3 Controllability, approximations, and optimal control

3.1 Introduction and examples

Control systems are distinguished from general dynamical systems by having inputs (and outputs), and, even more importantly, by new questions. E.g. instead of studying where a solution goes, one asks whether it is possible to get to a desired state by judiciously choosing the input. Just as for general dynamical systems, there are many different settings in which they may be studied, e.g. continuous versus discrete time (or even discrete event systems), with deterministic or stochastic dynamics, Systems on finite, discrete or continuous, on finite dimensional or infinite dimensional state spaces. The discussion here will be focused on continuous time, deterministic systems defined by smooth vector fields on finite dimensional spaces (smooth manifolds). The special case of linear systems has been studied in great depth starting in the 1960s and is at the root of most modern engineering, from telecommunications to space flight. The main mathematical tools are linear algebra, and the theory of functions of a complex variable, much functional analysis and stochastic processes. In the nonlinear setting, the classic Laplace transform is of little use, and tools from differential geometry naturally take the place that matrices and linear algebra occupy for linear systems.

The following notes will only be able to address a small subset of many fundamental questions in control. We just name a few:

- Modeling: We focus on simple systems that can be modeled by ordinary differential equations using first principles from mechanics.
- Regularity: We only state classical existence and uniqueness results, and make a few remarks about complications arising in feedback stabilization problems and optimal control.
- Realization: One traditional way is to consider control systems as abstract mapping from controls \( u(\cdot) \) to outputs \( y(\cdot) \). A fundamental question asks under which conditions such a mapping arises from a state-space system \( \dot{x} = f(x,u), y = h(x) \), and when it does, to construct a normal form from the data.
- Identification: Closely related is the problem of identifying a system that fits a set of, generally noisy, measured input-output pairs.
- For systems in state space form, a fundamental question that we will address is that of controllability: Does a given system have solutions that start at a given point \((t_0, x_0)\) and end at a desired point \((t_0 + T, x_1)\)? If so, the path-planning problem asks for constructing an input \( u(\cdot) \) that realizes such a transfer. Generally, one asks in addition that the corresponding solution curve stays close to a desired reference trajectory.
- The fundamental concept of automatic control is to replace the open loop controls \( u(t) \) by closed loop controls (or feedback controls \( u(x) \) or \( u(y) \) where \( y = h(x) \) is an output that generally only captures a small portion of the current state. A special case is that of feedback stabilization: finding a feedback law \( u(x) \) that renders a system \( \dot{x} = f(x,u(x)) \) asymptotically stable. Technically, discontinuities of the feedback law \( u(x) \) give rise to difficulties about even notions of existence of solutions. In the nonlinear setting, stabilizing feedback laws are severely restricted by topological constraints.
- Generally, for any task such as going from state \( x_0 \) to state \( x_1 \), there exist many possible controls that achieve this task. Thus it is natural to ask for an optimal way of achieving the task. We shall consider special cases of optimal control, demonstrate the dual relationship with controllability, and relate the Pontryagin Maximum Principle to the classical calculus of variations and to Hamiltonian mechanics.
• In a different way dual to controllability is the notion of observability which asks whether
one may infer the state of the system from observations $y(\cdot)$ over some time interval.
• Given two systems (or state-space realizations), a fundamental question is when these are
equivalent in a natural way (e.g. under diffeomorphisms of the state space, generally allowing
feedback transformations. For each orbit of equivalent systems, one likes to designate
a distinguished normal form. Closely related is the objective to construct approximat-
ing systems that are easier to analyze, but which preserve properties of critical interest.
In the sequel we will construct some nilpotent approximating systems that are better at
preserving controllability and stabilizability properties than simple linearizations.
• Once beyond the fundamental mathematical questions, the applied engineering sciences
introduce many more practical questions, starting with notions such as robustness which
address e.g. modeling errors and noisy environments and controls.

3.2 Some technical notes on existence and uniqueness
For most of the sequel we shall consider control systems that are defined by smooth vector fields
(in many cases even requiring analyticity). It is natural to start with piecewise smooth controls
(or even piecewise constant) controls. However, when studying optimality it is essential to be
able to make sense of controls that arise as limits of sequences. The natural setting is to require
that the controls are measurable (or integrable) functions of time.
The existence and uniqueness of solutions of initial value problems for control systems with such
regularity properties follow from the classical results by Carathéodory.

**Theorem 3.1 (Carathéodory)** Suppose $a, b > 0$ and $\Omega = \{(t, x) \in \mathbb{R}^{n+1}: |t - t_0| \leq a, \|x - x_0\| \leq b\}$ and $f: \Omega \rightarrow \mathbb{R}^n$ is continuous in $x$ for every fixed $t$ and measurable in $t$ for every fixed $x$.
If there exists a Lebesgue measurable function $m: [t_0 - a, t_0 + a] \rightarrow \mathbb{R}$ such that for all $(t, x) \in \Omega$, $\|f(t, x)\| \leq m(t)$ then there exists $\beta > 0$ and a continuous function $\phi: [t_0 - \beta, t_0 + \beta] \rightarrow \mathbb{R}^n$ such that $\phi(t_0) = x_0$ and for almost all $t \in (t_0 - \beta, t_0 + \beta)$ (i.e., there exists a set $M$ of measure zero such for all $t \in (t_0 - \beta, t_0 + \beta) \setminus M$) $\phi$ is differentiable at $t$ and $\phi'(t) = f(t, \phi(t))$.

A standard reference for this existence theorem is the classic textbook *Theory of Ordinary
The same text proves a number of extension and uniqueness results (pages 42-60). For our
purposes the following very simple version will suffice (a special case of theorem 2.1 (pages 48
and 49).

**Theorem 3.2** Suppose that in addition to the hypothesis of theorem 3.1 $f$ satisfies a local Lips-
chitz condition, i.e., there exists $L > 0$ such that for all $(t, x_1), (t, x_2) \in \Omega$, $\|f(t, x_1) - f(t, x_2)\| \leq L\|x_1 - x_2\|$ then the function $\phi$ satisfying the conclusion of theorem 3.1 is unique.

3.3 A few simple mechanical examples
The following mechanical systems serve for motivation and as examples on which to test the
theory and computational algorithms. For many more introductory examples, and various gen-
eralizations see the textbook *Nonholonomic Mechanics and Control* by A. Bloch, P. Crouch,
J. Baillieul, and J. Marsden [Springer 2007].

**Example 3.1** Modeling a penny that rolls in a plane without slipping and which is allowed to
rotate about its vertical axis. Let $(X_1, x_2) \in \mathbb{R}^2$ denote the point of contact of the penny with the
plane, $\phi \in S^1$ its orientation with respect to coordinate axes, and $\theta \in S^1$ the angle it has rotated about its horizontal axis from a fixed reference orientation.

$$x_1' = R \cos \phi \cdot \theta, \quad y_1' = R \sin \phi \cdot \theta.$$ (1)

Example 3.2 The following description for parallel parking a car is taken from The combinatorics of nonlinear controllability and noncommuting by M. Kawski [Abdus Salam ICTP Lect. Notes series Vol. 8 (2002) pp. 223-312].

Consider a mathematical model for the simpler bicycle (less controversial case as it does not require differentials to adjust for the different speeds of inside and outside wheels). Let $(x,y) \in \mathbb{R}^2$ denote the point of contact of the rear wheel with the plane (center of rear axle in the case of a car), $\theta \in S^1$ be the angle of the bicycle with the $x_1$-axis, and $\phi \in S^1$ be the angle of the front wheel(s) with the direction of the bicycle. An algebraic constraint captures that the distance between front and rear wheel is constant, equal to the length $L$. Thus the position of the front wheel (point of contact with plane) is $(x + L \cos \theta, y + L \sin \theta)$. The conditions that the wheels can slip neither forward nor sideways, each can only roll in the direction of the wheel is captured in

$$\begin{cases} 0 = \cos \theta \, dy - \sin \theta \, dx \\ 0 = \sin(\theta + \phi) \, d(x + L \cos \theta) - \cos(\theta + \phi) \, d(y + L \sin \theta) \end{cases}$$ (2)

Introducing the speed $v = \|x'^{2} + y'^{2}\|$ of the rear wheel (or of the center of the rear axle), write $x = v \cos \theta$ and $y = v \sin \theta$.


There are many generalizations of this problem – the easiest version being a ball of radius $a$ rolling a flat plane. Consider as control the motion of a movable second parallel plane which rests on top of the ball. Clearly the center of the ball moves at half the velocity of the top plane. The natural state-space is $\mathbb{R}^2 \times \text{SO}(3)$ and the kinematics is described by

$$x_1' = a \, u_1, \quad x_2' = a \, u_2, \quad \dot{\Omega} = \Omega (R_2 u_1 + R_1 u_2)$$ (3)

where $R_i \in \text{so}(3, \mathbb{R})$ describe the infinitesimal rotations about the $i$-th coordinate axes.
Example 3.4 (J. Hauser) The inverted pendulum and inverted double pendulum, on a cart or rotating platform, are popular test cases for engineering literature. The controllability of a double pendulum about the following special configuration was recently studied by J. Hauser.

Denoting by $\beta$ the ratio of link lengths the equations of motions are

$$\ddot{\theta} = u, \quad \ddot{\phi} = \sin \phi - \beta \cos(\phi - \theta) \dot{\theta}^2 + \beta u \sin(\phi - \theta)$$

Example 3.5 Known widely as the Brockett system, this system appears naturally in many contexts as the first nontrivial system that exhibits certain properties that are almost universal in more complicated systems. Structurally, e.g. from the point of Lie groups, it is well known under a different name (exercise): On $\mathbb{R}^3$ consider the 2-input system defined by:

$$\dot{x}_1 = u_1, \quad \dot{x}_2 = u_2, \quad \dot{x}_3 = x_1 u_2 - x_2 u_1.$$  \hspace{1cm} (4)

Further references

3.4 General remarks on controllability

The fundamental concept is very simple: A system is called controllable from a state \( p \) to a state \( q \) if there exists an admissible control \( u \) that steers the system from initial state \( p \) at time \( t_0 \) to the final state \( q = \Phi(t_0 + T,p,t_0,u) \). A system is called (globally) controllable if for every pair \( (p,q) \) of states it is controllable from \( p \) to \( q \). There are many specializations of this concept, ranging from finite state machines (or automata) to systems governed by partial differential equations. [References: S. Eilenberg, Automata, languages, and machines, and various recent papers by J.-M. Coron, on controllability of various PDEs.] Here we shall focus on systems with continuous states in \( \mathbb{R}^n \) or on finite dimensional manifolds \( M^n \), governed by controlled differential and difference equations.

**Exercise 3.1** Give an example of a control system with a finite state space, and define suitable notions of controllability for this kind of system. Discuss whether the controllability notions are symmetric with respect to time reversal.

A basic technical notion is that of a reachable set. For sake of clarity we shall assume here that the system is time-invariant, allowing us to generally assume initialization at time \( t_0 = 0 \).

Define the reachable set from the state \( p \) at time \( T \) as

\[
R_T(p) = \{ \Phi(T,p,u): u \in \mathcal{U} \text{ admissible control} \}
\]

It is convenient to also introduce

\[
R_{\leq T} = \bigcup_{t=0}^{T} R_t(p), \quad R_{\leq \infty} = \bigcup_{t \geq 0} R_t(p)
\]

The statement \( q \in R_p(T) \) may be read as “\( q \) is reachable from \( p \)”, or as “the system is controllable to \( q \) from \( p \)”. The system is reachable from \( p \) if the entire state-space \( X \subseteq R_{\leq \infty}(p) \) can be reached from \( p \). Analogously the system is \( q \)-controllable (“controllable to \( q \)” if \( q \in \bigcap_{p \in X} R_{\leq \infty}(p) \).

For many systems (but generally not for discrete-time systems) the reversal of time corresponds to interchanging the roles of reachability and controllability, and one commonly uses the term controllability when one technically analyzes reachability. In the sequel we shall investigate when the reachable set \( R_T(p) \) contains an equilibrium point or a reference trajectory in its interior, and still use the term controllability. For typical applications one usually studies the system obtained by replacing \( t \) by \(-t\).

Many special versions of the basic concept of controllability have been introduced. Important for us is that in the linear setting many of them coincide with each other, whereas in nonlinear settings one needs to very carefully distinguish between global and local controllability, and controllability in arbitrarily large or bounded, or arbitrarily small times, etc. A positive way to read this is that there are plenty of open research problems, each corresponding to a different meaningful setting in which to analyze controllability.

A central objective of the analysis is to obtain readily checkable conditions in terms of the systems data that allow one to algorithmically decide whether a system is controllable or not. Such conditions generally are only existence results (\( \exists u \text{ s.t. } \Phi(T,p,u) = q \)). Constructing controls that actually steer the system from a given state \( p \) to a given state \( q \) is generally a different problem which we shall revisit when discussing feedback stabilization.

For finite dimensional systems the classical result is the Kalman rank condition which asserts that the time-invariant system \( \dot{x} = Ax + Bu, \ x \in \mathbb{R}^n, \ \forall t, \ u(t) \in U = \mathbb{R}^m \) is globally controllable.
if and only if the rank of the compound matrix \([B, AB, A^2B, \ldots, A^{n-1}B]\) is \(n\). Remarkably, the same condition holds for discrete time systems with the corresponding data. Note that this condition applies to systems that impose no a-priori bound on the magnitude \(|u(\cdot)|\) of the control. The next subsection will study the Kalman rank condition in more detail.

**Exercise 3.2** Give an example of matrices \(A, B\) defining a linear system that is controllable for the set of unbounded controls \((U = \mathbb{R}^m)\), but is not controllable with bounded controls taking values in \(U = [-1, 1]^m\).

Different technical results are available for time-varying linear systems in which the matrices \(A = A(t)\) and \(B = B(t)\) are not necessarily constant. The next popular class of systems has a bilinear structure \(\dot{x} = Ax + \sum_{j=1}^{m} u_j B_j x\). While such systems at first sight seem to much resemble linear systems, in many ways they display many characteristics of general nonlinear systems. Indeed, many nonlinear systems can be lifted to bilinear systems (on a much higher dimensional state-space) – and therefore, in many ways, bilinear systems must be at least as complex as many nonlinear systems. However, what makes some bilinear systems particularly attractive to study is that their underlying state-space may be taken to have the structure of a (matrix-)Lie group, thus affording a particularly rich set of tools for studying them. In the last two decades many new results for controllability on Lie groups have been found, and this area continues to see much ongoing research.

In our course, we shall jump ahead and next concentrate on nonlinear systems that are affine in the control

\[
\dot{x} = f_0(x) + \sum_{j=1}^{m} u_j f_j(x)
\]

defined by a finite number of smooth vector fields \(f_0\) and \(f_i\) (which often are required to be smooth \(C^\infty\), or analytic \(C^\omega\)). These are a special case of fully nonlinear systems \(\dot{x} = f(x, u)\) for which fewer general results are available. In recent years, research has focused on hybrid systems which mix continuous and discrete states as they occur in systems with hysteresis or in switched systems. Again, this area provides many opportunities for new research projects, but goes far beyond our course.

In nonlinear systems, it is common practice to focus on the case of bounded controls. We shall usually assume that for all \(i\) and for all \(t\), \(|u_i(t)| \leq 1\), or \(\sum_{i=1}^{m} u_i^2 \leq 1\). More generally one may allow the controls \(u\) to take values in a convex, compact subset \(U \subseteq \mathbb{R}^m\). There may be many possible explanations for this distinction of working with bounded and unbounded controls when studying nonlinear and linear systems. Clearly, in practical applications there are almost always bounds on the admissible controls. However, the mathematical treatment may impose different conditions to be feasible. This distinction will become clearer later when we study optimal control problems which are closely related to controllability: Different system structures will suggest different objective functions which in turn determine the underlying sets of admissible controls.

Back to the affine system (7), the best studied notion is that of small-time local controllability (STLC) about an equilibrium point of the drift vector field \(f_0\) for the control set \(U = [-1, 1]^m\), with STLC defined as \(\forall T > 0, p \in \text{int} R_T(p)\). Algebraic conditions analogous to the Kalman rank condition, however, are known only for the weaker notion of accessibility defined as \(\forall T > 0, \emptyset \neq \text{int} R_T(p)\). For practical purposes, more interesting is the notion of STLC about a not necessarily stationary reference trajectory \(\Phi(\cdot, p, u^*)\), defined via \(\forall T > 0, \Phi(T, p, u^*) \in \text{int} R_T(p)\).
Exercise 3.3 Calculate the reachable sets $R_T(0)$ from $p = 0$ for the system $\dot{x} = y + u$, $\dot{y} = x$, subject to the constraint $\forall t, u(t) \in U = [0, 1]$. Discuss various notions of controllability for this system (incl. global, local, and small-time local controllability).

3.5 Linear systems

Consider the finite dimensional linear systems with no bounds on the control
\[ \dot{x} = Ax + Bu, \quad x \in \mathbb{R}^n, \quad u \in \mathbb{R}^m. \]  

Theorem 3.3 (Kalman Rank Condition) System \( (8) \) is globally controllable if and only if
\[ \text{rank} \begin{bmatrix} B, AB, A^2B, \ldots, A^{n-1}B \end{bmatrix} = n. \]  

The first key observation regarding the reachable sets of system \( (8) \) is that they are always affine subspaces on $\mathbb{R}^n$. Indeed, consider the variation of parameters formula for the solution curves of system \( (8) \) for any control $u(\cdot)$.
\[ \Phi(T, p, u) = e^{tA}p + \int_0^T e^{(t-s)A}Bu(s)ds \]  

It is immediate that for any constants $c_1, c_2$ and any controls $u_1, u_2$
\[ \Phi(T, p, c_1u_1 + c_2u_2) - e^{tA}p = c_1(\Phi(T, p, u_1) - e^{tA}p) + c_2(\Phi(T, p, u_2) - e^{tA}p) \]  

and thus $R_T(0) \subseteq \mathbb{R}^n$ is a subspace, and $R_T(p) = e^{tA}p + R_T(0)$ is a translate of a subspace. Due to the additive structure, it suffices to consider the reachable sets $R_T(0)$ for systems initialized at $\Phi(0) = 0$.

Exercise 3.4 What can you say in the case of a bounded control set $U = [1, 1]^m$?

Exercise 3.5 What can you say about the structure of the reachable sets for linear discrete time systems $x(t + 1) = Ax(t) + Bu(t)$?

Recall the definition of the matrix exponential as an absolutely convergent power series. Next, using the Cayley Hamilton theorem, there exist analytic functions $c_k : R \mapsto \mathbb{R}$ such that
\[ e^{tA} = \sum_{k=0}^{\infty} \frac{(tA)^k}{k!} = \sum_{k=0}^{n-1} c_k(t) \frac{A^k}{k!} \]  

Exercise 3.6 Prove this assertion. Calculate the $c_i$ for the case of $A = \begin{pmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$.

To prove the only-if part of theorem 3.3 assume to the contrary that rank$[B, AB, \ldots, A^{n-1}B] < n$. Thus there exists a nonzero vector $\xi \in \mathbb{R}^n$ such that $\xi^T \cdot [B, AB, A^2B, \ldots, A^{n-1}B] = 0$ and consequently also $\xi^T \cdot e^{tA} = 0$ for all $t$. This implies that
\[ \xi^T \cdot \Phi(T, 0, u) = \xi^T \cdot \int_0^T e^{(t-s)A}Bu(s)ds = \int_0^T (\xi^T \cdot e^{(t-s)A})Bu(s)ds = \int_0^T 0 \cdot u(s)ds = 0 \]  

and consequently $R(T,0) \subseteq \{\xi\}^\perp = \{x \in \mathbb{R}^n : x \perp \xi\}$ is contained in a proper subspace.

To prove the if-part of the theorem, one might try to compute the coefficients in a Taylor expansion of $\Phi(T,0,u)$ about $t=0$. But any straightforward such attempt will encounter derivatives of the control $u$, which generally is not even assumed to be continuous. One alternative is to approximate any control $u$ by a sequence of smooth controls, and then use analytic tools. This is more typical in the case of nonlinear systems. Here we follow [EDS p.89] and take advantage of the linear structure:

Suppose the system is not controllable. By above observations, the reachable set is contained in some proper subspace and hence there exists a nonzero vector $\xi \in \mathbb{R}^n$ such that $\xi^T \cdot \Phi(T,0,u) = 0$ for all $T$ and all $u$. Now analyze the trajectory corresponding to the control $u^*(t) = B^T e^{(T-t)A^T} \xi$

$$0 = \xi^T \cdot \Phi(T,0,u^*) = \xi^T \cdot \int_0^T e^{(t-s)A} B \cdot B^T e^{(T-s)A^T} \xi ds = \int_0^T \|B^T e^{(T-s)A^T} \xi\|^2 ds$$

and hence $B^T e^{(T-s)A^T} \xi \equiv 0$ for almost all $s \in [0,T]$. Now – with no control variables appearing anymore – the Taylor expansion about $s = T$

$$0 \equiv \xi^T e^{(T-s)A} B = \xi^T - \xi^T AB \cdot (s-T) + \frac{1}{2!} \xi^T A^2 B \cdot (s-T)^2 - \frac{1}{3!} \xi^T A^2 B \cdot (s-T)^3 \ldots$$

immediately yields the desired algebraic condition that for all $k \geq 0$, $\xi^T \cdot A^k B = 0$, and hence the compound matrix is rank deficient.

**Exercise 3.7** Explore whether some variation of the above argument may be applied to nonlinear systems. Explain what needs to be modified, and how. Or explain which arguments have little hope of generalization.

### 3.6 Nonlinear systems affine in the control

Among nonlinear systems we concentrate on systems that are affine in the control

$$\dot{x} = f_0(x) + \sum_{j=1}^m u_j f_j(x), \quad \forall t, \; x(t) \in \mathbb{R}^n, \; |u_j(t)| \leq 1$$

defined by a finite number of analytic vector fields $f_0$ and $f_j$. The controls are generally assumed to be measurable and bounded, although for many subsequent arguments it will suffice to consider controls that are piecewise constant. Meticulous analysis of the relation of the reachable sets $R_{T,pc}(p)$ and $R_{T,meas}(p)$ under various hypotheses on the smoothness of the data (e.g. $C^r, C^\infty, C^\omega$ and control bounds) may be found in numerous articles, especially [Grasse]. We will revisit this question in the context of optimal controls later in the course.

For piecewise constant controls $u: [0,T] \mapsto [-1,1]^m$ defined by the data $0 \leq t_1 \leq t_2 \leq \ldots t_r = T$ and control values $u(t) = u^k \in [-1,1]^m$ for $t_{k-1} \leq t < t_k$ we consider the *endpoint* map

$$(t_1, t_2, \ldots, t_r, u^1, \ldots, u^r) \mapsto u \mapsto \Phi(T,p,u) \in \mathbb{R}^n$$

While we do not a-priori limit the number $r$ of pieces (or switchings), effectively we will always consider only a finite number of switchings. In some sense the rationale is that we are looking for some sort of rank condition, and the range of the map is finite-dimensional. Hence, we hope for some *finite* derivative information to be sufficient to make decisions about controllability.
Indeed, the questions of whether small-time local controllability even for analytic systems is finitely determined is still an open question, see [Agrachev, open problems book vol. I]. Consequently, we consider this as a map from a subset of the finite dimensional space $\mathbb{R}^{2r}$ (rather than the infinite dimensional space $U = \{u: [0, T] \mapsto [-1, 1]^m \text{ measurable} \}$ of all admissible controls. The strategy, or game plan, is to use derivatives of this restriction of the endpoint map to this finite dimensional space to obtain some algebraic criteria similar to the Kalman rank condition of the linear setting. The objective is that the criteria be readily checkable in terms of algebraic computations involving the data $f_0$ and $f_j$, in particular, not involving the solving of any differential equations, or calculations explicitly involving any controls.

### 3.6.1 One basic variation, differential of the flow, and Lie derivatives

As a start consider the most simple variations of a reference control $u^*: [0, T] \mapsto U \subseteq \mathbb{R}$ that changes the value of the control only on some small subinterval $[t_0, t_0 + \varepsilon] \subseteq [0, T]$ by the amount $v$. Assuming that $u^* \pm v$ still lies in the set $U$ of admissible values of the control, the resulting endpoints may be considered as a curve in the state-space parameterized by the duration $\varepsilon$ of the variation, or by the size $v$ of the variation, or both. Using exponential notation also for the flows of nonlinear vector fields, this curve is

$$\gamma(\varepsilon, v) = e^{(T-t_0-\varepsilon)(f+u^*g)} e^{\varepsilon(f+u^*g+vg)} e^{t_0(f+u^*g)} (p)$$

As either one (or both) go to zero, the endpoints approach the endpoint of the reference trajectory. Differentiating with respect to either $\varepsilon$ or $v$ and evaluating at $\varepsilon = 0$ or $v = 0$ one obtains a variational vector at the endpoint. This provides infinitesimal information about the location of reachable set $R(T)$ relative to $\gamma(T)$. Note that in order to differentiate with respect to the size of the control variation one generally assumes that $u^* \in \text{int} U$, so that $u^* \pm v \in U$ for all sufficiently small values of $v$. On the other hand, one may consider a fixed $u^* + v \in U$ and only vary the times $\varepsilon$, even when $u^* \in \partial U$ is a boundary value of $U$.

For various computational and conceptual reasons, it turns out to be more convenient to pull back all such endpoints and variational vectors to the starting point – e.g. in view of desirable algebraic conditions in terms of (derivatives) of the data at the starting point. Thus we consider the curve (parameterized by either $v$ or $\varepsilon$)

$$\gamma_0(\varepsilon, v) = e^{-T(f+u^*g)} \gamma(\varepsilon, v) = e^{-(t_0+\varepsilon)(f+u^*g)} e^{\varepsilon(f+u^*g+vg)} e^{t_0(f+u^*g)} (p)$$

Note that even though can only move forward, not backward along drift field, in some expressions one still sees $e^{-\ell T}$. To calculate these derivatives is a matter of using the chain-rule, and differentiating flows with respect to the initial conditions.

$$\frac{\partial}{\partial \varepsilon} \big|_{\varepsilon=0} \gamma_0(\varepsilon, v) = \frac{\partial}{\partial \varepsilon} \big|_{\varepsilon=0} e^{-(t_0+\varepsilon)(f+u^*g)} e^{\varepsilon(f+u^*g+vg)} e^{t_0(f+u^*g)} (p), \quad \text{or}$$

$$\frac{\partial}{\partial v} \big|_{v=0} \gamma_0(\varepsilon, v) = \frac{\partial}{\partial v} \big|_{v=0} e^{-(t_0+\varepsilon)(f+u^*g)} e^{\varepsilon(f+u^*g+vg)} e^{t_0(f+u^*g)} (p)$$

The basic form of this expression, and a convenient and suggestive notation for the derivative is

$$\frac{\partial}{\partial \varepsilon} \big|_{\varepsilon=0} e^{-tX} e^{\varepsilon Y} e^{tX} = \left( e^{-tX} \right)_* Y e^{tX}$$

where $\Psi_*$ denotes the tangent map or differential of a map $\Psi$. In coordinates, $\Psi_* = D\Psi$ is the Jacobian matrix of partial derivatives of $\Psi$. In our case, $\Psi_t(p) = \Phi(t, p) = e^{tX}(p)$ is the flow of a vector field $X$ evaluated at time $t$ with initial point $p$, i.e. the solution of the initial value
problem $\dot{x} = X(x)$, $x(0) = p$. In coordinates, the differential $\Phi_\ast = (e^{tX})_\ast$ is the solution of the time-varying linear differential equation $M(t) = (DX)(\Phi(t,p)) \cdot M(t)$ with $M(0) = I \in \mathbb{R}^{n \times n}$ (the identity matrix).

**Exercise 3.8** Observe that $\Phi_\ast(t,(1,0)) = (\cos(t),\sin(t))$ is one solution the uncontrolled dynamical system $\dot{x} = -y$, $\dot{y} = x(2 - x^2 - y^2)$ with initial condition $x(0) = 1$, $y(0) = 0$.

Write down the variational equation $M(t) = (DF)(\Phi_\ast(t)) \cdot M(t)$ on $\mathbb{R}^{2 \times 2}$ obtained by linearizing the system of differential equations about $\Phi_\ast$. Solve the variational equation for the initial condition $M(0) = \text{Id}_{2 \times 2}$ and use this result to estimate $\Phi(t,(1.1,0.1))$ and $\Phi(t,(0.9,0.1))$ (you may have to resort to numerical integration).

**Exercise 3.9** For a fixed control $u^\ast$ and fixed initial condition, write down the variational equation for the linear system $\dot{x} = Ax + Bu^\ast$. Discuss the usefulness of the variational equation for linear differential equations, and speculate about what this means for the relation to local and global controllability of a linear system?

In terms of the differential of the flow we now may define the Lie derivative of a vector field. For comparison, we also define the Lie derivative of a smooth function in the same language:

\begin{align}
(X\varphi)(p) &= (L_X\varphi)(p) = \lim_{h \to 0} (\varphi(\Phi(h,p)) - \varphi(p)) \\
(L_X g)(p) &= \lim_{h \to 0} (g(p) - \Phi_\ast(h,p)(g(\Phi(-h,p))))
\end{align}

**Exercise 3.10** Give an example – e.g. of two vector fields $f$ and $g$ in the plane – to show that the Lie derivative $L_f g$ in general is NOT equal to the column vector $(L_f g_1, L_f g_2, \ldots L_f g_n)^T$ of directional derivatives of the components of $g = (g_1, g_2, \ldots, g_n)$.

Recall from APM 581 – here old type-up from control class kept for cross-references.

**Theorem 3.4 (Frobenius)** Suppose that $\Delta$ is a smooth $m$-dimensional regular distribution defined on a neighborhood $V$ of $p$. Then there exists an $m$-dimensional integral manifolds of $\Delta$ through every point $q \in V$ if and only if $\Delta$ is involutive on $V$.

We sketch the key ideas of a standard proof – for the detailed geometric version on manifolds see [Spivak vol.I]. The only if part is a direct consequence of the equality of mixed partial derivatives of smooth functions. Thus suppose that $\Delta$ is a smooth $m$-dimensional regular distribution defined on a neighborhood $V$ of $p \in \mathbb{R}^n$. After possibly reordering the coordinates and shrinking $V$, we may, without loss of generality, assume that there exist smooth vector fields $f_1, \ldots, f_m$ defined on $V$ that span $\Delta$ and which have the form $f_k = \frac{\partial}{\partial x_k} + \sum_{j=m+1}^n \beta_{kj}(x) \frac{\partial}{\partial x_j}$ for suitable smooth functions $\beta_{kj}$ defined on $V$.

**Exercise 3.11** Justify why it is possible to assume that $\Delta$ is spanned by such fields near $p$.

One immediately calculates that for any pair $i,j \leq m$, the Lie bracket of the vector fields $f_i$ and $f_j$ is of the form

\begin{equation}
[f_i, f_j] = \sum_{\ell=m+1}^n \alpha_{ij}^\ell(x) \frac{\partial}{\partial x_\ell}
\end{equation}
for suitable smooth functions $\alpha_{ij}: V \mapsto \mathbb{R}$. On the other hand, by the assumption of involutivity, there exist smooth functions $\tilde{c}_{ij}: V \mapsto \mathbb{R}$ such that

$$[f_i, f_j] = \sum_{\ell=1}^{m} c_{ij}^{\ell}(x) f_\ell = \sum_{\ell=1}^{m} c_{ij}^{\ell}(x) \frac{\partial}{\partial x_{\ell}} + \sum_{j=m+1}^{n} \tilde{c}_{ij}^{\ell}(x) \frac{\partial}{\partial x_{\ell}}$$

(26)

for suitable smooth functions $\tilde{c}_{ij}: V \mapsto \mathbb{R}$. Comparison of (25) with (26) establishes that for all $i, j, \ell \leq m$, $c_{ij}^{\ell} \equiv 0$. Consequently the vector fields $f_i$, $i \leq m$ commute pairwise and an integral manifold of $\Delta$ may be locally parameterized by the map $\gamma: (-\varepsilon, \varepsilon)^{m} \mapsto \mathbb{R}^{n}$ (for sufficiently small $\varepsilon > 0$) defined by

$$\gamma(t_1, \ldots, t_m) = e^{t_m f_m} \circ \ldots \circ e^{t_2 f_2} \circ e^{t_1 f_1}(p)$$

(27)

For such map it is always true that

$$\frac{\partial}{\partial t_n} \gamma(t) = f_n(e^{t_m f_m} \circ \ldots \circ e^{t_2 f_2} \circ e^{t_1 f_1}(p)) = f_n(\gamma(t))$$

(28)

but in general

$$\frac{\partial}{\partial t_j} \gamma(t) = (e^{t_m f_m})_{\ast} \circ \ldots \circ (e^{t_{j+1} f_{j+1}})_{\ast} \left( f_j(e^{t_j f_j} \circ \ldots \circ e^{t_2 f_2} \circ e^{t_1 f_1}(p)) \right) \neq f_j(\gamma(t))$$

(29)

for $j < m$. However, as argued above the vector fields $f_i$ and $f_j$ commute pairwise, and hence so do their local flows. Consequently, the last inequality in (29) actually is an equality, and hence the tangent space of the parameterized surface $\gamma$ at $\gamma(t) \in V$ is indeed equal to $\Delta(\gamma(t))$, thus establishing the theorem.

**Exercise 3.12** Consider the smooth distribution $\Delta: \mathbb{R}^3 \mapsto T\mathbb{R}^3$ spanned by the two vector fields $X = z \frac{\partial}{\partial y} - y \frac{\partial}{\partial z}$ and $Y = y \frac{\partial}{\partial z} - z \frac{\partial}{\partial y}$. Find the involutive closure of $\Delta$, that is, the distribution of smallest dimension that contains $\Delta$ and which is involutive. Suggestion: Calculate $Z = [X, Y]$ and analyze $[X, Z]$ and $[Y, Z]$. Then find all integral manifolds of this involutive closure.

### 3.6.2 Chow’s theorem and the Lie algebra rank condition

The previous subsection established one direction of a nonlinear analogue of the Kalman rank condition – that the reachable sets at any time are always contained in a subspace or submanifold (integral manifold) of dimension equal to the rank of some compound matrix. For the other direction, we very much relied on linearity, using a neat argument related to the controllability Gramian. There is little hope for any nonlinear analogue for this. Still we expect that at any time $T > 0$ the reachable sets to have relatively open interior in the integral manifolds described in the previous section. This is indeed true, but the straightforward argument quickly becomes unappealingly messy.

Intuitively, one expects that since (by definition!) one can approximate the flow $e^{t[f,g]}$ of the Lie bracket of two vector fields $f$ and $g$ by a concatenation of flows $e^{t_1 f}$ and $e^{t_2 g}$ in a suitable order, one also can approximate the flows of higher order iterated Lie brackets of $f$ and $g$ by, albeit more complicated, concatenations of flows $e^{t_1 f}$ and $e^{t_2 g}$. Assume that certain iterated Lie brackets $f^{\pi_j}$ of the system vector fields span the tangent space at $p$, and hence in a neighborhood of $p$. It is easy to show that then the map $\gamma(-\varepsilon, \varepsilon)^{n} \mapsto \mathbb{R}^{n}$ defined by

$$\gamma(t_1, \ldots, t_m) = e^{t_m f_m} \circ \ldots \circ e^{t_2 f_2} \circ e^{t_1 f_1}(p)$$

(30)
will be a diffeomorphism (and in particular, a bijection) onto its image for sufficiently small \( \varepsilon > 0 \). Combining this with some scheme of approximating the flows \( e^{tf_j} \) by concatenations of the flows of the original vector fields should yield the desired proof – but it is far from an appealing strategy.

Our textbook [E.D. Sontag, ] gives an elegant (but less constructive) argument, with key intermediate results summarized in lemma 4.28. The first idea is to define \( k \) as the largest integer for which there exists some \( k \)-tuple of vector fields \( g_i = \sum_{j=1}^{m} u_{ij} f_j \) (linear combinations of the systems vector fields, or elements of the distribution spanned by the fields \( f_j \)) for which the derivative of the map

\[
\sigma(t_1, \ldots, t_k) = e^{t_k g_k} \circ \cdots \circ e^{t_2 g_2} \circ e^{t_1 g_1}(p) \tag{31}
\]

has maximal rank \( k \). Arguing indirectly and using the involutivity one argues that each of the vector fields \( g_i \) must be tangent to the image of \( \sigma \). On the other hand, the tangent space of the image is an involutive distribution that must contain all the vector fields \( f_i \), and hence must also contain the Lie algebra generated by the \( f_i \). Consequently, the (relative) interiors of the sets reachable by concatenating flows of vector fields in the original distribution, and reachable by concatenating flows of vector fields in its involutive closure agree. For details please see the original argument in [E.D. Sontag, , pp.152-154].

In conclusion we have a nonlinear analogue of the Kalman Rank Condition for linear controllability. Recall that the system (7) is called accessible from \( p \) if for all possible times \( T > 0 \) the reachable set \( R_T(p) \) has nonempty (\( n \)-dimensional) interior.

**Theorem 3.5** The system (7) with analytic vector fields \( f_0, f_1, \ldots, f_m \) is accessible from \( p \) if and only if the Lie algebra rank condition (LARC) \( \dim L(F_0, f_1, \ldots, f_m)(p) = n \) is satisfied.

There are many fine points about this theorem. Clearly the necessity is only true under the assumption of analytic vector fields. (Just recall the Taylor expansion of the analytic function \( \phi: \mathbb{R} \to \mathbb{R} \) defined by \( \phi(0) = 0 \) and \( \phi(x) = \exp(-x^{-2}) \).) Moreover, it makes a difference whether one assumes that the LARC is satisfied at one point only, or in an open neighborhood of the initial point. Theorems covering the different settings are known under the names of Chow’s theorem and the Nagano/Sussmann theorem. See the bibliography in E.D. Sontag, for details.

In general, accessibility is weaker than even local controllability as is obvious in the simple example

\[
\begin{align*}
\dot{x}_1 &= u \\
\dot{x}_2 &= x_1^k \quad |u(t)| \leq 1
\end{align*} \tag{32}
\]

Here, \( f_0(x) = x_1^k \frac{\partial}{\partial x_2} \), \( f_1 = \frac{\partial}{\partial x_1} \), and up to the obvious relations in any Lie algebra, the only iterated Lie bracket that does not vanish at \( x(0) = p = 0 \) is \( \text{ad}^k f_1, f_0 \equiv (-)^k k! \frac{\partial}{\partial x_2} \). Thus, by the LARC, for any integer \( k \geq 1 \), this system is accessible from \( x(0) = 0 \). It is also clear that if \( k \) is an even integer, then for any control \( u \) and any \( T \) one always has that \( x_2(\Phi(T, u, (0, 0))) \geq 0 \) and hence \( x(0) \notin \text{int} R_T(p) \) for any \( T \geq 0 \), i.e. the system is not locally controllable from \( x(0) = 0 \).

On the other hand, if \( k \) is an odd integer, then one readily sees that for all \( T \geq 0 \), for all admissible controls \( u \) (and \( p = x(0) = (0, 0) \))

\[
\Phi(T, p, -u) = - \Phi(T, p, u) \tag{33}
\]
In this case, a simple calculation shows that for \( k \) an odd integer the system is indeed small-time locally controllable (STLC). Yet the observation made here is leads much further. The basic idea is that if a system is accessible, and for every control \( u : [0,t] \rightarrow U \) there exists a control \( u^- : [0,t] \rightarrow U \) such that \( \Phi(T,p,u^-) = -\Phi(T,p,u) \) then the system is STLC. In particular, if the system (7) has no drift term, i.e. \( f_0 \equiv 0 \) and the set \( U \) of admissible control values is symmetric about the origin (i.e. if \( u \in U \) then also \( -u \in U \)), then accessibility implies small-time local controllability. The key technical arguments rely on the convexity of high-order approximating cones (of high-order variational vectors) of the reachable set / endpoint map. This class may address some of these issues later in the context of optimal control.

We just mention one positive result which is very useful in practice, and which seems to be very suggestive. To facilitate the statement define the following subspaces of the Lie algebra generated by the system vector fields. \( S^1 = \text{span}_\mathbb{R}\{(\text{ad}^k f_0, f_1) : k \in \mathbb{Z}_0^+\} \), and inductively for \( k \geq 1 \), \( S^{k+1} = [S^1, S^k] = \text{span}_\mathbb{R}\{[v, w] : v \in S^1, w \in S^k\} \). With this terminology, we state the following sufficient condition for STLC of analytic single-input systems (\( m = 1 \))

**Theorem 3.6 (Hermes condition)** An analytic single-input systems of form (7) is STLC about \( p \in \mathbb{R}^n \) if it is accessible from \( p \in \mathbb{R}^n \) and if for every \( k \geq 0 \), \( S^{2k}(p) \subseteq S^{2k-1}(p) \).

For systems in the plane, i.e. \( n = 2 \), this theorem was first proved by Hermes in the early 1980s, and by Sussmann for general \( n \geq 2 \) a few years later. Several more general conditions (both necessary ones, and sufficient ones) in a similar spirit have been established (including generalizations of the Hermes condition to more than one input). However, conditions for STLC which are both necessary and sufficient, and at the same time are as algorithmic, algebraically computable have been elusive. Indeed, a modern formulation of the big question is whether (small-time) local controllability is finitely determined [Agrachev, open problems book, vol.I].

Roughly speaking, suppose that one establishes accessibility of a system of form (7) on an \( n \)-dimensional manifold by computing iterated Lie brackets of the vector fields \( f_i \) of length at most \( N \). The question is, whether there exist a global bound \( M \), depending on \( n \) and \( N \), such that STLC is determined by the values of all iterated Lie brackets of the vector fields \( f_i \) of length at most \( M \).

This questions naturally leads into the next topic, nilpotent approximating systems, which in some sense are systems which preserve iterated Lie brackets (at a point) up to a certain order, and typically set all higher order terms equal to zero.

### 3.7 Nilpotent approximating systems

It is routine to calculate solutions of linear dynamical systems using simple explicit formulas. On the other hand, even if only some amazingly minor nonlinear terms appear in a system of differential equations, there is generally no hope at all for any kind of a closed-form, explicit solution formula. This is nicely illustrated in the simplicity of the classical Lorenz system which exhibits chaotic dynamics.

\[
\begin{align*}
    \dot{x} &= -10x + 10y \\
    \dot{y} &= 28x - y - xz \\
    \dot{z} &= -\frac{8}{3}z + xy
\end{align*}
\]

(34)

Thus it is natural to search for approximating systems which reasonably well locally approximate the system dynamics and preserve key properties of interest, while permitting one to get some sort of solutions without excessive efforts.
Common basic choices in analysis are linearizations, or quadratic or higher order Taylor approximations of functions. In dynamical systems the first basic choices are local linearizations about fixed point or about a reference trajectory (as in the discussions preceding the introduction of the Lie derivative). In the case of affine control systems of the form \( \dot{x} = f_0(x) + \sum_{i=1}^m u_i f_i(x) \) with the origin an equilibrium point of the drift, i.e. \( f_0(0) = 0 \), a natural choice is the Jacobian linearization (in column vector notation)

\[
\dot{x} = Ax + Bu \quad \text{where} \quad A = \frac{\partial f}{\partial x}(0), \quad \text{and} \quad B = [f_1(0), \ldots, f_m(0)]
\] (35)

**Exercise 3.13** Calculate the Jacobian linearization of the parallel parking example (36), and show that this linearization is not controllable.

\[
\begin{align*}
\dot{\phi} &= u_1 \\
\dot{v} &= u_2 \\
\dot{\theta} &= v \tan \phi \quad \text{or} \quad f_0(z) = \begin{pmatrix} 0 \\ 0 \\ z_2 \tan z_1 \end{pmatrix}, \quad f_1(z) = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad f_2(z) = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}
\end{align*}
\] (36)

This example nicely illustrates that linearizations may simply be too crude, even locally, as they may preserve too few properties of interest. The next most simple approximation which comes to mind is to use bilinear systems. These are most useful when working in a neighborhood of a common equilibrium point of the drift \( f \) and of the controlled vector fields \( f_j \). There exists some nice literature available about this approach, which closely interfaces with systems on Lie groups. We shall not further pursue this approach here. The next step after this might be polynomial cascade systems of the form:

\[
\begin{align*}
\dot{x}_1 &= u \\
\dot{x}_2 &= p_2(x_1) \\
\dot{x}_3 &= p_3(x_1, x_2) \\
&\vdots \\
\dot{x}_s &= p_s(x_1, x_2, \ldots, x_{s-1})
\end{align*}
\] (37)

Systems of this form are attractive as for any given control \( u(t) \), i.e. a function of time, every resulting trajectory \( \Phi(t, x_0, u) \) can be computed by successive quadratures only – that is, by integrations of functions of time – integrating \( \dot{y} = f(t) \), no need to solve any (nonlinear) differential equations \( \dot{y} = f_1(y) \). Clearly there are functions of time which may not have nice formulas for their antiderivatives – but this is still much simpler than integrating general nonlinear differential equations such as (34).

Thus polynomial cascade systems satisfy at least one of the objectives – being amenable to analysis and to obtaining formulas for solution curves. In the sequel we will argue that they also preserve properties of interest when employed as suitably constructed approximating systems. They “often” preserve (small-time local) controllability. They preserve accessibility, and often have the same optimal controls. Moreover, stabilizing feedback laws of the approximating systems (when properly constructed) will also locally stabilize the original system.

However, from a geometric perspective this characterization (37) is unappealing as it depends on the choice of coordinates – even a simple linear coordinate change such as a rotation may completely mess up this structure. On the other end, it might be possible to bring a complicated-looking system into such nice form by a simple change of coordinates. A beautiful geometric
criterion that characterizes systems of this form, up to a change of coordinates are the following
(the definition of nilpotent comes after the theorem).

**Theorem 3.7** If a control system \( \dot{x} = f_0(x) + \sum_{i=1}^{m} u_i f_i(x) \) is in polynomial cascade form, the Lie algebra \( L(f_0, f_1, \ldots, f_m) \) is nilpotent.

**Theorem 3.8** (MK 1986) If a control system \( \dot{x} = f_0(x) + \sum_{i=1}^{m} u_i f_i(x) \) with analytic vector fields \( f_0 \) and \( f_i \) is such that the Lie algebra \( L(f_0, f_1, \ldots, f_m) \) is nilpotent, and it satisfies the Lie algebra rank condition (3.5) for accessibility at a point \( p \), then there exist local coordinates about \( p \) such that in these coordinates the system takes the form (37).

One actually can make stronger statements about the degrees of the polynomials in terms of the dimensions of certain subalgebras of \( L(f, f_1, \ldots f_m) \).

The proof of the first is an immediate consequence of some of the later discussions in this section. The proof of the second theorem is a fairly straightforward construction that obtains the new coordinates in terms of solutions of \( n \) differential equations. In some sense the argument is similar to proving that every analytic system of differential equations – including chaotic systems such as (34) – near every nonsingular point is equivalent – upon choosing suitable coordinates – to the system \( \dot{x}_1 = 1, \dot{x}_2 = 0, \ldots, \dot{x}_n = 0 \). Depending on one’s point of view, this may be considered an explicit construction, or useless in practice.

Nilpotency in the Lie algebra setting is closely related to the corresponding notions in linear algebra and groups. Recall that a matrix \( A \) is nilpotent if there exists an integer \( N \) such that \( A^N = 0 \). In the Lie algebra setting we define two sequences of subalgebras

\[
L^{(1)} = L, \quad \forall k \geq 1, \quad L^{(k+1)} = [L, L^{(k)}] = \text{span}_\mathbb{R}\{[v, w] : v, w \in L^{(k)}\} \tag{38}
\]

\[
L^{(0)} = L, \quad \forall k \geq 0, \quad L^{(k+1)} = [L^{(k)}, L^{(k)}] = \text{span}_\mathbb{R}\{[v, w] : v, w \in L^{(k)}\} \tag{39}
\]

The sequence \( \{L^{(k)}\}_{k=1}^\infty \) is called the **central descending series** of \( L \), and the sequence \( \{L^{(k)}\}_{k=1}^\infty \) is called the **derived series** of \( L \). A Lie algebra \( L \) is called **nilpotent**, or respectively **solvable**, if there exists an integer \( N \) such that \( L^{(N)} = \{0\} \), or \( L^{(N)} = \{0\} \), respectively.

**Exercise 3.14** Show that every nilpotent Lie algebra is solvable, and that if \( L \) is a solvable Lie algebra, then the first derived ideal \( L^{(1)} \) is a nilpotent Lie algebra.
In the following we shall exhibit an algorithmic procedure on how to construct nilpotent approximating systems. This procedure was independently presented by Hermes and Bella"ıche in the mid 1980s and early 1990s, but also known and used by Stefani, Bressan, Sussmann and others. A nice summary appeared in the 1991 SIAM Review [Hermes]. Suppose we are given analytic vector fields $f_0$ and $f_1$ on $\mathbb{R}^n$ which satisfy the Lie algebra rank condition (3.5) at the point $p \in \mathbb{R}^n$, i.e. $\dim L(f_0, f_1, \ldots, f_m)(p) = n$. We shall construct a new set of local coordinates about $p$ (that does not require any solution of differential equations), and new vector fields $\tilde{f}_0, \tilde{f}_1, \ldots, \tilde{f}_m$ such that in the new coordinates $p$ is the origin and the new system is of polynomial cascade form [37].

We first need to make a short technical digression related to notion such as the number of factors in a product or the first factor of a product. For illustration consider that $3 \cdot 4 = 12 = 2 \cdot 6 = 2 \cdot 3 \cdot 2$. Since all of these are equal, it does not make sense to talk about either the number of factors in $3 \cdot 4$ or the first factor of $3 \cdot 4$. On the other hand, one may consider formal products, that is finite sequences such as $(3, 4)$, $(2, 6)$, or $(2, 3, 2)$. For each of these there is a well defined length of the sequence and a well defined first term etc. In the Lie algebra setting, consider the vector fields $f(x, y) = \frac{\partial}{\partial x}$ and $g(x, y) = e^y \frac{\partial}{\partial y}$. Then $[f, g] = g = f = [f, [f, g]]$ and thus it does make sense to talk about the number of factors in the iterated Lie bracket $[f, [f, g]]$.

Instead we shall start with the set of binary labeled trees or the set parenthesized words (also called the free magma in [Bourbaki] generated by a set of pairwise distinct indeterminates $X_0, X_1, \ldots, X_m$. For each formal expression such as $(X_0, (X_1, X_0)), (X_0, X_1)$) we do have a well defined length, and a well-defines left (first) factor etc. For any set of analytic vector fields $f_0, f_1, \ldots, f_n$ there exists a natural evaluation map $\psi$ which sends each formal bracket (parenthesized word or labeled tree) $X^\sigma$ into the corresponding iterated Lie bracket $f^\pi$, by substituting $f_j$ for each $X_j$ and mapping parenthesized pairs to Lie brackets of vector fields. All of this can me made utterly precise, but that is not the thrust of our studies.

The first choice in the construction is to select weights $w(X_0) = \theta \in [0, 1]$ and, for simplicity here for all $i \geq 1$, $w(X_i) = 1$. Different choices of weights may yield different approximating systems which may suit different purposes. In many applications one considers only weights $\theta \in \{0, 1\}$. However, the endpoint value $\theta = 0$ may not yield nilpotent, only solvable approximating systems. Alternatively, one may consider arbitrarily small positive weights $\theta > 0$ that yield almost the same approximating systems, but with nilpotent Lie algebra. Extend the weight assignment to all formal brackets by recursively defining $w((X_\sigma, X^\pi) = w(X_\sigma) + w(X^\pi)$. For example, if $\theta = \frac{1}{2}$, then $w((X_1, X_0), ((X_1, X_0), X_1)) = 4$.

Since the system satisfies the Lie algebra rank condition (3.5), there exist formal brackets $X^{\pi_1}, X^{\pi_2}, \ldots, X^{\pi_n}$ such that $f^{\pi_1}(p), f^{\pi_2}(p), \ldots f^{\pi_n}(p) = 0$ such that are linearly independent. Among all possible such formal brackets, choose those such that the sequence of weights

$$r = (r_1, r_2, \ldots, r_n) = (w(X^{\pi_1}), w(X^{\pi_2}), \ldots, w(X^{\pi_n}))$$

(40)

is the smallest possible in lexicographical order. This step amounts to no more than calculating Lie brackets of vector fields, i.e. in coordinates computing Jacobian matrices of partial derivatives, matrix multiplication, and function evaluation.

Next perform a constant linear (affine) coordinate change such that $x(p) = 0$ and in the new coordinates for $1 \leq i \leq n$, $f^{\pi_i}(0) = \frac{\partial}{\partial x_i} \bigg|_p$. This step only involves finding the inverse of a constant nonsingular matrix and some matrix multiplication.
To facilitate the arguments introduce the notion of a group of dilations, that is a parameterized family of maps \( \Delta : \mathbb{R}^+ \times \mathbb{R}^n \to \mathbb{R}^n \) defined by
\[
\Delta_{r}(x) = (r_1 x_1, r_2 x_2, \ldots, r_n x_n)
\]
where \( r_i \in \mathbb{R}^+ \) are fixed exponents. In many applications one considers \( r_i \in \mathbb{Z}^+ \) and customarily labels that the coordinates and exponents such that \( 1 = r_1 \leq r_2 \leq \ldots \leq r_n \), but we shall allow \( r_i \in \mathbb{Q}^+ \) in the following. In the construction of the approximating system these exponents are the weights \( r_i = \nu(X^{n_i}) \). For a coordinate-free geometric version see [Kawski 1995].

A family of dilation induces graded structures on the vector spaces of polynomials and polynomial vector fields on \( \mathbb{R}^n \). In the spaces \( \mathcal{P} \) of all polynomial functions on \( \mathbb{R}^n \) and \( \mathcal{V} \) of all polynomial vector fields on \( \mathbb{R}^n \) define for \( m, k \in \mathbb{R} \) define the subspaces of \( \Delta \)-homogeneous polynomial functions and polynomial vector fields by
\[
\begin{align*}
H_m &= \{ p \in \mathcal{P} : \forall \varepsilon > 0, \ p \circ \Delta_{r} = \varepsilon^m \cdot p \} \\
\mathfrak{n}_k &= \{ f \in \mathcal{V} : \forall m \in \mathbb{Q}, \ \forall p \in H_m, \ \hat{f}p \in H_{m+k} \}
\end{align*}
\]
It is easy to see that e.g. the coordinate functions \( x_i \) and partial derivatives \( \frac{\partial}{\partial x_i} \) are homogeneous of degrees \( r_i \) and \( -r_i \), respectively.

**Lemma 3.9** Let \( \Delta \) be a family of dilations as above with exponents \( 1 = r_1 \leq \ldots \leq r_n \). Then a polynomial vector field \( f = p_1 \frac{\partial}{\partial x_1} + \ldots + p_n \frac{\partial}{\partial x_n} \in \mathfrak{n}_m \) if and only if for all \( 1 \leq i \leq n, \ p_i \in H_{m+r_i} \). Moreover, for all \( k < -r_n, \mathfrak{n}_k = \{0\} \), and for all \( m, k \in \mathbb{Q} \),
- If \( m \neq k \) then \( H_m \cap H_k = \{0\} \) and \( \mathfrak{n}_m \cap \mathfrak{n}_k = \{0\} \)
- \( H_m \cdot \mathfrak{n}_k \subseteq H_{m+k} \)
- \( \mathfrak{n}_m \cdot \mathfrak{n}_k \subseteq \mathfrak{n}_{m+k} \)
- \( H_m \cdot \mathfrak{n}_k \subseteq \mathfrak{n}_{m+k} \)

**Exercise 3.15** Show that each \( H_m \) and each \( \mathfrak{n}_k \) is indeed a subspace of \( \mathcal{P} \) or \( \mathcal{V} \), respectively. Prove the statements of lemma 3.9

Returning to the construction of the nilpotent approximating system, expand each component \( f_{ij}(x) = (f_j(x))(x_j) = (dx_j, f_i) \) (of the original vector fields in the new adapted coordinates) in a Taylor series about \( x = 0 \). Truncate the series at terms of order at most \( r_j \) for \( f_0 \) and at terms of order at most \( r_j - 1 \) for \( f_i, i \geq 1 \). Denote the resulting vector fields by \( \hat{f}_0, \hat{f}_1, \ldots, \hat{f}_m \).

**Exercise 3.16** Show that these vector fields \( \hat{f}_0, \hat{f}_1, \ldots, \hat{f}_m \) generate a nilpotent Lie algebra, and that they approximate the original systems in the sense that they preserve accessibility.

In general, the resulting vector fields \( \hat{f}_i \) are not homogeneous with respect to the dilation \( \Delta \), but they already have a desirable cascade polynomial structure. In general one has:
\[
\hat{f}_0 \in \bigoplus_{k \leq -\theta} \mathfrak{n}_k \quad \text{and for } i \geq 1, \quad \hat{f}_i \in \bigoplus_{k \leq -1} \mathfrak{n}_k.
\]

Technically, similar to the graded structures of \( \mathcal{P} \) and \( \mathcal{V} \) defined by \( H_m \) and \( \mathfrak{n}_k \), one considers the filtrations of \( \mathcal{P} \) and \( \mathcal{V} \) defined by the subspaces \( P_m = \bigoplus_{k \leq m} H_k \) and \( N_m = \bigoplus_{k \leq m} \mathfrak{n}_k \). These have similar properties to those listed in lemma 3.9. The main difference is that now if \( k \leq m \)
then \( P_m \cap P_k = P_k \) and \( P_m \cup P_k = P_m \), etc. However, for later arguments, especially for feedback stabilization it is convenient to work with truly homogeneous vector fields, and it is well worth to make the extra effort to fully adapt the coordinates. This may be achieved via a simple further coordinate change of triangular polynomial form which has a triangular polynomial inverse. This step is completely algorithmic, involving only differentiation, function evaluation at a point and matrix multiplication. The final result will be homogeneous vector fields \( \tilde{f}_j \) that do not exhibit geometrically irrelevant ghost terms. Not only are the resulting formulas simpler and nicer, this form also opens the door to using powerful results that relate the stability of systems nonlinear systems to that of homogeneous approximating systems.

For illustration, first consider the “desired” homogeneous system and two triangular polynomial coordinate changes \( y = \Phi(z) \) and \( x = \Psi(y) \)

\[
\begin{align*}
\dot{z}_1 &= = u \\
\dot{z}_2 &= = z_1 \\
\dot{z}_3 &= = z_1^{10} \\
\dot{z}_4 &= = z_1^{100}
\end{align*}
\begin{align*}
y_1 &= = z_1 \\
y_2 &= = z_2 \\
y_3 &= = z_3 \\
y_4 &= = z_4 + z_3^5
\end{align*}
\begin{align*}
x_1 &= = y_1 \\
x_2 &= = y_2 \\
x_3 &= = y_3 + y_2^3 \\
x_4 &= = y_4 + y_2^4
\end{align*}
\]

(44)

Exercise 3.17 Transform the given clean system in the \( z \)-coordinates to the equivalent system in the mildly mixed-up \( x \)-coordinates. Explicitly write out the differentials \( \Phi_* \) and \( \Psi_* \), and the inverses \( \Phi^{-1}(y) \) and \( \Psi^{-1}(x) \). Demonstrate that when evaluating the iterated Lie brackets of the system vector fields at the origin, the added off-diagonal terms do not show up at all.

It turns out to be relatively straightforward to eliminate such ghost-terms in the vector fields which have no geometric role for controllability etc, yet they potentially destroy the homogeneity properties of the vector fields, thereby preventing one from a battery of nice theorems for e.g. stability and stabilization.

Exercise 3.18 (Continuation of exercise 3.17) Starting with the system in the \( x \)-coordinates, reconstruct the coordinate changes by successively performing transformations of triangular form \( x^{(k+1)} = \Phi^{(k)}(x^{(k)}) \) that may also be defined geometrically by expressions of the form:

\[
x_k^{(i+1)} = x_k^{(i)} - \sum_j \frac{(x_i^{(i)})^j}{j!} \cdot L_j^{x_i} \bigg|_{x=0} (x_k^{(i)}) \quad \text{xxx Check this again xxx}
\]

(45)

where the sum ranges over all \( j \) such that \( jr_i < r_k - 1 \) or \( jr_i < r_k - \theta \). Use the illustrating example above to decide in which order to make the corrections (i.e. for \( i \) increasing or decreasing). Demonstrate that in the new “adapted” coordinates the truncated vector fields are homogeneous w.r.t. \( \Delta \).

We summarize the result in the following theorem, where we again use \( x = (x_1, \ldots, x_n) \) to now denote the final adapted coordinates obtained after at most \( n \) successive steps as outlined in exercise 3.18 and we denote the resulting vector fields in the new coordinates by \( \tilde{f}_0, \tilde{f}_1, \ldots \tilde{f}_m \).

Theorem 3.10 The vector fields \( \tilde{f}_0, \tilde{f}_1, \ldots \tilde{f}_m \) constructed above are \( \Delta \)-homogeneous vector fields (with the exponents \( r_i \) and adapted coordinates) such that \( f_0 \in \mathbb{P}_{-\theta} \) and for \( i \geq 1, \tilde{f}_i \in \mathbb{P}_{-1} \).
Theorem 3.11 If \( \theta > 0 \) then the Lie algebra \( L(\tilde{f}_0, \tilde{f}_1, \ldots, \tilde{f}_m) \) by the newly constructed vector fields is nilpotent. If \( \theta = 0 \), the Lie algebra \( L(S^1(\tilde{f}_0, \tilde{f}_1, \ldots, \tilde{f}_m)) \) is nilpotent where

\[
S^1(\tilde{f}_0, \tilde{f}_1, \ldots, \tilde{f}_m)) = \text{span}_\mathbb{R} \{(\text{ad}^j \tilde{f}_0, \tilde{f}_k) : j \geq 0, \ and \ 1 \leq k \leq n\}.
\]

Exercise 3.19 Prove theorem 3.11.

Exercise 3.20 Apply the algorithm described above to construct nilpotent approximating systems for different choices of \( \theta \), e.g. \( \theta \in \{0, \frac{1}{2}, 1\} \) for the model (36) for parallel parking a car. Verify that the approximating systems do indeed still satisfy the Lie algebra rank condition for accessibility.

Exercise 3.21 Show that if one chooses the weight \( \theta = w(X_0) = 0 \), then the approximating vector fields \( \tilde{f}_0, \ldots, \tilde{f}_m \) satisfy the Hermes condition [?] for small-time local controllability if and only if the original vector fields \( f_0, \ldots, f_m \) do.