2 Lie Algebras

In this chapter we fix our conventions and terminology and we provide a quick review of the notions in differential geometry and in Lie theory that will be used. Since algebraic geometry, mainly the geometry of Abelian varieties, will only show up later and since we will need to do in that case a little more than just a review, we defer that subject to Part II of the book.

2.1 Structures on Manifolds

Our manifolds will always be either real smooth or complex holomorphic. In both cases the algebra of functions on such a manifold M will be denoted by $\mathcal{F}(M)$. Thus $\mathcal{F}(M)$ is the algebra of smooth functions on M when M is a real manifold while $\mathcal{F}(M)$ is the algebra of holomorphic functions on M when M is a complex manifold; when $M = \mathbf{C}^n$ or a smooth affine variety we will often restrict ourselves to the polynomial functions on M (usually called *regular functions* on M). Since many of the basic definitions and constructions that are given below are algebraic they apply to complex (algebraic) manifolds as well as real manifolds, and we will just write "Let M be a manifold" when our definition or construction applies to the real as well as to the complex case. Similarly, the word "map" will stand for "smooth map" (resp. "holomorphic map" or "regular map") in the case of smooth manifolds (resp. holomorphic manifolds or (non-singular) algebraic varieties).

2.1.1 Vector Fields and 1-Forms

For a manifold M and a point $m \in M$ the (real or holomorphic) *tangent space* to M at m is denoted by T_mM and its dual space, the *cotangent space* to M at m, is denoted by T^*_mM. The tangent and cotangent spaces to M form the fibers of the *tangent bundle* TM, resp. the *cotangent bundle* T^*M. A *vector field* $\mathcal{V}$ is a section of the tangent bundle while a *1-form* ω is a section of the cotangent bundle; the values of $\mathcal{V}$ and ω at $m \in M$ are simply denoted by $\mathcal{V}(m)$ and $\omega(m)$, where $\mathcal{V}(m) \in T_mM$ and $\omega(m) \in T^*_mM$. The $\mathcal{F}(M)$-modules of vector fields and 1-forms on M will be denoted by $\mathfrak{X}(M)$ and $\Omega(M)$. We will find it convenient to denote the pairing between a vector space and its dual, such as T_mM and T^*_mM, by $\langle \cdot\,, \cdot \rangle$.

For example, if $\mathcal{V} \in \mathfrak{X}(M)$ and $\omega \in \Omega(M)$ then we may define a function $\omega(\mathcal{V}) \in \mathcal{F}(M)$ by setting

$$\omega(\mathcal{V})(m) := \langle \omega(m), \mathcal{V}(m) \rangle \tag{2.1}$$

for all $m \in M$. To a function $F \in \mathcal{F}(M)$ we may associate its differential $\mathsf{d}F \in \Omega(M)$, which is a 1-form, hence can be applied to vector fields on M. This is used to associate to every vector field $\mathcal{V}$ on M a *derivation* on $\mathcal{F}(M)$: for $F \in \mathcal{F}(M)$ we define $\mathcal{V}[F] \in \mathcal{F}(M)$ by

$$\mathcal{V}[F] := \mathsf{d}F(\mathcal{V}), \tag{2.2}$$

which means in view of (2.1) that

$$\mathcal{V}[F](m) = \langle \mathsf{d}F(m), \mathcal{V}(m) \rangle \tag{2.3}$$

for $m \in M$. Saying that $\mathcal{V}$ is a derivation on $\mathcal{F}(M)$ means that if $F, H \in \mathcal{F}(M)$ then

$$\mathcal{V}[FH] = \mathcal{V}[F]H + F\mathcal{V}[H],$$

an easy consequence of (2.2) and the Leibniz rule for differentials. It follows from (2.3) that $\mathcal{V}[F](m)$ depends on $\mathcal{V}(m)$ (and F) only; it is the derivative of F at m in the direction of $\mathcal{V}(m)$, hence it is legitimate to write it as $\mathcal{V}(m)[F]$. At $m \in M$ the derivation property then reads

$$\mathcal{V}(m)[FH] = (\mathcal{V}(m)[F])H(m) + F(m)(\mathcal{V}(m)[H]);$$

one says that $\mathcal{V}(m)$ defines a *derivation* on $\mathcal{F}(M)$ *at* m.

It is a fundamental fact that, conversely, every derivation on M corresponds to a unique vector field on M and that every derivation at m corresponds to a unique tangent vector at m. As a corollary, since the commutator of two derivations is a derivation we may define the *Lie bracket* $[\mathcal{V}_1, \mathcal{V}_2]$ of $\mathcal{V}_1$, $\mathcal{V}_2 \in \mathfrak{X}(M)$ as the vector field that corresponds to the derivation $\mathcal{V}_1 \circ \mathcal{V}_2 - \mathcal{V}_2 \circ \mathcal{V}_1$. This way, $\mathfrak{X}(M)$ becomes an infinite-dimensional Lie algebra. For $F \in \mathcal{F}(M)$ and for $\mathcal{V}_1, \mathcal{V}_2 \in \mathfrak{X}(M)$ one has

$$[F\mathcal{V}_1, \mathcal{V}_2] = F[\mathcal{V}_1, \mathcal{V}_2] - \mathcal{V}_2[F]\mathcal{V}_1. \tag{2.4}$$

Notice also that if $U \subseteq M$ is a coordinate neighborhood then a derivation on $\mathcal{F}(\mathcal{U})$ is completely determined once its effect on all elements x_i of a coordinate system $(x_1, \dots, x_n)$ on $\mathcal{U}$ is known, where $n := \dim M$. Indeed, since in terms of these coordinates

$$\mathsf{d}F = \sum_{i=1}^{n} \frac{\partial F}{\partial x_i} \mathsf{d}x_i$$

we have in view of (2.2) that

$$\mathcal{V}[F] = \sum_{i=1}^{n} \frac{\partial F}{\partial x_i} \mathcal{V}[x_i]. \tag{2.5}$$

When we are dealing with a fixed vector field $\mathcal{V}$ we often write $\dot{F}$ for $\mathcal{V}[F]$, where $F \in \mathcal{F}(M)$. In this notation, the coordinate expression (2.5) takes the form

$$\dot{F} = \sum_{i=1}^{n} \frac{\partial F}{\partial x_i} \dot{x}_i.$$

There is a one-to-one correspondence between vector fields on the coordinate neighborhood U and differential equations on U of the form

$$\begin{aligned} \frac{\mathsf{d}x_1}{\mathsf{d}t} &= f_1(x_1, \ldots, x_n), \\ &\vdots \\ \frac{\mathsf{d}x_n}{\mathsf{d}t} &= f_n(x_1, \ldots, x_n), \end{aligned} \tag{2.6}$$

where $f_i \in \mathcal{F}(U)$, for $i = 1, \ldots, n$. Indeed, given a vector field $\mathcal{V}$, define the functions f_i by $f_i := \mathcal{V}[x_i]$; given the functions f_i, define $\mathcal{V}[x_i] := f_i$ and extend $\mathcal{V}$ to a derivation on $\mathcal{F}(U)$ by using (2.5). Solutions to (2.6) are easily interpreted as parametrized curves in U, whose tangent vector at each point coincides with the value of $\mathcal{V}$ at that point; we will usually consider solutions that are defined on an open ball B_ϵ around ϵ, where $B_\epsilon := \{t \in \mathbf{C} \mid |t| < \epsilon\}$ in the holomorphic case and $B_\epsilon := \{t \in \mathbf{R} \mid |t| < \epsilon\}$ in the smooth real case. For that reason a solution $x(t) = (x_1(t), \ldots, x_n(t))$ to (2.6), defined on B_ϵ, and such that $x(0) = m$ is often called an *integral curve* of $\mathcal{V}$, starting at m. The well-known uniqueness and existence theorem for differential equations can (in the holomorphic case) be formulated in terms of vector fields and integral curves as follows.

Theorem 2.1 (Picard Theorem for ODE's). *Let $\mathcal{V}$ be a holomorphic vector field on an open subset U of $\mathbf{C}^n$ and let $m \in U$. There exists an integral curve of $\mathcal{V}$, starting at m; this integral curve $x(t; m)$ is unique in the sense that any two integral curves of $\mathcal{V}$ that start at m coincide on the intersection of their domains. Moreover, $x(t; m)$ depends in a holomorphic way on m.*

The analogous theorem for smooth vector fields on open subsets of $\mathbf{R}^n$ of course also holds. The theorem and its smooth analog imply that given a vector field $\mathcal{V}$ on an n-dimensional manifold M we can find for any $m \in M$ a coordinate neighborhood U of m, with coordinates $(x_1, \ldots, x_n)$, an open subset $U' \subseteq U$ and an $\epsilon > 0$, such that the solution $x(t; m)$ is defined for $(t, m) \in (B_\epsilon \times U')$. The map

$$\begin{aligned} \Phi : B_\epsilon \times U' &\to \quad U \\ (t, m) &\mapsto \Phi_t(m) := x(t; m) \end{aligned}$$

is called the *flow* of $\mathcal{V}$.

For a fixed $t \in B_\epsilon$ the map $\Phi_t : U' \to U$ is a biholomorphism (diffeomorphism, in the smooth case) from U' to $\Phi_t(U')$. It is customary to pretend that for small $|t|$ the local biholomorphism (diffeomorphism) Φ_t is global if we are in the case of a vector field on a manifold M and to write $\Phi_t : M \to M$, but of course — unless M is compact — the globalness needs not be true. For example, one writes the fundamental property that links directional derivatives to flows in the form

$$\mathcal{V}[F] = \frac{\mathsf{d}}{\mathsf{d}t}_{|t=0} \Phi_t^* F,$$

where $F \in \mathcal{F}(M)$ and $\Phi_t^* F = F \circ \Phi_t$. Theorem 2.1 and its smooth analog lead to the following theorem, that we will use often.

Theorem 2.2 (Straightening Theorem). *Let $\mathcal{V}$ be a vector field on a manifold M of dimension n and suppose that $\mathcal{V}(m) \neq 0$, where $m \in M$. Then there exist coordinates $x_1, \dots, x_n$ on a neighborhood U of m such that the restriction of $\mathcal{V}$ to U is the first coordinate vector field, i.e., $\mathcal{V}[F] = \partial F/\partial x_1$.*

In the same spirit we can, intuitively speaking, parameterize a neighborhood of an analytic hypersurface by local coordinates on the hypersurface on the one hand, and by the parameter which is going with any fixed vector field, on the other hand, assuming that the vector field is transversal to the divisor. Precisely, the following theorem holds (see Figure 2.1).

Theorem 2.3. *Let M be a complex manifold of dimension n and let $\mathcal{V}$ be a holomorphic vector field on M. Suppose that $\mathcal{D}$ is an analytic hypersurface of M and let m_0 be a smooth point of $\mathcal{D}$. If $\mathcal{V}$ is transversal to $\mathcal{D}$ at m_0, then there exist neighborhoods U and V of m_0 in $\mathcal{D}$, resp. in M, and there exists $\epsilon > 0$, such that the restriction of Φ to $B_\epsilon \times U$ is a biholomorphism onto V. In addition, if U is a coordinate neighborhood of m_0 in $\mathcal{D}$, with coordinates $x_2, \dots, x_n$ then V is a coordinate neighborhood of m_0 in M, with holomorphic coordinates $(t, x_2, \dots, x_n)$, where $\mathcal{V} = \frac{\partial}{\partial t}$ (on V). In the latter case,*

$$\mathcal{D} \cap V = \{m \in V \mid t(m) = 0\}.$$

It follows that, under the above transversality assumption, we can write any holomorphic function F on V locally as a series

$$F(t) = t^p (f^{(0)} + f^{(1)} t + \cdots), \tag{2.7}$$

where the coefficients $f^{(0)}, f^{(1)}, \dots$ of the series $F(t)$ are holomorphic functions on a neighborhood of m_0 in $\mathcal{D}$. By analyticity, the series $F(t)$ is actually convergent on an open neighborhood in M of an open dense subset of the irreducible component $\mathcal{D}'$ of $\mathcal{D}$ that contains m_0, and it coincides on this neighborhood with the function F. We call the series (2.7) the *Taylor series* of F with respect to $\mathcal{V}$, starting at $\mathcal{D}'$, and we denote it by $F(t; \mathcal{D}')$.

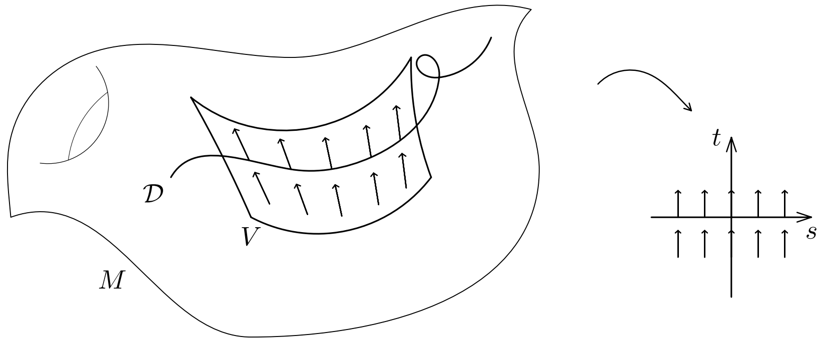


Fig. 2.1. When a holomorphic vector field $\mathcal{V}$ on a complex manifold M is transversal to a divisor $\mathcal{D}$ at $m_0 \in \mathcal{D}$ then a neighborhood V of m_0 in M admits holomorphic coordinates that come from coordinates $s = (s_1, \ldots, s_{n-1})$ on $\mathcal{D}$, plus the time coordinate t that goes with $\mathcal{V}$.

Since $\mathcal{V}$ is transversal to $\mathcal{D}'$ at m_0, the integer p in (2.7) is equal to $\mathrm{ord}_{\mathcal{D}'}(F)$, the order of vanishing of F along $\mathcal{D}'$, if $f^{(0)}$ is not identically zero on $\mathcal{D}'$. The restriction of F to $\mathcal{D}'$ is given by substituting 0 for t in this Taylor series, i.e., by the first coefficient of its Taylor series.

The same can be done for meromorphic functions on U: writing such a function F as the ratio G/H of two holomorphic functions we define the *Laurent series* of F with respect to $\mathcal{V}$, starting at $\mathcal{D}'$, denoted $F(t;\mathcal{D}')$, to be the quotient $G(t;\mathcal{D}')/H(t;\mathcal{D}')$. In this case the series converges for small, non-zero $|t|$. The Laurent series of F is still of the form (2.7), where

$$p = \mathrm{ord}_{\mathcal{D}'}(F) = \mathrm{ord}_{\mathcal{D}'}(G) - \mathrm{ord}_{\mathcal{D}'}(H)$$

is now any integer. Under the assumption that $f^{(0)}$ is not identically zero on $\mathcal{D}'$, it is still true that p is the order of vanishing of F along $\mathcal{D}'$, which in the case of negative p means that F has a pole of order $-p$ along $\mathcal{D}'$.

2.1.2 Distributions and the Frobenius Theorem

Instead of having a vector at every point of a manifold M, as is the case of a vector field on M, one may have a one-dimensional subspace of the tangent space to M, at every point of M. This is what is called a 1-dimensional *distribution* on M; a k-dimensional distribution Δ on M is then the datum of a k-dimensional subspace $\Delta(m)$ of T_mM for every $m \in M$. One says that Δ is smooth (or holomorphic) if there exist for every $m \in M$ smooth (or holomorphic) vector fields $\mathcal{V}_1, \ldots, \mathcal{V}_k$, on a neighborhood U of m, such that

$$\Delta(m) = \mathrm{span}\,\{\mathcal{V}_1(m), \ldots, \mathcal{V}_k(m)\}\,, \qquad \text{for any } m \in U.$$

The notion of an integral curve is easily adapted to the case of a k-dimensional distribution Δ: a k-dimensional connected immersed submanifold M' of M is called an *integral manifold* of Δ if $T_mM' = \Delta(m)$ for any $m \in M'$. In contrast to the case of integral curves, integral manifolds need not exist in general, even locally. One obstruction comes from the following fact: if $\mathcal{V}_1$ and $\mathcal{V}_2$ are two vector fields on M which are tangent to some submanifold M' (such as the candidate integral manifold) then their Lie bracket $[\mathcal{V}_1, \mathcal{V}_2]$ is also tangent to M'. In order to rephrase this in the language of distributions, let us say that a vector field $\mathcal{V}$ on $U \subseteq M$ is adapted to Δ on U if $\mathcal{V}(m) \in \Delta(m)$ for every $m \in U$. In these terms the obstruction reads: if $\mathcal{V}_1$ and $\mathcal{V}_2$ are adapted to Δ on some open subset U then $[\mathcal{V}_1, \mathcal{V}_2]$ is also adapted to Δ on U. One says that Δ is an *integrable distribution* if for any $\mathcal{V}_1$ and $\mathcal{V}_2$ that are adapted to Δ on an open subset U, their commutator $[\mathcal{V}_1, \mathcal{V}_2]$ is also adapted to Δ on U. The Frobenius Theorem says that the above obstruction to the existence of integral manifolds is the only one.

Theorem 2.4 (Frobenius). *Suppose that Δ is a (smooth or holomorphic) k-dimensional distribution on M. If Δ is integrable then there exists through any point $m_0 \in M$ a unique maximal integral manifold for Δ.*

The above version of the Frobenius Theorem is the analogue of Theorem 2.1 for distributions. For a short and elementary proof, which is immediately adapted to the holomorphic case, we refer to [42] or [108]. Another version of the Frobenius Theorem is the following.

Theorem 2.5 (Frobenius). *Under the conditions of Theorem 2.4 coordinates $x_1, \dots, x_n$ can be chosen in a neighborhood U of any point $m_0 \in M$ such that*

$$\Delta(m) = \operatorname{span}\left\{\frac{\partial}{\partial x_1}(m), \dots, \frac{\partial}{\partial x_k}(m)\right\}, \qquad \text{for any } m \in U.$$

In terms of these coordinates the integral manifold of $\Delta_{|U}$ through m_0 is given by the connected component of

$$\{m \in U \mid x_i(m) = x_i(m_0) \text{ for } i = k+1, \dots, m\}$$

that contains m_0.

It is clear that the latter version of the Frobenius Theorem generalizes the Straightening Theorem (Theorem 2.2). It implies that the maximal integral manifolds of an integrable distribution on M form the leaves of a foliation on M.

In applications it is sometimes necessary to consider the more general concept of a *singular distribution*, in which the dimension of the subspace of T_mM may vary with m. This happens for example when one considers the singular distribution associated with the Hamiltonian vector fields on a Poisson manifold, as we will see in Section 3.4. For a good account on singular distributions we refer to [107, Appendix 3].

2.1.3 Differential Forms and Polyvector Fields

We will make frequent use of k-forms and k-vector fields on manifolds and of the operations on and between them. For $k \in \mathbf{N}$ we denote the $\mathcal{F}(M)$-module of k-forms on a manifold M by $\Omega^k(M)$; an element of $\Omega^k(M)$ is by definition a section of $\bigwedge^k T^*M$, in particular $\Omega^0(M) = \mathcal{F}(M)$ and $\Omega^1(M) = \Omega(M)$. We let

$$\Omega^*(M) := \bigoplus_{k=0}^{n} \Omega^k(M),$$

where $n := \dim(M)$. An element of $\Omega^*(M)$ will be called a *differential form*. There is an $\mathcal{F}(M)$-bilinear map

$$\wedge : \Omega^*(M) \times \Omega^*(M) \to \Omega^*(M),$$

which associates to two differential forms ω, ω' their *wedge product* $\omega \wedge \omega'$. This operation makes $\Omega^*(M)$ into a graded associative algebra over $\mathcal{F}(M)$, called the *Grassmann algebra* of M. It is graded commutative, which means that for $\omega \in \Omega^k(M)$ and $\omega' \in \Omega^l(M)$ we have

$$\omega \wedge \omega' = (-1)^{kl} \omega' \wedge \omega.$$

The fact that a 1-form can be evaluated on a vector field to produce an element of $\mathcal{F}(M)$ generalizes in two ways to a k-form ω, where $k \geqslant 1$. We can evaluate ω on k vector fields $\mathcal{V}_1, \dots, \mathcal{V}_k$, giving $\omega(\mathcal{V}_1, \dots, \mathcal{V}_k) \in \mathcal{F}(M)$; from this point of view a k-form is an $\mathcal{F}(M)$-k-linear map on $\mathfrak{X}(M)$ with values in $\mathcal{F}(M)$. Or we can insert one vector field $\mathcal{V}$ as the first argument to ω, yielding a $(k-1)$-form, which is denoted by $\imath_{\mathcal{V}}\omega$; from this point of view a k-form is, for $k \geqslant 1$, an $\mathcal{F}(M)$-linear map $\mathfrak{X}(M) \to \Omega^{k-1}(M)$. Notice that it is from the former point of view natural to define a k-form by prescribing its value on all k-tuples of vector fields on M; however, one still needs to check besides skew-symmetry that the k-form is indeed $\mathcal{F}(M)$-k-linear. It is convenient to extend the above definition of $\imath_{\mathcal{V}}$ to all of $\Omega^*(M)$ by defining $\imath_{\mathcal{V}}\omega = 0$, for all 0-forms, i.e. functions, ω on M.

The differential is a linear map $\mathsf{d} : \Omega^*(M) \to \Omega^*(M)$ which maps k-forms to $(k+1)$-forms according to the following formula:

$$\begin{aligned} \mathsf{d}\omega\,(\mathcal{V}_0, \dots, \mathcal{V}_k) = &\sum_{i=0}^{k} (-1)^i \mathcal{V}_i \left[\omega(\mathcal{V}_0, \dots, \widehat{\mathcal{V}_i}, \dots, \mathcal{V}_k)\right] \\ &+ \sum_{i<j} (-1)^{i+j} \omega\left([\mathcal{V}_i, \mathcal{V}_j], \mathcal{V}_0, \dots, \widehat{\mathcal{V}_i}, \dots, \widehat{\mathcal{V}_j}, \dots, \mathcal{V}_k\right). \end{aligned} \tag{2.8}$$

As we just pointed out one has to verify that the right hand side of this formula is $\mathcal{F}(M)$-$(k+1)$-linear, but that is an easy consequence of (2.4).

We have that $\mathsf{d}\circ\mathsf{d}=0$, which implies that each *exact differential form* (an element of $\Omega^*(M)$ that is in the image of d) is a *closed differential form* (an element $\omega\in\Omega^*(M)$ for which $\mathsf{d}\omega=0$). On a coordinate neighborhood every closed differential form is exact, but this is false for general open subsets of manifolds. The differential of a wedge satisfies the graded Leibniz rule

$$\mathsf{d}(\omega\wedge\omega')=\mathsf{d}\omega\wedge\omega'+(-1)^k\omega\wedge\mathsf{d}\omega',$$

where ω is a k-form and ω' an l-form. The differential is not an $\mathcal{F}(M)$-linear map, as is seen from the following formula: if $F\in\mathcal{F}(M)$ and $\omega\in\Omega^l(M)$ then

$$\mathsf{d}(F\omega)=\mathsf{d}F\wedge\omega+F\mathsf{d}\omega,\qquad F\in\mathcal{F}(M),\ \omega\in\Omega^k(M).$$

We will also consider k-vector fields, mainly in the case $k=2,3$, in which cases we speak of a *bivector field* or a *trivector field*. A k-vector field is by definition a section of $\bigwedge^k TM$, hence we can evaluate any k-form ω on any k-vector field P: for P a k-vector field of the form $P=\mathcal{V}_1\wedge\ldots\wedge\mathcal{V}_k$, we let $\omega(P)=\omega(\mathcal{V}_1\wedge\ldots\wedge\mathcal{V}_k):=\omega(\mathcal{V}_1,\ldots,\mathcal{V}_k)$. Since vector fields correspond to derivations we have that a 2-vector field corresponds to a skew-symmetric *biderivation*, a 3-vector field corresponds to a skew-symmetric *triderivation*, and so on. Namely, if P is a k-vector field the value of the corresponding skew-symmetric k-derivation on k functions $F_1,\ldots,F_k\in\mathcal{F}(M)$ is denoted[1], resp. defined by

$$P[F_1\wedge\cdots\wedge F_k]:=(\mathsf{d}F_1\wedge\cdots\wedge\mathsf{d}F_k)(P) \tag{2.9}$$

and we have that P is completely specified on a coordinate neighborhood U once it is known on all k-tuples $(x_{i_1},\ldots,x_{i_k})$, with $1\leqslant i_1<i_2\cdots<i_k\leqslant n=\dim M$, where the k-tuples are taken from any chosen system of coordinates $(x_1,\ldots,x_n)$ on U. Explicitly, (2.5) admits the following generalization to arbitrary k-vector fields,

$$P[F_1\wedge\ldots\wedge F_k]=\sum_{i_1,\ldots,i_k=1}^{n}\frac{\partial F_1}{\partial x_{i_1}}\cdots\frac{\partial F_k}{\partial x_{i_k}}P[x_{i_1}\wedge\ldots\wedge x_{i_k}].$$

We denote the $\mathcal{F}(M)$-module of k-vector fields by $\mathfrak{X}^k(M)$, in particular $\mathfrak{X}^0(M)=\mathcal{F}(M)$ and $\mathfrak{X}^1(M)=\mathfrak{X}(M)$, and we let

$$\mathfrak{X}^*(M):=\bigoplus_{k=0}^{n}\mathfrak{X}^k(M),$$

where $n:=\dim(M)$. An element of $\mathfrak{X}^*(M)$ is called a *polyvector field*. We can define, as in the case of differential forms, a wedge product

$$\wedge:\mathfrak{X}^*(M)\times\mathfrak{X}^*(M)\to\mathfrak{X}^*(M),$$

which makes $\mathfrak{X}^*(M)$ into a graded associative algebra which is the covariant analogue of the Grassmann algebra.

[1] We use $P[F_1\wedge\ldots\wedge F_k]$ instead of $P[F_1,\ldots,F_k]$ to avoid confusion with the notation for Lie brackets. Moreover, this notation makes sense in view of (2.9).

2.1.4 Lie Derivatives

The most important operation on k-forms and on k-vector fields is the Lie derivative. For $\mathcal{V} \in \mathfrak{X}(M)$ and $\omega \in \Omega^k(M)$ we denote the *Lie derivative* of ω in the direction of $\mathcal{V}$ by $L_\mathcal{V}\omega$. For $F \in \mathcal{F}(M)$ and $\mathcal{V} \in \mathfrak{X}(M)$ the Lie derivative is given by $L_\mathcal{V}F := \mathcal{V}[F] = \mathsf{d}F(\mathcal{V})$, while for an arbitrary k-form ω (with $k > 0$) its Lie derivative $L_\mathcal{V}\omega$ is the k-form whose value on $\mathcal{V}_1 \dots, \mathcal{V}_k \in \mathfrak{X}(M)$ is given by

$$L_\mathcal{V}\omega(\mathcal{V}_1, \dots, \mathcal{V}_k) := \mathcal{V}[\omega(\mathcal{V}_1, \dots, \mathcal{V}_k)] - \sum_{i=1}^{k} \omega(\mathcal{V}_1, \dots, [\mathcal{V}, \mathcal{V}_i], \dots, \mathcal{V}_k); \quad (2.10)$$

again one checks that the right hand side of this formula is $\mathcal{F}(M)$-k-linear, so that $L_\mathcal{V}\omega$ is indeed a k-form.

The Lie derivative $L_\mathcal{V}\omega$ measures how ω changes in the direction of $\mathcal{V}$, hence $L_\mathcal{V}\omega = 0$ if and only if ω is constant on the integral curves of $\mathcal{V}$. This follows immediately from the following alternative (geometric!) definition,

$$L_\mathcal{V}\omega = \frac{\mathsf{d}}{\mathsf{d}t}_{|t=0} \Phi_t^* \omega,$$

where Φ denotes the (local) flow of $\mathcal{V}$ on M. The most useful expression for $L_\mathcal{V}\omega$ is given by the following formula, known as *Cartan's Formula*

$$L_\mathcal{V}\omega = \mathsf{d}\imath_\mathcal{V}\omega + \imath_\mathcal{V}\mathsf{d}\omega; \quad (2.11)$$

for example, Cartan's formula implies at once that $L_\mathcal{V}\omega$ is $\mathcal{F}(M)$-k-linear, i.e., that it is a k-form. We will also need the formula

$$\imath_{[\mathcal{V}_1,\mathcal{V}_2]}\omega = L_{\mathcal{V}_1}\imath_{\mathcal{V}_2}\omega - \imath_{\mathcal{V}_2}L_{\mathcal{V}_1}\omega, \quad (2.12)$$

which is valid for $\mathcal{V}_1$, $\mathcal{V}_2 \in \mathfrak{X}(M)$ and $\omega \in \Omega^*(M)$. Let us prove (2.12) when ω is a two-form (this is the only case which will be used), in which case both sides of (2.12) are one-forms. For any vector field $\mathcal{V}$ we have, in view of (2.10) applied to the one-form $\imath_{\mathcal{V}_2}\omega$,

$$L_{\mathcal{V}_1}\imath_{\mathcal{V}_2}\omega(\mathcal{V}) = \mathcal{V}_1[\imath_{\mathcal{V}_2}\omega(\mathcal{V})] - \imath_{\mathcal{V}_2}\omega([\mathcal{V}_1, \mathcal{V}]) = \mathcal{V}_1[\omega(\mathcal{V}_2, \mathcal{V})] - \omega(\mathcal{V}_2, [\mathcal{V}_1, \mathcal{V}]).$$

Applying (2.10) again, but now to the two-form ω we get

$$\imath_{\mathcal{V}_2}L_{\mathcal{V}_1}\omega(\mathcal{V}) = L_{\mathcal{V}_1}\omega(\mathcal{V}_2, \mathcal{V}) = \mathcal{V}_1[\omega(\mathcal{V}_2, \mathcal{V})] - \omega([\mathcal{V}_1, \mathcal{V}_2], \mathcal{V}) - \omega(\mathcal{V}_2, [\mathcal{V}_1, \mathcal{V}]).$$

It follows that for any vector field $\mathcal{V}$

$$L_{\mathcal{V}_1}\imath_{\mathcal{V}_2}\omega(\mathcal{V}) - \imath_{\mathcal{V}_2}L_{\mathcal{V}_1}\omega(\mathcal{V}) = \omega([\mathcal{V}_1, \mathcal{V}_2], \mathcal{V}) = \imath_{[\mathcal{V}_1,\mathcal{V}_2]}\omega(\mathcal{V}),$$

showing (2.12).

The Lie derivative $L_{\mathcal{V}}\mathcal{W}$ of a vector field $\mathcal{W}$ is given by $L_{\mathcal{V}}\mathcal{W} := [\mathcal{V}, \mathcal{W}]$, while the Lie derivative of an arbitrary k-vector field P is the k-vector field $L_{\mathcal{V}}P$, defined by

$$L_{\mathcal{V}}P[F_1 \wedge \ldots \wedge F_k] := \mathcal{V}[P[F_1 \wedge \ldots \wedge F_k]] - \sum_{i=1}^{k} P[F_1 \wedge \ldots \wedge \mathcal{V}[F_i] \wedge \ldots \wedge F_k], \tag{2.13}$$

where $F_1, \ldots, F_k \in \mathcal{F}(M)$. As in the case of the Lie derivative of a differential form, the Lie derivative $L_{\mathcal{V}}P$ of a k-vector field P on M also measures how P changes in the direction of $\mathcal{V}$, hence $L_{\mathcal{V}}P = 0$ if and only if P is constant on the integral curves of the vector field $\mathcal{V}$.

Example 2.6. In order to get familiar with the notations, let us verify that $L_{\mathcal{V}}\mathcal{W} = [\mathcal{V}, \mathcal{W}]$, for any $\mathcal{V}, \mathcal{W} \in \mathfrak{X}(M)$. In fact, if $F \in \mathcal{F}(M)$ then (2.13) implies

$$(L_{\mathcal{V}}\mathcal{W})[F] = \mathcal{V}[\mathcal{W}[F]] - \mathcal{W}[\mathcal{V}[F]] = (\mathcal{V} \circ \mathcal{W} - \mathcal{W} \circ \mathcal{V})[F] = [\mathcal{V}, \mathcal{W}][F],$$

proving our claim.

2.2 Lie Groups and Lie Algebras

Unless otherwise stated, all Lie groups and Lie algebras will be defined over $\mathbf{C}$. We use the standard convention that Lie groups are denoted by boldface capital letters ($\mathbf{G}, \mathbf{H}, \ldots$) and their Lie algebras by the corresponding gothic letters ($\mathfrak{g}, \mathfrak{h}, \ldots$). The main examples of Lie groups include linear groups, i.e., Lie subgroups of $\mathbf{GL}(n) = \mathbf{GL}(\mathbf{C}^n)$, the non-commutative group of all invertible $n \times n$ matrices (with coefficients in $\mathbf{C}$), where the group operation is given by the usual product of matrices. Similarly, the main examples of Lie algebras include matrix Lie algebras, i.e., Lie subalgebras of $\mathfrak{gl}(n) = \mathfrak{gl}(\mathbf{C}^n)$, the Lie algebra of all $n \times n$ matrices (with coefficients in $\mathbf{C}$), where the *Lie bracket* is given by the commutator of matrices. In fact, according to *Ado's Theorem*, every finite-dimensional Lie algebra is isomorphic to a matrix Lie algebra, but the corresponding theorem does not hold true for (finite-dimensional) Lie groups. $\mathfrak{gl}(n)$ is the Lie algebra of $\mathbf{GL}(n)$, and the Lie algebra of a given linear group can easily be realized as a matrix Lie algebra. This is done by using the exponential map exp, which is a natural local biholomorphism $\exp : U \subseteq \mathfrak{g} \to \mathbf{G}$ between an open neighborhood U of the origin 0 of any finite-dimensional Lie algebra $\mathfrak{g}$ and an open neighborhood of the unit element $e \in \mathbf{G}$. For $X \in \mathfrak{g} = T_e\mathbf{G}$ close to 0, $\exp X$ is the Lie group element $\Phi_1(e)$ where Φ denotes the flow of the left invariant vector field on $\mathbf{G}$, defined by X. For example, taking $\mathfrak{g} = \mathfrak{gl}(n)$ and $\mathbf{G} = \mathbf{GL}(n)$ the map exp is the usual exponential of matrices and the conditions that define a subgroup of $\mathbf{GL}(n)$ are easily translated in the conditions that define its Lie algebra.

Example 2.7. Consider the linear group $\mathbf{G}$ of all orthogonal $n \times n$ matrices. For $X \in \mathfrak{gl}(n)$ and for small $|t|$, consider the invertible matrix $\exp(tX) = \mathrm{Id}_n + tX + t^2X^2 + O(t^3)$. Orthogonality of $\exp(tX)$ yields

$$\begin{aligned} \mathrm{Id}_n &= \exp(tX)\exp(tX)^\top \\ &= (\mathrm{Id}_n + tX + t^2X^2 + O(t^3)) \left(\mathrm{Id}_n + tX^\top + t^2(X^2)^\top + O(t^3)\right) \\ &= \mathrm{Id}_n + t(X + X^\top) + O(t^2), \end{aligned}$$

hence the Lie algebra $\mathfrak{g}$ of $\mathbf{G}$ is given by the matrices X for which $X + X^\top = 0$, i.e., $\mathfrak{g}$ is the Lie algebra of all skew-symmetric $n \times n$ matrices.

More generally, Lie subalgebras of a finite-dimensional Lie algebra $\mathfrak{g}$ are in one-to-one correspondence with connected Lie subgroups of $\mathbf{G}$; notice however that a Lie subgroup needs not be closed in its ambient Lie group (consider the subgroup generated by a generic element of the complex torus $\mathbf{C}^2/\mathbf{Z}^2$).

The tangent space $T_g\mathbf{G}$ to $\mathbf{G}$ at an element $g \in \mathbf{G}$ is naturally identified with $\mathfrak{g}$: the left translation map $L_{g^{-1}} : \mathbf{G} \to \mathbf{G}$ maps g to e, and its differential maps $T_g\mathbf{G}$ to $\mathfrak{g}$. Similarly we can identify the cotangent spaces to $\mathbf{G}$ with $\mathfrak{g}^*$ (the dual vector space to $\mathfrak{g}$) and so on.

Example 2.8. Suppose that $\mathbf{G}$ is a linear group which is closed (as a topological subspace of $\mathbf{GL}(n)$). Elements of $T\mathbf{G}$ are then naturally represented by pairs of matrices (g, X), where a vector (g, X) acts by definition on $F \in \mathcal{F}(\mathbf{G})$ by

$$(g, X)[F] := \lim_{t \to 0} \frac{\overline{F}(g + tX) - F(g)}{t},$$

where $\overline{F}$ is any (holomorphic) extension of F to a small open neighborhood of g in its ambient space, the space of all $n \times n$ matrices (this can be done because $\mathbf{G}$ is closed). Then $\mathsf{d}L_{g^{-1}}(g, X) = (e, g^{-1}X)$, so that (g, X) gets naturally identified with the matrix $g^{-1}X$. The same result holds true when $\mathbf{G}$ is not closed, because a small neighborhood in $\mathbf{G}$ of any element g of $\mathbf{G}$ is closed in a small neighborhood of g in $\mathbf{GL}(n)$.

Since $\mathfrak{g}$ (resp. $\mathfrak{g}^*$) is a vector space, its tangent spaces are also naturally identified with $\mathfrak{g}$ (resp. $\mathfrak{g}^*$). These identification will be (ab)used in the sequel, often without further mention. A particular instance of this that will be used throughout the text is the following: if $F \in \mathcal{F}(\mathfrak{g}^*)$ and $\xi \in \mathfrak{g}^*$ then the differential of F at ξ is a linear map

$$\mathsf{d}F(\xi) : T_\xi\mathfrak{g}^* \to \mathbf{C}$$

which under the above identifications gets naturally identified with an element of $\mathfrak{g}$. Conversely, an element $X \in \mathfrak{g}$ will often be viewed (without changing notations) as a linear map $\mathfrak{g}^* \to \mathbf{C}$, a change of perspective that is made transparent by the notations, which simply read $\langle X, \xi \rangle = \langle \xi, X \rangle$ for $X \in \mathfrak{g}$ and $\xi \in \mathfrak{g}^*$.

Lie groups appear most often through their action (always taken to be a left action) on manifolds. By assumption, if

$$\chi : \mathbf{G} \times M \to M$$

is an action, then for each $g \in \mathbf{G}$ the map $\chi_g : M \to M$, defined by $m \mapsto \chi(g, m)$ is a biholomorphism, with inverse $m \mapsto \chi(g^{-1}, m)$. We will usually write $g{\cdot}m$ or gm for $\chi(g, m)$. The action allows us to associate to each element $X \in \mathfrak{g}$ a vector field $\underline{X}$, whose value at m, denoted $\underline{X}(m)$, is the derivation on $\mathcal{F}(M)$ at m given by

$$\underline{X}(m)[F] := \frac{\mathsf{d}}{\mathsf{d}t}_{|t=0} F((\exp tX) \cdot m),$$

for all $F \in \mathcal{F}(M)$. The vector field $\underline{X}$ is called the *fundamental vector field* corresponding to $X \in \mathfrak{g}$. Its flow is given by the action of the one-parameter group $\exp tX$. The fundamental vector fields describe infinitesimally the action of $\mathbf{G}$ on M and they span the tangent space to the orbits of $\mathbf{G}$ at every point of M.

The simplest action of a Lie group $\mathbf{G}$ on a vector space S is a *linear action*, which means that for each $g \in \mathbf{G}$ one has that $\chi_g \in \mathbf{GL}(S)$. Then one can view χ as a homomorphism $\chi : \mathbf{G} \to \mathbf{GL}(S)$ and one says that χ is a *representation* of $\mathbf{G}$ on S. A subspace $T \subseteq S$ is called an *invariant subspace* if χ_g leaves T stable, i.e. $\chi_g(T) \subseteq T$, for all $g \in \mathbf{G}$. Then χ induces a representation $\mathbf{G} \to \mathbf{GL}(T)$ which is called a *subrepresentation.* A representation χ of $\mathbf{G}$ on S is called an *irreducible representation* if $\dim S > 0$ and if χ does not admit a non-trivial (i.e., with T different from $\{0\}$ and T) subrepresentation. In general an invariant subspace T may or may not have a complement T' in S which is also an invariant subspace. One says that a representation χ is a *completely reducible representation* if any invariant subspace admits a complementary subspace which is also invariant. In that case one can describe χ as a direct sum of irreducible representations.

The above terminology applies equally to the case of Lie algebra representations, with the understanding that a *representation* of $\mathfrak{g}$ on S is a Lie algebra homomorphism $\mathfrak{g} \to \operatorname{End}(S)$, the Lie bracket in $\operatorname{End}(S)$ being the commutator of endomorphisms. One also says that S is a $\mathfrak{g}$-*module.* Using the fact that $\operatorname{End}(S)$ is the Lie algebra of $\mathbf{GL}(S)$ and using our convention that we identify all tangent spaces to $\mathbf{G}$ with $\mathfrak{g}$, every representation of $\mathbf{G}$ on S leads to a representation of $\mathfrak{g}$ on S by mapping $X \in \mathfrak{g} \mapsto \underline{X} \in \operatorname{End}(S)$. The fact that this gives indeed a representation follows from the formula

$$[\underline{X}, \underline{Y}] = \underline{[X, Y]} \qquad X, Y \in \mathfrak{g}.$$

The two most important examples are the adjoint and the coadjoint action (representation) of a Lie group $\mathbf{G}$ on its Lie algebra $\mathfrak{g}$, resp. on the dual $\mathfrak{g}^*$ of its Lie algebra.

For $g \in \mathbf{G}$, we define Ad_g to be the endomorphism of $\mathfrak{g}$ which is the derivative of the conjugation map $C_g : \mathbf{G} \to \mathbf{G} : h \mapsto ghg^{-1}$ at the identity, $\mathrm{Ad}_g := \mathsf{d}C_g(e)$. The *adjoint action* or *adjoint representation* of $\mathbf{G}$ on $\mathfrak{g}$ is then given by

$$\mathrm{Ad} : \mathbf{G} \to \mathbf{GL}(\mathfrak{g}) : g \mapsto \mathrm{Ad}_g \,.$$

For example, if $\mathbf{G}$ is a linear group and $\mathfrak{g}$ its (matrix) Lie algebra, then it follows, as in Example 2.8, that

$$\mathrm{Ad}_g X = \mathsf{d}C_g(e)(X) = gXg^{-1},$$

where $g \in \mathbf{G}$ and $X \in \mathfrak{g}$.

The representation of $\mathfrak{g}$ on itself which corresponds to the adjoint action is called the *adjoint representation* of the Lie algebra $\mathfrak{g}$ on itself and is denoted by ad; the image of $X \in \mathfrak{g}$ under ad will, for readability, be written as ad_X. By the above definition, ad_X is the fundamental vector field $\underline{X}$ on $\mathfrak{g}$ that corresponds to the adjoint action, viewed as an endomorphism of $\mathfrak{g}$ (by identifying all tangent spaces of $\mathfrak{g}$ to $\mathfrak{g}$). Explicitly, $\mathrm{ad}_X Y = [X,Y]$, for $Y \in \mathfrak{g}$.

We now turn to the coadjoint action. For $g \in \mathbf{G}$ we define Ad^*_g by duality:

$$\left\langle \mathrm{Ad}^*_g \xi, X \right\rangle = \left\langle \xi, \mathrm{Ad}_{g^{-1}} X \right\rangle,$$

where $\xi \in \mathfrak{g}^*$ and $X \in \mathfrak{g}$. The resulting map

$$\mathrm{Ad}^* : \mathbf{G} \to \mathbf{GL}(\mathfrak{g}^*) : g \mapsto \mathrm{Ad}^*_g$$

is called the *coadjoint action* or the *coadjoint representation* of $\mathbf{G}$ on $\mathfrak{g}^*$. Its orbits are called *coadjoint orbits* and they play an important role in what follows. The representation of $\mathfrak{g}$ on $\mathfrak{g}^*$ that corresponds to the coadjoint representation is denoted by ad^* and is called the *coadjoint representation* of $\mathfrak{g}$ on $\mathfrak{g}^*$. The relation between ad and ad^* is consequently given by

$$\langle \mathrm{ad}^*_X \xi, Y \rangle = \langle \xi, -\mathrm{ad}_X Y \rangle = \langle \xi, [Y,X] \rangle \,, \tag{2.14}$$

where $\xi \in \mathfrak{g}^*$ and $X, Y \in \mathfrak{g}$.

A function $H \in \mathcal{F}(\mathfrak{g})$ (resp. $H \in \mathcal{F}(\mathfrak{g}^*)$) is called Ad-*invariant* (resp. Ad^*-*invariant*) if $H(\mathrm{Ad}_g X) = H(X)$ for all $g \in \mathbf{G}$ and $X \in \mathfrak{g}$ (resp. $H\left(\mathrm{Ad}^*_g \xi\right) = H(\xi)$ for all $g \in \mathbf{G}$ and $\xi \in \mathfrak{g}^*$). The algebra of Ad-invariant functions on $\mathfrak{g}$ is denoted by $\mathcal{F}(\mathfrak{g})^{\mathbf{G}}$, while the algebra of Ad^*-invariant functions on $\mathfrak{g}^*$ is denoted by $\mathcal{F}(\mathfrak{g}^*)^{\mathbf{G}}$.

In the following lemma we describe two properties of Ad^*-invariant functions that we will use. The Ad-invariant functions have similar properties, that are easily written down, and are proven in the same way, but these properties will not be used explicitly used here.

Lemma 2.9. *Let $H \in \mathcal{F}(\mathfrak{g}^*)^{\mathbf{G}}$. For any $\xi \in \mathfrak{g}^*$ and for any $X \in \mathfrak{g}$ we have that $\langle \xi, [\mathsf{d}H(\xi), X] \rangle = 0$, i.e., one has for any $\xi \in \mathfrak{g}^*$ that*

$$\operatorname{ad}^*_{\mathsf{d}H(\xi)} \xi = 0.$$

Moreover, for any $g \in \mathbf{G}$ and $\xi \in \mathfrak{g}^*$ the following diagram is commutative.

$$\begin{array}{ccc} \mathfrak{g}^* & \xrightarrow{\mathsf{d}H} & \mathfrak{g} \\ {\scriptstyle \operatorname{Ad}^*_g}\downarrow & & \downarrow{\scriptstyle \operatorname{Ad}_g} \\ \mathfrak{g}^* & \xrightarrow[\mathsf{d}H]{} & \mathfrak{g} \end{array}$$

Proof. Let $X \in \mathfrak{g}$ and $\xi \in \mathfrak{g}^*$. If $H \in \mathcal{F}(\mathfrak{g}^*)^{\mathbf{G}}$ then $H(\xi) = H\left(\operatorname{Ad}^*_g \xi\right)$ for all $\xi \in \mathfrak{g}^*$ and $g \in \mathbf{G}$. Taking any $X \in \mathfrak{g}$ we therefore have that

$$\begin{aligned} \left\langle \operatorname{ad}^*_{\mathsf{d}H(\xi)} \xi, X \right\rangle &= -\left\langle \operatorname{ad}^*_X \xi, \mathsf{d}H(\xi) \right\rangle = -\operatorname{ad}^*_X \xi\, [H] \\ &= -\frac{\mathsf{d}}{\mathsf{d}t}_{|t=0} H\left(\operatorname{Ad}^*_{\exp tX} \xi\right) = -\frac{\mathsf{d}}{\mathsf{d}t}_{|t=0} H(\xi) = 0, \end{aligned}$$

showing the first property. In order to prove that the above diagram is commutative, differentiate for a fixed $g \in \mathbf{G}$ the identity $H = H \circ \operatorname{Ad}^*_g$ at $\xi \in \mathfrak{g}^*$. It gives

$$\mathsf{d}H(\xi) = \mathsf{d}H(\operatorname{Ad}^*_g \xi) \circ (\mathsf{d} \operatorname{Ad}^*_g)(\xi) = \mathsf{d}H(\operatorname{Ad}^*_g \xi) \circ \operatorname{Ad}^*_g,$$

because, for fixed g, the map Ad^*_g is a linear map. Thus, for $\eta \in \mathfrak{g}^*$, we have

$$\langle \mathsf{d}H(\xi), \eta \rangle = \left\langle \mathsf{d}H(\operatorname{Ad}^*_g \xi), \operatorname{Ad}^*_g \eta \right\rangle = \left\langle \operatorname{Ad}_{g^{-1}} \left(\mathsf{d}H(\operatorname{Ad}^*_g \xi)\right), \eta \right\rangle.$$

It follows that $\mathsf{d}H(\operatorname{Ad}^*_g \xi) = \operatorname{Ad}_g(\mathsf{d}H(\xi))$, which proves that the diagram is commutative. □

Lie algebras often come equipped with a non-degenerate symmetric bilinear form

$$\langle \cdot \,|\, \cdot \rangle : \mathfrak{g} \times \mathfrak{g} \to \mathbf{C}.$$

Such a form allows us to identify $\mathfrak{g}$ with $\mathfrak{g}^*$, simply by assigning to $X \in \mathfrak{g}$ the linear form $\hat{X}$ which maps $Y \in \mathfrak{g}$ to $\langle X \,|\, Y \rangle$, i.e.,

$$\langle \hat{X}, Y \rangle = \langle X \,|\, Y \rangle$$

for all $X, Y \in \mathfrak{g}$. Its inverse is the linear map $\mathfrak{g}^* \to \mathfrak{g} : \xi \mapsto X$, where X is the unique element of $\mathfrak{g}$ which satisfies $\langle X \,|\, Y \rangle = \langle \xi, Y \rangle$, for all $Y \in \mathfrak{g}$. A symmetric bilinear form $\langle \cdot \,|\, \cdot \rangle$ on $\mathfrak{g}$ will be called Ad-*invariant* when for any $g \in \mathbf{G}$ and for any $X, Y \in \mathfrak{g}$ one has

$$\langle \operatorname{Ad}_g X \mid \operatorname{Ad}_g Y \rangle = \langle X \,|\, Y \rangle.$$

Ad-invariance of $\langle \cdot \,|\, \cdot \rangle$ implies the following associativity-like rule: for any $X, Y, Z \in \mathfrak{g}$ one has

$$\langle \mathrm{ad}_Y X \,|\, Z \rangle = - \langle X \mid \mathrm{ad}_Y Z \rangle , \tag{2.15}$$

so that ad_Y is skew-symmetric with respect to $\langle \cdot \,|\, \cdot \rangle$. Ad-invariance of $\langle \cdot \,|\, \cdot \rangle$ also implies that for any $g \in \mathbf{G}$ the following diagram is commutative.

$$\begin{array}{ccc} \mathfrak{g} & \xrightarrow{\wedge} & \mathfrak{g}^* \\ {\scriptstyle \mathrm{Ad}_g}\downarrow & & \downarrow{\scriptstyle \mathrm{Ad}_g^*} \\ \mathfrak{g} & \xrightarrow[\wedge]{} & \mathfrak{g}^* \end{array} \tag{2.16}$$

Indeed, for any $X, Y \in \mathfrak{g}$ it follows from the definitions and from Ad-invariance that

$$\begin{aligned} \left\langle \mathrm{Ad}_g^* \hat{X}, Y \right\rangle &= \left\langle \hat{X}, \mathrm{Ad}_{g^{-1}} Y \right\rangle = \langle X \mid \mathrm{Ad}_{g^{-1}} Y \rangle \\ &= \langle \mathrm{Ad}_g X \,|\, Y \rangle = \left\langle \widehat{\mathrm{Ad}_g X}, Y \right\rangle . \end{aligned}$$

In words: upon identifying a Lie algebra with its dual (using an non-degenerate Ad-invariant symmetric bilinear form $\langle \cdot \,|\, \cdot \rangle$), the adjoint and coadjoint actions get identified. The datum of a bilinear form on $\mathfrak{g}$ leads to a notion of orthogonality: for any subset $A \subseteq \mathfrak{g}$ the $\langle \cdot \,|\, \cdot \rangle$-*orthogonal* of A is the subspace of $\mathfrak{g}$, defined by

$$A^\perp := \{ Y \in \mathfrak{g} \mid \langle X \,|\, Y \rangle = 0 \text{ for all } X \in A \} .$$

Example 2.10. Consider the subalgebras $\mathfrak{a}$ and $\mathfrak{b}$ of $\mathfrak{gl}(n)$ which consist respectively of the skew-symmetric and the upper triangular matrices. Obviously, $\mathfrak{gl}(n) = \mathfrak{a} \oplus \mathfrak{b}$ (direct sum of vector spaces). Consider on $\mathfrak{gl}(n)$ the non-degenerate symmetric bilinear form defined by $\langle X \,|\, Y \rangle := \mathrm{Trace}(XY)$. Then $\mathfrak{a}^\perp$ is the subspace of $\mathfrak{gl}(n)$ consisting of all symmetric matrices, while $\mathfrak{b}^\perp$ consists of all strictly upper triangular matrices.

The main example of an Ad-invariant symmetric bilinear form is the *Killing form* of $\mathfrak{g}$, which is defined by

$$\langle X \,|\, Y \rangle := \mathrm{Trace}(\mathrm{ad}_X \circ \mathrm{ad}_Y). \tag{2.17}$$

The Killing form of $\mathfrak{g}$ is non-degenerate if and only if $\mathfrak{g}$ is a semi-simple Lie algebra. Semi-simple Lie algebras and simple Lie algebras will be defined in the next section. Together with their infinite-dimensional analogues, the (twisted) affine Lie algebras, they will be the main types of Lie algebras encountered in this book.

2.3 Simple Lie Algebras

A non-empty subset $\mathfrak{h}$ of a Lie algebra $\mathfrak{g}$ is called an *ideal* when $[\mathfrak{g}, \mathfrak{h}] \subseteq \mathfrak{h}$. If $\mathfrak{g}$ contains no other ideals than 0 and itself and $\dim \mathfrak{g} > 1$ then $\mathfrak{g}$ is called a *simple Lie algebra.* A Lie algebra that is isomorphic to the direct sum of simple Lie algebras is called a *semi-simple Lie algebra.* Such a Lie algebra is characterized by the fact that its Killing form $\langle \cdot \,|\, \cdot \rangle$ (see (2.17)) is non-degenerate. Moreover, for any simple Lie algebra $\mathfrak{g}$ the Killing form is, up to a constant, the unique Ad-invariant symmetric bilinear form on $\mathfrak{g}$ which is non-degenerate.

2.3.1 The Classification

We describe in this paragraph the elements that appear in the classification and in the representation theory of simple Lie algebras, since we will need them in what follows. For proofs and details we refer to [87].

In this paragraph we assume that $\mathfrak{g}$ is a simple Lie algebra (over $\mathbf{C}$). Let $\mathfrak{h} \subseteq \mathfrak{g}$ be a *Cartan subalgebra* of $\mathfrak{g}$, i.e., $\mathfrak{h}$ is Abelian ($[\mathfrak{h}, \mathfrak{h}] = 0$) and self-normalizing ($x \in \mathfrak{g}$ and $[x, \mathfrak{h}] \subseteq \mathfrak{h}$ implies $x \in \mathfrak{h}$). The dimension of $\mathfrak{h}$ is called the *rank* of $\mathfrak{g}$, denoted $\operatorname{Rk} \mathfrak{g}$. It does not depend on the choice of $\mathfrak{h}$ because one shows that $\mathfrak{h}$ is unique up to an automorphism of $\mathfrak{g}$. For $X \in \mathfrak{h}$ the endomorphism $\mathrm{ad}_X : \mathfrak{g} \to \mathfrak{g}$ is diagonalizable and commutativity of $\mathfrak{h}$ implies that $\mathrm{ad}_{\mathfrak{h}}$ is a family of simultaneously diagonalizable endomorphisms of $\mathfrak{g}$, leading to a direct sum decomposition of $\mathfrak{g}$ into eigenspaces of $\mathrm{ad}_{\mathfrak{h}}$,

$$\mathfrak{g} = \mathfrak{h} \oplus \sum_{\alpha \in \Phi} \mathfrak{g}^{\alpha}, \tag{2.18}$$

where each subspace $\mathfrak{g}^{\alpha}$ can be shown to be one-dimensional. An element α of Φ is called a *root*, $\Phi \subseteq \mathfrak{h}^*$ is called a *root system* and the decomposition (2.18) is called the *root space decomposition.* A root $\alpha \in \Phi$ is a collection of eigenvalues of $\mathrm{ad}_{\mathfrak{h}}$ in the sense that if $E_\alpha \in \mathfrak{g}^{\alpha}$ then

$$[H, E_\alpha] = \langle \alpha, H \rangle\, E_\alpha \qquad \text{for all } H \in \mathfrak{h}.$$

Notice that 0 is not a root because we assumed that $\mathfrak{h}$ is self-normalizing. It can be shown that the root system Φ spans $\mathfrak{h}^*$ and that a (non-unique) basis Π for $\mathfrak{h}^*$ can be extracted from Φ, with the following property: any root $\alpha \in \Phi$ is a linear combination of elements of Π with coefficients in $\mathbf{Z}$ which are either all positive or all negative. Thus, $\Phi = \Phi^+ \cup \Phi^-$, where $\Phi^- = -\Phi^+$, and any element of Φ^+ is a linear combination of elements of Π with coefficients in $\mathbf{N}$. In particular, $\Pi \subseteq \Phi^+$ and the roots all belong to the lattice[2], generated by the simple roots, called the *root lattice.*

[2] This lattice is independent of the choice of simple roots Π since it is the smallest lattice in $\mathfrak{h}^*$ that contains all the roots.

We call Π a set of *simple roots* and we set $\Pi = (\alpha_1, \ldots, \alpha_l)$, where $l := \operatorname{Rk}\mathfrak{g}$. For $\alpha \in \Phi$ we define its *height* $|\alpha|$ by

$$|\alpha| = \sum_{i=1}^{l} a_i, \quad \text{where} \quad \alpha = \sum_{i=1}^{l} a_i \alpha_i.$$

It leads to a grading of $\mathfrak{g}$,

$$\mathfrak{g} = \bigoplus_k \mathfrak{g}_k, \qquad [\mathfrak{g}_k, \mathfrak{g}_l] \subseteq [\mathfrak{g}_{k+l}], \tag{2.19}$$

where $\mathfrak{g}_k$ is, for $k \neq 0$, the span of the eigenvectors of all E_α, with $|\alpha| = k$, and $\mathfrak{g}_0 := \mathfrak{h}$. One proves that the Killing form $\langle \cdot \,|\, \cdot \rangle$ restricts to a bilinear form on $\mathfrak{h}$, which is also non-degenerate, hence the isomorphism $\mathfrak{g} \to \mathfrak{g}^*$, induced by $\langle \cdot \,|\, \cdot \rangle$, leads to an isomorphism $\mathfrak{h} \to \mathfrak{h}^*$. We mainly use its inverse, $\mathfrak{h}^* \to \mathfrak{h}$, which maps $\lambda \in \mathfrak{h}^*$ to $h_\lambda \in \mathfrak{h}$, where

$$\langle \lambda \,, \cdot \rangle = \langle h_\lambda \,|\, \cdot \rangle \,.$$

We will use, for $i = 1, \ldots, l$, the abbreviation h_i for h_{α_i}. The l-tuple $(h_1, \ldots, h_l)$ forms a basis of $\mathfrak{h}$. For $\alpha \in \Phi$ we define the following useful normalization[3] of h_α:

$$H_\alpha := 2 \frac{h_\alpha}{\langle h_\alpha \,|\, h_\alpha \rangle}. \tag{2.20}$$

Each H_α is called a *coroot* and for $i = 1, \ldots, l$ the coroot which corresponds to α_i is denoted by H_i. The coroots appear in the following refinement of the root space decomposition.

Theorem 2.11 (Chevalley). *Let $\mathfrak{g}$ be a simple Lie algebra of rank l, let $\mathfrak{h}$ be a Cartan subalgebra with root system Φ, and let $\Pi = (\alpha_1, \ldots, \alpha_l)$ denote a system of simple roots with respect to $\mathfrak{h}$. Then there exists for every $\alpha \in \Phi$ a non-zero vector $E_\alpha \in \mathfrak{g}^\alpha$ (see (2.18)) such that for any $H, H' \in \mathfrak{h}$ and $\alpha, \beta \in \Phi$*

$$\begin{aligned} [H, H'] &= 0, \\ [H, E_\alpha] &= \langle \alpha, H \rangle \, E_\alpha, \\ [E_\alpha, E_\beta] &= \begin{cases} H_\alpha & \text{if } \alpha + \beta = 0, \\ 0 & \text{if } \alpha + \beta \notin \Phi \cup \{0\}\,, \\ N_{\alpha\beta} E_{\alpha+\beta} & \text{if } \alpha + \beta \in \Phi. \end{cases} \end{aligned}$$

Here, $N_{\alpha\beta} = \pm(p+1)$, where

$$p := \max \{ n \mid \beta - n\alpha \in \Phi \} \,.$$

The basis $(H_1, \ldots, H_l) \cup (E_\alpha)_{\alpha \in \Phi}$ of $\mathfrak{g}$ is called a Chevalley basis *of $\mathfrak{g}$, and each E_α is called a* root vector.

[3] It is a normalization in the sense that if we replace $\langle \cdot \,|\, \cdot \rangle$ by a non-zero multiple of itself, then H_α do not change, while the h_α get divided by that factor.

For computational purposes it is useful to know that for given $\alpha, \beta \in \Phi$ the set $\{n \mid \alpha - n\beta \in \Phi\}$, consists of (a finite number of) *consecutive* integers. The choice of sign for $N_{\alpha\beta}$ for all $\alpha, \beta \in \Phi$ is non-trivial, since the signs that correspond to the different values of α and β need to satisfy several non-trivial coherence conditions, but they can be determined algorithmically (see [67]). Notice that Chevalley's Theorem implies that, in terms of a Chevalley basis, all structure constants of $\mathfrak{g}$ are integers.

The Killing form $\langle \cdot \,|\, \cdot \rangle$ allows us to measure angles and lengths of roots in $\Phi \subseteq \mathfrak{h}^*$. To do this, let $\mathfrak{h}^*_{\mathbf{R}}$ denote the real vector space which is spanned by Φ and define $\langle \cdot \,|\, \cdot \rangle_{\mathfrak{h}^*}$ to be the bilinear form on $\mathfrak{h}^*_{\mathbf{R}}$ which corresponds to the Killing form via the isomorphism $h \mapsto h_\lambda$. Thus, for $\lambda, \mu \in \mathfrak{h}^*$ we have that $\langle \lambda \,|\, \mu \rangle_{\mathfrak{h}^*} = \langle h_\lambda \,|\, h_\mu \rangle$. It turns out that $\langle \cdot \,|\, \cdot \rangle_{\mathfrak{h}^*}$ is positive definite, making $\mathfrak{h}^*_{\mathbf{R}}$ into a genuine Euclidean space. For $\alpha \in \Phi$, let $s_\alpha : \mathfrak{h}^*_{\mathbf{R}} \to \mathfrak{h}^*_{\mathbf{R}}$ be the linear map defined by

$$s_\alpha(\lambda) = \lambda - 2\frac{\langle h_\alpha \,|\, h_\lambda \rangle}{\langle h_\alpha \,|\, h_\alpha \rangle}\alpha = \lambda - \langle \lambda, H_\alpha \rangle\, \alpha, \tag{2.21}$$

where $\lambda \in \mathfrak{h}^*_{\mathbf{R}}$. This linear map is the reflection in the hyperplane orthogonal to α, since it fixes all roots which are orthogonal to α and since $s_\alpha(\alpha) = -\alpha$.

The *Weyl group* $\mathbf{W}$ is the group generated by $\{s_\alpha \mid \alpha \in \Phi\}$. One shows the following properties of the Weyl group. Every non-trivial element of $\mathbf{W}$ permutes at least two elements of Φ hence $\mathbf{W}$ is finite. Moreover, $\mathbf{W}$ is generated by the l reflections that correspond to the elements $\alpha_1, \ldots, \alpha_l$ of Π. The root system Φ consists either of one $\mathbf{W}$-orbit, in which case all roots have the same length, or it consists of two $\mathbf{W}$-orbits, where roots from one $\mathbf{W}$-orbit have a length which is different from the length of the vectors in the other $\mathbf{W}$-orbit. The two $\mathbf{W}$-orbits are then distinguished by calling its elements *short roots* or *long roots*, according to their lengths. Among the long roots there is precisely one that has maximal height. It is called the *highest long root* (with respect to $\mathfrak{h}$ and Π). Similarly, there is among the short roots precisely one that has maximal height, the *highest short root.* When all roots have the same length then the highest long root and the highest short root of course coincide.

For $\alpha, \beta \in \Pi$ the fact that

$$s_\alpha(\beta) = \beta - 2\frac{\langle h_\alpha \,|\, h_\beta \rangle}{\langle h_\alpha \,|\, h_\alpha \rangle}\alpha$$

is a root, implies for $\alpha \neq \beta$ that $s_\alpha(\beta) \in \Phi^+$, hence that

$$a_{ij} := 2\frac{\langle h_i \,|\, h_j \rangle}{\langle h_j \,|\, h_j \rangle} = \langle h_i \,|\, H_j \rangle = \langle \alpha_i, H_j \rangle \tag{2.22}$$

is a non-positive integer for $i \neq j$, and equals 2 for $i = j$.

The numbers a_{ij} are called the *Cartan integers* and the matrix $A = (a_{ij})$ is called the *Cartan matrix* of $\mathfrak{g}$ (with respect to $\mathfrak{h}$ and Π). It is a fundamental result that there is a bijection between the triples $(\mathfrak{g}, \mathfrak{h}, \Pi)$, modulo conjugation in $\mathfrak{g}$ and their Cartan matrices, modulo conjugation by a permutation matrix. In fact, $\mathfrak{g}$ can be reconstructed from its Cartan matrix by a set of generators and relations (see [156, Chapter VI]).

If we denote by θ_{ij} the angle between α_i and α_j then

$$\cos\theta_{ij} = \frac{\langle \alpha_i \mid \alpha_j \rangle_{\mathfrak{h}^*}}{\sqrt{\langle \alpha_i \mid \alpha_i \rangle_{\mathfrak{h}^*}}\sqrt{\langle \alpha_j \mid \alpha_j \rangle_{\mathfrak{h}^*}}} = \frac{\langle h_i \mid h_j \rangle}{\sqrt{\langle h_i \mid h_i \rangle}\sqrt{\langle h_j \mid h_j \rangle}}$$

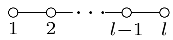
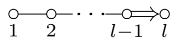
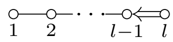
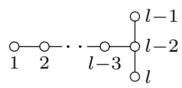
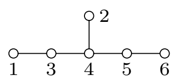
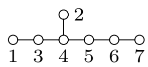
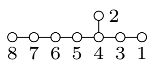
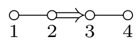
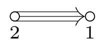
so that (2.22) implies that $4\cos^2\theta_{ij} = a_{ij}a_{ji}$. Letting $n_{ij} := a_{ij}a_{ji}$ we have that if $i \neq j$ then $0, 1, 2, 3$ are the only possible values for n_{ij}, since a_{ij} is a non-positive integer when $i \neq j$. The *Dynkin diagram* of $\mathfrak{g}$ is the graph with l nodes labeled by $1, \ldots, l$ such that the nodes i and j are joined with n_{ij} bonds. Notice that the integers n_{ij} do not contain enough information to determine the a_{ij}, i.e., to reconstruct the Cartan matrix: when $n_{ij} = 0$ then $a_{ij} = a_{ji} = 0$ and when $n_{ij} = 1$ then $a_{ij} = a_{ji} = -1$, but when $n_{ij} \in \{2, 3\}$ then there are two possibilities to assign the values -1 and $-n_{ij}$ to a_{ij} and a_{ji}. To resolve this ambiguity one adds an arrow to the double and the triple bonds in the Dynkin diagram which points to the shorter root ((2.22) shows that the two roots cannot have the same length). This way the Cartan matrix, and hence the whole structure of the simple Lie algebra, can be encoded in its Dynkin diagram. Analyzing the properties that root systems which come from a simple Lie algebra have and constructing all possible Dynkin diagrams that bear the corresponding properties one arrives at the well-known list of Dynkin diagrams, given in Table 2.1 (the labeling of the roots in the Dynkin diagram is the one that is used in most classical books on Lie algebras, in particular [37], [79] and [87]).

The coroots H_α, which were defined in (2.20), satisfy the axioms of a root system as well as the roots α, the *dual root system*, for which a system of simple roots can be chosen as $(H_1, \ldots, H_l)$. It leads to a natural duality on the set of simple Lie algebras, which at the level of the Cartan matrix amounts to $A \leftrightarrow A^\top$. As it turns out, this duality is trivial except that it permutes the Lie algebras $\mathfrak{b}_l$ and $\mathfrak{c}_l$.

If $\lambda \in \mathfrak{h}^*_{\mathbf{R}}$ has the property that $\langle \lambda, H_\alpha \rangle \in \mathbf{Z}$ for all $\alpha \in \Phi$ then λ is called a *weight* and the set of all weight vectors is a lattice in $\mathfrak{h}^*_{\mathbf{R}}$ which is denoted by Λ and which is called the *weight lattice*. Clearly $\Phi \subseteq \Lambda$. A basis for the lattice Λ can be constructed as follows: for $i = 1, \ldots, l$ let $\lambda_i \in \mathfrak{h}^*$ be such that $\langle \lambda_i, H_j \rangle = \delta_{ij}$, where $j = 1, \ldots, l$. Each of the basis vectors λ_i is called a *fundamental dominant weight*, or a *weight* for short. Since $\Phi \subseteq \Lambda$, (2.22) implies that

$$\alpha_i = \sum_{k=1}^{l} a_{ik}\lambda_k, \qquad i = 1, \ldots, l. \tag{2.23}$$

Table 2.1. Some data on simple Lie algebras. For each simple Lie algebra we list its rank, the order of its Weyl group $\mathbf{W}$, the determinant of its Cartan matrix A, the coefficients of highest long/short root in terms of the simple roots (only one is given if they are the same) and its Dynkin diagram. A label i in the Dynkin diagram refers to the root α_i. The Cartan matrix A is immediately written down from the Dynkin diagram.

$\mathfrak{g}$	Rank	$\#\mathbf{W}$	$\lvert A\rvert$	Highest long/short root	Dynkin diagram
$\mathfrak{a}_l$	$l \geqslant 1$	$(l+1)!$	$l+1$	$(1,1,\ldots,1)$	
$\mathfrak{b}_l$	$l \geqslant 2$	$2^l l!$	2	$(1,2,\ldots,2)/(1,1,\ldots,1)$	
$\mathfrak{c}_l$	$l \geqslant 3$	$2^l l!$	2	$(2,\ldots,2,1)/(1,2,\ldots,2,1)$	
$\mathfrak{d}_l$	$l \geqslant 4$	$2^{l-1} l!$	4	$(1,2,\ldots,2,1,1)$	
$\mathfrak{e}_6$	6	$2^7 3^4 5$	3	$(1,2,2,3,2,1)$	
$\mathfrak{e}_7$	7	$2^{10} 3^4 5\, 7$	2	$(2,2,3,4,3,2,1)$	
$\mathfrak{e}_8$	8	$2^{14} 3^5 5^2 7$	1	$(2,3,4,6,5,4,3,2)$	
$\mathfrak{f}_4$	4	$2^7 3^2$	1	$(2,3,4,2)/(1,2,3,2)$	
$\mathfrak{g}_2$	2	$2^2 3$	1	$(3,2)/(2,1)$	

It follows from this relation that the Cartan matrix describes the change of basis from the simple roots to the fundamental dominant weights, a property that will play a fundamental rôle in our study of the periodic Toda lattice (see Chapter 9).

In the four examples that follow we give a concrete representation of the classical Lie algebras, whose root systems are $\mathfrak{a}_l$, $\mathfrak{b}_l$, $\mathfrak{c}_l$ and $\mathfrak{d}_l$, together with a choice of root vectors which, supplemented with a basis of $\mathfrak{h}$, form a Chevalley basis. We only give a choice for root vectors corresponding to the roots E_{α_i}, $E_{-\alpha_i}$, where $i = 1, \dots, l$, and to plus and minus the highest long/short root, because the other root vectors will not be needed. The choices that we make are the most appropriate for our approach to the periodic Toda lattices (Chapter 9), and are taken from [36], where one also finds explicit expressions for the other root vectors. We denote by $\mathcal{E}_{ij}$ the square matrix (of the appropriate size) which has a 1 at position (i, j) and zeros elsewhere and Δ is the $l \times l$ matrix with 1's on the anti-diagonal and zeros elsewhere, $\Delta := \sum_{i=1}^{l} \mathcal{E}_{i,l-i+1}$. Notice that the condition $A\Delta = \Delta A^\top$ (resp. $A\Delta + \Delta A^\top = 0$) means that A is symmetric (resp. skew-symmetric) with respect to its anti-diagonal.

Example 2.12. $\mathfrak{a}_l$ is the root system of the semi-simple Lie algebra $\mathfrak{sl}(l+1)$ of all traceless matrices of size $l+1$. For E_{α_i} one chooses $\mathcal{E}_{i,i+1}$ and for the root vector corresponding to the highest (long = short) root α_0 one takes $E_{\alpha_0} := \mathcal{E}_{1,l+1}$. Then $E_{-\alpha_i} := E_{\alpha_i}^\top$ for $i = 0, \dots, l$.

Example 2.13. $\mathfrak{b}_l$ is the root system of the semi-simple Lie algebra of all block matrices of size $2l+1$ of the form

$$\begin{pmatrix} A & 2\Delta w & B \\ v^\top & 0 & w^\top \\ C & 2\Delta v & D \end{pmatrix}, \qquad \text{where} \qquad \begin{aligned} A\Delta + \Delta D^\top &= 0, \\ B\Delta + \Delta B^\top &= 0, \\ C\Delta + \Delta C^\top &= 0, \end{aligned}$$

and where $A, \dots, D$ are square matrices of size l, while v and w are column vectors on length l. For the root vectors of height ± 1 we choose

$$\begin{aligned} E_{\alpha_i} &:= \mathcal{E}_{i,i+1} - \mathcal{E}_{2l-i+1,2l-i+2}, & i &= 1, \dots, l-1, \\ E_{-\alpha_i} &:= \mathcal{E}_{i+1,i} - \mathcal{E}_{2l-i+2,2l-i+1}, & i &= 1, \dots, l-1, \\ E_{\alpha_l} &:= 2\mathcal{E}_{l,l+1} + \mathcal{E}_{l+1,l+2}, \\ E_{-\alpha_l} &:= \mathcal{E}_{l+1,l} + 2\mathcal{E}_{l+2,l+1}, \end{aligned}$$

while for the root vectors corresponding to the highest long/short roots we choose

$$\begin{aligned} E_{\alpha} &:= \mathcal{E}_{1,2l} - \mathcal{E}_{2,2l+1}, & &\alpha \text{ highest long root}, \\ E_{-\alpha} &:= \mathcal{E}_{2l,1} - \mathcal{E}_{2l+1,2}, & &\alpha \text{ highest long root}, \\ E_{\alpha} &:= 2\mathcal{E}_{1,l+1} + \mathcal{E}_{l+1,2l+1}, & &\alpha \text{ highest short root}, \\ E_{-\alpha} &:= \mathcal{E}_{l+1,1} + 2\mathcal{E}_{2l+1,l+1}, & &\alpha \text{ highest short root}. \end{aligned}$$

Example 2.14. $\mathfrak{c}_l$ is the root system of the semi-simple Lie algebra of all block matrices of size $2l$ of the form

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}, \qquad \text{where} \qquad \begin{aligned} A\Delta + \Delta D^\top &= 0, \\ B\Delta &= \Delta B^\top, \\ C\Delta &= \Delta C^\top, \end{aligned}$$

and where $A, \dots, D$ are square matrices of size l. For the root vectors of height ± 1 we choose

$$\begin{aligned} E_{\alpha_i} &:= \mathcal{E}_{i,i+1} - \mathcal{E}_{2l-i,2l-i+1}, & i &= 1, \dots, l-1, \\ E_{-\alpha_i} &:= \mathcal{E}_{i+1,i} - \mathcal{E}_{2l-i+1,2l-i}, & i &= 1, \dots, l-1, \\ E_{\alpha_l} &:= \mathcal{E}_{l,l+1}, \\ E_{-\alpha_l} &:= \mathcal{E}_{l+1,l}, \end{aligned}$$

while for the root vectors corresponding to the highest long/short roots we choose

$$\begin{aligned} E_\alpha &:= \mathcal{E}_{1,2l}, & &\alpha \text{ highest long root}, \\ E_{-\alpha} &:= \mathcal{E}_{2l,1}, & &\alpha \text{ highest long root}, \\ E_\alpha &:= \mathcal{E}_{1,2l-1} + \mathcal{E}_{2,2l}, & &\alpha \text{ highest short root}, \\ E_{-\alpha} &:= \mathcal{E}_{2l-1,1} + \mathcal{E}_{2l,2}, & &\alpha \text{ highest short root}. \end{aligned}$$

Example 2.15. $\mathfrak{d}_l$ is the root system of the semi-simple Lie algebra of all block matrices of size $2l$ of the form

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}, \qquad \text{where} \qquad \begin{aligned} A\Delta + \Delta D^\top &= 0, \\ B\Delta + \Delta B^\top &= 0, \\ C\Delta + \Delta C^\top &= 0, \end{aligned}$$

and where $A, \dots, D$ are square matrices of size l. For the root vectors of height ± 1 we choose

$$\begin{aligned} E_{\alpha_i} &:= \mathcal{E}_{i,i+1} - \mathcal{E}_{2l-i,2l-i+1}, & i &= 1, \dots, l-1, \\ E_{-\alpha_i} &:= \mathcal{E}_{i+1,i} - \mathcal{E}_{2l-i+1,2l-i}, & i &= 1, \dots, l-1, \\ E_{\alpha_l} &:= \mathcal{E}_{l-1,l+1} - \mathcal{E}_{l,l+2}, \\ E_{-\alpha_l} &:= \mathcal{E}_{l+1,l-1} - \mathcal{E}_{l+2,l}, \end{aligned}$$

while for the root vectors corresponding to the highest (long = short) root α we choose

$$\begin{aligned} E_\alpha &:= \mathcal{E}_{1,2l-1} - \mathcal{E}_{2,2l}, \\ E_{-\alpha} &:= \mathcal{E}_{2l-1,1} - \mathcal{E}_{2l,2}. \end{aligned}$$

2.3.2 Invariant Functions and Exponents

We have defined in Section 2.2 the algebra of Ad-invariant functions on $\mathfrak{g}$ and the algebra of Ad*-invariant functions on $\mathfrak{g}^*$. If $\mathfrak{g}$ is simple then these are isomorphic algebras since they correspond to each other under the isomorphism $\mathfrak{g} \to \mathfrak{g}^*$, defined by the Killing form. Moreover, they are isomorphic to a polynomial algebra which is generated by l homogeneous elements, since one proves that

$$\mathcal{F}(\mathfrak{g}^*)^{\mathbf{G}} \cong \mathcal{F}(\mathfrak{g})^{\mathbf{G}} \cong \mathbf{C}[I_1, \dots, I_l].$$

We define the *exponents* of $\mathfrak{g}$ to be the l integers $(m_1, \dots, m_l)$, where $m_i := (\deg I_i) - 1$. One has that $m_i + m_{l-i+1}$ is independent of i, and is equal to the so-called *Coxeter number* of $\mathfrak{g}$. It is a fundamental fact that the order of the Weyl group is given by

$$\#\mathbf{W} = \prod_{i=1}^{l}(m_i + 1).$$

For future use, we also give an alternative formula for the latter, namely let $\mathfrak{h}$ by a Cartan subalgebra of $\mathfrak{g}$ and let Π be a system of simple roots with respect to $\mathfrak{h}$, with Cartan matrix A. Then

$$\#\mathbf{W} = l! \prod_{i=1}^{l} \eta_i \det A, \tag{2.24}$$

where $\eta_1, \dots, \eta_l$ are the coefficients of the highest long root with respect to Π. For a proof of this (non-trivial) fact, see [37, Chapter VI no 2.4].

Below we give the list of exponents for all simple Lie algebras. Notice that a Lie algebra and its dual have the same exponents (i.e., for $\mathfrak{b}_l$ and $\mathfrak{c}_l$ they are the same; the other Lie algebras coincide with their duals, hence for those the statement is trivial).

The following proposition will play an important role in the study of Toda lattices (see Section 9.2).

Proposition 2.16. *Let $N := \operatorname{diag}(n_1, \dots, n_l)$, where the integers n_i are defined by $\sum_{\alpha \in \Phi_+} H_\alpha = \sum_{i=1}^{l} n_i H_i$ and consider the linear operator*

$$\begin{aligned} \Psi : \mathfrak{h} &\to \mathfrak{h} \\ X &\mapsto \sum_{i=1}^{l} n_i H_i \langle h_i \mid X \rangle, \end{aligned}$$

whose matrix is NA, in the basis $(H_1, \dots, H_l)$. The spectrum of Ψ, and hence of NA, is expressible in terms of the exponents of $\mathfrak{g}$ as follows:

$$\operatorname{Spec}(\Psi) = \{m_1(m_1 + 1), \dots, m_l(m_l + 1)\}.$$

Table 2.2. More data on simple Lie algebras: for each type we give the l exponents and the Coxeter number, which is the sum of the i-th and $(l+1-i)$-th exponents. The dual Weyl integers n_i are defined in Proposition 2.16.

$\mathfrak{g}$	Exponents	Coxeter	n_i or $(n_1, \dots, n_l)$
$\mathfrak{a}_l$	$1, 2, \dots, l$	$l+1$	$i(l-i+1)$
$\mathfrak{b}_l$	$1, 3, 5, \dots, 2l-1$	$2l$	$i(2l-i+1) - \delta_{il}\binom{l+1}{2}$
$\mathfrak{c}_l$	$1, 3, 5, \dots, 2l-1$	$2l$	$i(2l-i)$
$\mathfrak{d}_l$	$1, 3, \dots, 2l-3, l-1$	$2l-2$	$i(2l-i-1) - (\delta_{il} + \delta_{i,l-1})\binom{l}{2}$
$\mathfrak{e}_6$	$1, 4, 5, 7, 8, 11$	12	$(16, 22, 30, 42, 30, 16)$
$\mathfrak{e}_7$	$1, 5, 7, 9, 11, 13, 17$	18	$(34, 49, 66, 96, 75, 52, 27)$
$\mathfrak{e}_8$	$1, 7, 11, 13, 17, 19, 23, 29$	30	$2(46, 68, 91, 135, 110, 84, 57, 29)$
$\mathfrak{f}_4$	$1, 5, 7, 11$	12	$(16, 30, 42, 22)$
$\mathfrak{g}_2$	$1, 5$	6	$(10, 6)$

Proof. The proposition can be checked case by case by going through the list of simple Lie algebras (see Example 2.17 below). We give a representation theory proof, which was provided to us by Eric Sommers. Let (e, f, h) be an *S-triplet* for $\mathfrak{g}$, i.e., they are non-zero elements of $\mathfrak{g}$ which satisfy the standard $\mathfrak{sl}(2)$ commutation relations

$$[h, e] = 2e, \quad [h, f] = -2f, \quad [e, f] = h.$$

Let us denote the corresponding adjoint representation of $\mathfrak{sl}(2)$ on $\mathfrak{g}$ by χ. We will suppose that the S-triplet is a *principal S-triplet* which means that $\mathfrak{g}^e$, the *centralizer* of e (the subspace of all elements that commute with e), satisfies

$$\dim \mathfrak{g}^e = \min \{\mathfrak{g}^x \mid x \in \mathfrak{g}\}. \tag{2.25}$$

Then the representation χ decomposes in precisely $l = \operatorname{Rk} \mathfrak{g}$ irreducible sub-representations, $\chi_i : \mathfrak{sl}(2) \to \operatorname{End}(S_i)$ $(i = 1, \dots, l)$, where $\dim S_i = 2m_i + 1$. Decomposing S_i further into eigenspaces of $[h\,, \cdot]$ we can write for a fixed $1 \leqslant i \leqslant l$

$$S_i = \mathbf{C}v_{i,-m_i} \oplus \mathbf{C}v_{i,1-m_i} \oplus \cdots \oplus \mathbf{C}v_{i,m_i}$$

where the action of e, f and h is described by

$$\begin{aligned} [e, v_{i,j}] &= (m_i + j + 1)v_{i,j+1}, \\ [h, v_{i,j}] &= 2j v_{i,j}, \qquad\qquad j = -m_i, \dots, m_i, \\ [f, v_{i,j}] &= (m_i - j + 1)v_{i,j-1}, \end{aligned}$$

where $v_{i,-m_i-1} = v_{i,m_i+1} = 0$. Notice that

$$\mathfrak{h} = \bigoplus_{i=1}^{l} \mathbf{C} v_{i,0}$$

so that $\mathrm{ad}_f \circ \mathrm{ad}_e$ restricts to an endomorphism ψ of $\mathfrak{h}$, which is given by $\psi(v_{i,0}) = m_i(m_i + 1)v_{i,0}$ for $i = 1, \dots, l$. Thus, ψ has as eigenvalues the integers $m_i(m_i + 1)$, for $i = 1, \dots, l$.

We wish to relate ψ to Ψ. To do this we pick a particular principal S-triplet (all principal S-triplets are conjugate to each other). Choose a Chevalley basis $(H_1, \dots, H_l) \cup (E_\alpha)_{\alpha \in \Phi}$ for $\mathfrak{g}$ and let

$$h := \sum_{\alpha \in \Phi_+} H_\alpha = \sum_{i=1}^{l} n_i H_i \tag{2.26}$$

where the latter equality is a definition of the positive integers $n_1, \dots, n_l$. The element h, defined by (2.26), is called the *dual Weyl element* and the integers $n_1, \dots, n_l$ are called the *dual Weyl integers*. We define for $\alpha \in \Phi$ a reflection on $\mathfrak{h}$ in analogy with the reflection s_α on $\mathfrak{h}^*$, which was defined in (2.21). For $X \in \mathfrak{h}$ let

$$\sigma_\alpha(X) := X - \langle \alpha, X \rangle \, H_\alpha \tag{2.27}$$

It is the reflection with respect to the hyperplane orthogonal to the coroot H_α. Indeed, if X is orthogonal to H_α then $\langle \alpha, X \rangle = \langle h_\alpha \mid X \rangle = 0$, hence σ_α fixes X, while $\sigma_\alpha(H_\alpha) = -H_\alpha$ because $\langle \alpha, H_\alpha \rangle = 2$. For $1 \leqslant i \leqslant l$ the reflection σ_{α_i} permutes all coroots H_α, with $\alpha \in \Phi^+ \setminus \{\alpha_i\}$, so that

$$\sigma_{\alpha_i}\left(\sum_{\alpha \in \Phi^+} H_\alpha\right) = \sum_{\alpha \in \Phi^+} H_\alpha - 2H_i. \tag{2.28}$$

Combining (2.27) and (2.28) we find that

$$\langle \alpha_i, h \rangle = \left\langle \alpha_i, \sum_{\alpha \in \Phi^+} H_\alpha \right\rangle = 2, \qquad \text{for } i = 1, \dots, l, \tag{2.29}$$

which characterizes the dual Weyl element. Furthermore, let e and f be defined by

$$e := \sum_{i=1}^{l} E_{\alpha_i}, \qquad f := \sum_{i=1}^{l} n_i E_{-\alpha_i}.$$

e satisfies (2.25) and e, f, h satisfy the $\mathfrak{sl}(2)$ commutation relations, as follows from Chevalley's Theorem (Theorem 2.11) and (2.29). Thus, (e, f, h) is a principal S-triplet. For $k = 1, \ldots, l$ we have that

$$\begin{aligned}
\psi(H_k) &= \mathrm{ad}_f \circ \mathrm{ad}_e\, H_k \\
&= \left[\sum_{i=1}^{l} n_i E_{-\alpha_i}, \left[\sum_{j=1}^{l} E_{\alpha_j}, H_k\right]\right] \\
&= -\sum_{i,j=1}^{l} \left[n_i E_{-\alpha_i}, \langle \alpha_j, H_k \rangle\, E_{\alpha_j}\right] \\
&= \sum_{i=1}^{l} n_i H_i \langle \alpha_i, H_k \rangle \\
&= \sum_{i=1}^{l} n_i H_i \langle h_i \mid H_k \rangle \\
&= \Psi(H_k),
\end{aligned}$$

showing that $\psi = \Psi$. Since we have shown that the eigenvalues of ψ are the integers $m_i(m_i + 1)$, where $i = 1, \ldots, l$, this yields the announced eigenvalues for Ψ, and hence for the matrix NA. $\square$

Example 2.17. Let us verify Proposition 2.16 by direct computation for one of the simple Lie algebras, say for $\mathfrak{f}_4$. We find from the last columns of Tables 2.1 and 2.2 that

$$A = \begin{pmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -2 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{pmatrix} \qquad N = \begin{pmatrix} 16 & 0 & 0 & 0 \\ 0 & 30 & 0 & 0 \\ 0 & 0 & 42 & 0 \\ 0 & 0 & 0 & 22 \end{pmatrix}.$$

The exponents of $\mathfrak{f}_4$ are $(1, 5, 7, 11)$, as can be read of from the second column of Table 2.2, while the eigenvalues of

$$NA = \begin{pmatrix} 32 & -16 & 0 & 0 \\ -30 & 60 & -60 & 0 \\ 0 & -42 & 84 & -42 \\ 0 & 0 & -22 & 44 \end{pmatrix}$$

are given by $2 = 1.2,\ 30 = 5.6,\ 56 = 7.8,\ 132 = 11.12$, as follows from a direct computation. Clearly, this corresponds to the eigenvalues, predicted by Proposition 2.16.

2.4 Twisted Affine Lie Algebras

For any Lie algebra $\mathfrak{g}$ and for any element $g \in \mathbf{G}$, the linear map $\mathrm{Ad}_g : \mathfrak{g} \to \mathfrak{g}$ is an automorphism of $\mathfrak{g}$, which is called an *inner automorphism*. The group of outer automorphisms $\Gamma(\mathfrak{g})$ is by definition the group of all automorphisms, modulo the inner automorphisms. If $\mathfrak{g}$ is simple then any element of $\Gamma(\mathfrak{g})$ is represented by a (unique) automorphism of $\mathfrak{g}$ which is induced by an automorphism of the Dynkin diagram of $\mathfrak{g}$. Therefore, $\Gamma(\mathfrak{g})$ can be identified naturally with the group of automorphisms of the Dynkin diagram of $\mathfrak{g}$. By inspecting Table 2.1 one finds that only a few Dynkin diagrams admit a non-trivial automorphisms; those are given in Table 2.2.

Table 2.3. We list the simple Lie algebras $\mathfrak{g}$ which admit a non-trivial group $\Gamma(\mathfrak{g})$ of outer automorphisms. We give the possible values for the order of its elements.

$\mathfrak{g}$	Rank	$\Gamma(\mathfrak{g})$	order(ν)
$\mathfrak{a}_l$	$l > 1$	$\mathbf{Z}/2\mathbf{Z}$	2
$\mathfrak{d}_4$	$l = 4$	S_3	$2, 3$
$\mathfrak{d}_l$	$l > 4$	$\mathbf{Z}/2\mathbf{Z}$	2
$\mathfrak{e}_6$	6	$\mathbf{Z}/2\mathbf{Z}$	2

Let ν be an automorphism of $\mathfrak{g}$ which is induced by a diagram automorphism, and let us denote its order by m. Since $\nu^m = \mathrm{Id}_{\mathfrak{g}}$ each eigenvalue of ν has the form ε^i, where ε is a primitive m^{th} root of unity and $0 \leqslant i \leqslant m-1$. The eigenspace of ν which corresponds to this eigenvalue ε^i is denoted by $\mathfrak{g}_i$. Then the algebra $\mathfrak{g}$ admits the following finite grading:

$$\mathfrak{g} = \bigoplus_{i \in \mathbf{Z}_m} \mathfrak{g}_i \quad \text{and} \quad [\mathfrak{g}_i, \mathfrak{g}_j] \subseteq \mathfrak{g}_{i+j}.$$

We now define the (infinite-dimensional) *twisted affine Lie algebra* of $(\mathfrak{g}, \nu)$.

$$L(\mathfrak{g}, \nu) := \left\{ \sum_{j=M}^{N} \mathfrak{h}^j X_j \mid M, N \in \mathbf{Z} \text{ and } X_j \in \mathfrak{g}_{j \bmod m} \text{ for } M \leqslant j \leqslant N \right\}.$$

Notice that if we extend ν in the obvious way to elements of the form $X(\mathfrak{h}) = \sum_{j=M}^{N} \mathfrak{h}^j X_j$, then elements of $L(\mathfrak{g}, \nu)$ are characterized by the property $X(\varepsilon^p \mathfrak{h}) = \nu^p X(\mathfrak{h})$, for $p = 1, \ldots, m-1$. When $\nu = \mathrm{Id}_{\mathfrak{g}}$ then

$$L(\mathfrak{g}) := L(\mathfrak{g}, \mathrm{Id}_{\mathfrak{g}}) = \mathfrak{g} \otimes \mathbf{C}\left[\mathfrak{h}, \mathfrak{h}^{-1}\right]$$

the *affine Lie algebra* of $\mathfrak{g}$. The term *loop algebra* is also used. A natural Lie bracket on $L(\mathfrak{g}, \nu)$ is given by

$$\left[\sum_{i \leqslant N} \mathfrak{h}^i X_i, \sum_{j \leqslant M} \mathfrak{h}^j Y_j\right] = \sum_{k \leqslant M+N} \mathfrak{h}^k \left(\sum_{i+j=k} [X_i, Y_j]\right)$$

and the Killing form $\langle \cdot \,|\, \cdot \rangle$ on $\mathfrak{g}$ leads for every $k \in \mathbf{Z}$ to a non-degenerate symmetric form on $L(\mathfrak{g}, \nu)$, denoted by $\langle \cdot \,|\, \cdot \rangle_k$ which is defined by

$$\left\langle \sum_{i \leqslant N} \mathfrak{h}^i X_i \,\middle|\, \sum_{j \leqslant M} \mathfrak{h}^j Y_j \right\rangle_k := \sum_{i+j+k=0} \langle X_i \,|\, Y_j \rangle \,. \tag{2.30}$$

It is easy to see that each of the bilinear forms $\langle \cdot \,|\, \cdot \rangle_k$ on $L(\mathfrak{g}, \nu)$ is Ad-invariant. We will refer to $\langle \cdot \,|\, \cdot \rangle_0$ as the *Killing form* of $L(\mathfrak{g}, \nu)$.

Example 2.18. Consider the direct sum decomposition

$$L(\mathfrak{g}) = L(\mathfrak{g})_+ \oplus L(\mathfrak{g})_-,$$

where $L(\mathfrak{g})_+$ consists of those elements of $L(\mathfrak{g})$ which are polynomial in $\mathfrak{h}$, while $L(\mathfrak{g})_-$ consists of all elements of $L(\mathfrak{g})$ that are polynomial in $\mathfrak{h}^{-1}$, but without constant term. In terms of the orthogonality that is induced by the Killing form $\langle \cdot \,|\, \cdot \rangle_0$ we have that $L(\mathfrak{g})_+^{\perp}$ consists of those elements of $L(\mathfrak{g})$ which are polynomial in $\mathfrak{h}$, but without constant term, while $L(\mathfrak{g})_-^{\perp}$ consists of all elements of $L(\mathfrak{g})$ that are polynomial in $\mathfrak{h}^{-1}$.

It is possible to develop a theory of roots for twisted affine Lie algebras, which is analogous to the one for simple Lie algebras. We will first start with the easier case of (untwisted) affine Lie algebras. Let $\mathfrak{g}$ be a simple Lie algebra, let $\mathfrak{h}$ be a Cartan subalgebra with root system Φ and let $\Pi = \{\alpha_1, \dots, \alpha_l\}$ be a system of simple roots. By definition, a *root* of $L(\mathfrak{g})$ is a pair $(\alpha, i) \neq (0, 0)$, where $\alpha \in \Phi \cup \{0\}$ and $i \in \mathbf{Z}$; such pairs are added in the obvious way: $(\alpha, i) + (\beta, j) = (\alpha + \beta, i + j)$. We denote the set of all roots of $L(\mathfrak{g})$ by $\bar{\Phi}$ and we call $\bar{\Phi}$ the *root system* of $L(\mathfrak{g})$. Let α_0 denote minus the highest long root of $\mathfrak{g}$, and notice that α_0 is the unique root of $L(\mathfrak{g})$ which has the property that no decomposition of the form $(\alpha_0, 1) = (\alpha, 1) + (\beta, 0)$, with $\alpha \in \Phi$ and $\beta \in \Pi$ is possible. One calls $\bar{\alpha}_0 := (\alpha_0, 1)$ the *lowest root* of $L(\mathfrak{g})$. Define

$$\bar{\Pi} = \{\bar{\alpha}_0 = (\alpha_0, 1), \bar{\alpha}_1 = (\alpha_1, 0), \dots, \bar{\alpha}_l = (\alpha_l, 0)\} \,.$$

Using the fact that $-\alpha_0$ is the highest long root of $\mathfrak{g}$, it is easy to show that every root of $L(\mathfrak{g})$ can be written uniquely as a linear combination of the elements of $\bar{\Phi}$, where all coefficients belong to $\mathbf{Z}_+$ or they all belong to $\mathbf{Z}_-$.

Thus, $\bar{\Pi}$ is the natural analogue of Π, so we will call it a *system of simple roots* for $L(\mathfrak{g})$.

To a root $(\alpha, j) \in \bar{\Phi}$ we associate the vector $E_{(\alpha,j)} := E_\alpha \mathfrak{h}^j$, where E_α is the root vector that corresponds to α (see Theorem 2.11). We will call $E_{(\alpha,j)}$ the *root vector* which corresponds to (α, j). It follows easily from Theorem 2.11 that these root vectors satisfy the following relations: for any $\bar{\alpha} = (\alpha, j) \in \bar{\Phi}$,

$$\begin{aligned} [H, E_{\bar{\alpha}}] &= \langle \alpha, H \rangle \, E_{\bar{\alpha}}, \qquad H \in \mathfrak{h}, \\ [E_{\bar{\alpha}}, E_{-\bar{\alpha}}] &= H_\alpha. \end{aligned} \tag{2.31}$$

The *Cartan matrix* A of $L(\mathfrak{g})$ is constructed from the system of simple roots as before, namely

$$a_{ij} := \langle \alpha_i, H_j \rangle \,,$$

except that the indices i, j can now also take the value 0, besides the values $1, \dots, l$ (the numbering of the rows and columns of these bigger Cartan matrices starts from 0). The Cartan matrix of $L(\mathfrak{g})$ and its Dynkin diagram are easily computed from the one of $\mathfrak{g}$ and the coefficients $\xi_1, \dots, \xi_l$ of the highest long root (these coefficients are listed in Table 2.1). Indeed, one only needs to compute the first row and the first column of A, since the remaining block is precisely the Cartan matrix of $\mathfrak{g}$. In order to compute the first row of A, whose first element a_{00} is 2, it suffices to express that $\xi = (\xi_0 = 1, \xi_1, \dots, \xi_l)^\top$ is a (normalized) null-vector of A, which follows from the fact that

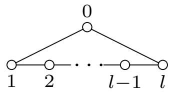
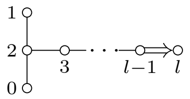
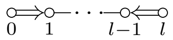
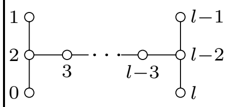
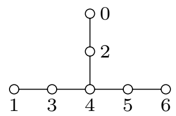
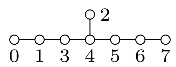
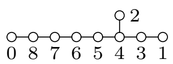
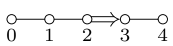
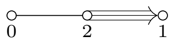
$$\alpha_0 = \sum_{i=1}^{l} \xi_i \alpha_i,$$

upon taking inner products with H_j, $0 \leqslant j \leqslant l$. Note that once again we have a duality between the system of roots $\left\{ \bar{\alpha} \mid \alpha \in \bar{\Pi} \right\}$ and coroots $\left\{ H_{\bar{\alpha}} \mid \alpha \in \bar{\Pi} \right\}$, which amounts to $A \leftrightarrow A^\top$, inducing a duality between the $L(\mathfrak{g}, \nu)$.

By a direct computation for each of the simple Lie algebras we find that there is in each case one non-zero entry in the first row, besides the leading 2. Therefore the same is true for the first column of A. That non-zero entry is then computed by expressing that the first element of $\xi^\top A$ is zero. The resulting matrices are given for each of the affine Lie algebras in Table 2.4; in this table the case $\mathfrak{a}_1^{(1)}$ and $\mathfrak{b}_2^{(1)}$ should be interpreted properly: the Cartan matrix of $\mathfrak{a}_1^{(1)}$ is $\begin{pmatrix} 2 & -2 \\ -2 & 2 \end{pmatrix}$, as follows from the fact that $\xi^\top = (1, 1)$.

We now turn to the case of twisted affine Lie algebras. Suppose that $\mathfrak{g}$ is a simple Lie algebra and that ν is an automorphism which corresponds to a non-trivial diagram automorphism of the Dynkin diagram of $\mathfrak{g}$. This means that $\mathfrak{g}$ is $\mathfrak{a}_l$ or $\mathfrak{d}_l$ or $\mathfrak{e}_6$ and the order of the automorphism ν is two, except in case $\mathfrak{d}_4$, for which we can also consider an automorphism of order 3.

Table 2.4. For each of the affine Lie algebras we give its Dynkin diagram, its Cartan matrix A, the normalized null-vectors ξ and $\hat{\xi}$ of $A^\top$ resp. of A (only one is given when they are the same) and the vector η that contains the coefficients of the highest weight vector.

$\mathfrak{g}$	Dynkin diagram	Cartan matrix	$\xi,\ \hat{\xi}$	η
$\mathfrak{a}_l^{(1)}$ $(l \geqslant 1)$		$\begin{pmatrix} 2 & -1 & & -1 \\ -1 & 2 & & \\ & & \ddots & -1 \\ -1 & & -1 & 2 \end{pmatrix}$	$\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$	$\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$
$\mathfrak{b}_l^{(1)}$ $(l>2)$		$\begin{pmatrix} 2 & 0 & -1 & 0 & & & \\ 0 & 2 & -1 & 0 & & & \\ -1 & -1 & 2 & -1 & & & \\ 0 & 0 & -1 & 2 & & & \\ & & & & \ddots & & \\ & & & & 2 & -1 & 0 \\ & & & & -1 & 2 & -2 \\ & & & & 0 & -1 & 2 \end{pmatrix}$	$\begin{pmatrix} 1 \\ 1 \\ 2 \\ \vdots \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 2 \\ \vdots \\ 2 \\ 1 \end{pmatrix}$	$\begin{pmatrix} 1 \\ 2 \\ \vdots \\ 2 \end{pmatrix}$
$\mathfrak{c}_l^{(1)}$ $(l>1)$		$\begin{pmatrix} 2 & -2 & 0 & 0 & & & \\ -1 & 2 & -1 & 0 & & & \\ 0 & -1 & 2 & -1 & & & \\ 0 & 0 & -1 & 2 & & & \\ & & & & \ddots & & \\ & & & & 2 & -1 & 0 \\ & & & & -1 & 2 & -1 \\ & & & & 0 & -2 & 2 \end{pmatrix}$	$\begin{pmatrix} 1 \\ 2 \\ \vdots \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$	$\begin{pmatrix} 2 \\ \vdots \\ 2 \\ 1 \end{pmatrix}$
$\mathfrak{d}_l^{(1)}$ $(l>3)$		$\begin{pmatrix} 2 & 0 & -1 & 0 & & & \\ 0 & 2 & -1 & 0 & & & \\ -1 & -1 & 2 & -1 & & & \\ 0 & 0 & -1 & 2 & & & \\ & & & & \ddots & & \\ & & & & 2 & -1 & -1 \\ & & & & -1 & 2 & 0 \\ & & & & -1 & 0 & 2 \end{pmatrix}$	$\begin{pmatrix} 1 \\ 1 \\ 2 \\ \vdots \\ 2 \\ 1 \\ 1 \end{pmatrix}$	$\begin{pmatrix} 1 \\ 2 \\ \vdots \\ 2 \\ 1 \\ 1 \end{pmatrix}$
$\mathfrak{e}_6^{(1)}$		$\begin{pmatrix} 2 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & -1 & 0 & 0 & 0 \\ -1 & 0 & 2 & 0 & -1 & 0 & 0 \\ 0 & -1 & 0 & 2 & -1 & 0 & 0 \\ 0 & 0 & -1 & -1 & 2 & -1 & 0 \\ 0 & 0 & 0 & 0 & -1 & 2 & -1 \\ 0 & 0 & 0 & 0 & 0 & -1 & 2 \end{pmatrix}$	$\begin{pmatrix} 1 \\ 1 \\ 2 \\ 2 \\ 3 \\ 2 \\ 1 \end{pmatrix}$	$\begin{pmatrix} 1 \\ 2 \\ 2 \\ 3 \\ 2 \\ 1 \end{pmatrix}$
$\mathfrak{e}_7^{(1)}$		$\begin{pmatrix} 2 & -1 & & & & & & \\ -1 & 2 & 0 & -1 & & & & \\ & 0 & 2 & 0 & -1 & & & \\ & -1 & 0 & 2 & -1 & & & \\ & & -1 & -1 & 2 & -1 & & \\ & & & & -1 & 2 & -1 & \\ & & & & & -1 & 2 & -1 \\ & & & & & & -1 & 2 \end{pmatrix}$	$\begin{pmatrix} 1 \\ 2 \\ 2 \\ 3 \\ 4 \\ 3 \\ 2 \\ 1 \end{pmatrix}$	$\begin{pmatrix} 2 \\ 2 \\ 3 \\ 4 \\ 3 \\ 2 \\ 1 \end{pmatrix}$
$\mathfrak{e}_8^{(1)}$		$\begin{pmatrix} 2 & 0 & & & & & & & -1 \\ 0 & 2 & 0 & -1 & & & & & \\ & 0 & 2 & 0 & -1 & & & & \\ & -1 & 0 & 2 & -1 & & & & \\ & & -1 & -1 & 2 & -1 & & & \\ & & & & -1 & 2 & -1 & & \\ & & & & & -1 & 2 & -1 & \\ & & & & & & -1 & 2 & -1 \\ -1 & & & & & & & -1 & 2 \end{pmatrix}$	$\begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 6 \\ 5 \\ 4 \\ 3 \\ 2 \end{pmatrix}$	$\begin{pmatrix} 2 \\ 3 \\ 4 \\ 6 \\ 5 \\ 4 \\ 3 \\ 2 \end{pmatrix}$
$\mathfrak{f}_4^{(1)}$		$\begin{pmatrix} 2 & -1 & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 \\ 0 & -1 & 2 & -2 & 0 \\ 0 & 0 & -1 & 2 & -1 \\ 0 & 0 & 0 & -1 & 2 \end{pmatrix}$	$\begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 3 \\ 2 \\ 1 \end{pmatrix}$	$\begin{pmatrix} 2 \\ 3 \\ 4 \\ 2 \end{pmatrix}$
$\mathfrak{g}_2^{(1)}$		$\begin{pmatrix} 2 & 0 & -1 \\ 0 & 2 & -1 \\ -1 & -3 & 2 \end{pmatrix}$	$\begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}$	$\begin{pmatrix} 3 \\ 2 \end{pmatrix}$

By definition a *root* of $L(\mathfrak{g}, \nu)$ is a pair $(\alpha, j) \neq (0,0)$, with $\alpha \in \mathfrak{h}^*$ and $j \in \mathbf{Z}$, such that the joint eigenspace

$$\{X \in \mathfrak{g}_j \mid [H, X] = \langle \alpha, H \rangle X \text{ for any } H \in \mathfrak{h}\}$$

is non-trivial (in the untwisted case this definition is equivalent to the one that we have given). We denote the set of all roots of $L(\mathfrak{g}, \nu)$ by $\bar{\Phi}$ and call it the *root system* of $L(\mathfrak{g}, \nu)$. There are two main differences with the case of (untwisted) affine Lie algebras. First, the Lie algebra $\mathfrak{g}_0$ is different from $\mathfrak{g}$, but it is still one of the simple Lie algebras; for each $(\mathfrak{g}, \nu)$ with $\nu \neq \mathrm{Id}_{\mathfrak{g}}$ the corresponding simple Lie algebra $\mathfrak{g}_0$ is given in Table 2.5. Second, the root α_0 which is used to define the lowest root $\bar{\alpha}_0 = (\alpha_0, 1)$ of $L(\mathfrak{g})$, takes now the following form

$$\begin{aligned} \alpha_0 &= -2(\text{highest short root of } \mathfrak{g}_0) && \text{if } \mathfrak{g} = \mathfrak{a}_{2l}, \\ \alpha_0 &= -(\text{highest short root of } \mathfrak{g}_0) && \text{otherwise.} \end{aligned} \tag{2.32}$$

The computation of the Cartan matrices is the same as in the case of the (untwisted) affine Lie algebras. The results are displayed in Table 2.5; in this table the case $\mathfrak{a}_2^{(2)}$ should be interpreted properly: its Cartan matrix is $\begin{pmatrix} 2 & -4 \\ -1 & 2 \end{pmatrix}$, as follows from the fact that $\xi^\top = (1, 2)$. A system of simple roots $\bar{\Phi}$ can be constructed also in the twisted case, where each element now belongs to $\mathfrak{g}_0$ (see [79, pp. 505–507] for explicit formulas). It leads, as before to the same formulas (2.31) for the simple roots $\bar{\alpha}$.

The Cartan matrices that we see in Tables 2.4 and 2.5 are characterized by a few of their properties, just as in the case of the Cartan matrix of a simple Lie algebra, yielding a different approach to affine Lie algebras. Start with a collection Π of $n+1$ non-zero vectors $\alpha_0, \dots, \alpha_n$ in $\mathbf{R}^n$ and let $\langle \cdot \mid \cdot \rangle$ be an inner product on $\mathbf{R}^n$. We will say that Π is an *indecomposable system of vectors* if Π cannot be split in two sets Π_1 and Π_2 such that $\langle \Pi_1 \mid \Pi_2 \rangle = 0$. The *Cartan matrix* of Π is by definition the $(n+1) \times (n+1)$ matrix A, which is defined by

$$a_{ij} := 2 \frac{\langle \alpha_i \mid \alpha_j \rangle}{\langle \alpha_j \mid \alpha_j \rangle}.$$

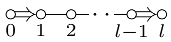
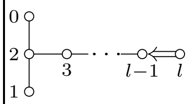
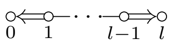
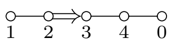
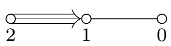
Then one has the following proposition.

Proposition 2.19. *Let Π be a collection of $n+1$ non-zero vectors $\alpha_0, \dots, \alpha_n$ in $(\mathbf{R}^n, \langle \cdot \mid \cdot \rangle)$ and denote its Cartan matrix by A. Suppose that Π and A satisfy the following three properties:*

(1) Π is an indecomposable system of vectors;
(2) Π spans $\mathbf{R}^n$;
(3) $a_{ij} \in \mathbf{Z}_-$ for $0 \leqslant i < j \leqslant n$.

Then A is the Cartan matrix of a (twisted) affine Lie algebra.

Table 2.5. For each of the twisted affine Lie algebras we give the type of $\mathfrak{g}_0$, its Dynkin diagram, its Cartan matrix A, the normalized null-vectors ξ and $\hat{\xi}$ of $A^\top$ resp. of A and the vector η that contains the coefficients of the highest weight vector of $\mathfrak{g}_0$.

$\mathfrak{g}$	$\mathfrak{g}_0$	Dynkin diagram	Cartan matrix	$\xi,\ \hat{\xi}$	η
$\mathfrak{a}_{2l}^{(2)}$ $(l \geqslant 1)$	$\mathfrak{b}_l$		$\begin{pmatrix} 2 & -2 & 0 & 0 & & & \\ -1 & 2 & -1 & 0 & & & \\ 0 & -1 & 2 & -1 & & & \\ 0 & 0 & -1 & 2 & & & \\ & & & & \ddots & & \\ & & & & 2 & -1 & 0 \\ & & & & -1 & 2 & -2 \\ & & & & 0 & -1 & 2 \end{pmatrix}$	$\begin{pmatrix} 1 \\ 2 \\ \vdots \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ \vdots \\ 1 \\ 1/2 \end{pmatrix}$	$\begin{pmatrix} 1 \\ 2 \\ \vdots \\ 2 \end{pmatrix}$
$\mathfrak{a}_{2l-1}^{(2)}$ $(l > 2)$	$\mathfrak{c}_l$		$\begin{pmatrix} 2 & 0 & -1 & 0 & & & \\ 0 & 2 & -1 & 0 & & & \\ -1 & -1 & 2 & -1 & & & \\ 0 & 0 & -1 & 2 & & & \\ & & & & \ddots & & \\ & & & & 2 & -1 & 0 \\ & & & & -1 & 2 & -1 \\ & & & & 0 & -2 & 2 \end{pmatrix}$	$\begin{pmatrix} 1 \\ 1 \\ 2 \\ \vdots \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 2 \\ \vdots \\ 2 \end{pmatrix}$	$\begin{pmatrix} 2 \\ \vdots \\ 2 \\ 1 \end{pmatrix}$
$\mathfrak{d}_{l+1}^{(2)}$ $(l > 1)$	$\mathfrak{b}_l$		$\begin{pmatrix} 2 & -1 & 0 & 0 & & & \\ -2 & 2 & -1 & 0 & & & \\ 0 & -1 & 2 & -1 & & & \\ 0 & 0 & -1 & 2 & & & \\ & & & & \ddots & & \\ & & & & 2 & -1 & 0 \\ & & & & -1 & 2 & -2 \\ & & & & 0 & -1 & 2 \end{pmatrix}$	$\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ \vdots \\ 2 \\ 1 \end{pmatrix}$	$\begin{pmatrix} 1 \\ 2 \\ \vdots \\ 2 \end{pmatrix}$
$\mathfrak{e}_6^{(2)}$	$\mathfrak{f}_4$		$\begin{pmatrix} 2 & 0 & 0 & 0 & -1 \\ 0 & 2 & -1 & 0 & 0 \\ 0 & -1 & 2 & -2 & 0 \\ 0 & 0 & -1 & 2 & -1 \\ -1 & 0 & 0 & -1 & 2 \end{pmatrix}$	$\begin{pmatrix} 1 \\ 1 \\ 2 \\ 3 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 4 \\ 3 \\ 2 \end{pmatrix}$	$\begin{pmatrix} 2 \\ 3 \\ 4 \\ 2 \end{pmatrix}$
$\mathfrak{d}_4^{(3)}$	$\mathfrak{g}_2$		$\begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -3 & 2 \end{pmatrix}$	$\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$	$\begin{pmatrix} 3 \\ 2 \end{pmatrix}$

The reader can verify easily by merely looking at the tables that, conversely, each of the Cartan matrices of the (twisted) affine Lie algebras satisfies the above three properties.

We end this section with some definitions, valid for a twisted or untwisted affine Lie algebras $L(\mathfrak{g},\nu)$. Let $\bar{\Pi}=\{\bar{\alpha}_0,\bar{\alpha}_1,\ldots,\bar{\alpha}_l\}$ denote a system of simple roots of $L(\mathfrak{g},\nu)$. The *height* $|\bar{\alpha}|$ of a root $\bar{\alpha}\in\bar{\Phi}$ is defined by

$$|\bar{\alpha}|=\sum_{i=0}^{l} m_i, \quad \text{where} \quad \bar{\alpha}=\sum_{i=0}^{l} m_i\bar{\alpha}_i.$$

This induces a grading of $L(\mathfrak{g},\nu)$, similar to the grading (2.19) of $\mathfrak{g}$, namely,

$$L(\mathfrak{g},\nu)=\bigoplus_{k\in\mathbf{Z}} L_k, \quad [L_k,L_l]\subseteq L_{k+l},$$

where L_k is the span of all $E_{\bar{\alpha}}$, with $|\bar{\alpha}|=k$. The grading also leads to a natural operation of *transpose*:

$$\left(H+\sum_{\bar{a}\in\bar{\Phi}} a_{\bar{\alpha}}E_{\bar{\alpha}}\right)^{\top}=H+\sum_{\bar{a}\in\bar{\Phi}} a_{\bar{\alpha}}E_{-\bar{\alpha}}, \quad H\in\mathfrak{g}_0.$$