconstrained Hamiltonian system if \( H \) is picked to be a Morse function. Our main result will be to prove that for compact and closed \( M \), the Poincaré polynomials of \( M \) and those of the connectivity components of \( C_{\text{gen}} \) are related in a way reminiscent of the Morse-Bott inequalities.

We will in fact give two different proofs of this result. We will first apply the generalization of Morse and Morse-Smale theory [8, 23, 29, 34] developed by C. Conley and E. Zehnder in [16] to an auxiliary gradient-like system that exhibits the same critical set as the physical system. Our second proof is based on comparing the Morse-Witten complexes of \( (M, H) \) and \( (C_{\text{gen}}, H|_{C_{\text{gen}}}) \), and gives an alternative geometrical explanation of same results.

We only address general constrained Hamiltonian systems as introduced in chapter one in this chapter. The special case of mechanical systems will be analyzed in chapter four.

The following definitions will be useful for the subsequent discussion. The zeros of the one form \( dH \) are called critical points of \( H \), and the value of \( H \) at a critical point is called a critical value. A level surface \( \Sigma_E \) that contains a critical value \( E \) of \( H \) is called a critical level surface. A level surface \( \Sigma_E \) that contains no critical points of \( H \) is called regular, and consequently, \( E \) is then called a regular value. Furthermore, we will use the Kähler metric \( g \)
and the almost complex structure $J$ constructed in chapter one, which are compatible with the symplectic structure $\omega$.

### 2.1 Local properties of the critical manifold

The critical set of the constrained Hamiltonian system \((1.12)\) is given by

$$C = \{a \in M \mid X_H^c(a) = 0\}, \quad (2.1)$$

and can, as one reads off from the representation \((1.13)\) of the equations of motion, also be described in terms of

$$C = \{a \in M \mid P\text{grad}H|_a = 0\}. \quad (2.2)$$

The first claim that we will prove is that generically, $C$ is a $2(n - k)$-dimensional submanifold of $M$. To this end, we use an open cover \(\{U_\alpha\}\) of $M$, that is sufficiently fine so that every $TU_\alpha$ is a trivial bundle. Then, picking one of its elements $U \subset M$, we introduce a local $g$-orthonormal family of vector fields \(\{X_i\}, i = 1, \ldots, 2k\), which spans those fibres of $V$ whose base points lie in $U$.

Clearly, $C \cap U$ is the set where all of the $C^\infty(U)$-functions

$$f_i \equiv g(X_i, P\text{grad}H) = g(PX_i, \text{grad}H) = X_i(H)$$

have the value zero. Notice here that $PX_i = X_i$, since $X_i$ is a section of $V$. From Sard's theorem, we know that the set of regular values of any smooth $f_i$ is dense in $f_i(U)$, hence one can generically assume that the value zero, if it is contained in $f_i(U)$, is regular for all $f_i$. This means that $df_i$ is generically nowhere zero on the level set $f_i^{-1}(0) \subset U$.

In case zero is a regular value that is contained in $f_i(U)$, the level set $f_i^{-1}(0)$ is an orientable, smooth hypersurface in $U$ [23]. If there exists an index $j$ for which $0 \not\in f_j(U)$, the intersection of $C$ with $U$ is empty; otherwise, one finds that

$$C \cap U = \cap_i f_i^{-1}(0).$$

If the one-forms $df_i$ are linearly independent everywhere on $C \cap U$, the intersection of the $2k$ hypersurfaces $f_i^{-1}(0)$ in $U$ is a $2(n - k)$-dimensional submanifold of $U$. Again, we can use Sard's theorem, which states that for any smooth mapping $f : U \to \mathbb{R}^{2k}$, the set of regular
points in $f(U) \subset \mathbb{R}^{2k}$ is dense [30], so that generically, 0 is a regular value (by definition, a regular point is an element $c \in f(U)$ for which the pushforward $f_* : TU \to T\mathbb{R}^{2k}$ has a full rank at every point in $f^{-1}(c)$). For a generic choice of $V$ and $H$, all of the above applies to the map $f \equiv (f_1, \ldots, f_{2k})$.

Applying this argument to every open neighborhood of the open cover $\{U_\alpha\}$ of $M$, one concludes that $C$ generically is a $2(n - k)$-dimensional submanifold of $M$.

However, instead of studying the generic case, we will use the fairly artificial, but more general assumption that all connected components of $C$ are submanifolds of $M$ with a constant dimension, and that they are, in addition, compact and closed. Notice that generically, subsets of $C$ that are not $2(n - k)$ dimensional submanifolds of $M$ do not even need to be manifolds at all.

Requiring that $C \cap U$ has a constant dimension is equivalent to requiring that the number of $df_i$'s which are linearly independent does not vary on the set $\cap_i f_i^{-1}(0)$. Denoting this number by $2k - d$, $\cap_i f_i^{-1}(0)$ consists of $2k - d$ transversely intersecting hyperplanes, which implies that it is a $2n - 2k + d$-dimensional submanifold of $U$. Thus, the dimension of every connected component of $C$ is, under the given assumptions, bounded from below by $2(n - k)$.

### 2.1.1 A generalized Hessian for the constrained system

We will here find a generalization of the Hessian, in order to make it available for the constrained system. It will turn out to be a useful item in the subsequent discussion.

Let us consider the case at hand, that is, a symplectic manifold $(M, \omega)$ together with a Hamiltonian $H : M \to \mathbb{R}$, and let $a$ denote a critical point of $H$. The matrix of second derivatives $H_{rs}(a)$ in a chart containing $a$ (‘, ‘ denotes the partial derivative $\frac{\partial}{\partial x^r}$) is called the Hessian of $H$ at $a$ relative to this chart.

The coordinate invariant definition of the Hessian is as follows. We let $\nabla$ denote the Levi-Civita connection of the Kähler metric $g$. In a local coordinate chart, its action on a one form $\theta = \theta_i dx^i$ is given by

$$\nabla_X \theta = (X^r \theta_i, r - \Gamma^s_{ri} X^r \theta_s) dx^i,$$

where $X = X^r \partial_r$ is a vector field. Its connection coefficients are the Christoffel symbols

$$\Gamma^i_{rs} = \frac{1}{2} g^{ij} (g_{rjs} + g_{rsj} - g_{rjs}).$$
Some important properties of $\nabla$ are for instance that $\nabla_\lambda \varphi = 0$ holds for all vector fields $\lambda$, and that it is torsion free, that is, the Christoffel symbols are symmetric in the lower indices.

The tensor $\nabla dH$ is regarded as the invariant version of the Hessian of $H$ [25], but it is of course not only defined at its critical points. It acts on pairs of vector field in terms of

$$\nabla (dH)(X, Y) \equiv \langle \nabla_X dH, Y \rangle$$

$$= (H_{ir} - \Gamma_{ri}^s H_{is}) X^r Y^s$$

in a local coordinate chart. It is a symmetric bilinear form, as one observes from the fact that the expression in the bracket is symmetric with respect to the lower indices. At the critical points of $H$, one has $H_{is} = 0$, and the non-vanishing term is determined by the matrix of second derivatives of $H$, which shows that both definitions of the Hessian agree.

A generalized Hessian for the constrained Hamiltonian system can easily be defined along these lines. Let us write $P^T$ for the dual projector associated to $P$ which acts on sections of the cotangent bundle $T^* M$. Given any one-form $\theta$, it satisfies

$$\langle P^T \theta, X \rangle = \langle \theta, PX \rangle,$$

which can be considered as a definition. The superscript should point out that in every coordinate chart, the matrix of $P^T$ is simply the transpose of the matrix of $P$.

The obvious coordinate invariant generalization of the Hessian is the tensor $\nabla (P^T dH)$. It acts as a bilinear form on pairs of vector fields in terms of

$$\nabla (P^T dH)(X, Y) \equiv \langle \nabla_X (P^T dH), Y \rangle$$

$$= ((P^T_{i} H_{j}), s - \Gamma_{ri}^s P^T_{i} H_{j}) X^r Y^s.$$

The second term in the bracket on the lower line is zero on $C$, because $P^T_{i} H_{j} = 0$ on $C$ (notice that $g(P \text{grad} H, \cdot) = P^T dH$, and by definition, $C$ consists of the zeros of $P \text{grad} H$). The non-vanishing term in $\nabla (P^T dH)$ on $C$ is determined by the matrix

$$K_{rs} \equiv (P^T_{i} H_{j}), s.$$

(2.3)

A straightforward calculation, where one uses $P^2 = P$ in its derived form, shows that

$$P^T_{i} K_{jk} = K_{ik}$$
holds everywhere on $C$. Obviously, the rank of $K$ is bounded from above by the rank of $P$ of value $2k$, corresponding to the rank of the vector bundle $V$.

The corank of $K|_a$ equals the dimension of the connectivity component of $C$ which contains $a$ (we remember that we assume that all connectivity components of $C$ are closed, compact manifolds with a constant dimension). This is because $K_{ik}$ is the local Jacobian matrix of the component vector $P_i H_j$ of $P^T dH$, so that for any $a' \in C$ close to $a$, Taylor's theorem implies that

$$(P_i H_j)(a') = (P_i H_j)(a) + K|_a(a' - a) + O(||a' - a||^2).$$

The left hand side and the first term on the right hand are both zero, since $C$ is so defined, thus $K(a' - a)/||a' - a||$ must converge to zero in the limit $||a' - a|| \to 0$. That is, the tangent space $T_a C$ is equal to ker$K|_a$, hence the corank of $K|_a$ is indeed the dimension of the component $C_i$.

### 2.1.2 An auxiliary gradient-like flow

The physical flow of the constrained Hamiltonian system is not a very powerful tool in the study of the global topology of $C$. The reasons will be explained below. It turns out that the auxiliary dynamical system on $M$ defined by

$$\dot{x} = - P \text{grad} H|_x$$

is much better suited. We will denote its flow by

$$\phi_t : M \to M.$$  \hfill (2.4)

Evidently, its orbits are tangent to $V$, and it is also clear that both $\Phi_t^c$ and $\phi_t$ exhibit the same critical set $C$.

The property that makes this auxiliary system so special is the fact that its flow is gradient-like. By definition, a flow is gradient-like if there exists a function $f : M \to \mathbb{R}$ that decreases strictly along all of its non-constant orbits. One can easily check that the Hamiltonian $H$ decreases strictly along the non-constant orbits of $\phi_t$, due to

$$\frac{d}{dt} H(x(t)) = \langle dH(x(t)) , \dot{x}(t) \rangle$$

$$= g(\text{grad} H , -P \text{grad} H)(x(t))$$

$$= -g(P \text{grad} H , P \text{grad} H)(x(t)) \leq 0.$$  \hfill (27)
At this point, we have used the fact that $P$ is symmetric with respect to the Kähler metric. This has been one of the main motivations for the construction of the compatible quadruple $(g,J,\omega,P)$.

It is immediately clear that $\phi_t$ generates no periodic trajectories, hence $C$ comprises all of its invariant sets. In contrast, invariant sets of $\Phi_t^C$ can contain periodic orbits.

### 2.1.3 The gradient-like flow in a vicinity of $C$

Let us briefly consider the orbits of $\phi_t$ in a small vicinity of $C$, mainly to gain some intuition. So we choose an arbitrary element $a \in C$ and local coordinates $x^i$, with the origin at $a$. In this chart, the equations of motion (2.4) are given by

$$\frac{d}{dt}x^i(t) = -(g^{ij}P^s_jH_s)(x(t)).$$

In linear approximation, this reduces to

$$\frac{d}{dt}x^i(t) = -(g^{ij}K_{js})(a)x^s(t),$$

where $K_{js}$ is the 'constrained Hessian' at $a$.

For notational convenience, the arguments 'a' will from now on be omitted, since all considerations in this and the following few subsections are strictly local, and are formulated relative to the same base point unless otherwise mentioned. $P$ will from now on denote the matrix of the indicated projector in the present chart. Moreover, we will write $K^2 = [g^{ij}K_{jk}]$ for the Jacobian matrix of $P\text{grad}H$ at $a$ in this chart, and use the notation

$$\mathcal{S} : S \mapsto S^d$$

to denote the mapping that associates to every matrix valued function $S_{rs}$ on $C$ the matrix valued function $S^d = [g^{ij}S_{jk}]$ (in components, it is the operation that simply raises an index of $S$ using the metric tensor). If $S$ is (anti-)symmetric as a matrix, the operator $S^d$ is (anti-)symmetric with respect to $g$, as one concludes from

$g(S^d_{ik}v^i,w) = g_{ij}g^{ik}S_{jk}v^jw^j = S_{ik}v^iw^j,$

which holds for all vectors $v$ and $w$ in $T_aC$ (the metric is here of course evaluated at $a$).

The linearized equations of motion at $a$ thus read

$$\frac{d}{dt}x = -K^d x.$$
Due to the $g$-orthogonality of $P$, $PK^t = K^t$ holds, as one observes from

$$P^i g^{kl} K_{lm} = g^{ik} P^l K_{im} = g^{ik} K_{km}.$$  

The constraints $\dot{x} = P \dot{x}$ are thus fulfilled in linear approximation, as it should be.

Let $\tilde{Q}$ denote the $g$-orthoprojector onto $\ker K^t = T_a \mathcal{C}$. Its complement $Q$ projects $T_a M$ orthogonally onto the fibre $N_a \mathcal{C}$ of the normal bundle of $\mathcal{C}$. The orbits of the linearized flow in the vicinity of $a$ are then given by

$$x(t) = \exp(-t PK^t Q)x_0 = \exp(-t PK^t Q)Qx_0 + \tilde{Q}x_0,$$

where the initial condition is determined by $x(0) = x_0$. The vector $\tilde{Q}x_0$ lies in $T_a \mathcal{C}$, and approximately connects $a$ with a critical point close to $a$, while $Qx_0$ lies in $N_a \mathcal{C}$. The insight here is that the gradient-like flow in the vicinity of $\mathcal{C}$ is determined by the operator $PK^t Q = K^t$.

Figure 2.1: The gradient-like flow in the vicinity of $\mathcal{C}_{gen}$. The axes denote the images of the respective projectors, and the dotted lines schematically account for the orbits of the linearized flow (assuming that $\mathcal{C}_{gen}$ is an attractor).
2.1.4 Local normal hyperbolicity

The critical manifold $\mathcal{C}$ is locally normal hyperbolic at the point $a \in \mathcal{C}$ with respect to the gradient-like flow $\phi_t$ if the stable and unstable manifolds intersect $\mathcal{C}$ transversely at $a$. This condition is satisfied if and only if the restriction of $K^s$ to the normal space $N_a \mathcal{C}$ is non-degenerate (that is, if the matrix $QK^sQ$ has the same rank as $K^s$), and if in addition, the spectrum of $QK^sQ$ purely consists of eigenvalues with non-vanishing real parts. The property of normal hyperbolicity is important for our later discussion of the topology of $\mathcal{C}$.

To investigate under which circumstances local normal hyperbolicity is given, we make the trivial observation that

$$\text{rank}\{QK^sQ\} \leq \text{rank}\{Q\} = \dim N_a \mathcal{C}.$$ 

Owing to $QK^sQ = QPK^sQ$, the condition of local normal hyperbolicity implies that

$$\text{rank}\{QP\} = \text{rank}\{Q\} \leq \text{rank}\{P\}.$$ 

Furthermore, $Q$ and $P$ are both symmetric with respect to $g$, so that the ranks of $QP$ and $PQ$, being a pair of mutually adjoint operators, are equal. Hence, local normal hyperbolicity is equivalent to the requirement that

$$\text{rank}\{QP\} = \text{rank}\{PQ\} = \text{rank}\{Q\}.$$ 

The mapping $QPQ : N_a \mathcal{C} \rightarrow N_a \mathcal{C}$ is an isomorphism. To see this, we notice that $QPQ = QP(QP)^*$ holds, where the star produces the adjoint with respect to $g$. Since for every linear operator $B$ on $T_a M$, the rank of $BB^*$ is the same as the rank of $B$, local normal hyperbolicity implies that

$$\text{rank}\{QPQ\} = \text{rank}\{Q\} = \dim N_a \mathcal{C}. \quad (2.5)$$

The assertion immediately follows.

On generic, $2(n-k)$-dimensional connectivity components of $\mathcal{C}$, the rank of $K^s$ is $2k$. In this special case, local normal hyperbolicity is equivalent to the condition that $PK^sP$ has a full rank. This is because

$$PK^sP = PK^sQP$$

has a full rank if and only if $QP$ does. The ranks of $Q$ and $P$ are both $2k$, therefore non-degeneracy of $QP$ implies that $QPK^s = QK^sQ$ also has maximal rank.
The geometric picture behind this argument is that local normal hyperbolicity expresses the fact that $Q$ projects the fibre $V_a$ surjectively onto the normal space $N_aC$, and that $P$ projects $N_aC$ injectively into $V_a$. If the rank of $K^t$ is maximal, the local dimension of $C$ is $2(n - k)$, and the orthogonal projections between $V_a$ and $N_aC$ are isomorphisms. In this specific case, $T_aC$ contains no linear subspace of $V_a$. A connected component of $C$ which is everywhere locally normal hyperbolic relative to $\phi_t$ is called normal hyperbolic.

2.2 Geometry of non-degenerate critical manifolds

We recall that $C$ is assumed to be a compact, boundary-free smooth submanifold of $M$, whose connected components have a constant dimension. We will prove that all generic, $2(n - k)$-dimensional connectivity components are necessarily normal hyperbolic. Moreover, we will prove that if $H$ is a Morse function on $M$, the $2(n - k)$ dimensional components of $C$ contain all critical points of $H$, but that there are no other extrema of the function $H|_C$. Finally, we will demonstrate that if a connectivity component of $C$ with a dimension larger than $2(n - k)$ is normal hyperbolic, it must be contained in an single energy level surface.

2.2.1 Morse functions and non-degenerate critical manifolds

Let us recall some standard definitions from Morse- and Morse-Bott theory that will be of use in the subsequent discussion.

The dimensions of the zero and negative eigenspaces of the Hessian of $f$ at a critical point $a$ are called the nullity and the index of the critical point $a$. If all critical points of $f : M \to \mathbb{R}$ have a zero nullity, $f$ is called a Morse function, and the index is then called the Morse index of $a$.

If the critical points of $f$ are not isolated, but elements of critical manifolds that are non-degenerate in the sense of Bott, $H$ is called a Morse-Bott function [7]. Throughout this text, we will assume that $H$ is a Morse function.

The critical manifold $C$ of the auxiliary gradient-like flow $\phi_t$ is called non-degenerate if every connected component is normal hyperbolic. The number of eigenvalues of $K^t$ with a negative real part on a non-degenerate connected component $C_t$ of $C$ will be called the index of $C_t$. 

31
2.2.2 Generic connectivity components are normal hyperbolic

We will now prove that all generic, \(2(n - k)\)-dimensional connectivity components of \(C\) are necessarily normal hyperbolic.

To this end, let us first consider an arbitrary connectivity component \(C_i\). Picking some chart at \(a \in C_i\), the eigenspaces of the matrix \(K^a\) (we will from here on continue omitting writing the argument 'a') in \(T_aM\) are contained in the image of \(P\), because for any eigenvalue \(\lambda\) and associated eigenspace \(E_\lambda\) of \(K^a\), one has \(K^a E_\lambda = PK^a E_\lambda = \lambda PE_\lambda\), since \(PK^a = K^a\). This implies that the eigenspaces and eigenvalues of \(K^a\) equal those of the matrix

\[
A^a \equiv PK^a P.
\]

Let us make a digression to briefly point out some properties of the matrix \(A_{rs}\), defined by \(A^2 = [g^{ir}A_{rs}]\), which will be useful later. We denote its symmetric part by \(A^+\), and its antisymmetric part by \(A^-\). Clearly, \(A^\pm\) is (anti-)selfadjoint with respect to the metric \(g|_a\) on \(T_aM\). Introducing the matrices

\[
R_{ij}^\pm := \frac{1}{2} H_{ij} P^r_i P^s_j \left( P_{r,s}^k \pm P_{s,r}^k \right) |_{a},
\]

a small computation shows that \(A^+ = R^+ + P^T D^2_a H P\), where \(D^2_a H\) denotes the Hessian of \(H\) at \(a\) in the present chart, and where the transpose of the matrix of \(P\) is taken relative to the Euclidean metric. The antisymmetric part of \(A\) is given by \(A^- = R^-\), and vanishes if \(V\) is integrable, as a consequence of the Frobenius condition (1.5).

Next, we consider a generic component \(C_i\) of \(C\), with a dimension \(2(n - k)\). We have proved above that \(C_i\) is locally normal hyperbolic in the vicinity of \(a\) if and only if \(\text{rank}\{PK^a P\} = 2k\). This in turn holds if and only if \(\text{rank}\{A^2\} = 2k\). Since the eigenspaces of \(K^a\) coincide with those of \(A^2\), lack of normal hyperbolicity (which is given if \(\text{rank}\{A^2\} < 2k\)) would imply that there are fewer non-zero eigenvalues of \(K^a\) than \(2k\), although the rank of \(K^a\) is \(2k\), which is a contradiction. It follows that the generic connectivity components of \(C\) are necessarily normal hyperbolic.

A loss of rank would not be associated with the vanishing of normal hyperbolicity, but with the emergence of a dimensional degeneracy (which is not admitted by our assumption that all connected components of \(C\) shall have a constant dimension). That is, a critical manifold which is not assumed to have a constant dimension everywhere may exhibit regions with a dimension \(2(n - k)\), which are necessarily normal hyperbolic, and regions of a higher dimension.

32
The above arguments applied to the generic components of $C$ do not hold for connectivity components with a dimension higher than $2(n-k)$. The reason is that although normal hyperbolicity implies that $PK^2P$ has a full rank, the converse is not true. One would have to prove the stronger condition that $QK^2Q$ has a full rank.

2.2.3 Non-degenerate critical manifolds and critical points of $H$

The Hamiltonian $H$ is, as we have pointed out, assumed to be a Morse function. We will now prove under the assumption that $C$ is non-degenerate, that every non-generic connectivity component is a submanifold of a single level surface of $H$, and that the generic connectivity components contain all critical points of $H$, but no other extrema of $H|_C$.

We notice first that obviously, every critical point of $H$ is also a zero locus of $P\nabla H$, and thus an element of $C$. Hence, every critical point of $H$ is an element of $C$.

We claim that more specifically, the critical points of every Morse function $H$ are elements of the generic, $2(n-k)$-dimensional connectivity components of $C$. To prove this, let $a$ denote a critical point of $H$, and let us pick a sufficiently small open neighborhood $U(a) \subset M$ so that $C \cap U(a)$ is connected. To determine the dimension of $C \cap U(a)$, we observe that at $a$, the matrix $K$ reduces to $K_{ik} = (P^r H_r)_{ik}$, because $H_r$ vanishes. By definition of a Morse function, the Hessian $H_{rk}$ is an invertible matrix. Hence, the rank of $K$ equals the rank of $P$, which is $2k$. We have proved above that the corank of $K$ equals the dimension of $T_aC$, thus $U(a) \cap C$ is a smooth, $2(n-k)$-dimensional submanifold of $U(a) \subset M$.

Conversely, we claim that every critical point of $H|_C$ on a non-degenerate component necessarily is a critical point of $H$. To prove this, let $C_i$ have any dimension $\geq 2n - 2k$. Normal hyperbolicity implies that $QP$ is an automorphism of the normal space $N_aC_i$ for arbitrary $a \in C_i$ (as a consequence of normal hyperbolicity, cf. (2.5)). Let $a$ denote an extremum of $H|_{C_i}$. Then,

$$\langle dH, v \rangle|_a = g(\nabla H, v)|_a = 0$$

holds for every vector $v \in T_aC_i$. Consequently, $\nabla H$ is a vector in $N_aC$, that is, $\nabla H = Q\nabla H$. Moreover, by definition of $C$, the condition $P\nabla H = 0$ is satisfied at $a \in C_i$, which in turn implies $QPQ\nabla H = 0$. But since $\nabla H \in N_aC_i$, and since $QPQ : N_aC_i \to N_aC_i$ is an isomorphism, $\nabla H = 0$ must necessarily be satisfied at $a$.

A corollary of this result is that if $C$ is non-degenerate, all critical points of $H$ are elements of the generic connectivity components. In addition, there are no extrema of $H|_C$
on non-generic connectivity components, as a consequence of which each of them must be a submanifold of a level surface of $H$.

If one does not require that $C$ is normal hyperbolic, but only assumes that every connected component of $C$ has a constant dimension, and that $H$ is a Morse function on $M$, it is still true that all components $C_i$ of dimension $2(n - k)$ are normal hyperbolic. Furthermore, it is also still true that no critical manifold of dimension larger than $2(n - k)$ contains any critical points of $H$, but without the requirement of normal hyperbolicity, $H$ may now exhibit conditional extrema on these components.

2.2.4 A note concerning integrable distributions

If the almost symplectic distribution $V$ is integrable, in case of which we call it symplectic, the above results imply that its integral manifolds intersect $C_{gen}$ transversely in finitely many points. This is because the dimension of an integral manifold is $2k$, while the dimension of $C_{gen}$ is $2(n - k)$. Therefore, on every integral manifold of a Hamiltonian system with integrable constraints, the equilibria are generically isolated points on the integral manifolds.

2.3 Global topology of non-degenerate critical manifolds

We are now prepared to address questions about the global topology of $C$. We will assume that $C$ is non-degenerate, closed, and compact. In this situation, a generalization of the Morse-Bott inequalities can be derived for $C$, which is based on an application of the theory of C. Conley and E. Zehnder for flows on Banach spaces with compact invariant sets. We begin this section with a short summary of the most crucial elements in the proof of the Conley-Zehnder (CZ) inequalities, and refer to [16] for the full proof. It is assumed that the reader is familiar with the basic concepts of homology and cohomology theory; for a fast introduction, consider for instance [25], and for more details, [12, 37].

2.3.1 Index pairs and relative (co-)homology

We will now roughly outline parts of the theory underlying the CZ inequalities. It is highly recommendable to consult the original publication [16] in order to grasp the concepts in full detail.
Let $C_i$ be a closed, compact component of $C$ that need not be non-degenerate. An index pair associated to $C_i$ is a pair of compact sets $(N_i, \tilde{N}_i)$ that possesses the following properties. The interior of $N_i$ contains $C_i$, and moreover, $C_i$ is the maximal invariant set under $\phi_t$ in the interior of $N_i$. $\tilde{N}_i$ is a compact subset of $N_i$ that has empty intersection with $C_i$, and the trajectories of all points in $N_i$ that leave $N_i$ at some time under the gradient-like flow $\phi_t$ intersect $\tilde{N}_i$. Therefore, $\tilde{N}_i$ is called the exit set of $N_i$.

Let us briefly recall the definition of relative homology and cohomology groups of a pair of manifolds, with coefficients in $\mathbb{R}$, $\mathbb{Z}$ or $\mathbb{Z}_2$, depending on whether it is orientable or not. The $p$-th relative homology group $H_p(X, A)$ of a pair of manifolds $A \subset X$ is defined in the following manner. Let $C_p(X)$ be the group of $p$-cycles of $X$. Since $C_p(A)$ is a subgroup of $C_p(X)$, the quotient group $C_p(X)/C_p(A)$ is well-defined. Let $B_p(X, A)$ denote the subgroup of $C_p(X)/C_p(A)$ which consists of boundaries. The associated quotient group is the $p$-th relative homology group $H_p(X, A)$.

The $p$-th relative de Rham cohomology group $H^p(X, A)$ consists of cohomology classes that are represented by closed $p$-forms on $X$ whose restriction to $A$ (via the pullback of the embedding $A \rightarrow X$) is exact. If $X$ and $A$ are both orientable, the coefficients of the relative homology and cohomology groups can be picked from $\mathbb{R}$ or $\mathbb{Z}$; otherwise, one uses $\mathbb{Z}_2$.

### 2.3.2 Computing the relative cohomology of an index pair

We will now determine the relative cohomology $H^*(N_i, \tilde{N}_i)$ of an index pair $(N_i, \tilde{N}_i)$ associated to a connectivity component $C_i$. It was proved in [16] that the homotopy type of the pointed space $N_i/\tilde{N}_i$ only depends on $C_i$, so that $H^*(N_i, \tilde{N}_i)$ is independent of the particular choice of index pairs (the space $N_i/\tilde{N}_i$ is obtained from collapsing the subspace $\tilde{N}_i$ of $N_i$ to a point). The equivalence class $[N_i/\tilde{N}_i]$ of pointed topological spaces under homotopy only depends on $C_i$, and is called the Conley index of $C_i$. In the present analysis, we will only consider flows exhibiting normal hyperbolic critical manifolds. In this special case, the result essentially reduces to a generalization of Morse-Bott theory.

Let us attempt to construct an index pair for $C_i$, for which the relative cohomology can be easily computed. To this end, consider, for some sufficiently small $\varepsilon_0 > 0$, a compact tubular $\varepsilon_0$-neighborhood $U$ of $C_i$ (of dimension $2n$), and let

$$W^{\varepsilon_0}_{in}(C_i) := (W^-(C_i) \cup C_i) \cap U$$
denote the intersection of the center unstable manifold of $C_i$ with $U$. $W^-(C_i)$ denotes the unstable manifold of $C_i$. Pick some small, positive $\epsilon < \epsilon_0$, and let $U_\epsilon$ be the compact tubular $\epsilon$-neighborhood of $W_u^{cu}(C_i)$ in $U$.

It is clear that letting $\epsilon$ continuously go to zero, a homotopy equivalence of tubular neighborhoods is obtained, which exhibit $W_u^{cu}(C_i)$ as a deformation retract. We let

$$ U_\epsilon^{out} := \partial U_\epsilon \cap \phi(I(U_\epsilon)) $$

denote the intersection of $\partial U_\epsilon$ with all orbits of the gradient-like flow that contain points in $U_\epsilon$. Then, evidently, $(U_\epsilon, U_\epsilon^{out})$ is an index pair for $C_i$, and by letting $\epsilon$ continuously go to zero, $U_\epsilon^{out}$ is homotopically retracted to $\partial W_u^{cu}(C_i)$.

Thus, homotopy invariance implies that the relative cohomology groups obey

$$ H^*(U_\epsilon, U_\epsilon^{out}) \cong H^*(W_u^{cu}(C_i), \partial W_u^{cu}(C_i)). $$

Due to the normal hyperbolicity of $C_i$ with respect to the gradient-like flow, $W_u^{cu}(C_i)$ has a constant dimension $n_i + \mu(C_i)$ everywhere, where $n_i = \dim C_i$. Therefore, one obtains from Lefschetz duality [17] that

$$ H^{n_i + \mu_i - p}(W_u^{cu}(C_i), \partial W_u^{cu}(C_i)) \cong H_p(W_u^{cu}(C_i) \setminus \partial W_u^{cu}(C_i)),$$

where $\mu_i = \mu(C_i)$. It is clear that $C_i$ is a deformation retract of the interior of $W_u^{cu}(C_i)$, so that the respective cohomology groups are isomorphic. Because the dimension of $C_i$ is $n_i$, Poincaré duality implies that

$$ H_p(W_u^{cu}(C_i) \setminus \partial W_u^{cu}(C_i)) \cong H_p(C_i) \cong H^{n_i - p}(C_i),$$

so that after the substitution $q = n_i - p$, we find

$$ H^{q + \mu_i}(U_\epsilon, U_\epsilon^{out}) \cong H^p(C_i), \quad (2.7) $$

which is the desired result.
2.3.3 The Conley-Zehnder inequalities for non-degenerate critical manifolds

Here follows a strongly simplified summary of the proof idea that leads to the CZ inequalities, which is in no sense complete; we refer to [16] for the proof. Consider a compact invariant set $I$ in $M$. A Morse decomposition of $I$ is a finite, disjoint family of compact, invariant subsets $\{M_1, \ldots, M_n\}$ that satisfies the following requirement on the ordering. For every $x \in I \setminus \bigcup_i M_i$, there exists a pair of indices $i < j$, such that

$$\lim_{t \to -\infty} \phi_t(x) \subset M_i \quad \lim_{t \to \infty} \phi_t(x) \subset M_j.$$

Such an ordering, if it exists, is called admissible, and the $M_i$ are called Morse sets of $I$.

Most essentially, it was proved in [16] that for every compact invariant set $I$ admitting an admissibly ordered Morse decomposition, there exists an increasing sequence of compact sets $N_i$ with $N_0 \subset N_1 \subset \ldots \subset N_m$, such that $(N_i, N_{i-1})$ is an index pair for $M_i$, and $(N_m, N_0)$ is an index pair for $I$.

Consider compact manifolds $A \supset B \supset C$. The exact sequence of relative cohomologies

$$\cdots \to H^k(A, B) \to H^k(A, C) \to H^k(B, C) \to H^{k+1}(A, B) \to \cdots$$

implies, in a standard fashion known from Morse theory, that, with $r_{i,p}$ being the rank of $H^p(N_i, N_{i-1})$, the polynomial identity

$$\sum_{i,p} t^p r_{i,p} = \sum_p B_p t^p + (1 + t)Q(t)$$

holds [25]. Here $B_j$ is the $j$-th Betti number of the index pair $(N_m, N_0)$ of $I$, and $Q(t)$ is a polynomial in $t$ (which, of course, is unrelated to the time coordinate) with non-negative integer coefficients. The details of this argument are given in [16]. These are the strong CZ inequalities; a corollary that straightforwardly follows, due to the positivity of the coefficients of $Q(t)$, is that

$$\sum_i r_{i,p} \geq B_p$$

holds. These are the weak Conley-Zehnder inequalities.

If the symplectic manifold $M$ is compact and closed, and if $C$ is assumed to be non-degenerate, the following results arise from application of the CZ inequalities. The invariant set $I$ can be chosen to be equal to $M$. So we let $N_m = M$ and $N_0 = \emptyset$ denote the top and bottom elements of the sequence defined above. Furthermore, we order the connected elements of $C$ according to the descending values of the maximum of $H$ attained on each $C_i$. 37
Then it follows that \( C \) becomes a Morse decomposition for \( M \). The homology groups of \( M \) are isomorphic to the relative homology groups of the index pair \((N_m, N_0)\). So the numbers \( B_p \) are the Betti numbers of the compact symplectic manifold \( M \).

The number \( r_{i,p} \) is the rank of the \( p \)-th relative cohomology group of the non-degenerate critical manifold \( C_i \), the index \( i \) being determined by the Morse decomposition. From (2.7), we deduce that

\[
r_{i,p} = \dim H^{i,p-\mu}(C_i);
\]

recall here that \( \mu_i \) is the index of \( C_i \). In other words, \( r_{i,p} \) is the \((p - \mu_i)\)-th Betti number of \( C_i \), written \( B_{i,p-\mu_i} \) in brief. Assuming that the number of connected components of \( C \) is finite, the Conley-Zehnder inequalities thus result in

\[
\sum_{i,p} t^{p+\mu_i} B_{i,p} = \sum_p B_p t^p + (1 + t) \mathcal{Q}(t), \tag{2.8}
\]

which implies that

\[
\sum_i B_{i,p-\mu_i} \geq B_p \tag{2.9}
\]

is satisfied. So the global topology of \( M \) enforces lower bounds on the ranks of the homology groups of the connected components of \( C \). Setting the variable \( t \) equal to \(-1\), one obtains that

\[
\sum_{i,p} (-1)^{p+\mu_i} B_{i,p} = \sum_i (-1)^{\mu_i} \chi(C_i) = \chi(M),
\]

where \( \chi \) denotes the Euler characteristic.

Our analysis is valid for constrained Hamiltonian systems with non-degenerate, closed, compact critical manifolds. However, in mechanical systems, the phase space is a symplectic manifold that is never compact, and the critical manifold is, as we will see, generally unbounded. Therefore, the arguments used here do not apply. But one can exploit the vector bundle structure of \( M \) and \( C \), which is typical for mechanical systems, to arrive at a result that is closely similar to (2.8).
2.3.4 A sharper result for generic connectivity components

One can actually prove a stronger result than (2.8) if one considers the special structure of the system at hand. In fact, we claim that if all connected components of $\mathcal{C}$ have a constant dimension,

$$\sum_{i,p: C_i \subset \mathcal{C}_{gen}} t^{\nu_p} \mu_p B_{i,p} = \sum_p B_p t^p + (1 + t) \bar{Q}(t)$$

(2.10)

holds, even if the non-generic components are not normal hyperbolic. Here, $\mathcal{C}_{gen}$ denotes the set of generic connectivity components of $\mathcal{C}$, and $\bar{Q}$ is a polynomial with non-negative integer coefficients.

To prove this claim, we construct an auxiliary, continuous vector field $X_\epsilon$ that can be regarded as a small deformation of the gradient-like vector field $P_{\text{grad}H}$. We pick a small positive, real number $\epsilon$, and consider the compact neighborhoods

$$U_\epsilon(C_i) = \{ x \in M | \text{dist}(x, C_i) \leq \epsilon \}$$

(2.11)

of connectivity components $C_i \subset \mathcal{C} \setminus \mathcal{C}_{gen}$ with a dimension higher than $2(n - k)$. Here, $\text{dist}$ denotes some Riemannian distance function, for instance the one induced by the Kähler metric $g$. We define $X_\epsilon$ by requiring that it shall be given by $P_{\text{grad}H}$ in $M \setminus U_\epsilon(C_i)$, and that inside of every $U_\epsilon(C_i)$ with $C_i \subset \mathcal{C} \setminus \mathcal{C}_{gen}$,

$$X_\epsilon |_{x} = P_{\text{grad}H} |_{x} + \epsilon h(x) \text{grad}H |_{x}$$

(2.12)

shall hold. We have here introduced a smooth function $h : U_\epsilon(C_i) \to [0, 1]$ obeying $h|_{C_i} = 1$ and $h|_{\partial U_\epsilon(C_i)} = 0$, that is strictly monotone along the the flow lines generated by $P_{\text{grad}H}$.

We have proved above that for all $C_i \subset \mathcal{C} \setminus \mathcal{C}_{gen}$, $\text{grad}H$ is strictly non-zero in $U_\epsilon(C_i)$. Now let us consider

$$g(X_\epsilon, \text{grad}H) = \left( \|P_{\text{grad}H}\|_g^2 \right)(x) + \epsilon h(x) \left( \|\text{grad}H\|_g^2 \right)(x).$$

Here we have used the $g$-symmetry of $P$, and $\|X\|_g^2 \equiv g(X, X)$. The first term on the right hand side is non-zero on the boundary of $U_\epsilon(C_i)$, while the second term vanishes. Moreover, the second term is non-zero everywhere in the interior of $U_\epsilon(C_i)$. Therefore, $X_\epsilon$ vanishes nowhere in $U_\epsilon(C_i)$.

This implies that the vector field $X_\epsilon$ is a deformation of $P_{\text{grad}H}$ that only exhibits $\mathcal{C}_{gen}$ as a critical set. Notice that the generic components cannot be removed like this, since they contain critical points of $H$. 

39
The scalar product of $X_\epsilon$ with $\nabla H$ is not only non-zero, but also strictly positive in every $U(\mathcal{C})$, which shows that $X_\epsilon$ is a gradient-like flow (the left hand side of the above expression corresponds to the time derivative of $H$ along orbits of $X_\epsilon$ in $U(\mathcal{C})$).

Clearly, the order of magnitude of $\|X_\epsilon - P\nabla H\|_g$ is at most $O(\epsilon)$ everywhere on $M$. Consequently, it is possible to pick $X_\epsilon$ arbitrarily close to $P\nabla H$ in the $\|\cdot\|_\infty$-norm on $\Gamma(TM)$ that is induced by $\|\cdot\|_g$.

Carrying out the Conley-Zehnder construction with respect to the flow generated by $X_\epsilon$ yields (2.10). This result does not require the assumption of normal hyperbolicity on $\mathcal{C}$.

2.4 Proof of the strong Conley-Zehnder inequalities using the Witten complex

We will now give a different proof of (2.10) based on the existence of the so-called Witten complex. To this end, let us first make some brief general remarks about Morse theory. A beautiful account on some milestones in its more recent history is given in [8].

The basic setting of Morse theory comprises a Morse function $f : M \to \mathbb{R}$ on a compact, closed manifold $M$. Using the gradient flow induced by the vector field $-\nabla f$ relative to some auxiliary metric, one gets the corresponding CZ inequalities in the manner demonstrated above. The only non-vanishing homology group of a point is the zeroth homology group, with a rank one. The index of an isolated critical point precisely is its Morse index, hence the CZ inequalities (2.10) reduce to

$$\sum_p t^p N_p = \sum_p t^p B_p + (1 + t)Q(t),$$

where $N_p$ is the number of critical points of $f$ with a Morse index $p$. $B_p$ is the $p$-th Betti number of $M$, and $Q(t)$ is a polynomial with non-negative integer coefficients. These are the strong Morse inequalities, and immediately imply the weak Morse inequalities $N_p \geq B_p$.

Morse theory in its classical version, which is not based on the Conley-Zehnder construction, is for instance presented in [25, 23].

If $f$ is not assumed to be a Morse, but more generally, a Morse-Bott function, the gradient flow of the vector field $-\nabla f$ can be used to derive the CZ inequalities for the critical manifolds of $f$. The first treatment of this situation is due to R. Bott [7].

Morse theory, which has been a classical topic in differential topology for more than sixty year now, experienced a tremendous revival in the early eighties, due to Witten’s completely new proof of the Morse inequalities [39]. His proof of the weak Morse inequalities
is based on a deformation and localization argument in Hodge theory, interpreted as super-symmetric quantum mechanics. The proof of the Morse-Bott inequalities using this (heat kernel) method has been given by J.-M. Bismut in [6].

Witten’s proof of the strong Morse inequalities is based on the differential complex constructed from the free module generated by the set of critical points of $f$, and graded by their Morse indices, together with a specific coboundary operator constructed from the gradient flow, which is nowadays coined the ‘(Morse)-Witten complex’ (although at least Milnor, Smale and Thom independently have arrived at some form of it already earlier [8, 19, 33]). The (co-)homology of the Witten complex is isomorphic to the singular (de Rham co-)homology of $M$, and thus straightforwardly implies the strong Morse inequalities.

A proof of the Morse and Morse-Bott inequalities that is based on the construction of the Witten complex is given in [4].

The Witten complex has been generalized to infinite dimensional systems by A. Floer [19], which has led to extremely fruitful applications. There is, for instance, a proof of the celebrated Arnol’d conjecture based on Floer homology. A beautiful account on this subject matter can be found in [24], and a survey is given in [41].

We will now discuss an alternative proof of (2.10) that is based on the Witten complex for non-degenerate Morse functions. To this end, the results of section 2.3 will be extensively used. The considerations that follow are not essential for the later chapters, and can be skipped in case of disinterest.

### 2.4.1 Definition of the Witten complex

Let $M$ be a compact, closed, orientable and smooth manifold, together with a Morse function $f : M \to \mathbb{R}$. We let $C^p$ denote the free $\mathbb{Z}$-module generated by the critical points of $f$ with a Morse index $p$. The set $C = \bigoplus_p C^p$ is the free $\mathbb{Z}$-module generated by the critical points of $f$, which is graded by their Morse indices. There exists a natural coboundary operator

$$\delta : C^p \to C^{p+1}$$

that obeys $\delta^2 = 0$, which we will define below. Most importantly, the cohomology of this differential complex is isomorphic to the de Rham cohomology of $M$,

$$\ker \delta / \operatorname{im} \delta \cong H^*(M, \mathbb{Z}).$$
A proof of this theorem has been given by Floer in [19] based on Conley-Zehnder theory. Other proofs can be found in [4, 33]. The original publication [39] uses physical arguments in connection to the quantum mechanical tunneling effect.

The coboundary operator is defined as follows [4, 8, 19, 39]. We denote the unstable and stable manifold of a critical point $a$ of $f$ under the gradient flow by $W^-_a$ and $W^+_a$, respectively, and assign an arbitrary orientation to every $W^-_a$. Since $M$ is assumed to be oriented, the orientation of $W^-_a$ at every critical point $a$ induces an orientation of $W^+_a$. The set of Morse functions for which the stable and unstable manifolds intersect transversely is dense in $C^\infty(M)$. Thus, we may generically assume that all $W^-_a$ and $W^+_a$ intersect transversely. The dimension of $W^-_a$ equals the Morse index $\mu(a)$ of $a$, and the dimension of the intersection

$$M(a, a') \equiv W^-_a \cap W^+_a$$

is given by $\max\{\mu(a) - \mu(a'), 0\}$.

In order to define the coboundary operator, let us consider pairs of critical points $a$ and $a'$, whose relative Morse index has the value 1, say $\mu(a) = p + 1$ and $\mu(a') = p$. It immediately follows that $M(a, a')$ is a finite collection of gradient lines that connect $a$ with $a'$.

The intersection of $M(a, a')$ with every regular level surface $\Sigma_c$ of $f$ with $f(\Sigma_c) = c$ lying between $f(a)$ and $f(a')$ is transverse, and consists of a finite collection of isolated points. The hypersurface $\Sigma_c$, being a level surface of $f$, is orientable, [23], so we pick the orientation, which, combined with the section $\text{grad} f$ of its normal bundle, shall agree with the orientation of $M$.

In addition, the intersection both of $W^-_a$ and $W^+_a$ with $\Sigma_c$ is transverse, and the submanifolds

$$W^-_{a,c} \equiv W^-_a \cap \Sigma_c \quad \quad W^+_{a',c} \equiv W^+_a \cap \Sigma_c$$

of $\Sigma_c$ are smooth, compact and closed. In addition, their dimensions add up to the dimension of $\Sigma_c$. To every point $b$ of the set $M(a, a') \cap \Sigma_c = W^-_{a,c} \cap W^+_{a',c}$, one assigns the number $\gamma(b) = 1$ if the induced orientation of $T_b \Sigma_c = T_b W^-_{a,c} \oplus T_b W^+_{a',c}$ agrees with the one picked for $\Sigma_c$, and $\gamma(b) = -1$ otherwise.

The sum

$$\langle a, \delta a' \rangle \equiv \sum_{b \in M(a, a') \cap \Sigma_c} \gamma(b)$$

42
is the intersection number $\#(W_{a,c}^- W_{a',c}^+)$ of the oriented submanifolds $W_{a,c}^-$ and $W_{a',c}^+$ of $\Sigma_c$, cf. for instance [23]. If for a pair $a$ and $a'$ of critical points with a relative Morse index 1, this intersection number is nonzero, we will say in this text that they are 'effectively connected' (by gradient lines).

The coboundary operator of the Witten complex is the $\mathbb{Z}$-linear map $\delta : C^p \to C^{p+1}$ defined by

$$\delta a' = \sum_{\mu(a') = p+1} \langle a, \delta a' \rangle a.$$ 

It satisfies $\delta^2 = 0$, and its cohomology is isomorphic to the de Rham cohomology of $M$. Proofs of this statement can be found in [4, 19, 39]. The image of a critical point of Morse index $p$ under the coboundary map consists of the critical points of Morse index $p + 1$ to which it is effectively connected. One also refers to the intersection number $\langle a, \delta a' \rangle$ as the matrix element of the $\delta$-operator with respect to the elements $a$ and $a'$ of $C$.

The existence of the Witten complex straightforwardly implies the strong Morse inequalities, as one deduces from the following fact. Let us denote the $p$-th cocycle group by $\mathbb{Z}^p \subset C^p$, which is defined as the intersection of $\ker \delta$ with $C^p$, and let $B^p \subset C^p$, the $p$-th coboundary group, denote the image of $C^{p-1}$ under $\delta$. Clearly, $B^p$ is a subset of $\mathbb{Z}^p$ because $\delta$ is nilpotent, therefore the set $H^p = \mathbb{Z}^p \setminus B^p$ is well-defined. It is the $p$-th cohomology group of the Witten complex. Since the cohomology of the Witten complex is isomorphic to the de Rham cohomology of $M$, the rank of $H^p$ coincides with the $p$-th Betti number $B_p(M)$ of $M$.

The image of $C^p$ in $C^{p+1}$ under $\delta$ is given by $B^{p+1}$. Denoting the preimage of $B^{p+1}$ in $C^p$ by $\delta^{-1}(B^{p+1})$, which is isomorphic to $B^{p+1}$, one has

$$C^p = H^p \oplus B^p \oplus \delta^{-1}(B^{p+1}),$$

so that

$$\dim C^p = B_p(M) + \dim B^p + \dim B^{p+1}.$$ 

The dimension of $C^p$ equals the number $N_p$ of critical points of $f$ with a Morse index $p$. Multiplying both sides of the equality sign with $t^p$, and summing over $p$, one finds

$$\sum t^p N_p = \sum t^p B_p(M) + \sum t^p (\dim B^p + \dim B^{p+1})$$

$$= \sum t^p B_p(M) + (1 + t) \sum t^{p-1} \dim B^p$$

(2.13)
(notice here that both $B^0$ and $B^{2n+1}$ are empty). These are the strong Morse inequalities, and the polynomial $Q(t)$ defined at the beginning of this section now has a very definitive interpretation:

$$Q(t) = \sum t^{p-1} \dim B^p.$$ 

Evidently, $\dim B^p$ is the number of critical points of Morse index $p$ that are effectively connected to critical points of Morse index $p - 1$ via gradient lines of $f$.

### 2.4.2 A relationship between the Witten complexes of $(M, H)$ and $(C_{\text{gen}}, H_{|c_{\text{gen}}})$

Certain results obtained in section 2.2 can be used to relate the Witten complexes of $(M, H)$ and $(C_{\text{gen}}, H_{|c_{\text{gen}}})$ to each other. In fact, we have proved that the generic connectivity components $C_{\text{gen}}$ of $C$ contain all critical points of $H$, but no other conditional extrema, and that they are necessarily normal hyperbolic with respect to $\phi_t$. Let us discuss the implications.

Let $A_i := \{a_{i,1}, \ldots, a_{i,m}\}$ denote the set of critical points of $H$ that are contained in $C_i$, and let $\mu(a_{i,r})$ be the associated Morse indices. Furthermore, let $H_i \equiv H_{|C_i}$ denote the restriction of the Hamiltonian to $C_i$. Obviously, the map $H_i : C_i \to \mathbb{R}$ is a Morse function, and its critical points are precisely the elements of $A_i$. Clearly, the number of negative eigenvalues of the Hessian of $H$ at any element of $A_i$, whose eigenspaces are normal to $b$, is the same as the index $\mu(C_i)$.

The Morse index of $a_{i,r}$ with respect to $H_i$ is hence given by $\mu(a_{i,r}) - \mu(C_i)$. The Witten complex associated to $C_i$ is defined in terms of the free $\mathbb{Z}$-module generated by the elements of $A_i$, which is graded by the Morse indices $p$ of the critical points of $H_i$,

$$C_i = \oplus_p C^p_i.$$ 

To define the coboundary operator $\delta_i : C^p_i \to C^{p+1}_i$, one uses the gradient flow on $C_i$ generated by $H_i$. One then understands that

$$\ker \delta_i / \im \delta_i \cong H^*(C_i, \mathbb{Z}).$$

Application of (2.13) shows that for every $C_i \in C_{\text{gen}},$

$$\sum_p t^p N_{i,p} = \sum_p t^p B_p(C_i) + (1 + t) \sum_p t^{p-1} \dim B^p_i,$$  

(2.14)

44
where $B^p_{i}$ is the $p$-th coboundary group of the Witten complex of $C_i$, and $N_{i,p}$ is the number of critical points of $H_i$ on $C_i$ whose Morse index is $p$.

Every critical point of $H$ lies on precisely one generic component $C_i$, hence the number $N_q$ of critical points of $H$ with a Morse index $q$ is given by

$$N_p = \sum_i N_{i,p}.$$

Multiplying both sides of (2.14) with $t^{\mu(C_i)}$, and summing over $i$, one obtains

$$\sum_{i,p;C_i \in \text{gen}} t^{\mu(C_i)} N_{i,p} = \sum_{i,p;C_i \in \text{gen}} t^{\mu(C_i)} B_p(C_i) + (1 + t) \sum_{C_i \in \text{gen}} t^{\mu(C_i) + p - 1 \dim B^p_{i}},$$

which becomes, after reindexing $\mu(C_i) + p \rightarrow q$,

$$\sum_q t^q N_q = \sum_{i,q;C_i \in \text{gen}} t^q B_{q-\mu(C_i)}(C_i) + (1 + t) \sum_{i,q;C_i \in \text{gen}} t^{q-1 \dim B^q_{i-\mu(C_i)}}.$$

Combining this result with the strong Morse inequalities

$$\sum_q t^q B_q(M) + (1 + t) \sum_q t^{q-1 \dim B^q},$$

one obtains that

$$\sum_{i,q;C_i \in \text{gen}} t^q B_{q-\mu(C_i)}(C_i) = \sum_q t^q B_q(M) + (1 + t) \sum_q t^{q-1 \dim B^q} - \sum_{C_i \in \text{gen}} \dim B^q_{i-\mu(C_i)}.$$

One observes that formula (2.10) implies that the polynomial which is multiplied by $(1 + t)$ has non-negative integer coefficients.

Conversely, if one can prove that for all $q$,

$$\dim B^q \geq \sum_{C_i \in \text{gen}} \dim B^q_{i-\mu(C_i)}$$

(2.15)

holds, one would also obtain an alternative proof of (2.10). It is remarkable that the left hand side is defined in terms of the operator $\delta$ associated to $(M, H)$, while the right hand side is defined in terms of the operators $\delta_i$ associated to all $(C_i, H_i)$ with $C_i \in \text{gen}$.

The quantity $\dim B^q_{i-\mu(C_i)}$ denotes the number of critical points of $H$ with a Morse index $q$ in $C_i$, which are effectively connected to critical points of Morse index $p + 1$ in $C_i$ via gradient lines of the Morse function $H_i$ on $C_i$. Therefore, the sum on the right hand side of (2.15) equals the number of those critical points of $H$ with a Morse index $q$, which are effectively connected to critical points of Morse index $q + 1$ via gradient lines of the functions $H \circ j_i$ on all generic $C_i$; here, $j_i : C_i \rightarrow M$ is the inclusion map.
We will now give a proof of the inequality (2.15) by relating the coboundary operators of the Witten complexes of \((M, H)\) and \((C_{\text{gen}}, H|_{C_{\text{gen}}})\) to each other. This will shed much light on the structure of these systems, and will amount to a different proof of the strong Conley-Zehnder inequalities (2.10).

The proof strategy is based on the construction of a particular homotopy of vector fields \(v_\sigma\), with \(\sigma \in [0, 1]\), that generate gradient-like flows. Their zeros will be independent of \(\sigma\), and hyperbolic. In particular, \(v_1\) will be given by \(\text{grad} H\), so that the zeros of \(v_\sigma\) are precisely the critical points of \(H\), and \(v_0\) will be a vector field that is tangent to every \(C_{\text{gen}}\). For every \(\sigma\), we will construct a coboundary operator via the one-dimensional integral curves of \(v_\sigma\) that connect its zeros. These coboundary operators are independent of \(\sigma\), and act on the free \(\mathbb{Z}\)-module \(C\) of the Witten complex associated to \((M, H)\). The desired estimate (2.15) then follows from a simple dimension argument.

Let us now turn to the construction of the vector field \(v_0\). As has been mentioned above, we require it to be gradient-like, and tangent to \(C_{\text{gen}}\). Furthermore, its zeros shall be hyperbolic, and identical to the critical points of \(H\). From this last condition follows that the dimension of any unstable manifold of the flow generated by \(-v_0\) equals the Morse index of the critical point of \(H\) from which it emanates.

To this end, we recall the vector field \(X_\epsilon\) from subsection 2.3.4, which has been obtained by deforming \(P_{\text{grad}H}\) in a manner that all elements \(C_i\) of \(C \setminus C_{\text{gen}}\) are removed. We introduce compact \(\epsilon\)-neighborhoods \(U_\epsilon(C_{\text{gen}})\) of the generic connected components of \(C\) in the same way as we did for (2.11).

It is now possible to extend the projector \(\bar{Q} : T_{C_{\text{gen}}} M \to TC_{\text{gen}}\) that has been introduced in 2.1 over the whole embedding space \(TU_\epsilon(C_{\text{gen}})\). To this end, we pick an arbitrary smooth distribution \(W\) over the base manifold \(U_\epsilon(C_{\text{gen}})\), whose fibres over \(C_{\text{gen}}\) shall coincide with the corresponding fibres of \(TC_{\text{gen}}\). Admitting a slight abuse of notation, we let \(\bar{Q}\) denote the \(g\)-orthogonal projector \(TM \to W\). Clearly, evaluating \(\bar{Q}\) in any \(a \in C_{\text{gen}}\) gives the projector \(\bar{Q}_a : T_a M \to T_a C_{\text{gen}}\) discussed in 2.1. Because \(W\) is smooth, \(\bar{Q}\) and its orthogonal complement \(Q\) are both smooth tensor fields.

We define the vector field \(v_0\) by requiring that it shall equal \(X_\epsilon\) in \(M \setminus U_\epsilon(C_{\text{gen}})\), and that for \(x\) in \(U_\epsilon(C_{\text{gen}})\), it shall be given by

\[ v_0(x) \equiv (P_{\text{grad}H})(x) + h(x)(\bar{Q}_{\text{grad}H})(x), \]
where $h : U_e(C_{gen}) \rightarrow [0,1]$ is a smooth function obeying $h|_{C_{gen}} = 1$ and $h|_{\partial U_e(C_{gen})} = 0$. In particular, $h$ shall be strictly monotonic along all non-constant trajectories of the flow generated by $P_{grad}H$, and the one form $dh$ shall vanish on $C_{gen}$.

It can now be easily verified that $v_0$ possesses all of the desired properties. That it generates a gradient-like flow can be seen from the fact that outside of $U_e(C_{gen})$, $g(\text{grad}H, v_0) = g(\text{grad}H, X_e)$ is strictly positive, as has been proved in 2.3.4. Inside of $U_e(C_{gen})$, one finds
\[
g(\text{grad}H, v_0) = ||P_{grad}H||_g^2 + h||Q_{grad}H||_g^2,
\]
due to the $g$-orthogonality both of $P$ and $Q$. The first term on the right hand side vanishes everywhere on $C_{gen}$, but nowhere else in $U_e(C_{gen})$. The second term reduces to $||Q_{grad}H||_g^2$ on $C_{gen}$. Since evidently, $Q_{grad}H|_{C_{gen}}$ is the gradient field of the Morse function $H|_{C_{gen}} : C_{gen} \rightarrow \mathbb{R}$ relative to the Riemannian metric on $T(C_{gen})$ that is induced by $g$, its zeros are precisely the critical points of $H$ on $C_{gen}$, and there are no other zeros apart from those. This shows that the right hand side of the above expression is strictly positive except at the critical points of $H$, where everything vanishes. Because the scalar product of $v_0$ with grad $H$ is strictly positive except at the critical points of $H$, it is clear that $-v_0$ generates a gradient-like flow $\psi_{0,t}$, so that $H$ is strictly decreasing along all non-constant orbits.

Furthermore, it is also clear from the given construction that $v_0$ is tangent to $C_{gen}$. Next, we prove that the zeros of $-v_0$ are hyperbolic, and that the number of negative eigenvalues of the Jacobian matrix at any zero equals the corresponding Morse index of $H$.

To this end, we pick a local chart at a critical point $a$ of $H$. The Jacobian matrix of $v_0$ at $a$ in this chart is given by
\[
\text{Jac}_a(v_0) = \text{Jac}_a(P_{grad}H) + \text{Jac}_a(Q_{grad}H)
\]
\[
= P_a(D^2_aH)^t + \bar{Q}_a(D^2_aH)^t \\
= (D^2_aH)^t + (P_a - Q_a)(D^2_aH)^t.
\]
(2.16)

Let us briefly explain this result. First of all, there is no dependency on the function $h$ because $dh|_{C_{gen}}$ is zero. We recall that $(D_a^2H)^t$ is defined as the matrix $[g^{ij}H_{jk}]$ in the given chart. Linearizing the vector fields $P_{grad}H$ and $Q_{grad}H$ at a critical point of $H$, all terms involving first derivatives of $H$ are zero. This explains the second line. The third line simply follows from $\bar{Q}_a = 1_{2n} - Q_a$.

Because it is more convenient to prove that the zeros of $v_\sigma$ are hyperbolic fixed points of $\psi_{\sigma,t}$ for arbitrary $\sigma \in [0,1]$ instead of focusing on the case $\sigma = 0$, let us discuss this
question for the homotopy of vector fields alluded to at the beginning of this subsection.

In fact, we are considering the one parameter family of vector fields defined by

\[ v_\sigma \equiv \sigma \nabla H + (1 - \sigma)v_0 \]

with \( \sigma \in [0,1] \). We claim that for arbitrary \( \sigma \), \(-v_\sigma\) generates a gradient-like flow \( \psi_{\sigma,t} \), in a manner that \( H \) is strictly decreasing along all non-constant orbits. Furthermore, we claim that the zeros of \( v_\sigma \) are hyperbolic fixed points of \( \psi_{\sigma,t} \) that do not depend on \( \sigma \). It then follows that the dimensions of the unstable manifolds are, for all \( \sigma \), given by the Morse indices of the critical points of \( H \) from which they emanate.

To prove these claims, we consider the scalar product

\[ g(\nabla H, v_\sigma) = \sigma \| \nabla H \|^2_g + (1 - \sigma)g(\nabla H, v_0). \]

The first term on the right hand side is obviously everywhere positive except at the critical points of \( H \), and the same fact has been proved above for the second term. Thus, \( H \) is strictly decreasing along all non-constant orbits of \( \psi_{\sigma,t} \), which proves that it is gradient-like.

The Jacobian of \( v_\sigma \) at a critical point of \( H \) is given by

\[
\begin{align*}
\text{Jac}_a(v_\sigma) &= \sigma (D^2_aH)^\sharp + (1 - \sigma)\text{Jac}_a(v_0) \\
&= (D^2_aH)^\sharp + (1 - \sigma)(P_a - Q_a)(D^2_aH)^\sharp \\
&= (12n + (1 - \sigma)(P_a - Q_a)(D^2_aH)^\sharp.
\end{align*}
\]

If we can prove that \( \text{Jac}_a(v_\sigma) \) is invertible for all \( \sigma \in [0,1] \), it follows that the number of negative eigenvalues is independent of \( \sigma \). To prove that this is indeed the case, we observe that because \((D^2_aH)^\sharp\) is invertible, one merely has to show that \((12n + (1 - \sigma)(P_a - Q_a))\) is invertible. This in turn is satisfied if \((Q_a - P_a)\) has no eigenvalues in \([1, \infty) \subset \mathbb{R}\).

Arguing by contradiction, we will prove that the spectrum of \((Q_a - P_a)\) has no intersection with \([1, \infty) \). To this end, assume that \( \kappa \in [1, \infty) \) is an eigenvalue associated to the eigenvector \( w \in T_aM \), so that

\[ (Q_a - P_a)w = \kappa w. \quad (2.17) \]

We multiply both sides of the equality sign with \( P_aQ_a \) from the left, and get

\[ P_aQ_aP_a w = (1 - \kappa)P_aQ_au. \]
On the other hand, multiplication from the left with $P_a$ gives

$$P_a Q_a w = (1 + \kappa) P_a w.$$ 

Combining these two results, one obtains

$$P_a Q_a P_a w = (1 - \kappa^2) P_a w,$$

which shows that $P_a w$ is an eigenvector of $P_a Q_a P_a : V_a \to V_a$ that belongs to the eigenvalue $(1 - \kappa^2)$. We have proved earlier that normal hyperbolicity of $C_{gen}$ implies that $P_a Q_a P_a$ is an automorphism of $V_a$. In addition, it is evident that $(1 - \kappa^2) \leq 0$.

Let us assume that $P_a w \neq 0$. Then,

$$g_\alpha(w, P_a Q_a P_a w) = g_\alpha(Q_a P_a w, Q_a P_a w) > 0$$

follows from the $g$-orthogonality of the projectors. On the other hand, we also have

$$g_\alpha(w, P_a Q_a P_a w) = (1 - \kappa^2) g_\alpha(w, P_a w) = (1 - \kappa^2) g_\alpha(P_a w, P_a w) \leq 0,$$

which is a contradiction. Hence, $P_a w = 0$ must be given. In this case, (2.17) reduces to $Q_a w = \kappa w$, and multiplication with $Q_a P_a$ from the left gives

$$Q_a P_a Q_a w = \kappa Q_a P_a w = 0.$$

Because $Q_a P_a Q_a : N_a C_{gen} \to N_a C_{gen}$ is an isomorphism, this implies that $Q_a w$ must be zero. Therefore, $w$ is cannot be contained in the intersection of the images of $P_a$ and $Q_a$. However, being an eigenvector that solves (2.17), it must be contained in this space. So we have arrived at a contradiction, which proves that $(1_2 + \sigma) (P_a - Q_a)$ is invertible for all $\sigma \in [0, 1]$.

The conclusion is that for all $\sigma \in [0, 1]$, the zeros of $v_\sigma$ are hyperbolic fixed points of $\psi_{\sigma,t}$, and that the dimensions of the corresponding unstable manifolds are given by the respective Morse indices of $H$.

Since $\frac{d}{dt}\psi_{\sigma,t} = -v_\sigma$ depends smoothly on $\sigma$, it follows that $\psi_{\sigma,t}$ is $C^\infty$ in $\sigma$. Furthermore, $\sigma$ smoothly parametrizes a homotopy of stable and unstable manifolds emanating from the critical points of $H$, which belong to the gradient-like flow $\psi_{\sigma,t}$. 

49
Because the fixed points of $\psi_{\sigma,t}$ are independent of $\sigma$, and since the corresponding dimensions of the unstable manifolds coincide with the Morse indices of the critical points of $H$, it makes sense to consider the free $\mathbb{Z}$-module

$$C = \oplus_p \mathbb{Z}^p$$

that is generated by the critical points of $H$, and graded by their Morse indices.

It is quite straightforward to define a coboundary operator on $C$ based on $\psi_{\sigma,t}$ for every $\sigma$. Picking a pair of critical points of $H$ with a relative Morse index 1, consider the unstable manifold $W_{\sigma,a}$ of $a$, and the stable manifold $W_{\sigma,a'}^+$ of $a'$ associated to $\psi_{\sigma,t}$, which are both smoothly parametrized by $\sigma$. The fact that $\sigma$ parametrizes a homotopy of such manifolds makes it clear that they naturally inherit an orientation from the one picked for $\sigma = 1$ in the definition of the Witten complex for $(M, H)$.

Let us then pick a regular energy surface $\Sigma_E$ for an arbitrary energy value $E$ between $H(a)$ and $H(a')$. The intersection of $W_{\sigma,a}^\pm$ with any regular energy level surfaces $\Sigma_E$ of $H$ is transverse, due to the fact that $H$ strictly decreases along all non-constant orbits generated by $-v_\sigma$.

$W_{\sigma,a}^- \cap \Sigma_E$ and $W_{\sigma,a'}^+ \cap \Sigma_E$ are oriented submanifolds of $\Sigma_E$, and smoothly parametrized by $\sigma$. Hence, they define two homotopies of manifolds in $\Sigma_E$. Their intersection number, being a homotopy invariant, is independent of $\sigma$, so that it equals the value found for $\sigma = 1$. This in turn implies that the coboundary operators obtained for arbitrary $\sigma$ are identical to the $\delta$-operator of the Witten complex given for $\sigma = 1$, since their respective matrix elements are equal.

To finally prove (2.15), we recall that all stable and unstable manifolds of $\psi_{0,t}$ are, by construction, either confined to some $C_i$, or otherwise, that they connect critical points lying on different $C_i$'s. This is because the stable and unstable manifolds of the flow generated by $P\text{grad}H$ only connect different connectivity components of $C_{\text{gen}}$.

Let us next consider pairs of critical points of $H$ with a relative Morse index 1 that lie on the same component $C_i \in C_{\text{gen}}$, and the corresponding stable and unstable manifolds of $\psi_{0,t}$ which are contained in $C_i$. Since $v_0|_{C_i}$ simply is the projection of $\text{grad}H|_{C_i}$ to $T C_i$, these stable and unstable manifolds are precisely those which we have used to define the Witten complex on $(C_i, H_i)$ (recall that $H_i$ is the restriction of $H$ to $C_i$, and that it is by itself a Morse function).
Now we pick all stable and unstable manifolds of $\psi_{0,t}$ which are contained in $C_{gen}$, and construct an operator acting on $C$ in the same manner we constructed the coboundary operator. It is again a coboundary operator, but instead of $\delta$ (which belongs to the pair $(M,H)$), it is given by

$$\bar{\delta} \equiv \oplus_i \delta_i.$$ 

In this formula, $\delta_i$ is the coboundary operator of the Witten complex associated to the pair $(G_i,H_i)$.

Finally, we denote by $P_i : C \rightarrow C_i$ the projection of the free $\mathbb{Z}$-module $C$ generated by all critical points of $H$ to the one generated by those critical points which lie in $C_i$. The above construction makes it evident that removing all integral lines of $-v_0$ that connect critical points on different connectivity components of $C_{gen}$, one gets $\delta_i = P_i \delta P_i$, so that

$$\bar{\delta} = \oplus_i P_i \delta P_i$$

(note that $\delta$ can be written as $\delta = \oplus_i \delta P_i$). If one included the missing integral lines, one would of course obtain $\delta$, as our homotopy argument has proved. This immediately makes the inequality

$$\dim(\text{im} \delta |_{C'}) \geq \dim(\text{im} \bar{\delta} |_{C'})$$

clear (recall that $\delta$ is a $\mathbb{Z}$-linear operator on $C$). We observe that this is precisely what is expressed in the inequality (2.15). Hence we have proved (2.10) by use of the existence of the Witten complex, and very essentially, by exploiting the special structure of the system at hand.

The observations in this concluding section should have elucidated some of the interesting structure that is inherent in the particular constrained Hamiltonian systems in discussion. One of our main intentions was to point out relationships between the critical points of the systems at the presence and at the absence of constraints.
CHAPTER 3

CONSTRAINED HAMILTONIAN MECHANICS

We introduce constrained Hamiltonian mechanical systems in this chapter, and relate them to the constrained Hamiltonian systems that were studied in chapter one.

Hamiltonian mechanics is defined on the cotangent bundle $T^*Q$ of a Riemannian manifold $Q$ (which we assume to be orientable) with its natural symplectic structure. The majority of all technically relevant constrained Hamiltonian mechanical systems at the absence of controls can be described by linear conditions imposed on their generalized velocities, that is, the tangent vectors to their orbits in $Q$. This is the type of systems we will study.

The time evolution of such a system is determined by Hölder’s variational principle, whose associated Euler-Lagrange equations are a set of differential-algebraic relations. We will prove the existence and uniqueness of solutions, and the fact that they are all confined to a submanifold $\mathcal{M}_{\text{phys}} \subset T^*Q$, which we will refer to as the 'physical sheet'.

We will study the physical flow of the constrained mechanical system on $\mathcal{M}_{\text{phys}}$ by introducing a new, auxiliary constrained Hamiltonian system on $T^*Q$ of the type discussed in chapter one. We will prove that such an auxiliary system can be defined in a way that $\mathcal{M}_{\text{phys}}$ is a (non-asymptotically) stable invariant manifold, and that its flow generates the correct physical dynamics on $\mathcal{M}_{\text{phys}}$. Moreover, the auxiliary system can be introduced in a manner that its equilibria are trivially related to those of the physical system, and that it inherits eventual symmetries of the physical system.

Possible applications of these results will be pointed out, but a detailed discussion is postponed until chapter five. The physical and geometrical aspects of the subject matter will have an equal importance in the subsequent discussion, therefore we will mix the languages of analytical dynamics and differential geometry. Hopefully, this will not give rise to any confusion. Our notation has, to a large extend, been adopted from [10].
3.1 A minisurvey of Hamiltonian mechanics

Let us very briefly summarize some basic aspects of Hamiltonian mechanics. For a detailed introduction to this topic, we refer for instance to [1, 2, 3, 26, 28].

A classical Hamiltonian mechanical system consists of a smooth Riemannian manifold \( Q \) of dimension \( n \) with a smooth Riemannian metric \( h \), together with a smooth function \( u : Q \to \mathbb{R} \) which is called the potential energy. A curve is a smooth embedding \( q : \mathbb{R} \to Q \). For every curve \( q \) parametrized by \( t \in \mathbb{R} \), the quadratic form \( h(q, q) \) on \( TQ \) is called the kinetic energy of the curve at \( q(t) \) (the dot is an abbreviation for the derivative \( \frac{d}{dt} \)). The matrix of \( h \) in a local coordinate chart is, in the multibody systems literature, referred to as the mass matrix, because it is determined by the mass distribution in the system.

The metric \( h \) defines an isomorphism \( h : TQ \to T^*Q \), and a unique associated dual Riemannian metric \( h^* \) on the cotangent bundle in the following way. Letting \( \theta_X \) denote the one form that is related to a given vector field \( X \) by the requirement that \( \langle \theta_X, Y \rangle = h(X, Y) \) holds for all vector fields \( Y \), the Riemannian dual metric on \( T^*Q \) is determined by the condition that \( h(X, Y) = h^*(\theta_X, \theta_Y) \) holds for all vector fields \( X \) and \( Y \). In a local chart where the matrix of \( h \) is given by \( M \), the matrix of \( h^* \) is given by \( M^{-1} \).

Every cotangent bundle is a vector bundle, and the map
\[
\pi : T^*Q \to Q
\]
\[
T^*Q \ni q \mapsto q
\]
is called the base point projection. In a local bundle chart, points in \( T^*Q \) are represented by pairs \((q, p)\). \( q = (q^1, \ldots, q^n) \) parametrizes points on the base manifold \( Q \), and its components are called generalized coordinates. \( p = (p_1, \ldots, p_n) \) are fibre coordinates in \( T^*_q Q \); its components are usually referred to as canonical momenta.

There exists a natural symplectic structure associated to every cotangent bundle which is compatible with the vector bundle structure. As a matter of fact, since one forms on \( Q \) are smooth mappings from the zero section \( Q \subset T^*Q \) to \( T^*Q \), it is clear that every one form on \( Q \) induces a pullback that acts on elements of \( \Lambda^*(T^*Q) \), the Grassmann algebra over \( T^*Q \). In particular, there exists a unique one form \( \theta_0 \) on \( T^*T^*Q \) so that \( \psi^*\theta_0 = \psi \) is satisfied for every one form \( \psi : Q \to T^*Q \). In a local bundle chart, it is given by
\[
\theta_0 = \sum p_i dq^i.
\]
The proof of this result can, for instance, be found in [1]. The two form $\omega = -d\theta_0$ is closed and non-degenerate, and defines, in local coordinates, the particular symplectic structure

$$\omega = \sum_i dq^i \wedge dp_i$$
on $T^*Q$. It is the unique symplectic structure for which $\psi^*\omega = -d\psi$ holds for every one form $\psi : Q \to T^*Q$. Hence, the pair $(T^*Q, \omega)$ is a symplectic manifold whose symplectic structure is consistent with the vector bundle structure in the manner presently described, so that every bundle chart is a Darboux chart (here, the symplectic structure is an exact form, which is, as mentioned at the beginning of chapter one, only possible since $T^*Q$ is not compact and closed).

The Hamiltonians $H : T^*Q \to \mathbb{R}$ which we will consider have the form

$$H(q,p) = \frac{1}{2} h^*_q(p,p) + u(q). \quad (3.1)$$

We recall from chapter one that the Hamiltonian vector field $X_H$ associated to $H$ is defined by $i_{X_H} \omega = -dH$. In the present bundle chart, $X_H$ is therefore given by

$$X_H = \sum_i (H_{p_i} \partial_{q^i} - H_{q^i} \partial_{p_i}).$$

Motions of the physical system are described by curves $x : \mathbb{R} \to T^*Q$ with $x = (q^i, p_i)$ obeying the Hamiltonian equations of motion

$$\dot{q}^i = H_{p_i}, \quad \dot{p}_j = -H_{q^j}. \quad (3.2)$$

We denote the Hamiltonian flow by $\Phi_t$, as we did in chapter one.

The equations of motion can be derived from an action principle. The action of the Hamiltonian system is a functional on the space of smooth curves $\gamma : I \subset \mathbb{R} \to T^*Q$. Using a coordinate $t$ on $\mathbb{R}$, one has a basis one form $dt$ for sections of $T^*\mathbb{R}$, and the action is defined as

$$\mathcal{I}[\gamma] = \int_I (\gamma^* \theta_0 - H \circ \gamma dt). \quad (3.3)$$

By definition of the symplectic one form $\theta_0$, this can locally be written as

$$\mathcal{I}[\gamma] = \int_I dt \left( \sum p_i(t) \dot{q}^i(t) - H(q(t), p(t)) \right),$$
where \( \dot{\gamma} = \sum (\ddot{q}^i \partial_{q^i} + \dot{p}_i \partial_{p_i}) \). Let

\[
c := (\pi \circ \gamma) : I \to Q
\]
denote the projection of \( \gamma \) to the base manifold, and assume that the length of \( c(I) \) is so small that there exist solutions of the Hamiltonian equations (3.2) by which they are connected.

The action principle states that among all curves \( \gamma : I \to T^*Q \), for which the end points of the projection \( c(\partial I) \) are being kept fixed, those that extremize \( \mathcal{I} \) are physical orbits of the system. Curves of this kind are called stationary solutions (of the variational problem). In general, there are infinitely many physical orbits \( \gamma \) that are projected to the same end points \( c(\partial I) \) on \( Q \), but which differ in the energy values \( H(\gamma) \) (see for instance chapter nine in [2]).

To be precise, consider a curve \( \gamma : I \to T^*Q \). A one-parameter family of curves \( \gamma_s : I \to T^*Q \) with \( s \in [0,1] \), so that \( (\pi \circ \gamma_s)(\partial I) \) is independent of \( s \), is called a variation of \( \gamma \) if it depends smoothly on \( s \), and satisfies \( \gamma_0 = \gamma \). It is common to write \( \delta \) for the operator \( \frac{\partial}{\partial s} \big|_{s=0} \), and to denote the components of \( \frac{\partial}{\partial s} \big|_{s=0} \gamma_s \) (which is a section of \( T_{\gamma(I)}M \)) by \( (\delta q^i, \delta p_i) \). Then, \( \delta q^i|_{\partial I} = 0 \) expresses the fact that \( (\pi \circ \gamma_s)(\partial I) \) is independent of \( s \).

Let us now consider curves \( \gamma \) for which \( \frac{\partial}{\partial s} \big|_{s=0} \mathcal{I}[\gamma_s] = 0 \) holds with respect to all variations \( \gamma_s \) with \( \gamma_0 = \gamma \). Applying partial integration once, one has

\[
0 = \delta \mathcal{I}[\gamma_s] = \sum p_i \delta q^i|_{\partial I} + \int_I \sum \left( \dot{p}_i \delta q^i - q^i \delta p_i - H_{q^i} \delta q^i - H_{p_i} \delta p_i \right)
\]
in the present bundle chart. The boundary term vanishes because of \( \delta q^i|_{\partial I} = 0 \). The variations \( \delta p_i \) and \( \delta q^i \) are, apart from the condition \( \delta q^i|_{\partial I} = 0 \), fully arbitrary. Therefore, vanishing of this integral is granted if and only if (3.2) holds, that being the coordinate expression for \( \dot{\gamma}(t) = X_H \big|_{\gamma(t)} \). This proves that the Hamiltonian equations of motion are the Euler-Lagrange equations of this variational principle.

### 3.2 Constrained Hamiltonian mechanical systems

We will now consider Hamiltonian mechanical systems with Pfaffian constraints. The starting point is a Hamiltonian mechanical system, for instance a multibody system, whose positions are parametrized by points on a Riemannian manifold \( Q \) as described at the beginning of this chapter, where the metric tensor \( h \) is determined by its mass distribution. We also assume that it moves under the influence of an external potential energy \( u : Q \to \mathbb{R} \).
Assume now that one implements additional linkages and joints between the individual rigid body constituents. As a technical example, assume that \( q^{(i)} \) and \( q^{(j)} \) denote the coordinate vectors describing the center of mass positions of the \( i \)'th and \( j \)'th rigid body with respect to a local Euclidean chart; fixing their center of masses to the ends of a rigid, massless rod of length \( r \) (in the Euclidean metric) can be described as the requirement that

\[
\| q^{(i)} - q^{(j)} \|^2 = r^2
\]

(with respect to the Euclidean norm) shall hold, or respectively in differentiated form, that

\[
\langle \langle q^{(i)}; \dot{q}^{(i)} \rangle \rangle + \langle \langle q^{(j)}; \dot{q}^{(j)} \rangle \rangle = 0
\]

(where \( \langle \langle \cdot, \cdot \rangle \rangle \) denotes the Euclidean scalar product) shall be satisfied for every motion of the system.

Any technically relevant multibody system can be modelled as an initial, 'free' multibody system, upon which one imposes additional links and joints. We will only consider links or joints that can be described as a linear kinematical condition on the generalized velocities. Links and joints of this sort define a smooth, rank \( k \) distribution \( W \) over \( Q \), and the complete multibody system, which contains them, is described by curves \( q(t) \) in \( Q \) which satisfy \( \dot{q} \in W_q \). The equations of motion of the complete system, which is referred to as a constrained mechanical or multibody system, is determined by the action principle of H"older. Let us now make these ideas precise.

A Hamiltonian mechanical system with Pfaffian constraints is a triple \((T^*Q, H, W)\), where \((T^*Q, H)\) is a Hamiltonian mechanical system of the type introduced in the previous section, and where \( W \) is a smooth, rank \( k \) distribution over \( Q \). An admissible curve is a curve \( \gamma : I \subset \mathbb{R} \to T^*Q \) for which the projection \( c = \pi \circ \gamma \) is tangent to the distribution \( W \), that is, \( Tc(I) \subset W_{c(I)} \). Let \( \alpha : TQ \to W \) denote the orthogonal bundle projector with respect to \( h \). Denote by \( W^* \) the image of \( W \) under \( h : TQ \to T^*Q \), which is called the dual distribution associated to \( W \) (and \( h \)).

The associated orthogonal bundle projection \( T^*Q \to W^* \) with respect to \( h^* \) will be denoted by \( \alpha^T \), because in every standard basis, its matrix is the transpose of the matrix of \( \alpha \). The symbol \( \alpha^{(T)} \) will thus interchangeably be used both for the projectors, and for their matrices relative to a coordinate chart. The context will always make it clear which of the two interpretations is being referred to. Accordingly, \( \beta^{(T)} : T^{(o)}Q \to W_{\beta^*(o)} \) will stand for the
associated complementary projections, defined by $\beta^{(T)} = \text{id} - \alpha^{(T)}$, where $\text{id}$ is the identity tensor. The distribution $W^*_\beta$ is the complement of $W^*$ in $T^*Q$, and is called the *annihilator* of $W$ (because $W$ is annihilated by every one form that is a section of $W^*_\beta$).

We refer the reader to [14] for the theory of constrained mechanical systems in the Lagrangian picture, based on the use of bundle projections.

### 3.2.1 The Hölder variational principle

The correct dynamical equations for Hamiltonian mechanical systems with Pfaffian constraints are obtained from the Hölder variational principle [3]. This principle should be considered as a physical law of the same level as the action principle for Hamiltonian mechanical systems at the absence of constraints.

Let $\gamma_s : I \to T^*Q$, with $s \in [0,1]$, be a smooth one parameter family of curves for which the end points $c_s(\partial I)$ are independent of $s$ (where $c_s := \pi \circ \gamma_s$). An *admissible variation* is a smooth one parameter family $\gamma_s$, with $s \in [0,1]$, of curves for which $\frac{\partial}{\partial s}(\pi \circ \gamma_s)$ lies in $W$.

Let us now consider a pair of points $q_1$ and $q_2$ in $Q$ that can be joined by smooth admissible curves, and the action (3.3)

$$I[\gamma_s] = \int_I \left( \sum p_i(t) \dot{q}^i(t) - H(q(t),p(t)) \right) dt,$$

where $\gamma_s$ is an admissible variation of the admissible curve $\gamma_0$ with $(\pi \circ \gamma_s)(\partial I) = \{q_1,q_2\}$ and $\delta q^i|_{\partial I} = 0$. The Hölder principle states that if

$$0 = \delta I[\gamma_s] = \sum p_i \delta q^i|_{\partial I} + \int_I \sum \left( \dot{p}_i \delta q^i - \dot{q}^i \delta p_i - H_{q^i} \delta q^i - H_{p^i} \right),$$

holds for all admissible variations of this type, $\gamma_0$ is a physical orbit. The boundary term vanishes because of $\delta q^i|_{\partial I} = 0$.

The variation vector $\delta q$ can be written in the form

$$\delta q(t) = \sum_{r=1}^k f_r(q(t)) e_r(q(t)),$$

where $e_r$ is an orthonormal family of vector fields over $c(I)$ that spans $W_{c(I)}$, and where $f_r : c(I) \to \mathbb{R}$ are arbitrary functions with $f_r(c(\partial I)) = 0$. Because $f_r$ and $\delta p$ are fully arbitrary, the terms in $\delta I[\gamma_s]$ which depend on $\delta q$, and the ones which depends on $\delta p$ must vanish independently; so the term involving $\delta q$ must satisfy

$$0 = \int_I dt \langle \dot{p} + H_{q^i}(q,p), e_r \rangle f_r$$

58
for all test functions $f_r$. This implies that $\langle (\dot{p} + H_p), e_r \rangle$ vanishes for all $r = 1, \ldots, k$, which is equivalent to the statement that $\alpha^T(p + H_p) = 0$. Because of $\beta(q)\dot{q} = 0$, which holds since $\gamma_0$ is an admissible curve, one obtains for the remaining term in $\delta I[\gamma_0]$

$$0 = \int_I dt \langle (\dot{q} - H_p(q,p)), \delta p \rangle = \int_I dt \langle (\dot{q} - \alpha(q)H_p(q,p)), \alpha^T(q)\delta p \rangle + \int_I dt \langle \beta(q)H_p(q,p), \beta^T(q)\delta p \rangle.$$

Because $\delta p$ is fully arbitrary, its components in the images of $\alpha^T(q)$ and $\beta^T(q)$ can be varied independently; hence, the same argument as above implies that $\dot{q} = \alpha(q)H_p(q,p)$ and $\beta(q)H_p(q,p) = 0$ must be satisfied.

The Euler-Lagrange equations of the Hölder variational principle are thus given by

$$\dot{q} = \alpha(q)H_p(q,p) \quad (3.4)$$
$$\alpha^T(q)\dot{p} = -\alpha^T(q)H_q(q,p) \quad (3.5)$$
$$0 = \beta(q)H_p(q,p). \quad (3.6)$$

A smooth path $(q(t), p(t))$ in $T^*Q$ that satisfies this system of differential relations is called a physical orbit of the constrained system (and correctly describes its time evolution). The third line defines an at least $n + k$-dimensional (the rank of $\beta(q)$ is $n - k$) submanifold $\mathcal{M}_{phys} := \{(q,p) \mid \beta(q)H_p(q,p) = 0\}$ of $T^*Q$ that contains all physical orbits, which will be referred to as the physical sheet. Recall that $M_{ij} := h_{ij}$ is the matrix of the Riemannian metric $h$ in the present chart. For the Hamiltonian (3.1), relation (3.6) gives

$$0 = \beta(q)M^{-1}(q)p = M^{-1}(q)\beta^T(q)p,$$

as a consequence of the orthogonality of $\beta$ with respect to $h$. Therefore, (3.6) is equivalent to the condition $\beta^T(q)p = 0$. Next, one has to prove that the system of differential relations obtained from the Euler-Lagrange equations of the Hölder principle indeed has solutions.

To this end, notice that the tangent bundle of every curve $\gamma : \mathbb{R} \to \mathcal{M}_{phys}$, viewed as an embedded 1-manifold, naturally lies in $T\mathcal{M}_{phys}$. If the differential relations (3.4) ~ (3.6) necessitate that there exists a unique section $X$ of $T\mathcal{M}_{phys}$ so that $\dot{\gamma}(t) = X|_{\gamma(t)}$ is satisfied, physical orbits do exist and are unique. In order to prove that there is indeed such an $X$, let us view $\mathcal{M}_{phys}$ as an embedded submanifold of $TT^*Q$, and pick a local bundle chart
with coordinates \((q, p)\) in \(T^*Q\) that intersects \(M_{phys}\). The differentials of the \(n\) component functions of \(\beta^T(q)p\) on \(M_{phys}\) span the rank \(n - k\) normal bundle of \(M_{phys}\).

Every section \(Y\) of \(TM_{phys}\) is locally represented by a rank \(n + n\) component vector \((Y(q), Y(p))\) in the present bundle chart, and is annihilated by the one forms \(d(\beta^T(q)p)_i\) for \(i = 1, \ldots, n\) (of which only \(n - k\) are linearly independent), where everything is evaluated on \(M_{phys}\). This is expressed by the condition

\[
\langle (\beta^T(q)p)_{i, q}, Y(q) \rangle + \langle (\beta^T(q)p)_{i, p}, Y(p) \rangle = \langle (\beta^T(q)p)_{i, q}, Y(q) \rangle + (\beta^T(q)p)_{i, p} = 0,
\]

which shows that the projection \(\beta^T(q)Y(p)\) is completely determined by the components \(Y(q)\), such that the mere knowledge of \((Y(q), \alpha^T(q)Y(p))\) on \(M_{phys}\) suffices to let one reconstruct the whole section \(Y\) of \(TM_{phys}\). Consequently, the right hand sides of (3.4) and (3.5) determine a unique section \(X\) of \(TM_{phys}\), so that every curve \(\gamma : \mathbb{R} \rightarrow M_{phys}\) that satisfies (3.4) and (3.5) must obey \(\dot{\gamma}(t) = X|_{\gamma(t)}\). Consequently, \(X\) generates the physical orbits of the constrained mechanical system, and from the existence and uniqueness theorem of ordinary differential equations, one infers their existence and uniqueness.

Finally, let us determine the equilibria of this dynamical system on \(M_{phys}\). These are obtained by setting \(\dot{q} = 0\) and \(\dot{p} = 0\) in (3.4) ~ (3.6), whereupon one arrives at

\[
p = 0 \quad \alpha^T(q)u(q) = 0.
\]

The critical set of the constrained Hamiltonian mechanical system at hand is thus given by

\[
C_Q := \{q \in Q|\alpha^T(q)u(q) = 0\}, \quad (3.7)
\]

which is generically an \(n - k\)-dimensional submanifold of the configuration manifold \(Q\), as an argument involving Sard’s theorem shows (the rank of \(\alpha^T(q)\) is \(k\)), cf. the beginning of chapter two. The relationship between the equilibria of the dynamical system (3.12) and \(C_Q\) is, unfortunately, not simple enough to be technically practicable, as will be argued below.

### 3.3 Continuations away from the physical sheet and issues on the numerical simulation of constrained Hamiltonian mechanical systems

So far, it has been clarified that the Hölder variational principle produces a system of differential relations which imply the existence of a unique vector field on \(M_{phys}\) that generates the physical flow of the constrained Hamiltonian mechanical system. Instead of trying to
construct this specific vector field on $\mathcal{M}_{\text{phys}}$ (which is not necessary for the engineering applications intended later), we will introduce an auxiliary constrained Hamiltonian system along the line of chapter one with a flow $\tilde{\Phi}_t \in \text{Diff}(T^*Q)$ ($t \in \mathbb{R}$ denotes the time coordinate). Its salient features will be that it exhibits $\mathcal{M}_{\text{phys}}$ as an invariant manifold, and that the orbits on $\mathcal{M}_{\text{phys}}$ generated by $\tilde{\Phi}_t$ are precisely the physical ones of the constrained system. A dynamical system $\tilde{\Phi}_t \in \text{Diff}(T^*Q)$ of this kind will be referred to as a continuation away from the physical sheet (of the constrained mechanical system on $\mathcal{M}_{\text{phys}}$).

For the numerical simulation of mechanical systems of the type at hand, it is particularly interesting that one can understand so-called 'redundant formulations' in the framework presently described (by definition, a redundant formulation describes the dynamics of a mechanical system in a manner that more coordinates are employed to express the equations of motion than there are physical degrees of freedom; in our case, the $2n$ phase space coordinates of the free system are used to describe the dynamics of the constrained system which only has $n + k$ degrees of freedom [36]). We will not delve into questions of numerical mathematics, but only remark that DAE's (differential algebraic equations) are very often used to numerically simulate the time evolution of a constrained system. If the constraints are non-holonomic, elimination of the Lagrange multipliers in the DAE's produces a dynamical system defined on the whole $T^*Q$ (if the constraints are holonomic, the situation can be slightly more complicated; in fact, one sometimes finds a condition that forces the orbits of the constrained system to stay on an integral manifold in $Q$). To study the numerically simulated orbits based on such a DAE formulation, one can hence focus on the dynamical system on $T^*Q$ to which they are equivalent. We will not write down the DAE systems related to the dynamical systems introduced below, but only remark that they exist.

In trying to design a numerical integration routine for a given constrained mechanical system, one desires a number of positive attributes as prerequisites, such as 'numerical stability' in the sense that all orbits of $\tilde{\Phi}_t$ should remain close to $\mathcal{M}_{\text{phys}}$ for all times if the initial condition is picked close to $\mathcal{M}_{\text{phys}}$. It is also advantageous if the equilibria of the auxiliary system $\tilde{\Phi}_t$ can be used to precisely determine the location of physical equilibria on $\mathcal{M}_{\text{phys}}$. Moreover, it is certainly desirable that the total energy as well as eventual symmetries shall be conserved by $\tilde{\Phi}_t$. In the following discussion, we will first present a continuation away from the physical sheet of the type presented in chapter one that turns out to be insufficient with respect to these requirements, which will make clear some of the
problems involved in finding a suitable auxiliary system on $T^*Q$. Then, we will discuss a particular continuation of this kind which satisfies all of the above.

Throughout this thesis, we will not address detailed issues of numerical mathematics (such as the precise construction and discussion of DAE solvers etc.). The literature on the numerical mathematics of constrained multibody systems is very extensive, and the theoretical considerations made here should merely add an interpretation in the light of constrained Hamiltonian dynamics. We emphasize that since this important area of problems is very broad, it should be clear that we will only be able to address a very small fraction of important questions.

### 3.3.1 A preliminary attempt

Let us make a first attempt at constructing a dynamical system on $T^*Q$, for which $\mathcal{M}_{\text{phys}}$ is an invariant manifold, and which satisfies (3.4), (3.5) on $\mathcal{M}_{\text{phys}}$. We first have to state the necessary conditions that must be obeyed by its orbits $(q(t), p(t))$. On $\mathcal{M}_{\text{phys}}$, the components $\dot{q}$ and $\alpha^T(q)\dot{p}$ of the tangent vector are precisely determined by (3.4) and (3.5). To determine the component $\beta^T(q)\dot{p}$, we take the time derivative of $\beta^T(q)p$ along an orbit $(q(t), p(t))$ of this dynamical system, which gives

$$\frac{d}{dt} \left( \beta^T(q(t)) p(t) \right) = \beta^T(q(t)) \dot{p}(t) - \left( \frac{d}{dt} \alpha^T(q(t)) \right) p(t)$$

(3.8)

(using $\beta_{,j} = -\alpha_{,j}$). This expression must vanish on $\mathcal{M}_{\text{phys}}$, that is, for $\beta^T(q)p = 0$, to guarantee that $\mathcal{M}_{\text{phys}}$ is an invariant manifold. Multiplying equation (3.8) with $\alpha^T(q(t))$ from the left gives

$$\alpha^T(q(t)) \frac{d}{dt} \left( \beta^T(q(t)) p(t) \right) = \left( \frac{d}{dt} \alpha^T(q(t)) \right) \beta^T(q(t)) p(t),$$

which follows from

$$\alpha^T(q) \alpha^T_j(q) = \alpha^T_j(q) \beta^T(q)$$

(3.9)

(this is obtained from $((\alpha^T)^2)_{,j} = \alpha^T_{,j}$). This expression vanishes on $\mathcal{M}_{\text{phys}}$. Multiplying (3.8) by $\beta^T(q(t))$ from the left gives

$$\beta^T(q(t)) \frac{d}{dt} \left( \beta^T(q(t)) p(t) \right) = \beta^T(q(t)) \dot{p}(t) - \left( \frac{d}{dt} \alpha^T(q(t)) \right) \alpha^T(q(t)) p(t),$$
(using $\beta^T\alpha^T = \alpha^T_0\alpha^T$) which vanishes if one requires that

$$\beta^T(q(t))\dot{p}(t) = \left( \frac{d}{dt} \alpha^T(q(t)) \right) \alpha^T(q(t))p(t) \quad (3.10)$$

is satisfied.

For the particular Hamiltonian (3.1), (3.4) can be written as

$$\dot{q} = \alpha(q)M^{-1}(q)p. \quad (3.11)$$

Inserting (3.11) for $\dot{q}$ in (3.10), and adding the resulting expression for $\beta^T(q)\dot{p}$ to (3.5), one arrives at the system of ordinary differential equations on $T^*Q$ given by

$$\dot{q} = \alpha(q)M^{-1}(q)p$$

$$\dot{p} = -\alpha^T(q)H_{q}(q,p) + \Lambda(q,\alpha^T(q)p)\alpha^T(q)p, \quad (3.12)$$

where we have defined

$$\Lambda^i_k(q,p) := \alpha^i_k(q)M^{im}(q)p_m$$

(along orbits of this dynamical system, one has $\Lambda(q(t),\alpha^T(q(t))p(t)) = \frac{d}{dt}\alpha^T(q(t))$).

The dynamical system (3.12) is well-defined on the whole $T^*Q$, and by construction, it is clear that $M_{phys}$ is an invariant manifold of the associated flow. Therefore, every orbit with an initial condition contained in $M_{phys}$ is a physical orbit. We emphasize that every vector field on $T^*Q$ that agrees with (3.12) on $M_{phys}$, and for which $M_{phys}$ is an invariant manifold of the corresponding flow, can in principle be used to study or numerically simulate the physical orbits of the constrained Hamiltonian mechanical system on $M_{phys}$.

We should now verify that (3.12) is a constrained Hamiltonian system of the type introduced in chapter one. Along all of its orbits,

$$\beta(q)\dot{q} = 0 \quad \beta^T(q)\dot{p} - \Lambda(q,\alpha^T(q)p)M(q)\dot{q} = 0$$

is satisfied (here, $M(q)\dot{q}$ has been substituted for $\alpha^T(q)p$). The kernel of the matrix

$$L(q,p) := \begin{pmatrix} \beta(q) & 0 \\ -\Lambda(q,\alpha^T(q)p)M(q) & \beta^T(q) \end{pmatrix}$$

locally determines a distribution $V_{(1)}$ of rank 2$k$ in $TT^*Q$ (the rank of $\beta$ is $n-k$), to which all orbits of (3.12) are tangent.

Let us show that $V_{(1)}$ is almost symplectic. To this end, one must prove that the restriction of the symplectic structure to $V_{(1)}$ is non-degenerate. This is satisfied if and only
if its symplectic complement $V^\perp_{(1)}$ in $TT^*Q$ is almost symplectic, which in turn holds if the rank of $J_L^L^T$ equals the rank of $L$. Direct calculation gives

$$L(q,p)J_L^L^T(q,p) = \begin{pmatrix} 0 & \beta(q) \\ \beta^T(q) & \kappa(q,p) \end{pmatrix},$$

with

$$\kappa(q,p) := \beta^T(q)M(q)\Lambda^T(q,\alpha^T(q)p) - \Lambda(q,\alpha^T(q)p)M(q)\beta(q),$$

hence its rank is $2n - 2k$. This implies that $V_{(1)}$ is almost symplectic.

Moreover, we claim that

$$P_{(1)}(q,p) := \begin{pmatrix} \beta(q) & 0 \\ \kappa(q,p) & \beta^T(q) \end{pmatrix}$$

is the matrix of the $\omega$-skew selfadjoint bundle projector $TT^*Q \to V^\perp_{(1)}$, with $V^\perp_{(1)}$ denoting the symplectic complement of $V_{(1)}$ in $TT^*Q$. To prove that it is a projector, one only needs the relationship

$$\beta^T(q)\Lambda(q,p)\beta(q) = 0,$$

which follows from $\beta^T(q)\alpha^T_j(q)\beta(q) = 0$, due to (3.9). Therefore, it is immediately clear that the Hamiltonian $H$ is an integral of motion for (3.12). A small calculation shows that $\dot{x} = P_{(1)}(x)JH_{\dot{x}}(x)$ for $x = (q,p)$ is indeed equivalent to (3.12).

However, there are two adverse properties of this dynamical system. Firstly, the critical set of (3.12) is fairly complicated to determine, and to analyze. Secondly, orbits that start close to $M_{phys}$ will in most cases move away from $M_{phys}$. A little more precisely, the equilibria of (3.12) are given by

$$\alpha^T(q)p = 0, \quad \alpha^T(q)H_{\dot{q}}(q,\beta^T(q)p) = 0.$$

There are no restrictions to the values of the covector $\beta^T(q)p$, therefore, there is no very obvious relationship between equilibria away from $M_{phys}$, where $\beta^T(q)p$ is not zero, to those on $M_{phys}$. Furthermore, (3.12) has been constructed in a manner that

$$\frac{d}{dt} \left( \beta^T(q(t))p(t) \right) = -\left( \frac{d}{dt} \alpha^T(q(t)) \right) \beta^T(q(t))p(t)$$

holds, as one easily verifies from (3.8) and (3.10). This shows that even if the norm of $\beta^T(q)p$ is very small, its value will in general fluctuate severely at locations where $\|\frac{d}{dt} \alpha^T(q(t))\|$ (relative to some matrix norm) is large – this is the case at points in $Q$ where $W$ varies strongly along the orbit. Hence, one has to expect substantial numerical errors from numerical simulations based on implementing (3.12) if $W$ is not sufficiently mildly behaved. As an example, approximate singularities of $W$ can emerge from impact-like reactive forces in the links between rigid parts, due to critical values of the system parameters [11].
3.3.2 A more refined auxiliary dynamical system

We will now construct a continuation away from the physical sheet that is significantly more suitable for numerical purposes than our previous attempt, if adequately implemented. The subsequent discussion is inspired by an earlier publication on this problem [36]; other closely related works on the topic, with a strong emphasis on numerical issues, are for instance [20, 32]. In all of these papers, the authors have numerically simulated holonomically constrained mechanical test systems, for which continuations similar to those considered below have produced good results.

The auxiliary system that we intend to construct is a constrained Hamiltonian system of the type introduced in chapter one. We recall that such a system is determined by the introduction of a symplectic distribution $V$ over the symplectic manifold $(T^*Q, \sum dq^i \wedge dp_i)$. The distribution we want to work with is defined as follows. Let us pick a smooth, orthonormal family of one forms $\{\zeta_1, \ldots, \zeta_{n-k}\}$ which locally spans the annihilator $W_\beta$ of $W$, and let us write the defining relation $\beta^T(q)p = 0$ for $M_{\text{phys}}$ in the form

$$f_I(q, p) := h^*_q(p, \zeta_I(q)) = 0$$

(3.13)

with $I = 1, \ldots, n - k$, where the maps $f_I : T^*Q \to \mathbb{R}$ are smooth. We require that the level sets

$$M_\mu := \{(q, p) \mid f_I(q, p) = \mu_I, I = 1, \ldots, n - k\},$$

with $\mu := (\mu_1, \ldots, \mu_{n-k})$, are invariant manifolds, and it is clear that the leaf $M_0$ corresponds to $M_{\text{phys}}$. We also require that the condition $\beta(q)\dot{q} = 0$ shall additionally be satisfied along all of its orbits. This defines $V$, as we will see.

On $M_0$, this system exhibits the physical flow of the constrained mechanical system. (This is because we have proved above that the physical flow on $M_{\text{phys}}$ can be continued into $T^*Q$ by a constrained Hamiltonian system of the type introduced in chapter one. The orbits of this dynamical system satisfy $\beta(q)\dot{q} = 0$ on the invariant manifold $M_{\text{phys}} = M_0$, and its vector field is the image of the Hamiltonian vector field under the $\omega$-skew selfadjoint projector $P(1)(q, p)$. The dynamical system that we are about to construct is also obtained from a projection of the free Hamiltonian vector field $X_H$, whose flow exhibits $M_{\text{phys}} = M_0$ as an invariant manifold, together with $\beta^T(q)\dot{q} = 0$. Therefore, evaluating the projectors of both continuations on $M_{\text{phys}}$ gives the same result, so both vector fields must be identical on $M_{\text{phys}}$.)

65
There are, as will be shown below, two main characteristics of this auxiliary dynamical system that make it more suitable for applications than the previous one. Firstly, its critical manifold simply is a rank $n - k$ vector bundle over the critical manifold $C_\mathcal{Q}$ of the physical system on $\mathcal{M}_{\text{phys}}$, and much easier to analyze. Secondly, the function $v(q, p) := \|\beta^T(q)p\|^2$ (where the norm is defined via $h^*$) is an additional integral of motion, apart from $H$. The function $v : T^*Q \to \mathbb{R}^+$ is positive definite away from $\mathcal{M}_{\text{phys}}$, and has a (degenerate) minimum of value zero on $\mathcal{M}_{\text{phys}}$. Hence, it is a Lyapunov function for $\mathcal{M}_{\text{phys}}$, and the invariant manifold $\mathcal{M}_{\text{phys}}$ is (critically, but not asymptotically) stable. We will now prove these facts.

The orbits $\gamma : \mathbb{R} \to T^*Q$ of the constrained Hamiltonian system we intend to construct exhibit $\mathcal{M}_\mathcal{L}$ as invariant manifolds, thus

$$\langle df_I, \dot{\gamma} \rangle = 0$$

is satisfied for all $I$. The additional condition $\beta(q)\dot{q} = 0$ on $TQ$ can be lifted to a condition on $TT^*Q$, by considering the one forms $\zeta_I$ on $Q$ as one forms on $T^*Q$, whose coefficients relative to the basis one forms $dp_i$ are all zero. This allows one to write the condition

$$\langle \zeta_I, \dot{\gamma} \rangle = 0$$

in $TT^*Q$, which shall hold for all $J$. Let $V$ denote the distribution over $T^*Q$ given by the intersection of the kernels of all $\zeta_I$'s and $df_I$'s.

We claim that $V$ is almost symplectic. To prove this, notice that $V$ is almost symplectic if and only if its symplectic complement $V^\perp$ is almost symplectic; the latter condition is easier to check here. In order to prove almost symplecticity for $V^\perp$, let us pick a family of spanning vector fields $(Y_1, \ldots, Y_{2k})$ for $V^\perp$ given by

$$\omega(Y_I, \cdot) = \zeta_I \quad \omega(Y_{I+k}, \cdot) = df_I,$$

(3.14)

with $I = 1, \ldots, k$, and $\omega = -df(p_idq^i)$. It spans $V^\perp$ because all sections $X$ of $V$ satisfy $\omega(Y_I, X) = \langle \zeta_I, X \rangle = 0$. In local bundle coordinates of $T^*Q$, we have

$$df_I = f_{I,I^*}(q, p) dq^i + M^{ij}_I(q) \zeta_I(q) dp_j \quad \zeta_I = \zeta_{I^*} dq^i.$$

The requirement of almost symplecticity of $V^\perp$ is the same as demanding that $C_{IJ} = \omega(Y_I, Y_J)$ shall be the components of a smooth $GL(2k)$-valued function on $T^*Q$. In the
present notation, capital indices range from 1 to $k$ if they label one forms, and from 1 to $2k$ if they label vector fields.

We adopt the notation used in [10], and denote the $k$ by $n$ matrix with $I$-th row vector $(\zeta_{Ii}(q))$ by $E^T$. Furthermore, we denote the matrix with $I$-th row vector $(f_{I,i}(q,p))$ by $D^T$. We also introduce the $2k$ by $2n$ matrix

$$Z(q,p) := \begin{pmatrix} E^T(q) & 0 \\ D^T(q,p) & E^T(q)M^{-1}(q) \end{pmatrix}.$$  

The conditions (3.14) are expressed by $Z(q,p)x = 0$ in the present chart, where $x = (q,p)$. The symplectic structure is locally represented by the symplectic standard matrix $J$ (1.1), and one easily verifies that the row vectors of the matrix $Z(q,p)J^{-1}$ are the components of the vectors $Y_I$; hence, the $2k$ by $2k$ matrix $[C_{ij}]$ (that is, the matrix $C$ introduced in 1.3.2.), is here given by

$$C(q,p) = Z(q,p)JZ^T(q,p) = \begin{pmatrix} 0 & G(q) \\ -G(q) & S(q,p) \end{pmatrix},$$

where we have introduced

$$G(q) := E^T(q)M^{-1}(q)E(q)$$

$$S(q,p) := D^T(q,p)M^{-1}(q)E(q) - E^T(q)M^{-1}(q)D(q,p).$$

The matrix $G(q)$ is invertible, because the row vectors of $E^T(q)$ are linearly independent, which implies that $C(q,p)$ is invertible. Therefore, the distribution $V^\perp$, together with its symplectic complement $V$, are almost symplectic, as claimed.

In a next step, let us construct the $\omega$-skew selfadjoint bundle projector $P : TT^*Q \to V^\perp$. The recipe presented in 1.3.2 requires one to determine the inverse of the matrix $C(q,p)$, which is here given by

$$C^{-1}(q,p) = \begin{pmatrix} G^{-1}(q)S(q,p)G^{-1}(q) & -G^{-1}(q) \\ G^{-1}(q) & 0 \end{pmatrix}.$$  

The $I$-th column vector of the matrix $Z(q,p)J^{-1}$ is given by the components of the vector field $Y_I$. Hence, using the construction presented in 1.3.2, we see that the matrix representation of $P$ is, in this local bundle chart of $T^*Q$, given by

$$\bar{P}(q,p) = JZ^T(q,p)C^{-1}(q,p)Z(q,p).$$
We remark that the construction presently carried out for $\tilde{P}$ can also be applied to the orthogonal bundle projector $\beta : TQ \to W$, where $\omega$ is replaced by the Riemannian metric $h$ on $Q$. The rôle of $C(q,p)$ is then attributed to $G(q)$, which is invertible, as we have already proved. In the same fashion as in 1.3.2, one shows that in the given bundle chart, the matrix of the projector $\beta(q)$ is

$$\beta(q) = M^{-1}(q)E(q)G^{-1}(q)E^T(q).$$

For more details, we refer to [10]. A straightforward calculation lets one conclude that

$$P(q,p) = \begin{pmatrix} \beta(q) & 0 \\ T(q,p) & \beta^T(q) \end{pmatrix},$$

where we have defined

$$T(q,p) := E(q)G^{-1}(q)D_T(q,p)\alpha(q) - \alpha^T(q)D(q,p)G^{-1}(q)E^T(q). \quad (3.15)$$

Because the family of one-forms $\zeta_I$ has been picked to be orthonormal, $G(q)$ is simply the $n-k$ dimensional unit matrix. The matrix of the projector $P : TT^*Q \to V$ is obtained from $P = 1_{2n} - \tilde{P}$,

$$P(q,p) = \begin{pmatrix} \alpha(q) & 0 \\ -T(q,p) & \alpha^T(q) \end{pmatrix}$$

in the present chart.

The identity

$$P(x)J = J P^T(x),$$

expresses the $\omega$-skew selfadjointness of $P$ in the present Darboux chart (in invariant notation, that would be $\omega(PX,Y) = \omega(X, PY)$ for all sections $X$, $Y$ of $TT^*Q$). The equations of motion $\dot{x} = P(x)X_H(x)$ of the constrained system, with $X_H$ denoting the Hamiltonian vector field of $H$, can be written as $\dot{x} = P(x)JH_{\varphi}(x)$, or in components of the bundle chart,

$$\begin{pmatrix} \dot{q} \\ \dot{p} \end{pmatrix} = \begin{pmatrix} 0 & \alpha(q) \\ -\alpha^T(q) & -T(q,p) \end{pmatrix} \begin{pmatrix} H_\varphi(q,p) \\ H_\varphi(q,p) \end{pmatrix}. \quad (3.16)$$

The matrix on the right hand side of the equality sign is antisymmetric, of rank $2k$. Due to its antisymmetry, the energy $H$ is constant along all orbits of this dynamical system. By construction, $df_I(PX_H) = 0$ is satisfied; thus, $f_I$ are integrals of motion, as intended.

Next, we claim that the function $v(q,p) := \|\beta^T(q)p\|^2$ is an integral of motion (the norm is defined via $h^*$) for (3.16). To prove this, we write $\beta^T(q)p$ in the form

$$\beta^T(q)p = \sum \rho_I(q,p)\zeta_I(q)$$
with \( \rho_I(q,p) := h^*(\zeta_I(q), \beta^T(q)p) \), recalling the fact that the orthonormal family of one forms \( \zeta_I \) spans the image of \( \beta^T(q) \) for every \( q \in Q \). It is evident that the functions \( \rho_I : T^*Q \to \mathbb{R} \) are equal to the functions \( f_I \) (3.13), so that on each level set \( M_{\mu} \), one simply has

\[
\beta^T(q)p = \sum \mu_I \zeta_I(q).
\]

The function

\[
v(q,p) := h^*(\beta^T(q)p, \beta^T(q)p)
\]

has the constant value \( \sum \mu_I^2 \) on \( M_{\mu} \), and is an integral of motion of (3.16), because the level sets \( M_{\mu} \) are, by construction, invariant manifolds. The value of \( v(q,p) \) on the physical sheet \( M_{\text{phys}} \) is zero, and away from \( M_{\text{phys}} \), \( v(q, \cdot) \) is positive. In other words, \( M_{\text{phys}} \) is the (degenerate) minimum of \( v \).

This makes \( v \) a Lyapunov function for \( M_{\text{phys}} \), and proves that \( M_{\text{phys}} \), as an invariant manifold, is stable under the flow of (3.16). Of course, \( M_{\text{phys}} \) is only critically stable (that is, the eigenvalues of \( PX_H \) transverse to \( M_{\text{phys}} \) are purely imaginary), but not asymptotically stable; because of energy conservation, one cannot expect anything better than critical stability. It would, of course, be possible to construct a continuation away from the physical sheet for which \( M_{\text{phys}} \) is an asymptotically stable invariant manifold, but we will not discuss dissipative systems here.

Finally, let us determine the critical set \( \mathcal{C} \) of the constrained Hamiltonian system (3.16). To this end, we consider its vector field on a single leaf \( M_{\mu} \), to which it is, by construction, tangent. Recall that \( \zeta_I \) is an orthonormal family of one forms, so that \( G(q,p) \) is the \( n-k \)-dimensional unit matrix, implying that (3.16) is given by

\[
\begin{align*}
\dot{q} &= \alpha(q)M^{-1}(q)p \\
\dot{p} &= -\alpha^T(q)\partial_q H(q,p) - E(q)D^T(q,p)\alpha(q)M^{-1}(q)p + \alpha^T(q)D(q,p)E^T(q)M^{-1}(q)p.
\end{align*}
\]

The second line can be written as

\[
\dot{p} = -\alpha^T(q)\partial_q H(q,p) - E(q)D^T(q,p)\dot{q} + \alpha^T(q)D(q,p)f(q,p)
\]

with

\[
f(q,p) := (f_1(q,p), \ldots, f_{n-k}(q,p))^T = E^T(q)M^{-1}(q)p.
\]
The invariance of $M_\mu$ under the flow is expressed by $\frac{d}{dt} f_I(q(t), p(t)) = 0$ along all orbits of (3.16), which is equivalent to $D^T(q, p) \dot{q} + E^T(q) M^{-1}(q) \dot{p} = 0$, from which one obtains that

$$\dot{p} = -\alpha^T(q) \partial_q H(q, p) + E(q) E^T(q) M^{-1}(q) \dot{p} + \alpha^T(q) \partial_q \left( \frac{1}{2} f^T(q, p) f(q, p) \right).$$

The second term on the right hand side simply corresponds to $\beta^T(q) \dot{p}$, and the scalar function $f^T(q, p) f(q, p)$ equals $v(q, p)$. Notice that the Hamiltonian (3.1) decomposes into

$$H(q, p) = H(q, \alpha^T(q)p) + \frac{1}{2} v(q, p),$$

owing to the fact that $\alpha^T$ and $\beta^T$ are orthogonal with respect to $h^*$. Therefore, one finds that

$$\dot{p} = -\alpha^T \partial_q H(q, \alpha^T(q)p) + \beta^T \dot{p}.$$  

Setting $\dot{q} = 0$ and $\dot{p} = 0$ in (3.16), one understands that the equilibria of (3.16) are determined by the conditions

$$0 = \alpha^T(q)p \quad \quad 0 = -\alpha^T(q) \partial_q H(q, 0),$$

which impose no restriction on the values of the quantity $\beta^T(q)p$. The critical manifold of the present constrained Hamiltonian system is thus given by the set

$$\mathcal{C} = \left\{ (q, p) \mid q \in \mathcal{C}_Q; \alpha^T(q)p = 0; \beta^T(q)p = \sum_I \mu_I \zeta_I(q) \right\},$$

which is a vector bundle over the base space $\mathcal{C}_Q$, (3.7), whose fibres equal those of the annihilator $W^*_\beta$ of $W$ on $\mathcal{C}_Q$. Every equilibrium $(q_0, p_0)$ of the artificial constrained Hamiltonian system $P X_H$ on $T^*Q$ defines a unique equilibrium $q_0$ of the physical system on $\mathcal{M}_{phys}$.

The conclusion so far is that the continuation away from the physical sheet defined by (3.16) is a good candidate for numerical simulation routines because (1) the physical sheet is a stable invariant manifold under its flow, and (2) its critical set is trivially related to the critical set $\mathcal{C}_Q$ of the physical system, which may also apply to other possible continuations away from the physical sheet. There is, however, an additional strength that makes it somewhat special, as will be explained next.
3.4 A note on constrained Hamiltonian mechanical systems with symmetries

We will finally dedicate a few sentences to the topic of constrained Hamiltonian mechanical systems which exhibit symmetries. We will show that systems of the discussed type can be suitably continued away from the physical sheet by a system of the class (3.16) so that the symmetries are also continued. To be more precise, assume a constrained Hamiltonian mechanical system to be given that is defined in terms of the Hamiltonian (3.1), and Pfaffian constraints determined by a smooth rank \( k \) distribution \( W \) over \( Q \), as before.

A constrained Hamiltonian mechanical system that exhibits symmetries is defined as follows. Let \( G \) be a Lie group, and let

\[
\psi: G \times Q \rightarrow Q
\]

\[(g, q) \mapsto \psi_g(q),\]

be a group action on the configuration manifold (that is, \( \psi_g \in \text{Diff}(Q) \) for all \( g \in G \), \( \psi_e = \text{id} \), where \( e \) denotes the unit in \( G \), and \( \psi_{g_1} \psi_{g_2} = \psi_{g_1 g_2} \)) so that:

1. The Riemannian metrics \( h \) and \( h^* \) are invariant with respect to \( \psi \), that is, \( h \circ \psi_g = h \), and \( h^* \circ \psi_g = h^* \) for all \( g \in G \).
2. The potential energy is invariant with respect to \( \psi \), that is, \( u \circ \psi_g = u \) for all \( g \in G \).
3. The distribution \( W \) and its dual \( W^* \) are invariant with respect to \( \psi \), that is, \( \psi_g W = W \) and \( \psi_g^* W^* = W^* \) for all \( g \in G \).

It is always possible to pick the local orthonormal family of one forms \( \zeta_I \) (which span the annihilator of \( W \)) in a manner that \( \psi_g^* \zeta_I = \zeta_I \) is satisfied for all \( g \in G \) in a vicinity of the unit element \( e \), so that the functions \( f_I(q, p) = h^*_g(\zeta_I, p) \) and their level sets \( M_g \) are invariant under the group action. The group action on \( Q \) induces, via the pullback map, the group action

\[
\Psi := \psi^*: G \times T^*Q \rightarrow T^*Q
\]

on the phase space \( T^*Q \). A group action on a symplectic manifold which leaves the symplectic structure invariant is said to act \textit{symplectically} on it. As a matter of fact, \( \Psi_g \) acts symplectically on \((T^*Q, \omega = -d\theta_0)\) with \( \theta_0 = \sum p_i dq^i \) (for a proof, consider for instance
Then, lifting $\zeta$ to one forms on $T^*Q$, one has $\Psi^*_g \zeta = \zeta$, and $f_I \circ \Psi_g = f_I$ for all $g \in G$. This implies that 

$$\Psi^*_g df_I = df_I,$$

so that the common kernel of $\zeta$ (lifted to one forms on $T^*Q$) and $df_I$, which has been defined as the distribution $V$ over $T^*Q$, satisfies

$$\Psi^*_g V = V$$

for all $g \in G$. Invariance of the Hamiltonian with respect to the group action is expressed by $H \circ \Psi_g = H$, and hence

$$\Psi^*_g dH = dH$$

for all $g \in G$. One can locally introduce a family of spanning one forms $X_r$ for $V$ which satisfy $\Psi_g X_r = X_r$, and carry out the construction for the bundle projector $P : TT^*Q \to V$ as presented in 1.3.2. This only involves linear algebra applied to $\omega$ and $X_r$, which are both invariant under the group action, so that the $\omega$-skew selfadjoint bundle projector $P$ satisfies $P \circ \Psi_g = P$, and that its dual satisfies $P^T \circ \Psi_g = P^T$ for all $g \in G$.

The vector field $X_c^f = PX_H$ of the constrained Hamiltonian system satisfies

$$\omega(X_c^f, \cdot) = -dH(P\cdot).$$

Application of $\Psi_g^*$ on both sides of the equality sign leads to the conclusion that

$$\Psi_g^* X_c^f = X_c^f$$

is satisfied for all $g \in G$. In other words, the vector field of the constrained Hamiltonian system $X_c^f$ is invariant under the group action $\Psi$ on $T^*Q$ that has, as we recall, been obtained from lifting the group action $\psi$ on $Q$.

The analysis of constrained Hamiltonian systems with symmetries is a whole research field of its own, and we will only say this much about that topic. In particular, notice that we strictly focus on equilibria of constrained Hamiltonian systems, and that we do not consider the phenomenon of steady motions at all (this interesting class of topics is for instance discussed in [26, 42], where a nice set of mechanical examples is presented, such as the celtic stone).
CHAPTER 4
EQUILIBRIA IN CONSTRAINED HAMILTONIAN MECHANICS

In this chapter, two topics are addressed. Firstly, we study the topology of the critical manifold $C$ of the auxiliary dynamical system \((3.16)\), which has been defined in section 3.4. It is a constrained Hamiltonian system of the type of chapter one, and continues the vector field of the physical constrained Hamiltonian mechanical system away from the physical sheet $\mathcal{M}_{phys} \subset T^*Q$ over the whole $T^*Q$. The critical manifold $C$ is in this case \textit{never} compact, but nevertheless, the fact that both $T^*Q$ and $C$ are vector bundles allows one can be used to derive results that are in full analogy to those in chapter two. The method is again based on a suitable application of the Conley-Zehnder construction, but here, one considers $Q$, which we assume to be closed, compact, instead of $T^*Q$.

Secondly, we discuss the stability of equilibrium solutions by use of methods that benefit substantially from the projector formalism developed in the previous chapters.

4.1 The topology of the critical manifold

This section addresses the global topology of the critical set of the constrained Hamiltonian mechanical system \((3.16)\), defined by

$$C = \left\{(q,p) \mid q \in C_Q; \alpha^T(q)p = 0; \beta^T(q)p = \sum \mu_i \zeta_i(q) \right\}.$$ 

Here, $\zeta_i$ is an orthonormal spanning family of one forms for the rank $n - k$ annihilator of the rank $k$ distribution $W$, and

$$C_Q = \left\{ q \in Q \mid \alpha^T(q)u_q(q) = 0 \right\}$$

is the critical set of the physical system on $\mathcal{M}_{phys}$. We recall that generically, $C_Q$ is a smooth $n - k$-dimensional submanifold of $Q$. Evidently, $C$ is the smooth rank $n - k$ vector bundle

$$C = W^*_\beta |_{C_Q}$$
over the base manifold $C_Q$, whose fibres are given by those of the annihilator $W^*_{\beta}$ of $W$.

The arguments and results demonstrated in chapter two can be straightforwardly applied to the present problem. First of all, we claim that $C_Q$ is normal hyperbolic with respect to the gradient-like flow $\psi_t$ generated by

$$\dot{q} = -\alpha(q)\text{grad}_h u(q),$$

and that it contains all critical points of the Morse function $u$, but no other conditional extrema of $u|_{C_Q}$ apart from those (it is gradient-like because along all of its non-constant orbits, $\frac{d}{dt}u(t) = -h(\alpha\text{grad}_h u, \alpha\text{grad}_h u)|_{q(t)} < 0$ holds, as $\alpha$ is an orthoprojector with respect to the Riemannian metric $h$). The fact that this is true can be proved by substituting $M \rightarrow Q$, $H \rightarrow u$, $P \rightarrow \alpha$, $g \rightarrow h$, and $C \rightarrow C_Q$ in chapter two, and by applying the arguments used there. Hence, letting $\mu_i$ denote the index of the connected component $C_{Q_i}$, defined as the dimension of its unstable manifold, the CZ inequalities (2.10) imply that for a compact, closed $Q$, the topological formula

$$\sum_{i,p} t^{p+\mu_i} \dim H^p(C_{Q_i}) = \sum_p t^p \dim H^p(Q) + (1 + t)Q(t)$$

holds, where $C_{Q_i}$ are the connected components of $C_Q$. Here, $H^p$ denotes the $p$-th de Rham cohomology group with suitable coefficients, and $Q(t)$ is a polynomial with non-negative integer coefficients. The argument using the Witten complexes associated to $(Q,u)$ and to $(C_Q, u|_{C_Q})$ to derive (4.1) can also be carried out in the manner explained in section 2.4.

Our next issue is to discuss the global topology of $C$. Clearly, $C$ is not a compact submanifold of $T^*Q$, hence the Conley-Zehnder inequalities of the second chapter, which were derived for compact, closed, generic critical manifolds, do not apply. However, since both $C$ and $T^*Q$ are vector bundles of a particular type, one can nevertheless prove a result that is closely related to (2.10). Our claim is that given the above stated properties of $C_Q$, the generalized CZ inequalities

$$\sum_{i,p} t^{p+\mu_i} \dim H^p_c(C_i) = \sum_p t^p \dim H^p_c(T^*Q) + (1 + t)Q(t)$$

are valid. In this formula, the connected components $C_i$ of $C$ are vector bundles whose base manifolds are the connected components $C_{Q_i}$ of $C_Q$, and the numbers $\mu_i$ are the indices of $C_{Q_i}$ with respect to $\psi_t$. The polynomial $Q(t)$ exhibits non-negative integer coefficients, and $H^*_c$ denotes the de Rham cohomology based on differential forms with compact supports.
To prove (4.2), we use the fact that the base space of any vector bundle is a deformation retract the vector bundle. Hence, $C_Q$, being the zero section of $C$, is a deformation retract of $C$, and likewise, $Q$ is a deformation retract of $T^*Q$. Since the de Rham cohomology groups are invariant under retraction, one infers the equality

$$H^p_c(C_i) \cong H^p(C_{Q_i}) \quad H^p_c(T^*Q) \cong H^p(Q).$$

Hence, formula (4.2) is equivalent to the assertion that

$$\sum_{i,p} t^{p+\mu_i} \dim H^p(C_{Q_i}) = \sum_p t^p \dim H^p(Q) + (1 + t) \mathcal{Q}(t). \quad (4.3)$$

But this has just been proved by application of CZ theory to the flow $\psi_t$ on $Q$. The weak Conley-Zehnder inequalities derived from this result are hence given by

$$\sum_i \dim H^{p-\mu_i}(C_{Q_i}) \geq B_p,$$

where $B_p$ is the $p$-th Betti number of $Q$. In particular, for the special value $t = -1$, one obtains

$$\sum_{i,p} (-1)^{p+\mu_i} \dim H^p(C_{Q_i}) = \sum_i (-1)^{\mu_i} \chi(C_{Q_i}) = \chi(Q),$$

where $\chi$ evaluates the Euler characteristic of the closed, compact manifold in its argument. This concludes our brief discussion of the global topology of the critical manifold $C_Q$.

### 4.2 Stability of equilibria in constrained Hamiltonian systems

From this section on, we will focus on the stability of equilibria in general and mechanical constrained Hamiltonian systems, and will always assume that the generic situation $C = C_{\text{gen}}$ is at hand.

We will first discuss general constrained Hamiltonian systems as introduced in chapter one, and will approach the case of asymptotic behaviour with a standard application of the center manifold theorem. We will analyze the critically stable case only approximately, that is, in the context of averaging theory, but no theorem will be proved. The methods required to prove the averaging results belong into KAM and Nekhoroshev theory, and are beyond the scope of this thesis. Finally, these general results will be applied to constrained Hamiltonian mechanical systems.
The equations of motion in a neighborhood of a critical point

Thus, let a general constrained Hamiltonian system of the kind described in chapter one be given. We will neither assume that the symplectic manifold $M$ nor that the critical manifold $C_{\text{gen}}$ is compact, but we assume that $C = C_{\text{gen}}$. The first issue to be addressed is the construction of a local coordinate chart in which the equations of motion are represented in a particularly transparent form.

Let $a \in C_{\text{gen}}$ be an equilibrium of the constrained system, and pick some small neighborhood $U(a)$ together with an associated Darboux chart, whose origin lies at $a$. In this chart, the equations of motion are given by

$$\frac{d}{dt}x = P(x)JH_{\omega}(x) =: X_H^\omega(x),$$

where $x = (x_1, \ldots, x^n, x_{n+1}, \ldots, x_{2n})$ is the vector of coordinate components, and where $\mathcal{J}$ is the symplectic standard matrix. $H_{\omega}$ denotes the component vector of the partial derivatives of $H$, while $P(x)$ is the matrix of the $\omega$-skew selfadjoint projector $P : TM \to V$, with $P(x)\mathcal{J} = \mathcal{J}P^T(x)$. The right hand side of (4.4) can be written in the form $\mathcal{J}P^T(x)H_{\omega}(x)$.

Next, let us introduce coordinates

$$u := P_0 x, \quad \bar{y} := P_0 x$$

that are mutually orthogonal with respect to the Kähler metric $g|_0$ in $\mathbb{R}^{2n} \cong T_0M$. Thus, $\mathbb{R}^{2n}$ factors into $\mathbb{R}^{2n-2k} \times \mathbb{R}^{2k}$, because the rank of $P$ is $2k$, with $\bar{P}$ denoting its complementary projector.

We claim that locally, there exists a function $F : \mathbb{R}^{2n-2k} \to \mathbb{R}^{2k}$, $u \mapsto F(u)$, for which $C_{\text{gen}}$ is its graph. This implies that every point $z \in C_{\text{gen}} \cap U(a)$ can be uniquely represented by a pair of vectors $(u_z, F(u_z))$, where $u_z = P_0 z$. To prove this claim, let us first show that the images of $P_0$ and $T_0C_{\text{gen}}$ together span $T_0M \cong \mathbb{R}^{2n}$. Indeed, writing $\bar{Q}_0$ for the orthoprojector $T_0M \to T_0C_{\text{gen}}$ as before, we observe that invertibility of the matrix

$$P_0 + \bar{Q}_0 = 1_{2n} + (P_0 - Q_0)$$

implies that $T_0M$ equals the image of $P_0 + \bar{Q}_0$. However, invertibility of this matrix has been proved during the discussion of the induced Witten complexes on $C_i$ in section 2.3.4. Since $P_0 + \bar{Q}_0$ is invertible, the rank of $\bar{P}_0\bar{Q}_0 = \bar{P}_0(P_0 + \bar{Q}_0)$ attains its maximal value $2(n - k)$, which is the rank of $\bar{P}_0$. This implies that $\bar{P}_0$ projects $T_0M$ surjectively onto its image.
Consequently, there exists a linear map $G : \mathbb{R}^{2n-2k} \to \mathbb{R}^{2k}$ for which $T_0M$ is the graph $(u, G(u))$. The embedded submanifold $C_{\text{gen}}$ thus admits a parametrization $(u, F(u))$ with $F(u) = G(u) + O(||u||^2)$, as long as $||u||$ is sufficiently small. This proves our claim.

Next, we introduce the function $y(\tilde{y}, u) := \tilde{y} - F(u)$, and apply the coordinate transformation

$$\Phi : (u, \tilde{y}) \mapsto (z, y).$$

The Jacobi matrix of its inverse $\Phi^{-1}$ at $(z, y)$ is

$$D\Phi^{-1}(z, y) = \begin{pmatrix} \bar{P}_0 + DF(u(z, y))\bar{P}_0 & 0 \\ -DF(u(z, y))\bar{P}_0 & P_0 \end{pmatrix},$$

where it follows from the definition of $F$ that $P_0DF(u) = DF(u)$. The equations of motion are now represented by

$$\frac{d}{dt}z = (\bar{P}_0 + DF(u(z, y))\bar{P}_0) X(z, y)$$
$$\frac{d}{dt}y = (P_0 - DF(u(z, y))\bar{P}_0) X(z, y),$$

where $X(z, y)$ stands for $X^c_H(\Phi^{-1}(z, y))$. In this chart, every coordinate $(z, 0)$ denotes a point on $C_{\text{gen}}$; thus, by definition of $C_{\text{gen}}$, it follows that $X(z, 0) = 0$ for all $z$.

![Figure 4.1: The present coordinate system in the vicinity of $a$.](image)

The Taylor series of $\Phi^{-1}$ relative to the origin of the coordinate system gives

$$\Phi^{-1}(z, y) = \Phi^{-1}(0, 0) + (D\Phi^{-1}(0, 0))(z, y) + O(||x||^2).$$
Since the present coordinate system has been picked in a manner that $F(0) = 0$, one understands that $\Phi^{-1}(0,0) = (0,0)$. Furthermore, Taylor expansion of $X_H^c(u, \tilde{y})$ relative to the origin yields

$$X_H^c(y, z) = (DX_H^c(0,0))(u, \tilde{y}) + \tilde{R}(u, \tilde{y}),$$

where $\tilde{R}(u, \tilde{y})$ is a quadratic Taylor remainder term.

Consequently, one finds

$$X(z, y) = (DX_H^c(0,0))(DF^{-1}(0,0))(z, y) + R(y, z)$$

with a quadratic Taylor remainder term $R(y, z)$. It follows from $P_0DX_H^c(0,0) = 0$ that the equations of motion in the $(z, y)$-coordinate system are given by

$$\frac{d}{dt}z = \left(P_0 + DF(u(z,y))P_0\right)R(z, y) =: Z(z, y) \tag{4.5}$$

$$\frac{d}{dt}y = P_0DX_H^c(0,0)P_0y - P_0(DX_H^c(0,0))(DF(u(z,y)))(z, y) + \tilde{Y}(z, y),$$

where $Z(z, y)$ and $\tilde{Y}(z, y)$ are quadratic Taylor remainder terms. Finally, since the kernel of $DX_H^c(0,0)$ is the tangent space $T_0C_{gen}$, it coincides with the image of $DF(0)$, so that $(DX_H^c(0,0))DF(0) = 0$.

Therefore, the second line in this system of ordinary differential equations can be written as

$$\frac{d}{dt}y = P_0DX_H^c(0,0)P_0y + Y(z, y), \tag{4.6}$$

with a quadratic Taylor remainder term $Y(z, y)$. Since every pair $(z, 0)$ labels a point on $C_{gen}$, it follows that $Z(z, 0) = 0$ and $Y(z, 0) = 0$ are fulfilled. In conclusion, the system of ordinary differential equations (4.6) and (4.5) exhibits $C_{gen} \cap U(a)$ as a center manifold that is parametrized by $(z, y = 0)$.

The linearly approximated equations of motion

$$\frac{d}{dt}y(t) = P_0DX_H^c(0,0)P_0y(t) \quad \frac{d}{dt}z(t) = 0, \tag{4.7}$$

generate orbits

$$y(t) = e^{tP_0}DX_H^c(0,0)P_0y(0) \quad z(t) = z(0).$$

Obviously, the matrix $P_0DX_H^c(0,0)P_0$ determines the entire dynamics of the linearized system (recall that it is identical to $P_0J_0A_0$, where $A_0$ has been defined in 2.2.2).
4.2.2 Asymptotic behaviour

The linearized equations of motion and their solution suggest that the existence of non-vanishing real parts of the eigenvalues of

\[ P_0 J A_0 := P_0 J DX_H^0(0,0) P_0 \]

should give rise to orbits that either get closer or move away from \( C_{\text{gen}} \) exponentially in time, and along the respective eigendirections.

An easy application of the center manifold theorem proves that this is indeed so. In fact, if the real parts of all \( 2k \) eigenvalues of the matrix \( P_0 J A_0 = P_0 J DX_H^0(0,0) P_0 \) are non-zero, the center manifold theorem implies that there exists a coordinate transformation \( (y, z) \to (\tilde{y}, \tilde{z}) \), so that (4.6) and (4.5) can be written as

\[
\begin{align*}
\frac{d}{dt} \tilde{y}(t) &= P_0 J A_0 \tilde{y}(t) + \tilde{Y}(\tilde{y}(t), \tilde{z}(t)) \\
\frac{d}{dt} \tilde{z}(t) &= 0
\end{align*}
\]

[42], where \( \tilde{Y}(0, \tilde{z}) = 0 \) for all \( \tilde{z} \). Thus, for every initial value \( \tilde{z}_0 \), one has a differential equation for \( \tilde{y} \) that is parametrized by \( \tilde{z}(t) = \tilde{z}_0 \); the equilibrium 0 (which is the point \( a \) in the given chart) is asymptotically unstable if there are eigenvalues with a positive real part, and asymptotically stable if all eigenvalues have a negative real part.

If the constraints are holonomic, asymptotic stability is impossible, because the eigenvalues always come in pairs or quadruples with both positive and negative real parts. However, in case of non-holonomic constraints, there is no obvious obstruction to the existence of asymptotically stable equilibria, since the flow map is not symplectic. It would certainly be interesting to construct an example of an asymptotically stable equilibrium in a non-holonomic constrained Hamiltonian system, but unfortunately, I do not have one at hand.

Let us make a small digression on how to study the case of 'weak non-holonomy', that is, the situation in which terms that originate from the non-integrability of the constraints can locally be considered as small perturbations of a holonomic system. To this end, we recall the definition of the matrices

\[ R^\pm_{ij} \equiv \frac{1}{2} H_{k} P^r_i P^s_j \left( P^k_{r,s} \pm P^k_{s,r} \right) \]

from (2.6) in 2.2.2, and the fact that

\[ A_0 = R^- + R^+ + P_0^T D^2_0 H P_0. \]  

\[ \text{(4.8)} \]
The antisymmetric part of \( A_0 \) is given by \( A_0^- = R^- \), and vanishes if \( V \) is holonomic, as one reads off from the Frobenius condition (1.5). Hence, if the distribution \( V \) is holonomic, only the symmetric part of \( A_0 \) is present, which we have denoted by \( A_0^+ \). The spectrum of \( JA_0^+ \) consists of pairs of eigenvalues with opposite signs and equal multiplicities, because \( JA_0^+ \) is the Hamiltonian vector field of the Hamiltonian \( \frac{1}{2}(x, A_0^+ x) \).

If \( V \) is non-integrable, the term \( JA_0^- \) does not vanish, and its influence on the spectrum of \( JA_0 \) can be studied by considering the one parameter family of matrices \( J(A_0^+ + \kappa A_0^-) \), for \( \kappa \in [0, 1] \). The eigenvalues of \( JA_0 \) can, for sufficiently small \( \kappa \), be written as power series in \( \kappa \). The corresponding coefficients can be iteratively determined by use of elementary methods in the perturbation theory of linear operators. The explicit formulae can be looked up in any textbook on quantum mechanics.

4.2.3 Critically stable equilibrium solutions

The case of critical stability is at hand if the spectrum of \( P_0J A_0 \) is purely imaginary. An exhaustive analysis of this situation would require methods of KAM and Nekhoroshev theory, and is beyond the scope of this thesis. We will content ourselves with an application of averaging theory, to point out the relevant questions, and some approximative results.

Let us assume that the matrix \( JA_0 \) is diagonalizable over the complex numbers. Carrying out the diagonalization requires one to find a well-defined continuation of the given dynamical system, defined in \( \mathbb{R}^{2n} \), into \( \mathbb{C}^{2n} \). To this end, we assume that the vector fields \( Y(y, z) \) and \( Z(y, z) \) in (4.6) and (4.5) are analytic with respect to the variables \( (y, z) \), so that they possess a unique analytical continuation into a complex vicinity of \( a \in \mathcal{C}_{\text{gen}} \). The complex dynamical system

\[
\frac{d}{dt} y(t) = P_0J A_0 y(t) + Y(y(t), z(t)) \quad \frac{d}{dt} z(t) = Z(y(t), z(t))
\]

is well-defined under these circumstances, where \( (y, z) \) is now understood as a vector in the vicinity of the origin of \( \mathbb{C}^{2k} \times \mathbb{C}^{2n-2k} \). The continuation of \( \mathcal{C}_{\text{gen}} \) into the complex domain is defined by the common zeros of \( Y(0, z) \) and \( Z(0, z) \) for complex \( z \).

By assumption, the spectrum of \( JA_0 \) consists of \( 2k \) non-zero, purely imaginary eigenvalues \( i\omega_r \in i\mathbb{R} \), where \( i \) denotes the imaginary unit. There exists a complex linear transformation

\[
\Psi : \mathbb{C}^{2n} \to \mathbb{C}^{2n}
\]

80
that diagonalizes $J A_0$. By a slight abuse of notation, we denote the coordinates in the eigenbasis again by $(y, z)$. The equations of motion then reduce to

$$\frac{dy(t)}{dt} = \text{diag}(i\omega)y(t) + Y(y(t), z(t)) \quad \frac{dz(t)}{dt} = Z(y(t), z(t)), \quad (4.9)$$

where $\omega$ denotes the vector of 'frequencies' $\omega_r$. This notation is unambiguous, because in this chapter, all considerations are strictly local, and the symplectic structure will always be represented by the matrix $J$. The solutions of the $\mathbb{R}$-valued physical system are obtained from application of the inverse transformation $\Psi^{-1}$ to those complex solutions $(y(t), z(t))$ whose initial conditions $(y(0), z(0))$ are contained in the linear subspace $\Psi(\mathbb{R}^{2n}\subset \mathbb{C}^{2n}) \subset \mathbb{C}^{2n}$. In other words, the physical system is conjugated by the linear automorphism $\Psi$ of $\mathbb{C}^{2n}$, and the equations of motion are solved in its image.

Next, we introduce polar coordinates $(I, \phi)$ and $(J, \theta)$ in terms of

$$y_r = e^{i\phi_r} I^r \quad z_s = e^{i\theta_s} J^s$$

with $r = 1, \ldots, 2k$ and $s = 1, \ldots, 2n - 2k$. In particular, we have $I \in \mathbb{R}^{2k}$, $J \in \mathbb{R}^{2n-2k}$, $\phi \in [0, 2\pi]^{2k} = T^{2k}$ (the 2$k$-dimensional torus), and $\theta \in [0, 2\pi]^{2n-2k} = T^{2n-2k}$. For brevity, vectors $(e^{i\phi_r} v^r)$ and $(e^{i\theta_s} w^s)$ will subsequently be denoted by $e^{i\phi} v$ and $e^{i\theta} w$, respectively, where $v \in \mathbb{R}^{2k}$ and $w \in \mathbb{R}^{2n-2k}$.

We insert this expression into the complexified equations of motion for the components $\dot{y}$ (the dot abbreviates $\frac{d}{dt}$), and obtain

$$e^{i\phi} \dot{I} + i\dot{\phi} e^{i\phi} I = \text{diag}(i\omega) e^{i\phi} I + Y(e^{i\phi} I, e^{i\theta} J),$$

so that one easily finds

$$\dot{I} = \text{Re}\{e^{-i\phi} Y(e^{i\phi} I, e^{i\theta} J)\}$$

$$\dot{\phi} = \omega + \text{Im}\{e^{-i\phi} \text{diag}(\partial_I) Y(e^{i\phi} I, e^{i\theta} J)\}. \quad (4.10)$$

In the same manner, one arrives at

$$\dot{J} = \text{Re}\{e^{-i\theta} Z(e^{i\phi} I, e^{i\theta} J)\}$$

$$\dot{\theta} = \text{Im}\{e^{-i\theta} \text{diag}(\partial_J) Z(e^{i\phi} I, e^{i\theta} J)\}. \quad (4.11)$$

The set of ordinary differential equations (4.10) and (4.11) determines the dynamics of the complexified system, where $(I, J) \in \mathbb{R}^{2n}$ lies in a small vicinity of the origin.
Next, we fix a small parameter \( \epsilon := \|J(0)\| \), where the norm is induced by \( g_0 := g|_0 \), by the assumption that \( \|J(0)\| \leq O(\epsilon^2) \) shall be satisfied. Then, we apply the rescaling transformation \( I \rightarrow \epsilon I \) and \( J \rightarrow \epsilon^2 J \), so that the new coordinates \((I, J)\) have a norm of the order \( O(1) \). Because \( Y(y, z) \) and \( Z(y, z) \) are analytical functions of \((y, z)\), they can be written in terms of convergent power series with respect to the variables \( e^{i\phi} I \) and \( e^{i\theta} J \) for all \((I, J) \in \mathbb{R}^{2n}\) in a sufficiently small vicinity of the origin. Carrying out these expansions on the right hand sides of the four equations in (4.10) and (4.11), one arrives at the system of ordinary differential equations

\[
\begin{align*}
\dot{i} & = \sum_{|m|+|p|\geq 2} \epsilon^{|m|+2|p|-1} F_{mp}(I, J) e^{i((m,\phi)-\phi')} e^{i(p,\theta)} \quad (4.12) \\
\dot{j} & = \sum_{|m|+|p|\geq 2} \epsilon^{|m|+2|p|-2} G_{mp}(I, J) e^{i((m,\phi)-\phi')} e^{i(p,\theta)} \quad (4.13) \\
\dot{\phi} & = \omega + \sum_{|m|+|p|\geq 2} \epsilon^{|m|+2|p|-1} \Phi_{r,mp}(I, J) e^{i((m,\phi)-\phi')} e^{i(p,\theta)} \quad (4.14) \\
\dot{\theta} & = \sum_{|m|+|p|\geq 2} \epsilon^{|m|+2|p|-2} \Theta_{s,mp}(I, J) e^{i((m,\phi)-\phi')} e^{i(p,\theta)}, \quad (4.15)
\end{align*}
\]


where \( m \in \mathbb{Z}^{2k} \) and \( p \in \mathbb{Z}^{2n-2k} \) are multiindices, and \(|m| := \sum |m_r| \) and \(|p| := \sum |p_s|\). Evidently, the right hand sides here are Fourier expansions relative to the \(2\pi\)-periodic variables \( \phi \) and \( \theta \), and every Fourier coefficient labelled by a pair of indices \((m, p)\) is a homogenous polynomial of degree \(|m|\) in \( I \), and of degree \(|p|\) in \( J \).

In the limit \( \epsilon \to 0 \), one arrives at the associated 'unperturbed' system of ordinary differential equations

\[
\dot{i} = 0 \quad \dot{j} = 0 \quad \dot{\phi} = \omega \quad \dot{\theta} = 0. \quad (4.16)
\]

Under the assumption that all components of \( \omega \) are rationally independent, averaging can be applied with respect to the variable \( \phi \) in order to obtain an approximation to the long time behaviour of the perturbed system.

Put very briefly, the method of averaging is based on the following idea (for an introduction, consider for instance [3]). Consider a system of ordinary differential equations of the form

\[
\dot{I} = \epsilon f(I, \phi) \quad \quad \dot{\phi} = \omega + \epsilon g(I, \phi),
\]

where \( I \in \mathbb{R}^m \) and \( \phi \in \mathbb{T}^n \) (which are here unrelated to the variables above), and where the functions \( f \) and \( g \) are periodic with respect to the angular variable \( \phi \). The parameter \( \epsilon \) is
assumed to be small; \( I \) is a 'slow', and \( \phi \) is a 'fast' variable. For \( \epsilon = 0 \), it is clear that if the components of the vector \( \omega \) are rationally independent, \( \phi(R) \) is dense in the torus \( \mathbb{T}^n \).

Averaging is an expression for the transition to a new dynamical system that is given by

\[
\dot{\bar{I}} = \epsilon \frac{1}{(2\pi)^n} \int_{\mathbb{T}^n} d^r \phi f(\bar{I}, \phi),
\]

where \( \bar{I} \) denotes the averaged coordinate that replaces \( I \). If the components of \( \omega \) satisfy certain incommensurability conditions, it can be proved that \( \|I(t) - \bar{I}(t)\| \leq O(\epsilon) \) for \( t \in [0, \frac{1}{\epsilon}] \), or even longer times, so that the long time behaviour of \( I(t) \) is very well approximated by the averaged solutions.

The only quantities in (4.12) \~ (4.15) that survive the averaging operation

\[
f(\phi) \rightarrow \frac{1}{(2\pi)^{2k}} \int_{\mathbb{T}^{2k}} d^{2k} \phi f(\phi)
\]

with respect to \( \phi \) are those for which the exponents in the \( \phi \)-dependent exponential functions vanish.

Now we recall that in the initial equations of motion (4.9), the functions \( Y(y, z) \) and \( Z(y, z) \) vanish for \( y = 0 \). The leading terms of their power series in \( (y, z) \) are at least homogenous of degree 1 in \( y \), and thus involve terms \( e^{i(m, \phi)} \) with \( |m| \geq 1 \), but no terms with \( |m| = 0 \). A quick inspection of the right hand sides of (4.10) and (4.11) assures us of the fact that (4.12) and (4.14) yield terms that survive the averaging process, but not (4.13) and (4.15). Therefore, averaging the perturbed equations of motion (4.12) \~ (4.15) with respect to \( \phi \) gives

\[
\dot{\bar{I}} = \epsilon^2 \bar{F}(\bar{I}, \bar{J}, \bar{\theta}) \quad \text{where the bars account for averaged variables.}
\]

Returning to the initial coordinate chart for real \( (y, z) \), let us now assume that the quadratic form \( \langle y, D^2_H y \rangle \) is positive definite. We claim that the averaging result then suggests that \( a \) is stable. To this end, notice that Taylor expansion of the Hamiltonian (which is an integral of motion for the constrained system) relative to the base point 0 gives

\[
H(x) = H(0) + \langle \partial_y |_0 H, u \rangle + \frac{1}{2} \langle y, D^2_H y \rangle + O(\epsilon^3),
\]

and recall that \( u = \bar{P}_0 x \). \( C_{\text{gen}} \) can locally be described as the graph \( (u, F(u)) \) of some mapping \( F : \mathbb{R}^{2n-2k} \rightarrow \mathbb{R}^{2n} \) that satisfies \( F(0) = 0 \), so that the assumption \( \|z\| \leq O(\epsilon^2) \) implies that
\[ \|u\| \leq O(\epsilon^2). \] The quantity \( H(x) - H(0) = O(\epsilon^2) \) is an integral of motion; thus, if the term \( \langle y, D^2_H y \rangle \) is a positive definite quadratic form, \( \|y\| \) has an order of magnitude \( O(\epsilon) \) as long as \( \|u\| = O(\epsilon^2) \) remains valid. The averaged equations of motion imply that \( \|\tilde{z}\| = \|\tilde{J}\| = O(\epsilon^2) \) is time-independent, so that due to \( \tilde{u} = \tilde{P}_0 \tilde{z} \), \( \|\tilde{u}\| \) is also time-independent.

Hence, averaging, which has eliminated the fast oscillations parametrized by \( \phi(t) \), suggests that the equilibrium at the origin of the coordinate system is stable. The assumptions that have been made here are analyticity of the vector fields \( Y(y, z) \) and \( Z(y, z) \), rational independence of the entries of \( \omega \), and positive definiteness of the quadratic form \( \langle y, D^2_H y \rangle \).

I suspect that it should be possible, by use of methods from KAM and Nekhoroshev theory, to prove a theorem that verifies the averaging result; however, given the limited scope of this thesis, I content myself here with these approximation arguments.

4.2.4 Resonances and instabilities: a special case

Let us briefly illustrate a typical phenomenon that may occur in case the components of \( \omega \) are not rationally independent. To this end, let us consider the specific situation where at least one eigenvalue of \( P_0 \mathcal{J} A_0 \) exhibits a multiplicity larger than one.

The time evolution of linearized orbits \( y(t) \) in the the present chart is determined by

\[ y(t) = \exp(tP_0 \mathcal{J} A_0) y(0). \]

There exist real vectors \( X_r \) associated to the eigenvalues \( i\omega_r \) of \( P_0 \mathcal{J} A_0 \), by use of which \( y(t) \) can be expanded into the harmonic series

\[ y(t) = \sum_k (a_k X_k \sin \omega_k t + b_k X_k \cos \omega_k t), \]

where the real constants \( a_k \) and \( b_k \) are determined from the initial condition. Next, let us consider the approximative time evolution of the variable \( z(t) \) to lowest non-trivial perturbation order, and recall that \( z(t) \) is bounded if and only if the variable \( u(t) = \tilde{P}_0 z(t) \) is bounded. To study the time evolution of \( u(t) \), it is convenient to consider the constraint conditions \( P(x) \dot{x} = 0 \). Taylor expansion of \( P(x) \) relative to the origin of the coordinate system yields

\[ P_0 \dot{x} + D_0 \tilde{P}(x) \dot{x} + O(\|x\|^2) = 0, \]

where in components, \([D_0 \tilde{P}(x) \dot{x}]^i = \tilde{P}_0^j i_0 x^j x^k \).
Since $P$ and $\bar{P}$ add up to the $n$-dimensional unit matrix, it is clear that $P_{i,k}^j = -P_{i,k}^j$ must hold, so that the constraint relations are, to second order perturbation approximation, given by

$$\dot{\hat{u}}^i = P_{i,k}^j |_{Q} y^j \hat{y}^k.$$  

Here, we have substituted the first order approximation result for $y(t)$. Consequently, one has, to second order approximation,

$$\dot{\hat{u}}^i = \Delta_{r,s}^+ y^r \hat{y}^s + \Delta_{r,s}^- y^r \hat{y}^s,$$

where $\Delta_{r,s}^\pm := \frac{1}{2} P^i_r P^k_s (P^l_{r,s} \pm P^l_{s,r})$. Evidently, the matrices $\Delta^+_{r,s}$ are symmetric, and the matrices $\Delta^-_{r,s}$ are antisymmetric with respect to the lower indices.

Integration of (4.18) with respect to the time variable yields

$$z^j(t) = \int_0^t d\tau \left( \Delta_{r,s}^+ y^r(\tau) \dot{y}^s(\tau) + \Delta_{r,s}^- y^r(\tau) \dot{y}^s(\tau) \right).$$

The first term in the bracket can be integrated by quadrature, hence it is bounded, because by assumption, $y(t)$ is bounded. If instabilities arise, they can only occur due to the second term in the bracket. Evidently, it involves the matrices $\Delta^-_{r,s}$, whose vanishing implies the integrability of $V$ according to the Frobenius condition.

This is a consequence of the fact that $z(t)$ can diverge for oscillatory, bounded $y(t)$ if and only if the distribution is non-integrable (if $V$ is integrable, the integral manifolds can locally be parametrized by $z = z(y)$, so that $z(t) = z(y(t))$ is bounded if $y(t)$ is bounded). So it remains to be clarified in which cases of non-integrable $V$ divergences are indeed expected to arise.

A typical contribution to the second term on the right hand side of (4.19) is for example the expression

$$a_i a_j \int_0^t d\tau \Delta_{r,s}^+ X^r_i X^s_j \omega_i \sin \omega_1 \tau \cos \omega_j \tau;$$

it is obvious that in the full expression, all possible combinations of sine and cosine functions appear. Because trigonometric integrals of the form $\int_0^t d\tau \sin^2 \tau$ and $\int_0^t d\tau \cos^2 \tau$ produce an unbounded term $\frac{t}{2}$, a divergent term will in general arise if there are eigenfrequencies $\omega_k = \omega_j$ with $k \neq j$, or in other words, if there are eigenvalues of $\mathcal{J} A_0$ with a multiplicity higher than one.

If there exists an eigenvalue of multiplicity $r$, so that the associated terms in the harmonic series of $y(t)$ are spanned by the real vectors $X_{i_1}, \ldots, X_{i_r}$, then $u(t)$ will generically
diverge in the directions of the Lie bracket commutators $[P(x)X_{ik}, P(x)X_{il}]_{x=0,...}$, which span a subspace in the first flag element $V_1$ of $V$ at the origin, cf. subsection 1.3.1. The time to exit an $\epsilon$-neighborhood of 0 is in this case of the order $O(\frac{1}{\epsilon})$.

This simple illustration describes a mechanism that causes divergences linear in time if eigenfrequencies of nonminimal multiplicity are present. The cases of general rational dependencies between the frequencies of the unperturbed system can presumably be analyzed by use of methods stemming from KAM and Nekhoroshev theory.

### 4.3 The stability of equilibria in constrained Hamiltonian mechanical systems

We will now discuss the stability of equilibria of constrained Hamiltonian mechanical systems of the type introduced in chapter three.

For this purpose, we use the continuation away from the physical sheet that is defined by (3.16), which has the advantage that there is no dynamics transverse to the physical sheet $\mathcal{M}_{phys}$ in the vicinity of any physical equilibrium. The stability analysis of general constrained Hamiltonian systems given above applies without restriction to the constrained Hamiltonian mechanical system (3.16).

We have shown in chapter three that the matrix of the $\omega$-skew selfadjoint bundle projector of this system is given by

$$P(q,p) = \begin{pmatrix} \alpha(q) & 0 \\ -T(q,p) & \alpha^T(q) \end{pmatrix}. $$

Furthermore, we recall that the Hamiltonian is given by

$$H = \frac{1}{2} h^*(p,p) + u(q), $$

where the matrix of $h^*$ is $M^{-1}(q)$. Let $a := (c, 0) \in T^*Q$ denote an equilibrium of the constrained system on the physical sheet $\mathcal{M}_{phys}$ with $c \in C_Q$, so that $P^T H_{(q,p)}|(c,0) = 0$. The stability criteria which have been derived above require one to compute the matrix $A_a$, as defined in (2.6) and (4.8).

To this end, let us to begin with determine the matrix of second derivatives $D^2_aH$, which is, at the point $a = (c, 0)$, given by

$$D^2_aH = \begin{pmatrix} D^2_cu & 0 \\ 0 & M^{-1} \end{pmatrix}. $$

86
Furthermore, we get

\[ P_a = \begin{pmatrix} \alpha_c & 0 \\ 0 & \alpha^T_c \end{pmatrix}, \]

and we must also determine the matrix \( R_{(c,0)} = R^+ + R^- \) from (2.6). One finds

\[ R_{(c,0)} = \begin{pmatrix} \tilde{R}_c & 0 \\ 0 & \tilde{R}^T_c \end{pmatrix}, \]

where \( \tilde{R}_c \) is the \( n \times n \) matrix given by \( \tilde{R}_{c,ij} = u_{ij}^s \alpha^j_k \alpha^i_l l_{i,c} \).

Straightforward, but slightly tedious regrouping of the components of \( R_{(a,0)} \) with respect to indices in \( \{1, \ldots, n\} \) and \( \{n+1, \ldots, 2n\} \) according to (2.6) makes it clear that the term \( T(q,p) \) in \( P(q,p) \), together with its partial derivatives, does not contribute because it vanishes at \( (c,0) \). Finally, inserting all of these pieces into the formula \( A_a = R + P_a^T D^2_H P_a \), we get

\[ P_a J A_a = \begin{pmatrix} 0 & \alpha_c M^{-1}_c \alpha^T_c + \tilde{R}_c^T \\ -(\alpha^T_c D^2_H u \alpha_c + \tilde{R}_c) & 0 \end{pmatrix}. \]

Next, we have to determine those eigenvalues of this matrix that belong to eigenvectors tangent to \( M_{phys} \); the eigenvalues associated to eigendirections normal to \( M_{phys} \) account for the stability of the invariant manifold \( M_{phys} \), but have no relevance for the physical system. The projection onto \( T_a M_{phys} \) is given by \( \text{diag}(1_n, \alpha^T_c) \) (recall that \( p = 0 \) in \( a = (c, 0) \)), while the projection normal to \( M_{phys} \) is given by \( \text{diag}(0_n, \beta^T_c) \). It is clear that \( P_a J A_a \) has no non-zero eigenvalues normal to \( M_{phys} \), because the auxiliary dynamical system (3.16) admits a complete foliation of \( T^*Q \) into invariant submanifolds; this has been explained in chapter three.

The spectrum of \( P_a J A_a \) determines the stability of the equilibrium \( c \in Q \) of the physical system. The results derived above for general constrained Hamiltonian systems imply that if there exist eigenvalues with a positive real part, \( c \) is asymptotically unstable; if all eigenvalues have a negative real part, \( c \) is asymptotically stable (but I do not know any example of an asymptotically stable equilibrium in a Hamiltonian mechanical system with non-holonomic constraints, although there are no obvious obstructions to its existence).

If all eigenvalues are purely imaginary and rationally independent, averaging theory suggests that \( c \) is stable if the spectrum of the matrix \( P_a^T D^2_H P_a \) relative to the Euclidean metric is strictly positive (but we haven’t rigorously proved this last statement). Notice that the spectrum of \( P_a^T D^2_H P_a \) is strictly positive if and only if the spectrum of \( \alpha^T_c D^2_H u \alpha_c \) is strictly positive (this is because \( \alpha_c M^{-1}_c \alpha^T_c \) has a strictly positive spectrum).
4.4 A note on the shape of non-holonomic e-balls

The emergence of a 'fast' direction tangent to the distribution $V$ and a 'slow' direction transverse to it is typical for non-holonomic systems. This phenomenon is closely related to the 'shape' of 'non-holonomic e-balls'. Let us briefly address some basic facts about this interesting class of problems, which belongs to the field in mathematics coined sub-Riemannian geometry.

Let $M$ be a Riemannian manifold with a metric tensor $g$ (for instance, the Kähler metric constructed in the first chapter associated to a symplectic manifold $M$). The Riemannian metric on $M$ is defined by the distance function

$$d_R(x, y) = \inf_{\gamma} \left\{ l_g(\gamma) \mid \gamma : [0, 1] \to M; \gamma(0) = x, \gamma(1) = y \right\},$$

where $x$ and $y$ are two arbitrary points on $M$, and $l_g(\gamma)$ is the length of the trajectory $\gamma$, measured with the metric tensor $g$ in terms of

$$l_g(\gamma) = \int_{[0,1]} (g(\dot{\gamma}, \dot{\gamma}))^{1/2} dt.$$

At the presence of non-holonomic Pfaffian constraints, another notion of distance can in addition be introduced, which is referred to as the Carnot-Caratheodory (CC) metric. Picking any pair of points $x$ and $y$ that are connectible via admissible paths, the CC distance is defined by

$$d_{CC}(x, y) = \inf_{\gamma} \left\{ l_g(\gamma) \mid \gamma : [0, 1] \to M; \gamma(0) = x, \gamma(1) = y; \dot{\gamma} \in V_{\gamma} \right\}.$$

If a pair of points is not connectible by any admissible path, the convention is that their CC distance is infinite. A beautiful treatment of the connections between the CC and Riemannian structures, and of numerous other fascinating issues is given in [22].

If $M$ is connected and Riemannian, one infers from the Hopf-Rinow theorem that any two points can be connected by geodesics of finite length in the $d_R$-metric. The generalization of the Hopf-Rinow theorem for case of non-holonomic systems is the Rashevsky-Chow theorem, which states that if $V$ is strongly bracket generating, any two points on $M$ have a finite CC distance, that is, any two points can be connected by an admissible curve of finite Riemannian length.
On a Riemannian manifold, an open Riemannian $\epsilon$-ball around a point $x$ is defined as the subset of $M$ that consists of points $y$ with $d_R(x, y) < \epsilon$. The analogous definition in terms of CC metrics is as follows. An open non-holonomic, or CC $\epsilon$-ball is defined as the set

$$B_{CC}(x, \epsilon) = \{ y \in M \mid d_{CC}(x, y) < \epsilon \}.$$  

The shape of non-holonomic $\epsilon$-balls is discussed in much detail in [22]. Most essentially, one notes that $B_{CC}(x, \epsilon)$ is contained in a '$\epsilon \times \epsilon^2$-neighborhood' of $x$ of the following kind. Let $N$ be a $2k$ dimensional submanifold of $M$ that is tangent to the distribution $V$ at $x$, and pick the Riemannian $\epsilon$-ball $B^{(N)}_\epsilon(x)$ around $x$ in $N$. Then, the $\epsilon \times \epsilon^2$-neighborhood of interest is the tubular $\epsilon^2$-neighborhood of $B^{(N)}_\epsilon(x)$ in $M$.

Along the line of similar arguments given in [22], one understands this by considering any admissible trajectory $c : [0, \epsilon] \to M$ with $c(0) = x$ and $\| \dot{c}(0) \| = 1$; here the norm is taken with respect to the Riemannian structure. Then, $c$ is tangent to $N$ at $x$, and since the length of $c([0, \epsilon])$, if measured with the Riemannian metric, is approximately $\approx \epsilon$, the distance between $c(\epsilon)$ and $N$ has an order of magnitude $O(\epsilon^2)$, which follows from the Taylor rest term theorem. Therefore, there is a number $C \approx 1$ so that the CC $\epsilon$-ball is contained in the $\epsilon^2$-neighborhood of $B^{(N)}_\epsilon(x)$ in $M$.

These are the only remarks we place here to point out relations of the problems at hand to natural questions of sub-Riemannian geometry. This field between pure and applied mathematics is a rapidly developing area with fascinating applications in geometrical control theory. Apart from [5, 38], it is recommendable to consult the classical works of R. Brockett, H. Sussman, and R. Montgomery, and the authors cited by them.
CHAPTER 5

APPLICATIONS, ILLUSTRATIONS AND EXAMPLES

We will now illustrate the results derived in the previous chapters with examples, and discuss several possible practical applications. The presentation will be given in a significantly more intuitive style, and will involve very few technicalities. We believe that although the emphasis of this thesis has not been to analyze any particular technical system, engineers might find our results very useful for comprehending particular aspects about the qualitative structure of constrained Hamiltonian systems.

5.1 Constrained Hamiltonian multibody systems

Let us briefly recall the setting of a constrained Hamiltonian multibody system in the standard language of the multibody systems literature.

The configuration space is a smooth $n$-dimensional manifold $Q$, which is locally represented by coordinates $q = (q^1, \ldots, q^n)$ in $\mathbb{R}^n$ that one refers to as generalized coordinates. If $Q$ is topologically non-trivial, one will generically not be able to describe an entire orbit on $Q$ in a single coordinate system, but one will have to employ a number of coordinate changes.

In this thesis, we have considered Hamiltonians of the standard form $H = T + u$, where

$$T = \frac{1}{2} p^T M^{-1}(q) p$$

is the kinetic energy, and $u(q)$ is the potential energy. The quantities $p = (p_1, \ldots, p_n)$ denote the generalized canonical momenta, and $M(q)$ is the mass matrix of the system (which will henceforth be referred to as the free system).

We assume that Pfaffian constraints are imposed on this mechanical system by requiring that the generalized velocities obey the linear kinematical condition

$$E^T(q) \dot{q} = 0,$$
where $E^T$ denotes a $n-k$ by $n$ matrix of full rank. We also introduce the "mass-orthogonal" projectors $\alpha(q)$ and $\beta(q)$ associated to the admissible and forbidden directions, defined by the kernel and cokernel of $E^T(q)$, respectively.

Mass-orthogonality is expressed by $\alpha^T M = M \alpha$, respectively $\beta^T M = M \beta$. The explicit formula for $\beta(q)$ is given by

$$\beta(q) = M^{-1}(q)E(q) \left( E^T(q)M^{-1}(q)E(q) \right)^{-1} E^T(q),$$

and $\alpha(q)$ is obtained from $\alpha(q) = 1_n - \beta(q)$.

Pfaffian constraints are called holonomic, or integrable, if $Q$ admits a complete foliation into $k$-dimensional integral manifolds, each of which is everywhere tangent to the admissible directions defined by $\alpha(q)$ (respectively $E^T(q)$). The question whether some given constraints are integrable or not is answered by the theorem of Frobenius, which, on the level of orthoprojectors, states that the constraints are holonomic if and only if

$$\alpha_i^r \alpha_j^s (\alpha_{r,s}^{k} - \alpha_{s,r}^k) = 0$$

is satisfied.

5.1.1 The continuation away from the physical sheet

In chapter three, we have discussed in much detail how the equations of motion of the constrained system are obtained from the variational principle of Hölder.

We have proved for Hamiltonian mechanical systems with Pfaffian velocity constraints, that all physical orbits are confined to the submanifold

$$\mathcal{M}_{phys} := \{(q,p) \mid \beta^T(q)p = 0\}$$

of $T^*Q$, which we have called the physical sheet. In order to study the physical dynamics on $\mathcal{M}_{phys}$, and for the purpose of numerically simulating physical orbits, it is very convenient to consider the auxiliary dynamical system (3.16), which defines a flow on the whole $T^*Q$. It possesses the particular structure of a general constrained Hamiltonian system of the type introduced and analyzed in the first two chapters.

The physical sheet $\mathcal{M}_{phys}$ is a stable (but not asymptotically stable) invariant manifold under the associated flow, which precisely produces the physical flow on $\mathcal{M}_{phys}$. We have also demonstrated that symmetries of the physical system on $\mathcal{M}_{phys}$ extend to symmetries
of (3.16). Furthermore, if \((q_0, p_0)\) is an equilibrium of (3.16), then \(q_0\) is automatically an equilibrium of the physical system.

It has been noted in numerous publications on the numerical mathematics of constrained multibody systems that this or very similar auxiliary dynamical systems (which are, however, usually written in terms of DAE's, not as ODE's) are suitable for numerical simulation routines [20, 32, 36]. Issues of numerical mathematics have not been considered in this thesis; we must refer the reader to the existing, vast body of excellent literature on the topic, to which we have little to add (geometrical transparency does not imply numerical efficiency; very different classes of questions and problems are being addressed in the development of state-of-the-art numerical integrators than those presented in this thesis).

The equilibria of the physical constrained system are points on the critical set

\[ C_Q := \{ q \in Q \mid \alpha^T(q)\nu(q) = 0 \}, \]

which is generically a smooth, \( n - k \)-dimensional submanifold of \( Q \). If the constraints are holonomic, each integral manifold in \( Q \) intersects the critical manifold \( C_Q \), if it is generic, in a finite set of isolated points. In other words, equilibria of holonomic systems, whose orbits are necessarily confined to a \( k \) dimensional invariant submanifold of \( Q \), are generically isolated (on this integral manifold).

5.2 Geometry and topology of the critical manifold

One of the central topics in this thesis has been to investigate geometrical and topological properties of \( C_Q \). We have already said that generically, \( C_Q \) is a smooth submanifold of \( Q \), whose dimension equals the rank of \( E^T(q) \), which is given by \( n - k \). The notion of genericity has the meaning that even in case the situation, which is described as being generic, is not given, an arbitrarily small deformation of \( E^T(q) \) or other system parameters suffices to produce it. We have referred to non-generic situations as being exceptional.

Practically, this implies that given any out of the blue constrained multibody system, the likelihood that \( C_Q \) is a smooth, \( n - k \)-dimensional submanifold of \( Q \) is overwhelmingly larger than the likelihood that it is exceptional. In case \( C_Q \) is exceptional, it will in general not even be a manifold. The proof of the genericity statements is based on the application of Sard's theorem at the beginning of chapter two. Technically, the relevance of this result can be expressed as a construction rule.
In fact, any set $\mathcal{P}$ of design parameters or constraints that produces an exceptional critical set $\mathcal{C}_Q$ in a constrained Hamiltonian multibody system should be avoided. Exceptional equilibria are not stable under small perturbations of $\mathcal{P}$. Therefore, any mechanism built according to specifications that produce exceptional critical sets is very likely to exhibit equilibria that are substantially different from the intended ones, due to unavoidable manufacturing tolerances.

This suggests that one should aim at finding constraints and parameters, for which $\mathcal{C}_Q$ is a smooth, $n - k$-dimensional submanifold of $Q$. The discussion in chapters two and four show that if $\mathcal{C}_Q$ is generic, it contains all critical points of the Morse function $u$, but no other extrema of the function $u|_{\mathcal{C}_Q}$.

In addition, $\mathcal{C}_Q$ is then necessarily normal hyperbolic with respect to the gradient-like flows $\psi^\pm_t$ that is generated by the auxiliary dynamical system

$$\dot{q} = \pm (M^{-1} \alpha^T u_q)(q)$$

on $Q$. Normal hyperbolicity describes the fact that all stable and unstable manifolds of $\mathcal{C}_Q$ intersect $\mathcal{C}_Q$ transversely. These flows are gradient-like, which means that $u$ increases, respectively decreases strictly monotonically along all non-constant orbits, depending on the sign $+$ or $-$. The dimension of the unstable manifold of a connectivity component of $\mathcal{C}_Q$ is called its index.

Every orbit of the gradient-like flow $\psi^\pm_t$ starts and ends on a pair of non-identical connectivity components of $\mathcal{C}_Q$, and is a separatrix solution (or 'instanton'). That is, every orbit needs a very long time to leave a small vicinity of $\mathcal{C}_Q$, and to tend to $\mathcal{C}_Q$ again. However, it moves very fast between two connectivity components of $\mathcal{C}_Q$, and thus the notion of an 'instanton': Its 'lifetime' is restricted to a brief instant somewhere between past and future infinity.

An interesting question to ask is in which way it might be possible to use this property to numerically determine $\mathcal{C}_Q$. Simple forward integration does not suffice because $\mathcal{C}_Q$ is a set of measure zero in $Q$, whose connectivity components will be missed by most instantons if their index is $> 0$. Almost all instantons will end up at connectivity components with index zero.
5.2.1 Numerically determining the critical manifold

Knowledge about the location of equilibria is very crucial for the design of a constrained multibody system. It is particularly desirable to know whether a chosen set of parameters and constraints has the effect that the critical sets are generic or not.

Real technical systems are usually very large, so that their equilibria must usually be constructed by means of numerical methods. The method of analysis that we have employed in chapters two and four to investigate global topological properties of generic critical manifolds inspires the following way to proceed.

We claim that if \( u \) is a Morse function, whose critical points are all known, and if \( Q \) is compact and closed, it is possible to numerically construct all generic connectivity components of \( C_Q \).

This is because generic components of \( C_Q \) are smooth, \( n - k \)-dimensional submanifolds of \( Q \) containing all critical points of \( u \), and no other critical points of \( u|_{C_Q} \). This information can be exploited to find sufficiently many points on \( C_Q \), so that suitable spline interpolation permits the approximate reconstruction of an entire connectivity component. To this end, one picks a vicinity of a critical point \( a \) of \( u \), and uses a fixed point solver to determine neighboring zeros of \( |a^Tu_q(q)|^2 \), which are, of course, elements of \( C_Q \).

Given a critical point \( a \) in a region which is of technical interest, and picking initial conditions for the fixed point solver in a small vicinity of \( a \) that lie in the kernel of \( \text{Jac}_a(au_q) \), one gets new points on the same connectivity component of \( C_Q \). Iterating this procedure, pieces of \( C_Q \) of arbitrary size can be determined.

If all critical points of \( u \) are a priori known, one can proceed like this to construct all connectivity components of \( C_Q \) that contain critical points of \( u \). In this case, one is guaranteed to have found all of the generic components of \( C \) if the numerically determined connectivity components are closed, compact, and contain all critical points of \( u \).

We remark that the determination of the critical points of a Morse function \( u : Q \rightarrow \mathbb{R} \) is a difficult numerical task about which we have little to say here. We can only say that if one attempts at numerically finding the critical points of \( u \), one should recall that their total number is bounded from below by the sum of the Betti numbers of \( Q \), as a result from the Morse inequalities. This is a lower bound for the minimal number of critical points of \( u \) on \( Q \).

Another remark is that there might exist critical points \( a \) at which \( \text{Jac}_a(au_q) \) has a reduced rank. Then of course, \( C_Q \) is not generic. If there are such exceptional critical points in a technically relevant region of \( Q \), the system parameters and the constraints should be locally modified in order to eliminate them.
5.2.2 The weak Conley-Zehnder inequalities

If both $Q$ and $C_Q$ are smooth, compact, orientable and boundary free manifolds, there are very precise lower bounds on the ranks of the homology groups of $C_Q$ enforced by those of $Q$. The full discussion of these topological relations is in chapters two and four, and can be looked up there.

Because we have made a substantial effort in discussing that topic in those parts of this thesis, we will illustratively explain some of the basic ideas for the two-dimensional case. All smooth, compact, orientable, and boundary-free 2-dimensional manifolds are surfaces with an integer number of holes, which is referred to as their genus. For instance, the two-sphere $S^2$ has genus zero, the two-torus $T^2$ has genus one, the outer surface of two tori glued together via a connecting cylinder has genus two, and so on. In fact, it can be proved that the genus classifies all compact, closed, orientable two dimensional surfaces.

Figure 5.1: Some orientable 2-dimensional surfaces and their genera.

Let us now consider the possible closed, compact submanifold of a surface $\Sigma_g$. The zero dimensional submanifolds are points, the one dimensional submanifolds are embedded circles $S^1$, and the two dimensional embedded submanifolds are the manifolds themselves. If two points can be joined by a connected line, for which they define the boundary, they are said to be homologous. It is a general fact that homology is an equivalence relation. The number of inequivalent points in $\Sigma_g$ under homology is called the zeroth Betti number of $\Sigma_g$. It is clear that it simply denotes the number of connectivity components of $\Sigma_g$.

If a pair of non-intersecting embedded circles in $\Sigma_g$ can be joined by a compact two dimensional embedded surface in $\Sigma_g$, for which they define the boundary, they are likewise
homologous. The number of equivalence classes of embedded, non-contractible, homologous circles is referred to as the first Betti number of $\Sigma_g$. For $g = 0$, every circle on the two sphere $S^2$ can be continuously deformed into any other embedded circle. The 'swept out set' between a pair of non-intersecting embedded circles in $S^2$ is a compact surface bounded by them; therefore, all embedded circles in $S^2$ are homologous to each other, and of course contractible. Thus, the first Betti number of $S^2$ is zero. In case of $g = 1$, the first Betti number of the two torus $T^2$ is two. This is because all non-intersecting embedded circles around the 'small radius' are homologous to each other, and all non-intersecting embedded circles around the 'big radius' are homologous to each other. However, every circle around the small radius intersects each circle around the big radius; hence, they cannot be homologous. So one concludes that the number of inequivalent embedded circles in $T^2$ under homology is two, as claimed. Generally, one can prove that the first Betti number of $\Sigma_g$ is given by $2g$.

Finally, one considers pairs of embedded closed, compact surfaces in $\Sigma_g$ that can be joined by an embedded compact three dimensional submanifold. Since $\Sigma_g$ contains no three dimensional submanifolds, the only equivalence class of surfaces under homology is represented by $\Sigma_g$ itself, if it is connected. Otherwise, the number of inequivalent embedded surfaces under homology, which is the second Betti number, is given by the number of connectivity components of $\Sigma_g$, and so equals the zeroth Betti number. For every $n$-dimensional smooth, compact, closed manifold, the $n-k$-th Betti number equals the $k$-th Betti number; this relationship is called Poincaré duality.

We are now prepared to formulate one of the main results of chapter two. Let $\mu_i$ denote the dimension of the unstable manifold of the $i$-th connectivity component of $C_Q$ (which, as we recall, is called its index), and let $B_k(C_i)$ be its $k$-th Betti number. Furthermore, let $B_p(Q)$ denote the $p$-th Betti number of $Q$. Then, the equation (4.1) in chapter four implies the 'weak Conley-Zehnder inequalities'

$$\sum_i B_{p-\mu_i}(C_{Qi}) \geq B_p(Q)$$

for $p = 0, \ldots, n$.

That is, one collects all connectivity components $C_{Qi}$ of $C_Q$, and sums over their $p-\mu_i$-th Betti numbers. Then, the result is bounded from below by the $p$-th Betti number of the configuration manifold $Q$. The interpretation of this fact is that it is not possible that the critical manifold of a constrained Hamiltonian multibody system of the given type is 'topologically simple' if the configuration manifold is 'topologically contrived'.
5.2.3 An illustrative mechanical example

Let us now present a mechanical example to illustrate the implications of the weak Conley-Zehnder inequalities. To this end, we assume that $Q$ is some genus $g$ surface $\Sigma_g$, and let the Hamiltonian of the free system describe a mass point that moves on $\Sigma_g$ under the influence of the gravitational force (let $\Sigma_g$ be positioned in a way that the component of the gravitational force vector tangent to $\Sigma_g$ only vanishes at isolated points in $\Sigma_g$).

Furthermore, we assume constraints to be given that can be characterized by a smooth field $V$ of directional vectors (a rank 1 distribution) on $\Sigma_g$. Here, we run into a problem; the considerations in chapters two and four have assumed that $V$ is everywhere non-degenerate. However, for two dimensional closed, compact surfaces, this is only possible for the torus $T^2$ with $g = 1$. For all other $g$, every vector field necessarily has singularities which are enforced by the topology.

Figure 5.2: The torus $T^2$ standing on the table; the four thick dots are the equilibria of the free system. The thick lines around the small radius of the torus denote the components of $C_Q$. The dashed lines are the directions along which the mass point is permitted to move.
Hence, consider a torus $T^2$ which stands vertically on a horizontal table, and a mass point moving on its surface under the gravitational attraction of the earth. For the free system, there are precisely four isolated equilibria in $T^2$. Let the constraints be characterized by the requirement that the mass point shall not exhibit any horizontal velocity components parallel to the table during its motion. Then, it necessarily moves on embedded vertical circles in $T^2$.

The highest point of every such circle is an unstable equilibrium, and the lowest point is a stable equilibrium. The set of all unstable equilibria $C_1$ of the constrained system is a circle around the small radius of the standing torus, which contains the two upper equilibria of the free system. The set of all stable equilibria $C_2$ is a circle around the small radius of the torus that contains the two lower equilibria of the free system.

For this example, $C_Q$ apparently has two connectivity components $C_1$ and $C_2$ with indices $\mu(C_1) = 1$ and $\mu(C_2) = 0$. The weak Conley-Zehnder inequalities (2.9)

$$\sum_i B_{p-\mu_i}(C_i) \geq B_p(T^2)$$

imply that

$$B_0(C_2) \geq B_0(T^2) \quad B_1(C_2) + B_0(C_1) \geq B_2(T^2) \quad B_1(C_1) \geq B_2(T^2)$$

must be satisfied. Since $B_0(C_i) = B_1(C_i) = 1$, $B_0(T^2) = B_2(T^2) = 1$, and $B_1(T^2) = 2$, one verifies that in this example, the inequality signs can even be replaced by equality signs.
5.3 The stability of equilibrium solutions

We have derived a set of stability criteria for constrained Hamiltonian systems in chapter four, which apply to the multibody systems that are considered here.

Let \( a = (c, 0) \) with \( c \in \mathcal{C}_Q \) denote an equilibrium of a given constrained Hamiltonian mechanical system. According to our analysis in chapter four, one proceeds by determining the spectrum of the matrix

\[
\left( \begin{array}{cc}
0 & \alpha_c M_c^{-1} \alpha_c^T + \tilde{R}_c \\
-(\alpha_c^T D_c^2 u \alpha_c + \tilde{R}_c) & 0
\end{array} \right)
\]

The \( n \) by \( n \) matrix \( \tilde{R}_c \) is given by

\[
\tilde{R}_{c;jk} = u_i \alpha_j \alpha_k \alpha_{i,j} \Big|_{c}.
\] (5.1)

If there are eigenvalues with a positive real part, \( c \) is asymptotically unstable, and if all eigenvalues have a negative real part, \( c \) is asymptotically stable. These cases can be treated by use of the center manifold theorem in the manner demonstrated in chapter four.

It is much more difficult to analyze the case in which all eigenvalues are purely imaginary. We have only contented ourselves with an application of averaging theory, but have not arrived at any theorem that would control this case. Our averaging results suggest that if all of the imaginary eigenvalues are rationally independent, \( c \) is stable if the spectrum of \( \alpha_c^T D_c^2 u \alpha_c \) is strictly positive.

We remark for practical purposes that

\[
\alpha_c^T D_c^2 u \alpha_c + \tilde{R}_c = \alpha_c^T \text{Jac}_c(\alpha^T u_{\partial}) \alpha_c
\]

where ‘Jac’ computes the Jacobian matrix. It is usually easier to first compute the right hand side, and then to derive the corresponding expression for \( \tilde{R}_c \).

If \( \lambda \) is an eigenvalue of (5.1), then \(-\lambda^2\) obviously is an eigenvalue of the product of the two matrices in the off-diagonal blocks in arbitrary product order. Hence, if all eigenvalues of the matrix \((\alpha_c M_c^{-1} \alpha_c^T + \tilde{R}_c)^{(\alpha_c^T D_c^2 u \alpha_c + \tilde{R}_c)}\) are strictly positive and rationally independent, and if in addition, the spectrum of \( \alpha_c^T D_c^2 u \alpha_c \) is strictly positive, averaging theory suggests that \( c \) is stable.

We have, in chapter four, demonstrated that if there are imaginary eigenvalues with an algebraic multiplicity higher than one, one will in general observe solutions that diverge linearly in time. This ‘diffusive’ instability is very typical for non-holonomic systems, but does not occur if the constraints are integrable.
5.3.1 An illustration: A guided rolling disk on the 2-sphere

Figure 5.3: The mechanism in discussion. The disk rolls without sliding on the inner surface of the sphere; the spring acts in the axial direction along the rod, but also exhibits bending stiffness. If the rod is positioned in the horizontal plane, the disk lies on the horizontal plane for the value $\psi = 0$. The thick vertical circle is the critical manifold $C_Q$.

Let us consider another mechanical example, in order to illustrate our proposed method to analyze $C_{gen}$ and the stability of critical points under the physical flow.

The model consists of a thin disk of radius $r$ and mass $m$ that rolls on the inner surface of a 2-sphere of radius $R + r$, where $R > r$. The disk is guided by a massless rod of length $R$ that is linked to a frictionless rotational joint at the center of the sphere.
There shall be a linear spring that is attached to the rod, which produces both an axial and a bending torque. The axial angle shall be denoted by \( \psi \) (the position of the disk for \( \psi = 0 \) is described in the text below Figure 5.3). In spherical coordinates \((\phi, \theta)\), the relaxed position is given when \( \phi = 0 \). This accounts for the fact that the indicated torsional spring has a non-vanishing bending stiffness.

The center of mass of the disk moves on a sphere of radius \( R \), and the kinetic energy stemming from its translation is given by

\[
T_{tr} = \frac{mR^2}{2} (\dot{\phi}^2 + \dot{\theta}^2).
\]

The rotational kinetic energy is given by

\[
T_{rot} = \frac{1}{2} \frac{mR^2}{2} R^2 (\dot{\phi}^2 + \dot{\theta}^2) + \frac{1}{4} \frac{mR^2}{2} \psi^2.
\]

In the coordinate system \((\phi, \theta, \psi)\), the mass matrix therefore reads

\[
\begin{pmatrix}
\frac{3mR^2}{2} & 0 & 0 \\
0 & m_R^2 & 0 \\
0 & 0 & \frac{m_R^2}{4}
\end{pmatrix}.
\]

The potential energy is the sum of the gravitational potential center of mass, and the deformation energies of the spring in the \( \phi \) and \( \psi \) direction, that is,

\[
u = mgR \sin \theta + \frac{c_\phi}{2} \phi^2 + \frac{c_\psi}{2} \psi^2
\]

\((c_i\) are the corresponding spring constants).

There is a non-holonomic constraint condition that is obtained from the requirement that the disk shall roll on the sphere without sliding. This is expressed by

\[
\cos \psi \dot{\theta} + \sin \psi \dot{\phi} = 0.
\]

The matrix \( E^T(q) \) introduced at the beginning of this chapter to define the Pfaffian velocity constraints is thus given by \((\sin \psi, \cos \psi, 0)\), so that \( E^T M^{-1} E = \frac{2}{3mR^2} \).

The matrices of the orthoprojectors \( \beta \) and \( \alpha \) are straightforwardly obtained to be given by

\[
\beta(\phi, \theta, \psi) = \begin{pmatrix}
\sin^2 \psi & \sin \psi \cos \psi & 0 \\
\sin \psi \cos \psi & \cos^2 \psi & 0 \\
0 & 0 & 0
\end{pmatrix},
\]

\[
\alpha(\phi, \theta, \psi) = \begin{pmatrix}
\cos^2 \psi & -\sin \psi \cos \psi & 0 \\
-\sin \psi \cos \psi & \sin^2 \psi & 0 \\
0 & 0 & 1
\end{pmatrix}.
\]
The critical set of this mechanical system consists of all points where \( \alpha^T u_{\alpha} = 0 \) is satisfied. In this example, \( u_{\alpha} = (c_\phi \phi, mgR \cos \theta, c_\psi \psi) \). Therefore, one arrives at

\[
C_Q = \{ (\phi, \theta, \psi) | \phi = 0, \psi = 0 \},
\]

which is the thick vertical circle of radius \( R + r \) in Figure 5.3.

The Jacobian of \( \alpha^T u_{\alpha} \) on \( C_Q \) is given by

\[
\text{Jac}_c(\alpha^T u_{\alpha}) = \begin{pmatrix}
c_\phi & 0 & -mg \cos \theta \\
0 & 0 & mg \cos \theta \\
0 & 0 & c_\psi
\end{pmatrix},
\]

and since \( \alpha \) is given by \( \text{diag}(1, 0, 1) \) everywhere on \( C_Q \), one finds that

\[
\alpha_c^T D_c^2 u_\alpha + \bar{R}_c \iff \begin{pmatrix}
c_\phi & 0 & -mg \cos \theta \\
0 & 0 & c_\psi
\end{pmatrix}, \tag{5.2}
\]

(where we have cancelled the zero row and column) so that from

\[
\alpha_c^T D_c^2 u_\alpha \iff \begin{pmatrix}
c_\phi & 0 \\
0 & c_\psi
\end{pmatrix},
\]

which is positive definite, one immediately obtains that

\[
\bar{R}_c \iff \begin{pmatrix}
0 & -mg \cos \theta \\
0 & 0
\end{pmatrix}
\]

holds.

Hence, we find that

\[
\alpha_c M_c^{-1} \alpha_c^T \bar{R}_c \iff \begin{pmatrix}
\frac{2}{3mR^2} & 0 \\
-\frac{2g \cos \theta}{3mR^2} & \frac{4g^2}{m^2 R^2} + \frac{m^2 g^2 \cos^2 \theta}{2}
\end{pmatrix},
\]

(where we again have cancelled the zero row and column) and multiplying this matrix with the right hand side of (5.2) yields

\[
\Omega_c(\theta) := \begin{pmatrix}
\frac{2c_\phi}{3mR^2} & \frac{-2g \cos \theta}{3mR^2} \\
-mgc_\phi \cos \theta & \frac{4c_\phi^2}{mR^2} + \frac{m^2 g^2 \cos^2 \theta}{2}
\end{pmatrix}.
\]

It is clear that the spectrum of \( \Omega_c(\theta = -\frac{\pi}{2}) \) is strictly positive, so that the lowest point on the sphere is suggested to be stable according to averaging theory if the 'torsional eigenfrequency' \( \frac{4c_\phi}{mR^2} \) and the 'pendular eigenfrequency' \( \frac{2c_\phi}{3mR^2} \) are rationally independent.

One obtains the same result for \( \theta = +\frac{\pi}{2} \), which implies that according to averaging theory, the highest point on the sphere is stable if the lowest one is. Furthermore, a small
calculation shows that \( \det \Omega_c(\theta) = \frac{8c_0^3c_\phi}{3m^2r^4K^2} \), and that \( \text{Tr} \Omega_c(\theta) = \frac{2c_0}{3mK^2} + \frac{4c_\phi}{m^2} + m^2g^2 \cos^2 \theta \). The zeros of the characteristic polynomial

\[
\chi_{\Omega_c(\theta)}(\rho) = \rho^2 - \text{Tr} \Omega_c(\theta) \rho + \det \Omega_c(\theta)
\]

are the two eigenvalues of \( \Omega_c(\theta) \), and are given by

\[
\rho_{1,2}(\theta) = \frac{1}{2} \left( \text{Tr} \Omega_c(\theta) \pm \sqrt{\left(\text{Tr} \Omega_c(\theta)\right)^2 - 4\det \Omega_c(\theta)} \right).
\]

One can straightforwardly verify that both of them are positive for all \( \theta \in [-\frac{\pi}{2}, \frac{\pi}{2}] \). According to the averaging results from chapter four, those points on \( C_0 \) for which the ratio \( \frac{\rho_1(\theta)}{\rho_2(\theta)} \) is irrational are suggested to be stable.

The purpose of this example has been to demonstrate the way to proceed in order to study the stability of a Hamiltonian multibody system with non-holonomic constraints. In our opinion, the presented formalism, which is based on the use of mass-orthogonal projectors, renders the calculations very short and transparent, and allows one to go very far by analytical means.
Seite Leer / Blank leaf
Bibliography


Contents

1 CONSTRANGED HAMILTONIAN SYSTEMS ON SYMPLECTIC MANIFOLDS 7

1.1 A minisurvey of Hamiltonian dynamics .......................... 7
   1.1.1 The Darboux theorem .............................................. 9
   1.1.2 Symplectomorphisms ............................................. 9
   1.1.3 Poisson structures and energy conservation ................. 10

1.2 Induced Hamiltonian systems on symplectic submanifolds ........... 11

1.3 Almost symplectic distributions and associated natural structures 13
   1.3.1 Integrability and non-integrability ............................. 14
   1.3.2 Existence of a ω-skew selfadjoint bundle projector .......... 15
   1.3.3 Further auxiliary structures ................................... 17

1.4 Definition of the constrained Hamiltonian system ................. 18
   1.4.1 Integrable distributions and symplecticity .................. 20

2 THE GEOMETRY AND TOPOLOGY OF THE CRITICAL MANIFOLD 23

2.1 Local properties of the critical manifold .......................... 24
   2.1.1 A generalized Hessian for the constrained system ............ 25
   2.1.2 An auxiliary gradient-like flow ................................. 27
   2.1.3 The gradient-like flow in a vicinity of C ..................... 28
   2.1.4 Local normal hyperbolicity .................................... 30

2.2 Geometry of non-degenerate critical manifolds .................... 31
   2.2.1 Morse functions and non-degenerate critical manifolds ....... 31
   2.2.2 Generic connectivity components are normal hyperbolic .... 32
   2.2.3 Non-degenerate critical manifolds and critical points of H ... 33
   2.2.4 A note concerning integrable distributions .................. 34

2.3 Global topology of non-degenerate critical manifolds ............ 34
2.3.1 Index pairs and relative (co-)homology .................................. 34
2.3.2 Computing the relative cohomology of an index pair ................. 35
2.3.3 The Conley-Zehnder inequalities for non-degenerate critical manifolds 37
2.3.4 A sharper result for generic connectivity components .................. 39

2.4 Proof of the strong Conley-Zehnder inequalities using the Witten complex ................................................................. 40
  2.4.1 Definition of the Witten complex ........................................... 41
  2.4.2 A relationship between the Witten complexes of \((M, H)\) and \((C_{\text{gen}}, H|_{C_{\text{gen}}})\) 44
  2.4.3 Proof of inequality (2.15) ..................................................... 46

3 CONSTRAINTED HAMILTONIAN MECHANICS ................................. 53
  3.1 A minisurvey of Hamiltonian mechanics .................................. 54
  3.2 Constrained Hamiltonian mechanical systems ............................ 56
    3.2.1 The Hölder variational principle ...................................... 58
  3.3 Continuations away from the physical sheet and issues on the numerical simulation of constrained Hamiltonian mechanical systems 60
    3.3.1 A preliminary attempt .................................................... 62
    3.3.2 A more refined auxiliary dynamical system ......................... 65
  3.4 A note on constrained Hamiltonian mechanical systems with symmetries ............................................................... 71

4 EQUILIBRIA IN CONSTRAINED HAMILTONIAN MECHANICS ................. 73
  4.1 The topology of the critical manifold .................................... 73
  4.2 Stability of equilibria in constrained Hamiltonian systems ............ 75
    4.2.1 The equations of motion in a neighborhood of a critical point ... 76
    4.2.2 Asymptotic behaviour .................................................... 79
    4.2.3 Critically stable equilibrium solutions ............................. 80
    4.2.4 Resonances and instabilities: a special case ....................... 84
  4.3 The stability of equilibria in constrained Hamiltonian mechanical systems ............................................................... 86
  4.4 A note on the shape of non-holonomic \(\varepsilon\)-balls ................... 88
5 APPLICATIONS, ILLUSTRATIONS AND EXAMPLES

5.1 Constrained Hamiltonian multibody systems ........................................... 91
  5.1.1 The continuation away from the physical sheet .................................. 92

5.2 Geometry and topology of the critical manifold ....................................... 93
  5.2.1 Numerically determining the critical manifold .................................... 95
  5.2.2 The weak Conley-Zehnder inequalities ............................................. 96
  5.2.3 An illustrative mechanical example ................................................. 98

5.3 The stability of equilibrium solutions .................................................... 100
  5.3.1 An illustration: A guided rolling disk on the 2-sphere ......................... 101
Seite Leer / Blank leaf
CURRICULUM VITAE

I was born on the 6th of October 1968 in Winterthur ZH, as the son of Dr. Ing. ETH Yian Nian Chen, a mechanical engineer, and of Shu Hsia (Susie) Chen-Chao, a writer in the Chinese language. I attended elementary and high school in Winterthur, and obtained a Matura Typus B degree from the Kantonsschule Rychenberg in 1987. From October 1987 until spring 1992, I studied mechanical engineering at ETH Zürich. The diploma thesis was written under the supervision of Prof. Dr. H. Brauchli. Between February 1992 and August 1998, I worked as an assistant at the institute of mechanics. I began a second study in theoretical physics at ETH Zürich in spring 1992, which I concluded in August 1996 with a diploma thesis under the supervision of Prof. Dr. J. Fröhlich. Since September 1998, I have been employed at the institute of theoretical physics at ETH Zürich, and am completing a PhD thesis in mathematical physics under his supervision.