Lecture 9. THE SCHRÖDINGER REPRESENTATION

9.1 Let V be a real finite-dimensional vector space with a skew-symmetric nondegenerate bilinear form $S : V \times V \to \mathbb{R}$. For example, $V = \mathbb{R}^n \oplus \mathbb{R}^n$ with the bilinear form

$$S((\mathbf{u}', \mathbf{v}'), (\mathbf{u}'', \mathbf{v}'')) = \mathbf{u}'\mathbf{v}'' - \mathbf{u}''\mathbf{v}'. \tag{9.1}$$

We define the *Heisenberg group* $\tilde{V}$ as the central extension of V by the circle $\mathbb{C}_1^* = \{z \in \mathbb{C} : |z| = 1\}$ defined by S. More precisely, it is the set $V \times \mathbb{C}_1^*$ with the group law defined by

$$(x, \lambda) \cdot (x', \lambda') = (x + x', e^{iS(x,x')}\lambda\lambda'). \tag{9.2}$$

There is an obvious complexification $\tilde{V}_{\mathbb{C}}$ which is the extension of $V_{\mathbb{C}} = V \otimes \mathbb{C}$ with the help of $\mathbb{C}^*$.

We shall describe its Schrödinger representation. Choose a complex structure on V

$$J : V \to V$$

(i.e., an $\mathbb{R}$-linear operator with $J^2 = -I_V$) such that
(i) $S(Jx, Jx') = S(x, x')$ for all $x, x' \in V$;
(ii) $S(Jx, x') > 0$ for all $x, x' \in V$.

Since $J^2 = -I_V$, the space $V_{\mathbb{C}} = V \otimes \mathbb{C} = V + iV$ decomposes into the direct sum $V_{\mathbb{C}} = A \oplus \bar{A}$ of eigensubspaces with eigenvalues i and $-i$, respectively. Here the conjugate is an automorphism of $V_{\mathbb{C}}$ defined by $v + iw \to v - iw$. Of course,

$$A = \{v - iJv : v \in V\}, \quad \bar{A} = \{v + iJv : v \in V\}. \tag{9.3}$$

Let us extend S to $V_{\mathbb{C}}$ by $\mathbb{C}$-linearity, i.e.,

$$S(v + iw, v' + iw') := S(v, v') - S(w, w') + i(S(w, v') + S(v, w')).$$

Since

$$S(v \pm iJv, v' \pm iJv') = S(v, v') - S(Jv, Jv') \pm i(S(v, Jv') + S(Jv, v'))$$

$$= \pm i(S(Jv, J^2v') + S(Jv, v')) = \pm i(S(Jv, -v') + S(Jv, v'))$$
$$= \pm i(-S(Jv, v') + S(Jv, v')) = 0, \tag{9.4}$$

we see that A and $\bar{A}$ are complex conjugate isotropic subspaces of S in $V_{\mathbb{C}}$. Observe that

$$S(Jv, w) = S(J^2v, Jw) = S(-v, Jw) = -S(v, Jw) = S(Jw, v), \quad v, w \in V.$$

This shows that $(v, w) \to S(Jv, w)$ is a symmetric bilinear form on V. Thus the function

$$H(v, w) = S(Jv, w) + iS(v, w) \tag{9.5}$$

is a Hermitian form on (V, J). Indeed,

$$H(w, v) = S(Jw, v) + iS(w, v) = S(J^2w, Jv) - iS(v, w) = S(Jv, w) - iS(v, w) = \overline{H(v, w)},$$

$$H(iv, w) = H(Jv, w) = S(J^2v, w) + iS(Jv, w) = -S(v, w) + iS(Jv, w) =$$
$$= i(S(Jv, w) + iS(v, w)) = iH(v, w).$$

By property (ii) of S, this form is positive definite.

Conversely, given a complex structure on V and a positive definite Hermitian form H on V, we write

$$H(v, w) = Re(H(v, w)) + iIm(H(v, w))$$

and immediately verify that the function

$$S(v, w) = Im(H(v, w))$$

is a skew-symmetric real bilinear form on V, and

$$S(iv, w) = Re(H(v, w))$$

is a positive definite symmetric real bilinear form on V. The function $S(v, w)$ satisfies properties (i) and (ii).

We know from (9.4) that A and $\bar{A}$ are isotropic subspaces with respect to $S : V_{\mathbb{C}} \to \mathbb{C}$. For any $a = v - iJv, b = w - iJw \in A$, we have

$$S(a, \bar{b}) = S(v - iJv, w + iJw) = S(v, w) + S(Jv, Jw) - i(S(Jv, w) - S(v, Jw))$$

$$= 2S(v, w) - i(S(Jv, w) + S(Jw, v)) = 2S(v, w) - 2iS(Jv, w) = -2iH(v, w).$$

We set for any $a = v - iJv, b = w - iJw \in A$

$$\langle a, \bar{b} \rangle = 4H(v, w) = 2iS(a, \bar{b}). \tag{9.6}$$

9.2 The standard representation of $\tilde{V}$ associated to J is on the Hilbert space $\tilde{S}(A)$ obtained by completing the symmetric algebra $S(A)$ with respect to the inner product on $S(A)$ defined by

$$(a_1a_2 \cdots a_n | b_1b_2 \cdots b_n) = \sum_{\sigma \in \Sigma_n} \langle a_1, \bar{b}_{\sigma(1)} \rangle \cdots \langle a_n, \bar{b}_{\sigma(n)} \rangle. \tag{9.7}$$

We identify A and $\bar{A}$ with subgroups of $\tilde{V}$ by $a \to (a, 1)$ and $\bar{a} \to (\bar{a}, 1)$. Consider the space $Hol(\bar{A})$ of holomorphic functions on $\bar{A}$. The algebra $S(A)$ can be identified with the subalgebra of polynomial functions in $Hol(\bar{A})$ by considering each a as the linear function $z \to \langle a, z\rangle$ on $\bar{A}$. Let us make $\bar{A}$ act on $Hol(\bar{A})$ by translations

$$\bar{a} \cdot f(z) = f(z - \bar{a}), \quad z \in \bar{A},$$

and A by multiplication

$$a \cdot f(z) = e^{\langle a,z\rangle} f(z) = e^{2iS(a,z)} f(z).$$

Since

$$[a, \bar{b}] = [(a, 1), (\bar{b}, 1)] = (0, e^{2iS(a,\bar{b})}) = (0, e^{\langle a,\bar{b}\rangle}),$$

we should check that

$$a \circ \bar{b} = e^{\langle a,\bar{b}\rangle} \bar{b} \circ a.$$

We have

$$\bar{b} \circ a \cdot f(z) = e^{\langle a,z-\bar{b}\rangle} f(z - \bar{b}) = e^{-\langle a,\bar{b}\rangle} e^{\langle a,z\rangle} f(z - \bar{b}),$$

$$a \circ \bar{b} \cdot f(z) = e^{\langle a,z\rangle} f(z - \bar{b}) = e^{\langle a,\bar{b}\rangle} \bar{b} \circ a \cdot f(z),$$

and the assertion is verified.

This defines a representation of the group $\tilde{V}_{\mathbb{C}}$ on $Hol(\bar{A})$. We get representation of $\tilde{V}$ on $Hol(\bar{A})$ by restriction. Write $v = a + \bar{a}, \quad a \in A$. Then

$$(v, 1) = (a + \bar{a}, 1) = (a, 1) \cdot (\bar{a}, 1) \cdot (0, e^{-iS(a,\bar{a})}). \tag{9.8}$$

This gives

$$v \cdot f(z) = e^{-iS(a,\bar{a})} a \cdot (\bar{a} \cdot f(z)) = e^{-\frac{1}{2}\langle a,\bar{a}\rangle} e^{\langle a,z\rangle} f(z - \bar{a}). \tag{9.9}$$

The space $\tilde{S}(A)$ can be described by the following:

Lemma 1. *Let W be the subspace of $\mathbb{C}^A$ spanned by the characteristic functions χ_a of $\{a\}, a \in A$, with the Hermitian inner product given by*

$$\langle \chi_a, \chi_b\rangle = e^{\langle a,b\rangle},$$

where χ_a is the characteristic function of $\{a\}$. Then this Hermitian product is positive definite, and the completion of W with respect to the corresponding norm is $\tilde{S}(A)$.

Proof. To each $\xi \in A$ we assign the element $e^a = 1 + a + \frac{a^2}{2} + \ldots$ from $\tilde{S}(A)$. Now we define the map by $\chi_a \to e^a$. Since

$$\langle e^a | e^b\rangle = \langle \sum_{n=1}^{\infty} \frac{a^n}{n!} \Big| \sum_{n=1}^{\infty} \frac{b^n}{n!}\rangle = \sum_{n=1}^{\infty} (\frac{1}{n!})^2 \langle a^n | b^n\rangle$$

$$= \sum_{n=1}^{\infty} (\frac{1}{n!})^2 n! \langle a, \bar{b} \rangle^n = \sum_{n=1}^{\infty} \frac{1}{n!} \langle a, \bar{b} \rangle^n = e^{\langle a, \bar{b} \rangle}, \tag{9.10}$$

we see that the map $W \to \tilde{S}(A)$ preserves the inner product. We know from (9.7) that the inner product is positive definite. It remains to show that the functions e^a span a dense subspace of $\tilde{S}(A)$. Let F be the closure of the space they span. Then, by differentiating e^{ta} at $t = 0$, we obtain that all a^n belong to F. By the theorem on symmetric functions, every product $a_1 \cdots a_n$ belongs to F. Thus F contains $\tilde{S}(A)$ and thus must coincide with $\tilde{S}(A)$.

We shall consider e^a as a holomorphic function $z \to e^{\langle a, z \rangle}$ on $\bar{A}$.

Let us see how the group $\tilde{V}$ acts on the basis vectors e^a. Its center $\mathbb{C}_1^*$ acts by

$$(0, \lambda) \cdot e^a = e^{\lambda a}. \tag{9.11}$$

We have

$$\bar{b} \cdot e^a = \bar{b} \cdot e^{\langle a, z \rangle} = e^{\langle a, z - \bar{b} \rangle} = e^{-\langle a, \bar{b} \rangle} e^{\langle a, z \rangle} = e^{-\langle a, \bar{b} \rangle} e^a,$$

$$b \cdot e^a = b \cdot e^{\langle a, z \rangle} = e^{\langle b, z \rangle} e^{\langle a, z \rangle} = e^{a+b}.$$

Let $v = \bar{b} + b \in V$, then

$$(v, 1) = (b + \bar{b}, 1) = (b, 1) \cdot (\bar{b}, 1) \cdot (0, e^{-iS(b, \bar{b})}).$$

Hence

$$v \cdot e^a = e^{-iS(b, \bar{b})} b \cdot (\bar{b} \cdot e^a) = e^{-\frac{1}{2}\langle b, \bar{b} \rangle} e^{-\langle a, \bar{b} \rangle} e^{a+b} = e^{-\frac{1}{2}\langle b, \bar{b} \rangle - \langle a, \bar{b} \rangle} e^{a+b}. \tag{9.12}$$

In particular, using (9.10), we obtain

$$||v \cdot e^a||^2 = \langle e^a | e^a \rangle = e^{-\langle b, \bar{b} \rangle - 2\langle a, \bar{b} \rangle} ||e^{a+b}||^2 = e^{-\langle b, \bar{b} \rangle - 2\langle a, \bar{b} \rangle} \langle e^{a+b} | e^{a+b} \rangle =$$

$$= e^{-\langle b, \bar{b} \rangle - 2\langle a, \bar{b} \rangle} e^{\langle a+b, \bar{a}+\bar{b} \rangle} = e^{-\langle b, \bar{b} \rangle - 2\langle a, \bar{b} \rangle} e^{\langle a, \bar{a} \rangle + 2\langle a, \bar{b} \rangle + \langle b, \bar{b} \rangle} = e^{\langle a, \bar{a} \rangle} = ||e^a||^2.$$

This shows that the representation of $\tilde{V}$ on $\tilde{S}(A)$ is unitary. This representation is called the *Schrödinger representation* of $\tilde{V}$.

Theorem 1. *The Schrödinger representation of the Heisenberg group $\tilde{V}$ is irreducible.*

Proof. The group $\mathbb{C}^*$ acts on $\tilde{S}(A)$ by scalar multiplication on the arguments and, in this way defines a grading:

$$\tilde{S}(A)_k = \{f(z) : f(\lambda \cdot z) = \lambda^k f(z)\}, \quad k \geq 0.$$

The induced grading on $S(A)$ is the usual grading of the symmetric algebra. Let W be an invariant subspace in $\tilde{S}(A)$. For any $w \in W$ we can write

$$w(z) = \sum_{k=0}^{\infty} w_k(z), \quad w_k(z) \in S(A)_k. \tag{9.13}$$

Replacing $w(z)$ with $w(\lambda z)$, we get

$$w(\lambda z) = \sum_{k=0}^{\infty} w_k(\lambda z) = \sum_{k=0}^{\infty} \lambda^k w_k(z) \in W.$$

By continuity, we may assume that this is true for all $\lambda \in \mathbb{C}$. Differentiating in t, we get

$$\frac{dw(\lambda z)}{d\lambda}(0) = \lim_{t\to 0} \frac{1}{t}[w((\lambda+t)z) - w(\lambda z)] = w_1(z) \in W.$$

Continuing in this way, we obtain that all w_k belong to W. Now let $\tilde{S}(A) = W \perp W'$, where W' is the orthogonal complement of W. Then W' is also invariant. Let $1 = w + w'$, where $w \in W$, $w' \in W'$. Then writing w and w' in the form (9.13), we obtain that either $1 \in W$, or $1 \in W'$. On the other hand, formula (9.12) shows that all basis functions e^a can be obtained from 1 by operators from $\tilde{V}$. This implies $W = \tilde{S}(A)$ or $W' = \tilde{S}(A)$. This proves that $W = \{0\}$ or $W = \tilde{S}(A)$.

9.3 From now on we shall assume that V is of dimension $2n$. The corresponding complex space (V, J) is of course of dimension n. Pick some basis $e_1, \ldots, e_n$ of the complex vector space (V, J) such that $e_1, Je_1, \ldots, e_n, Je_n$ is a basis of the real vector space V. Since $e_i + iJe_i = i(Je_i + iJJe_i)$ we see that the vectors $e_i + iJe_i, i = 1, \ldots, n$, form a basis of the linear subspace $\bar{A}$ of $V_{\mathbb{C}}$. Let us identify $\bar{A}$ with $\mathbb{C}^n$ by means of this basis. Thus we shall use $\mathbf{z} = (z_1, \ldots, z_n)$ to denote both a general point of this space as well as the set of coordinate holomorphic functions on $\mathbb{C}^n$. We can write

$$\mathbf{z} = \mathbf{x} + i\mathbf{y} = (x_1, \ldots, x_n) + i(y_1, \ldots, y_n)$$

for some $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. We denote by $\mathbf{z}\cdot\mathbf{z}'$ the usual dot product in $\mathbb{C}^n$. It defines the standard unitary inner product $\mathbf{z} \cdot \bar{\mathbf{z}}'$. Let

$$||\mathbf{z}||^2 = |z_1|^2 + \ldots |z_n|^2 = (x_1^2 + y_1^2) + \ldots + (x_n^2 + y_n^2) = ||\mathbf{x}||^2 + |\mathbf{y}||^2$$

be the corresponding norm squared.

Let us define a unitary isomorphism between the space $L^2(\mathbb{R}^n)$ and $\tilde{S}(A)$. First let us identify the space $\tilde{S}(A)$ with the space of holomorphic functions on on $\bar{A}$ wich are square-integrable with respect to some gaussian measure on $\bar{A}$. More precisely, consider the measure on $\mathbb{C}^n$ defined by

$$\int_{\mathbb{C}^n} f(\mathbf{z}) d\mu = \int_{\mathbb{C}^n} f(\mathbf{z}) e^{-\pi||\mathbf{z}||^2} d\mathbf{z} := \int_{\mathbb{R}^{2n}} e^{-\pi(||\mathbf{x}||^2 + ||\mathbf{y}||^2)} d\mathbf{x} d\mathbf{y}. \tag{9.14}$$

Notice that the factor π is chosen in order that

$$\int_{\mathbb{C}^n} e^{-\pi||\mathbf{z}||^2} d\mathbf{z} = \Big(\int_{\mathbb{R}} e^{-\pi(x^2+y^2)} dx dy\Big)^n = \Big(\int_0^{2\pi}\int_{\mathbb{R}} e^{-r^2} r dr d\theta\Big)^n = \Big(\int_{\mathbb{R}} e^{-t} dt\Big)^n = 1.$$

Thus our measure μ is a probability measure (called the *Gaussian measure* on $\mathbb{C}^n$).

Let

$$H_n = \{f(\mathbf{z}) \in \mathrm{Hol}(\mathbb{C}^n) : \int\limits_{\mathbb{C}^n} |f(\mathbf{z})|^2 e^{-\pi||\mathbf{z}||^2} d\mathbf{z} \quad \text{converges}\}.$$

Let

$$f(\mathbf{z}) = \sum_{\mathbf{i}} a_{\mathbf{i}} \mathbf{z}^{\mathbf{i}} = \sum_{i_1,\ldots,i_n=0} a_{i_1,\ldots,i_n} z_1^{i_1} \cdots z_n^{i_n}$$

be the Taylor expansion of $f(\mathbf{z})$. We have

$$\int\limits_{\mathbb{C}^n} |f(\mathbf{z})|^2 d\mu = \lim_{r\to\infty} \int\limits_{\max\{|z_i|\}\leq r} |f(\mathbf{z})|^2 d\mu =$$

$$= \sum_{\mathbf{i},\mathbf{j}=1}^{\infty} a_{\mathbf{i}} \bar{a}_{\mathbf{j}} \lim_{r\to\infty} \int\limits_{\max\{|z_i|\}\leq r} \mathbf{z}^{\mathbf{i}} \bar{\mathbf{z}}^{\mathbf{j}} d\mu = \sum_{\mathbf{i},\mathbf{j}} a_{\mathbf{i}} \bar{a}_{\mathbf{j}} \int\limits_{\mathbb{C}^n} \mathbf{z}^{\mathbf{j}} d\mu$$

Now, one can show that

$$\int\limits_{\mathbb{C}^n} \mathbf{z}^{\mathbf{i}} \bar{\mathbf{z}}^{\mathbf{j}} d\mu = (i)^n \int\limits_{\mathbb{C}^n} \mathbf{z}^{\mathbf{i}} \bar{\mathbf{z}}^{\mathbf{j}} e^{-\pi \mathbf{z}\bar{\mathbf{z}}} d\mathbf{z} d\bar{\mathbf{z}} = (i)^n \int\limits_{\mathbb{C}^n} \mathbf{z}^{\mathbf{i}} \bar{\mathbf{z}}^{\mathbf{j}} e^{-\pi \mathbf{z}\bar{\mathbf{z}}} d\mathbf{z} d\bar{\mathbf{z}} = 0 \quad \text{if} \quad \mathbf{i} \neq \mathbf{j},$$

and

$$\int\limits_{\mathbb{C}^n} \mathbf{z}^{\mathbf{i}} \bar{\mathbf{z}}^{\mathbf{i}} d\mu = \prod_{k=1}^{n} \int\limits_{\mathbb{R}} \int\limits_{\mathbb{R}} (x_k^2 + y_k^2)^{i_k} e^{-\pi(x_k^2+y_k^2)} dx dy = \prod_{k=1}^{n} \int\limits_{0}^{2\pi} \int\limits_{\mathbb{R}} r^{2i_k} e^{-\pi r^2} r dr d\theta =$$

$$= \prod_{k=1}^{n} \pi \int\limits_{\mathbb{R}} t^{i_k} e^{-\pi t} dt = \prod_{k=1}^{n} i_k!/\pi^{i_k} = \mathbf{i}! \pi^{-|\mathbf{i}|}.$$

Here the absolute value denotes the sum $i_1 + \ldots + i_n$ and $\mathbf{i}! = i_1! \cdots i_n!$. Therefore, we get

$$\int\limits_{\mathbb{C}^n} |f(\mathbf{z})|^2 d\mu = \sum_{\mathbf{i}} |a_{\mathbf{i}}|^2 \pi^{-|\mathbf{i}|} \mathbf{i}!.$$

From this it follows easily that

$$\int\limits_{\mathbb{C}^n} f(\mathbf{z}) \bar{\phi}(\mathbf{z}) d\mu = \sum_{\mathbf{i}} a_{\mathbf{i}} \bar{b}_{\mathbf{i}} \pi^{-|\mathbf{i}|} \mathbf{i}!,$$

where $\bar{b}_{\mathbf{i}}$ are the Taylor coefficients of $\phi(\mathbf{z})$. Thus, if we use the left-hand-side for the definition of the inner product in H_n, we get an orthonormal basis formed by the functions

$$\phi_{\mathbf{i}}(\mathbf{z}) = (\frac{\pi^{|\mathbf{i}|}}{\mathbf{i}!})^{\frac{1}{2}} \mathbf{z}^{\mathbf{i}}. \tag{9.15}$$

Lemma 2. *H_n is a Hilbert space and the ordered set of functions $\psi_{\mathbf{i}}$ is an orthonormal basis.*

Proof. By using Taylor expansion we can write each $\psi(\mathbf{z}) \in H_n$ as an infinite series

$$\psi(\mathbf{z}) = \sum_{\mathbf{i}} c_{\mathbf{i}}\phi_{\mathbf{i}}. \tag{9.16}$$

It converges absolutely and uniformly on any bounded subset of $\mathbb{C}^n$. Conversely, if ψ is equal to such an infinite series which converges with respect to the norm in H_n, then $c_{\mathbf{i}} = (\psi, \phi_{\mathbf{i}})$ and

$$||\psi||^2 = \sum_{\mathbf{i}} |c_p|^2 < \infty.$$

By the Cauchy-Schwarz inequality,

$$\sum_{\mathbf{i}} |c_{\mathbf{i}}||\phi_{\mathbf{i}}| \leq (\sum_{\mathbf{i}} |c_{\mathbf{i}}|^2)^{\frac{1}{2}} e^{\pi||\mathbf{z}||^2/2}.$$

This shows that the series (9.16) converges absolutely and uniformly on every bounded subset of $\mathbb{C}^n$. By a well-known theorem from complex analysis, the limit is an analytic function on $\mathbb{C}^n$. This proves the completeness of the basis $(\psi_{\mathbf{i}})$.

Let us look at our functions e^a from $\tilde{S}(\bar{A})$. Choose coordinates $\zeta = (\zeta_1, \ldots, \zeta_n)$ in A such that, for any $\zeta = (\zeta_1, \ldots, \zeta_n) \in A, z = (z_1, \ldots, z_n) \in \bar{A}$,

$$\langle \zeta, z\rangle = \pi(\bar{\zeta}_1 z_1 + \ldots + \bar{\zeta}_n z_n) = \pi\bar{\zeta} \cdot \mathbf{z}.$$

We have

$$e^{\zeta} = e^{\pi(\zeta_1 z_1 + \ldots \zeta_n z_n)} = e^{\pi\zeta\cdot\mathbf{z}} = \sum_{n=1}^{\infty} \frac{\pi^n(\zeta\cdot\mathbf{z})^n}{n!} = \sum_{n=1}^{\infty} \frac{\pi^n}{n!}\Big(\sum_{|\mathbf{i}|=n} \frac{n!}{\mathbf{i}!}\zeta^{\mathbf{i}}\mathbf{z}^{\mathbf{i}}\Big) = \sum_{\mathbf{i}} \phi_{\mathbf{i}}(\zeta)\phi_{\mathbf{i}}(\mathbf{z}).$$

Comparing the norms, we get

$$(e^{\pi\zeta\cdot\mathbf{z}}, e^{\pi\zeta\cdot\mathbf{z}})_{H_n} = \sum_{\mathbf{i}} |\phi_{\mathbf{i}}(\zeta)|^2 = \sum_{\mathbf{i}} \frac{\pi^{|\mathbf{i}|}}{\mathbf{i}!}|\zeta_1|^{2i_1}\ldots|\zeta_n|^{2i_n} =$$

$$= \sum_{n=1}^{\infty} \frac{\pi^n||\zeta||^{2n}}{n!} = e^{\pi\bar{\zeta}\cdot\zeta} = e^{\langle\bar{\zeta},\zeta\rangle} = \langle e^{\zeta}|e^{\zeta}\rangle_{\tilde{S}(\bar{A})}.$$

So our space $\tilde{S}(A)$ is maped isometrically into H_n. Since its image contains the basis functions (9.16) and $\tilde{S}(A)$ is complete, the isometry is bijective.

9.4 Let us find now an isomorphism between H_n and $L^2(\mathbb{R}^n)$. For any $\mathbf{x} \in \mathbb{R}^n, \mathbf{z} \in \mathbb{C}^n$, let

$$k(\mathbf{x}, \mathbf{z}) = 2^{\frac{n}{4}} e^{-\pi||\mathbf{x}||^2} e^{2\pi i\mathbf{x}\cdot\mathbf{z}} e^{\frac{\pi}{2}||\mathbf{z}||^2}. \tag{9.17}$$

Lemma 3.

$$\int_{\mathbb{R}^n} k(\mathbf{x},\mathbf{z})\overline{k(\mathbf{x},\zeta)}d\mathbf{x} = e^{\pi\bar{\zeta}\cdot\mathbf{z}}.$$

Proof. We have

$$\int_{\mathbb{R}^n} k(\mathbf{x},\mathbf{z})\overline{k(\mathbf{x},\zeta)}d\mathbf{x} = 2^{\frac{n}{2}}e^{\frac{\pi}{2}(||\mathbf{z}||^2+||\zeta||^2)}\int_{\mathbb{R}^n} e^{-2\pi||\mathbf{x}||^2}e^{-2\pi i\mathbf{x}\cdot(\zeta-\mathbf{z})}d\mathbf{x} =$$

$$= e^{\frac{\pi}{2}(||\mathbf{z}||^2+||\zeta||^2)}\frac{1}{\sqrt{(2\pi^n}}\int_{\mathbb{R}^n} e^{-||\mathbf{y}||^2/2}e^{-i\sqrt{\pi}\mathbf{y}\cdot(\zeta-\mathbf{z})}d\mathbf{y} = e^{\frac{\pi}{2}(||\mathbf{z}||^2+||\zeta||^2)}F(\sqrt{\pi}(\mathbf{z}-\zeta)).$$

where $F(\mathbf{t})$ is the Fourier transform of the function $e^{-||\mathbf{y}||^2/2}$. It is known that $\widehat{e^{-x^2/2}} = e^{-t^2/2}$. This easily implies that $F(\mathbf{t}) = e^{-||\mathbf{t}||^2/2}$. Plugging in $\mathbf{t} = \sqrt{\pi}(\mathbf{z}-\zeta)$, we get the assertion.

Lemma 4. *Let*

$$k(\mathbf{x},\mathbf{z}) = \sum_{\mathbf{i}} h_{\mathbf{i}}(\mathbf{x})\phi_{\mathbf{i}}(\mathbf{z})$$

be the Taylor expansion of $k(\mathbf{x},\mathbf{z})$ (recall that $\phi_{\mathbf{i}}$ are monomials in $\mathbf{z}$). Then $(h_{\mathbf{i}}(\mathbf{x}))_{\mathbf{i}}$ forms a complete orthonormal system in $L^2(\mathbb{R}^n)$. In fact,

$$h_{\mathbf{i}}(x_1,\ldots,x_n) = h_{i_1}(x_1)\ldots h_{i_n}(x_n),$$

where $h_k(x_i) = H_k(\sqrt{2\pi}x_i)$ with $H_k(x)$ being a Hermite function.

Proof. We will check this in the case $n=1$. Since both $k(\mathbf{x},\mathbf{z})$ and $\phi_{\mathbf{i}}$ are products of functions in one variable, the general case is easily reduced to our case. Since $k(x,z) = 2^{\frac{1}{4}}e^{\pi x^2}e^{-2\pi(x-iz/2)^2}$, we get

$$h_i(x) = \int_{\mathbb{R}} k(x,z)\phi_i(x)dx = 2^{\frac{1}{4}}(\pi/i!)^{1/2}\int_{\mathbb{R}} e^{\pi x^2}e^{-2\pi(x-iz/2)^2}x^i dx.$$

We omit the computation of this integral which lead to the desired answer.

Now we can define the linear map

$$\Phi : L^2(\mathbb{R}^n) \to \mathrm{Hol}(\mathbb{C}^n), \quad \psi(\mathbf{x}) \to \int_{\mathbb{R}^n} k(\mathbf{x},\mathbf{z})\psi(\mathbf{x})d\mathbf{x}. \tag{9.18}$$

By Lemma 3 and the Cauchy-Schwarz inequality,

$$|\Phi(\psi)(\mathbf{z})| \le \Big(\int_{\mathbb{R}^n} k(\mathbf{x},\mathbf{z})\overline{k(\mathbf{x},\mathbf{z})}d\mathbf{x}\Big)||\psi(\mathbf{x})|| = e^{(\pi||\mathbf{z}||^2/2)}||\psi(\mathbf{x})||.$$

This implies that the integral is uniformly convergent with respect to the complex parameter $\mathbf{z}$ on every bounded subset of $\mathbb{C}^n$. Thus $\Phi(\psi)$ is a holomorphic function; so the map is well-defined. Also, by Lemma 4, since $\phi_{\mathbf{i}}$ is an orthonormal system in $\mathrm{Hol}(\mathbb{C}^n)$ and $h_{\mathbf{i}}$ is an orthonormal basis in $L^2(\mathbb{R}^n)$, we get

$$\int_{\mathbb{C}^n} |\Phi(\psi)(\mathbf{z})|^2 d\mu = \sum_{\mathbf{i}} |(h_{\mathbf{i}}, \bar{\psi})|^2 = ||\psi||^2 < \infty.$$

This shows that the image of Φ is contained in H_n, and at the same time, that Φ is a unitary linear map. Under the map Φ, the basis $(h_{\mathbf{i}}(\mathbf{x}))_{\mathbf{i}}$ of $L^2(\mathbb{R}^n)$ is mapped to the basis $(\phi_{\mathbf{i}})_{\mathbf{i}}$ of H_n. Thus Φ is an isomorphism of Hilbert spaces.

9.5 We know how the Heisenberg group $\tilde{V}$ acts on H_n. Let us see how it acts on $L^2(\mathbb{R}^n)$ via the isomorphism $\Phi : L^2(\mathbb{R}^n) \cong H_n$.

Recall that we have a decomposition $V_{\mathbb{C}} = A \oplus \bar{A}$ of $V_{\mathbb{C}}$ into the sum of conjugate isotropic subspaces with respect to the bilinear form S. Consider the map $V \to A, v \to v - iJv$. Since $Jv - iJ(Jv) = i(v - iJv)$, this map is an isomorphism of complex vector spaces $(V, J) \to A$. Similarly we see that the map $v \to v + iJv$ is an isomorphism of complex vector spaces $(V, -J) \to \bar{A}$. Keep the basis $e_1, \ldots, e_n$ of (V, J) as in 9.3 so that $V_{\mathbb{C}}$ is identified with $\mathbb{C}^n$. Then $e_i - iJe_i, i = 1, \ldots, n$ is a basis of A, $e_i + iJe_i, i = 1, \ldots, n$, is a basis of $\bar{A}$, and the pairing $A \times \bar{A} \to \mathbb{C}$ from (9.6) has the form

$$\langle(\zeta_1, \ldots, \zeta_n), (z_1, \ldots, z_n)\rangle = \pi(\sum_{i=1}^n \zeta_i z_i) = \pi\zeta \cdot \mathbf{z}.$$

Let us identify (V, J) with $\mathbb{C}^n$ by means of the basis (e_i). And similarly let us do it for A and $\bar{A}$ by means of the bases $(e_i - iJe_i)$ and $(e_i + iJe_i)$, respectively. Then $V_{\mathbb{C}}$ is identified with $A \oplus \bar{A} = \mathbb{C}^n \oplus \mathbb{C}^n$, and the inclusion $(V, J) \subset V_{\mathbb{C}}$ is given by $\mathbf{z} \to (\mathbf{z}, \bar{\mathbf{z}})$.

The skew-symmetric bilinear form $S : V_{\mathbb{C}} \times V_{\mathbb{C}} \to \mathbb{C}$ is now given by the formula

$$S((\mathbf{z}, \mathbf{w}), (\mathbf{z}', \mathbf{w}')) = \frac{\pi}{2i}(\mathbf{z} \cdot \mathbf{w}' - \mathbf{w} \cdot \mathbf{z}').$$

Its restriction to V is given by

$$S(\mathbf{z}, \mathbf{z}') = S((\mathbf{z}, \bar{\mathbf{z}}), (\mathbf{z}', \bar{\mathbf{z}}')) = \frac{\pi}{2i}(\mathbf{z}\cdot\bar{\mathbf{z}}' - \bar{\mathbf{z}}\cdot\mathbf{z}') = \frac{\pi}{2i}(2i\mathrm{Im})(\mathbf{z}\cdot\bar{\mathbf{z}}') = \pi\mathrm{Im}(\mathbf{z}\cdot\bar{\mathbf{z}}') = \pi(\mathrm{yx}' - \mathrm{xy}'),$$

where $\mathbf{z} = \mathbf{x} + i\mathbf{y}, \mathbf{z}' = \mathbf{x}' - i\mathbf{y}'$. Let

$$V_{re} = \{\mathbf{z} \in V : \mathrm{Im}(\mathbf{z}) = 0\} = \{\mathbf{z} = \mathbf{x} \in \mathbb{R}^n\},$$
$$V_{im} = \{\mathbf{z} \in V : \mathrm{Re}(\mathbf{z}) = 0\} = \{\mathbf{z} = i\mathbf{y} \in i\mathbb{R}^n\}.$$

Then the decomposition $V = V_{re} \oplus V_{im}$ is a decomposition of V into the sum of two maximal isotropic subspaces with respect to the bilinear form S.

As in (9.8), we have, for any $v = \mathbf{x} + i\mathbf{y} \in V$,

$$v = \mathbf{x} + i\mathbf{y} = a + \bar{a} = (\mathbf{x} + i\mathbf{y}, 0) + (0, \mathbf{x} - i\mathbf{y}) \in A \oplus \bar{A}.$$

The Heisenberg group V acts on $\mathrm{Hol}(\bar{A})$ by the formulae

$$\begin{aligned}\mathbf{x} + i\mathbf{y} \cdot f(\mathbf{z}) &= e^{-S(a,\bar{a})}a \cdot \bar{a}f(\mathbf{z}) = e^{-S(a,\bar{a})}e^{\langle\bar{a},\mathbf{z}\rangle}f(\mathbf{z} - \bar{a}) = \\ &= e^{-\frac{\pi}{2}a\cdot\bar{a}}e^{\pi a\cdot\mathbf{z}}f(z - \bar{a}) = e^{-\frac{\pi}{2}(\mathbf{x}\cdot\mathbf{x}+\mathbf{y}\cdot\mathbf{y})}e^{\pi(\mathbf{x}+i\mathbf{y})\cdot\mathbf{z}}f(\mathbf{z} - \mathbf{x} + i\mathbf{y}). \end{aligned} \tag{9.19}$$

Theorem 2. *Under the isomorphism $\Phi : H_n \to L^2(\mathbb{R}^n)$, the Schrödinger representation of $\tilde{V}$ on H_n is isomorphic to the representation of $\tilde{V}$ on $L^2(\mathbb{R}^n)$ defined by the formula*

$$(\mathbf{v} + i\mathbf{u}, t)\psi(\mathbf{x}) = te^{\pi i\mathbf{v}\cdot\mathbf{u}}e^{-2\pi i\mathbf{x}\cdot\mathbf{v}}\psi(\mathbf{x} - \mathbf{u}). \tag{9.20}$$

Proof. In view of (9.9) and (9.10), we have

$$(\mathbf{v} + i\mathbf{u}) \cdot \Phi(\psi(\mathbf{x})) = \int_{\mathbb{R}^n} (\mathbf{v} + i\mathbf{u})k(\mathbf{x}, \mathbf{z})\psi(\mathbf{x})d\mathbf{x}$$

$$= \int_{\mathbb{R}^n} e^{-\frac{\pi}{2}(\mathbf{u}\cdot\mathbf{u}+\mathbf{v}\cdot\mathbf{v})}e^{\pi(\mathbf{v}+i\mathbf{u})\cdot\mathbf{z}}k(\mathbf{z} - \mathbf{v} + i\mathbf{u})\psi(\mathbf{x})d\mathbf{x}$$

$$= \int_{\mathbb{R}^n} e^{-\frac{\pi}{2}(\mathbf{u}\cdot\mathbf{u}+\mathbf{v}\cdot\mathbf{v})}e^{\pi(\mathbf{v}+i\mathbf{u})\cdot\mathbf{z}}e^{2\pi i\mathbf{x}(-\mathbf{v}+i\mathbf{u})}e^{\frac{\pi}{2}(2\mathbf{z}(-\mathbf{v}+i\mathbf{u})+(-\mathbf{v}+i\mathbf{u})\cdot(-\mathbf{v}+i\mathbf{u}))}k(\mathbf{x}, \mathbf{z})\psi(\mathbf{x})d\mathbf{x}$$

$$= \int_{\mathbb{R}^n} e^{-\pi\mathbf{u}\cdot\mathbf{u}+2\pi i\mathbf{u}\cdot\mathbf{z}-2\pi i\mathbf{x}\mathbf{v}-2\pi\mathbf{x}\mathbf{u}-\pi i\mathbf{v}\cdot\mathbf{u}}k(\mathbf{x}, \mathbf{z})\psi(\mathbf{x})d\mathbf{x}.$$

$$\Phi((\mathbf{v} + i\mathbf{u}) \cdot \psi(\mathbf{x})) = \int_{\mathbb{R}^n} k(\mathbf{x}, \mathbf{z})e^{\pi i\mathbf{v}\cdot\mathbf{u}}e^{-2\pi i\mathbf{x}\cdot\mathbf{v}}\psi(\mathbf{x} - \mathbf{u})d\mathbf{x}$$

$$= \int_{\mathbb{R}^n} k(\mathbf{t} + \mathbf{u}, \mathbf{z})e^{\pi i\mathbf{v}\cdot\mathbf{u}}e^{-2\pi i(\mathbf{t}+\mathbf{u})\cdot\mathbf{v}}\psi(\mathbf{t})d\mathbf{t} = \int_{\mathbb{R}^n} k(\mathbf{x} + \mathbf{u}, \mathbf{z})e^{\pi i\mathbf{v}\cdot\mathbf{u}}e^{-2\pi i(\mathbf{x}+\mathbf{u})\cdot\mathbf{v}}\psi(\mathbf{x})d\mathbf{x}$$

$$= \int_{\mathbb{R}^n} e^{-\pi\mathbf{u}\cdot\mathbf{u}-2\pi\mathbf{u}\mathbf{x}+2\pi i\mathbf{u}\cdot\mathbf{z}-\pi i\mathbf{v}\cdot\mathbf{u}-2\pi i\mathbf{x}\mathbf{v}}k(\mathbf{x}, \mathbf{z})\psi(\mathbf{x})d\mathbf{x}.$$

By comparing, we observe that

$$(\mathbf{v} + i\mathbf{u}) \cdot \Phi(\psi(\mathbf{x})) = \Phi((\mathbf{v} + i\mathbf{u}) \cdot \psi(\mathbf{x})).$$

This checks the assertion.

9.6 It follows from the proof that the formula (9.20) defines a representation of the group $\tilde{V}$ on $L^2(\mathbb{R}^n)$. Let us check this directly. We have

$$(\mathbf{v} + i\mathbf{u}, t)(\mathbf{v}' + \mathbf{u}', t') \cdot \psi(\mathbf{x}) = ((\mathbf{v} + \mathbf{v}') + i(\mathbf{u} + \mathbf{u}'), tt'e^{i\pi(\mathbf{u}\cdot\mathbf{v}'-\mathbf{v}\cdot\mathbf{u}')} \cdot \psi(\mathbf{x}) =$$

$$= tt'e^{i\pi(\mathbf{u}\cdot\mathbf{v}'-\mathbf{v}\cdot\mathbf{u}')}e^{\pi i(\mathbf{v}+\mathbf{v}')\cdot(\mathbf{u}+\mathbf{u}')}e^{-2\pi i\mathbf{x}\cdot(\mathbf{v}+\mathbf{v}')}\psi(\mathbf{x} - \mathbf{u} - \mathbf{u}') =$$

$$= tt'e^{i\pi(\mathbf{v}\cdot\mathbf{u}+\mathbf{v}'\cdot\mathbf{u}'+2\mathbf{v}'\cdot\mathbf{u})}e^{-2\pi i\mathbf{x}\cdot(\mathbf{v}+\mathbf{v}')}\psi(\mathbf{x} - \mathbf{u} - \mathbf{u}').$$

$$(\mathbf{v} + i\mathbf{u}, t) \cdot ((\mathbf{v}' + \mathbf{u}', t') \cdot \psi(\mathbf{x})) = (\mathbf{v} + i\mathbf{u}, t) \cdot (t'e^{i\pi\mathbf{u}'\mathbf{v}'}e^{-2\pi i\mathbf{x}\cdot\mathbf{v}'}\psi(\mathbf{x} - \mathbf{u}')) =$$

$$= tt'e^{i\pi\mathbf{uv}}e^{-2\pi i\mathbf{x}\cdot\mathbf{v}}e^{i\pi\mathbf{u}'\mathbf{v}'}e^{-2\pi i(\mathbf{x}-\mathbf{u})\cdot\mathbf{v}'}\psi(\mathbf{x}-\mathbf{u}-\mathbf{u}') =$$

$$= tt'e^{i\pi(\mathbf{v}\cdot\mathbf{u}+\mathbf{v}'\cdot\mathbf{u}'+2\mathbf{v}'\cdot\mathbf{u})}e^{-2\pi i\mathbf{x}\cdot(\mathbf{v}+\mathbf{v}')}\psi(\mathbf{x}-\mathbf{u}-\mathbf{u}').$$

So this matches.

Let us alter the definition of the Heisenberg group $\tilde{V}$ by setting

$$(\mathbf{z},t)\cdot(\mathbf{z}',t') = (\mathbf{z}+\mathbf{z}', tt'e^{2\pi i\mathbf{x}\cdot\mathbf{y}'}),$$

where $\mathbf{z} = \mathbf{x}+i\mathbf{y}, \mathbf{z}' = \mathbf{x}'+i\mathbf{y}'$. It is immediately checked that the map $(\mathbf{z},t)\to(\mathbf{z},te^{\pi\mathbf{x}\cdot\mathbf{y}})$ is an isomorphism from our old group to the new one. Then the new group acts on $L^2(\mathbb{R}^n)$ by the formula

$$(\mathbf{v}+i\mathbf{u},t)\cdot\psi(\mathbf{x}) = te^{-2\pi i\mathbf{x}\cdot\mathbf{v}}\psi(\mathbf{x}-\mathbf{u}), \tag{9.21}$$

and on $\mathrm{Hol}(\bar{A})$ by the formula

$$(\mathbf{v}+i\mathbf{u},t)\cdot f(\mathbf{z}) = te^{-\frac{\pi}{2}(\mathbf{u}\cdot\mathbf{u}+\mathbf{v}\cdot\mathbf{v})+\pi\mathbf{z}\cdot(\mathbf{v}+i\mathbf{u})+i\pi\mathbf{v}\cdot\mathbf{u}}f(\mathbf{z}-\mathbf{u}) =$$

$$= e^{\pi((\mathbf{z}-\frac{1}{2}\mathbf{v})\cdot\mathbf{v}-\frac{1}{2}\mathbf{u}\cdot\mathbf{u})+i\pi(\mathbf{z}+\mathbf{v})\cdot\mathbf{u}}f(\mathbf{z}-\mathbf{u}). \tag{9.22}$$

This agrees with the formulae from **[Igusa]**, p. 35.

9.7 Let us go back to Lecture 6, where we defined the operators $V(\mathbf{v})$ and $U(\mathbf{u})$ on the space $L^2(\mathbb{R}^n)$ by the formula:

$$U(\mathbf{u})\psi(\mathbf{q}) = e^{i(u_1P_1+\ldots+u_nP_n)}\psi(\mathbf{q}) = \psi(\mathbf{q}-\hbar\mathbf{u}),$$

$$V(\mathbf{v})\psi(\mathbf{q}) = e^{i(v_1Q_1+\ldots+v_nQ_n)}\psi(\mathbf{q}) = e^{i\mathbf{v}\cdot\mathbf{q}}\psi(\mathbf{q}).$$

Comparing this with the formula (9.19), we find that

$$U(\mathbf{u})\psi(\mathbf{q}) = i\hbar\mathbf{u}\cdot\psi(\mathbf{q}),\quad V(\mathbf{v})\psi(\mathbf{q}) = -\frac{1}{2\pi}\mathbf{v}\cdot\psi(\mathbf{q}),$$

where we use the Schrödinger representation of the (redefined) Heisenberg group $\tilde{V}$ in $L^2(\mathbb{R}^n)$. The commutator relation (6.18)

$$[U(\mathbf{u}),V(\mathbf{v})] = e^{-i\hbar\mathbf{u}\cdot\mathbf{v}}$$

agrees with the commutation relation

$$[i\hbar\mathbf{u},-\frac{1}{2\pi}\mathbf{v}] = [(i\hbar\mathbf{u},1),(-\frac{1}{2\pi}\mathbf{v},1)] = (0,e^{2\pi i(\mathbf{u}\cdot-\frac{1}{2\pi}\mathbf{v})}) = (0,e^{-i\mathbf{u}\cdot\mathbf{v}}).$$

In Lecture 7 we defined the Heisenberg algebra $\mathcal{H}$. Its subalgebra $\mathcal{N}$ generated by the operators a and a^* coincides with the Lie subalgebra of self-adjoint (unbounded) operators in $L^2(\mathbb{R})$ generated by the operators $P = i\hbar\frac{d}{dq}$ and $Q = q$. If we exponentiate these operators we find the linear operators e^{ivQ}, e^{iuP}. It follows from the above that they form a group isomorphic to the Heinseberg group $\tilde{V}$, where $\dim V = 2$. The Schrödinger representation of this group in $L^2(\mathbb{R})$ is the exponentiation of the representation of $\mathcal{N}$ described in Lecture 7. Adding the exponent e^{iH} of the Schrödinger operator $H = \omega a a^* - \frac{\hbar\omega}{2}$ leads to an extension

$$1 \to \tilde{V} \to G \to \mathbb{R}^* \to 1.$$

Here the group G is isomorphic to the group of matrices

$$\begin{pmatrix} 1 & x & y & z \\ 0 & w & 0 & y \\ 0 & 0 & w^{-1} & -x \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Of course, we can generalize this to the case of n variables. We consider similar matrices

$$\begin{pmatrix} 1 & \mathbf{x} & \mathbf{y} & z \\ 0 & wI_n & 0 & \mathbf{y}^t \\ 0 & 0 & w^{-1}I_n & -\mathbf{x}^t \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

where $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n, z, w \in \mathbb{R}$. More generally, we can introduce the *full Heisenberg group* as the extension

$$1 \to \tilde{V} \to \tilde{G} \to \mathrm{Sp}(2n, \mathbb{R}) \to 1,$$

where $\mathrm{Sp}(2n, \mathbb{R})$ is the symplectic group of $2n \times 2n$ matrices X satisfying

$$X \begin{pmatrix} 0_n & -I_n \\ I_n & 0_n \end{pmatrix} X^t = \begin{pmatrix} 0_n & -I_n \\ I_n & 0_n \end{pmatrix}.$$

The group $\tilde{G}$ is isomorphic to the group of matrices

$$\begin{pmatrix} 1 & \mathbf{x} & \mathbf{y} & z \\ 0 & A & B & \mathbf{y}^t \\ 0 & C & D & -\mathbf{x}^t \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

where $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$ is an $n \times n$-block presentation of a matrix $X \in \mathrm{Sp}(2n, \mathbb{R})$.

Exercises.

1. Let V be a real vector space of dimension $2n$ equipped with a skew-symmetric bilinear form S. Show that any decomposition $V_{\mathbb{C}} = A \oplus A$ into a sum of conjugate isotropic subspaces with the property that $(a, b) \to 2iS(a, \bar{b})$ is a positive definite Hermitian form on A defines a complex structure J on V such that $S(Jv, Jw) = S(v, w), S(Jv, w) > 0$.

2. Extend the Schrödinger representation of $\tilde{V}$ on $\tilde{S}(\bar{A})$ to a projective representation of the full Heisenberg group in $\mathbb{P}(\tilde{S}(\bar{A}))$ preserving (up to a scalar factor) the inner product in $\tilde{S}(\bar{A})$. This is called the *metaplectic representation* of order n. What will be the corresponding representation in $\mathbb{P}(L^2(\mathbb{R}^n))$?

3. Consider the natural representation of $SO(2)$ in $L_2(\mathbb{R}^2)$ via action on the arguments. Desribe the representation of $SO(2)$ in H_2 obtained via the isomorphism $\Phi : H_2 \cong L_2(\mathbb{R}^2)$.

4. Consider the natural representation of $SU(2)$ in H_2 via action on the arguments. Describe the representation of $SU(2)$ in $L_2(\mathbb{R}^2)$ obtained via the isomorphism $\Phi : H_2 \cong L_2(\mathbb{R}^2)$.

Lecture 10. ABELIAN VARIETIES AND THETA FUNCTIONS

10.1 Let us keep the notation from Lecture 9. Let Λ be a lattice of rank $2n$ in V. This means that Λ is a subgroup of V spanned by $2n$ linearly independent vectors. It follows that Λ is a free abelian subgroup of V and thus it is isomorphic to $\mathbb{Z}^{2n}$. The orbit space V/Λ is a compact torus. It is diffeomorphic to the product of $2n$ circles $(\mathbb{R}/\mathbb{Z})^{2n}$. If we view V as a complex vector space (V,J), then T has a canonical complex structure such that the quotient map

$$\pi : V \to V/\Lambda = T$$

is a holomorphic map. To define this structure we choose an open cover $\{U_i\}_{i\in I}$ of V such that $U_i \cap U_i + \gamma = \emptyset$ for any $\gamma \in \Lambda$. Then the restriction of π to U_i is an isomorphism and the complex structure on $\pi(U_i)$ is induced by that of U_i.

The skew-symmetric bilinear form S on V defines a complex line bundle L on T which can be used to embed T into a projective space so that T becomes an algebraic variety. A compact complex torus which is embeddable into a projective space is called an *abelian variety*.

Let us describe the construction of L. Start with any holomorphic line bundle L over T. Its pre-image $\pi^*(L)$ under the map π is a holomorphic line bundle over the complex vector space V. It is known that all holomorphic vector bundles on V are (holomorphically) trivial. Let us choose an isomorphism $\phi : \pi^*(L) = L \times_T V \to V \times \mathbb{C}$. Then the group Λ acts naturally on $\pi^*(L)$ via its action on V. Under the isomorphism ϕ it acts on $V \times \mathbb{C}$ by a formula

$$\gamma \cdot (v,t) = (v+\gamma, \alpha_\gamma(v)t), \quad \gamma \in \Lambda, v \in V, t \in \mathbb{C}. \tag{10.1}$$

Here $\alpha_\gamma(v)$ is a non-zero constant depending on γ and v. Since the action is by holomorphic automorphisms, $\alpha_\gamma(v)$ depends holomorphically on v and can be viewed as a map

$$\alpha : \Lambda \to \mathcal{O}(V)^*, \quad \gamma \to \alpha_\gamma(v),$$

where $\mathcal{O}(V)$ denotes the ring of holomorphic functions on V and $\mathcal{O}(V)^*$ is its group of invertible elements. It follows from the definition of action that

$$\alpha_{\gamma+\gamma'}(v) = \alpha_\gamma(v+\gamma')\alpha_{\gamma'}(v). \tag{10.2}$$

Denote by $Z^1(\Lambda, \mathcal{O}(V)^*)$ the set of functions α as above satisfying (10.2). These functions are called *theta factors* associated to Λ. Obviously $Z^1(\Lambda, \mathcal{O}(V)^*)$ is a commutative group with respect to pointwise multiplication. For any $g(v) \in \mathcal{O}(V)^*$ the function

$$\alpha_\gamma(v) = g(v+\gamma)/g(v) \tag{10.3}$$

belongs to $Z^1(\Lambda, \mathcal{O}(V)^*)$. It is called the *trivial theta factor*. The set of such functions forms a subgroup $B^1(\Lambda, \mathcal{O}(V)^*)$ of $Z^1(\Lambda, \mathcal{O}(V)^*)$. The quotient group is denoted by $H^1(\Lambda, \mathcal{O}(V)^*)$. The reader familiar with the notion of group cohomology will recognize the latter group as the first cohomology group of the group Λ with coefficients in the abelian group $\mathcal{O}(V^*)$ on which Λ acts by translation in the argument.

The theta factor $\alpha \in Z^1(\Lambda, \mathcal{O}(V)^*)$ defined by the line bundle L depends on the choice of the trivialization ϕ. A different choice leads to replacing ϕ with $g \circ \phi$, where $g : V \times \mathbb{C} \to V \times \mathbb{C}$ is an automorphism of the trivial bundle defined by the formula $(v,t) \to (v, g(v)t)$ for some function $g \in \mathcal{O}(V)^*$. This changes α_γ to $\alpha_\gamma(v)g(v+\gamma)/g(v)$. Thus the coset of α in $H^1(\Lambda, \mathcal{O}(V)^*)$ does not depend on the trivialization ϕ. This defines a map from the set $\mathrm{Pic}(T)$ of isomorphism classes of holomorphic line bundles on T to the group $H^1(\Lambda, \mathcal{O}(V)^*)$. In fact, this map is a homomorphism of groups, where the operation of an abelian group on $\mathrm{Pic}(T)$ is defined by tensor multiplication and taking the dual bundle.

Theorem 1. *The homomorphism*

$$\mathrm{Pic}(T) \to H^1(\Lambda, \mathcal{O}(V)^*)$$

is an isomorphism of abelian groups.

Proof. It is enough to construct the inverse map $H^1(\Lambda, \mathcal{O}(V)^*) \to \mathrm{Pic}(T)$. We shall define it, and leave to the reader to verify that it is the inverse.

Given a representative α of a class from the right-hand side, we consider the action of Λ on $V \times \mathbb{C}$ given by formula (10.2). We set L to be the orbit space $V \times \mathbb{C}/\Lambda$. It comes with a canonical projection $p : L \to T$ defined by sending the orbit of (v,t) to the orbit of v. Its fibres are isomorphic to $\mathbb{C}$. Choose a cover $\{U_i\}_{i\in I}$ of V as in the beginning of the lecture and let $\{W_i\}_{i\in I}$ be its image cover of T. Since the projection π is a local analytic isomorphism, it is an open map (so that the image of an open set is open). Because T is compact, we may find a finite subcover of the cover $\{W_i\}_{i\in I}$. Thus we may assume that I is finite, and $\pi^{-1}(W_i) = \coprod_{\gamma\in\Lambda}(U_i+\gamma)$. Let $\pi_i : U_i \to W_i$ be the analytic isomorphism induced by the projection π. We may assume that $\pi_i^{-1}(W_i \cap W_j) = \pi_j^{-1}(W_i \cap W_j) + \gamma_{ij}$ for some $\gamma_{ij} \in \Lambda$, provided that $W_i \cap W_j \neq \emptyset$. Since each $U_i \times \mathbb{C}$ intersects every orbit of Λ in $V \times \Lambda$ at a unique point (v,t), we can identify $p^{-1}(W_i)$ with $W_i \times \mathbb{C}$. Also

$$W_i \times \mathbb{C} \supset p_i^{-1}(W_i \cap W_j) \cong p_j^{-1}(W_i \cap W_j) \subset W_j \times \mathbb{C},$$

where the isomorphism is given explicitly by $(v,t) \to (v, \alpha_{\gamma_{ij}}(v)t)$. This shows that L is a holomorphic line bundle with transition functions $g_{W_i,W_j}(x) = \alpha_{\gamma_{ij}}(\pi_i^{-1}(x))$. We also leave to the reader to verify that replacing α by another representative of the same class in H^1 changes the line bundle L to an isomorphic line bundle.

10.2 So, in order to construct a line bundle L we have to construct a theta factor. Recall from (9.11) that V acts on $\tilde{S}(A)$ by the formula

$$v \cdot e^a = e^{-\langle b,\bar{b}\rangle/2-\langle a,\bar{b}\rangle}e^{a+b},$$

where $v = b + \bar{b}$. Let us identify V with A by means of the isomorphism $v \to b = v - iJv$. From (9.6) we have

$$\langle v - iJv, w + iJw\rangle = 4H(v,w) = 4S(Jv,v) + 4iS(v,w).$$

Thus

$$e^{-\langle b,\bar{b}\rangle/2-\langle a,\bar{b}\rangle} = e^{-2H(\gamma,\gamma)-4H(v,\gamma)}.$$

Set

$$\alpha_\gamma(v) = e^{-2H(\gamma,\gamma)-4H(v,\gamma)}.$$

We have

$$\alpha_{\gamma+\gamma'}(v) = e^{-2H(\gamma+\gamma',\gamma+\gamma')-4H(v,\gamma+\gamma')} = e^{-2H(\gamma,\gamma)-2H(\gamma',\gamma')-4ReH(\gamma,\gamma')-4H(v,\gamma)-4H(v,\gamma')}$$

$$= e^{4iImH(\gamma,\gamma')-2H(\gamma,\gamma)-4H(v+\gamma',\gamma)-2H(\gamma',\gamma')-4H(v,\gamma')} = e^{-4iImH(\gamma,\gamma')}\alpha_\gamma(v+\gamma')\alpha_{\gamma'}(v).$$

We see that condition (10.2) is satisfied if, for any $\gamma,\gamma' \in \Lambda$,

$$ImH(\gamma,\gamma') = S(\gamma,\gamma') \in \frac{\pi}{2}\mathbb{Z}.$$

Let us redefine α by replacing it with

$$\alpha_\gamma(v) = e^{-\pi H(\gamma,\gamma)-2\pi H(v,\gamma)} \tag{10.4}$$

Of course this is equivalent to multiplying our bilinear form S by $-\frac{\pi}{2}$. Then the previous condition is replaced with

$$S(\Lambda \times \Lambda) \subset \mathbb{Z}. \tag{10.5}$$

Assuming that this is true we have a line bundle L_α. As we shall see in a moment it is not yet the final definition of the line bundle associated to S. We have to adjust the theta factor (10.4) a little more. To see why we should do this, let us compute the first Chern class of the obtained line bundle L_α.

Since V is obviously simply connected and $\pi : V \to T$ is a local isomorphism, we can idenify V with the universal cover of T and the group Λ with the fundamental group of T. Since it is abelian, we can also identify it with the first homology group $H_1(T,\mathbb{Z})$. Now, we have

$$H_k(T,\mathbb{Z}) \cong \bigwedge^k(H_1(T,\mathbb{Z})$$

because T is topologically the product of circles. In particular,

$$H^2(T,\mathbb{Z}) = Hom(H_2(T,\mathbb{Z}),\mathbb{Z}) = \bigwedge^2(H_1(T,\mathbb{Z}))^* = (\bigwedge^2\Lambda)^*. \tag{10.6}$$

This allows one to consider $S : \Lambda\times\Lambda \to \mathbb{Z}$ as an element of $H^2(T,\mathbb{Z})$. The latter group is where the first Chern class takes its value.

Recall that we have a canonical isomorphism between $\mathrm{Pic}(T)$ and $H^1(T,\mathcal{O}_T^*)$ by computing the latter groups as the Čech cohomology group and assigning to a line bundle L the set of the transition functions with respect to an open cover. The exponential sequence of sheaves of abelian groups

$$0 \longrightarrow \mathbb{Z} \longrightarrow \mathcal{O}_T \stackrel{e^{2\pi i}}{\longrightarrow} \mathcal{O}_T^* \longrightarrow 1 \tag{10.7}$$

defines the coboundary homomorphism

$$H^1(T,\mathcal{O}_T^*) \to H^2(T,\mathbb{Z}).$$

The image of $L\in\mathrm{Pic}(T)$ is the first Chern class $c_1(L)$ of L. In our situation, the coboundary homomorphism coincides with the coboundary homomorphism for the exact sequence of group cohomology

$$H^1(\Lambda,\mathcal{O}(V)) \to H^1(\Lambda,\mathcal{O}(V)^*) \stackrel{\delta}{\longrightarrow} H^2(\Lambda,\mathbb{Z}) \cong H^2(T,\mathbb{Z})$$

arising from the exponential exact sequence

$$1 \longrightarrow \mathbb{Z} \longrightarrow \mathcal{O}(V) \stackrel{e^{2\pi i}}{\longrightarrow} \mathcal{O}(V)^* \longrightarrow 1. \tag{10.8}$$

Here the isomorphism $H^2(\Lambda,\mathbb{Z}) \cong H^2(T,\mathbb{Z})$ is obtained by assigning to a $\mathbb{Z}$-valued 2-cocycle $\{c_{\gamma,\gamma'}\}$ of Λ the alternating bilinear form $\tilde{c}(\gamma,\gamma') = c_{\gamma,\gamma'} - c_{\gamma',\gamma}$. The condition for a 2-cocycle is

$$c_{\gamma_2,\gamma_3} - c_{\gamma_2+\gamma_1,\gamma_3} + c_{\gamma_1,\gamma_2+\gamma_3} - c_{\gamma_1,\gamma_2} = 0. \tag{10.9}$$

It is not difficult to see that this implies that $\tilde{c}$ is an alternating bilinear form.

Let us compute the first Chern class of the line bundle L_α defined by the theta factor (10.4). We find $\beta_\gamma(v) : \Lambda\times V \to \mathbb{C}$ such that $\alpha_\gamma(v) = e^{2\pi i\beta_\gamma(v)}$. Then, using

$$1 = \alpha_{\gamma+\gamma'}(v)/\alpha_{\gamma'}(v+\gamma')\alpha_{\gamma'}(v),$$

we get

$$c_{\gamma,\gamma'}(v) = \beta_{\gamma+\gamma'}(v) - \beta_{\gamma'}(v+\gamma') - \beta_{\gamma'}(v) \in \mathbb{Z}.$$

By definition of the coboundary homomorphism $\delta(L_\alpha)$ is given by the 2-cocycle $\{c_{\gamma,\gamma'}\}$. Returning to our case when $\alpha_\gamma(v)$ is given by (10.4), we get

$$\beta_\gamma(v) = \frac{i}{2}(H(\gamma,\gamma) + 2H(v,\gamma)),$$

$$c_{\gamma,\gamma'}(v) = \beta_{\gamma+\gamma'}(v) - \beta_{\gamma'}(v+\gamma') - \beta_{\gamma'}(v) = \frac{i}{2}(2ReH(\gamma,\gamma') - 2H(\gamma',\gamma)) = -ImH(\gamma,\gamma').$$

Thus

$$c_1(L_\alpha) = \{c_{\gamma,\gamma'} - c_{\gamma',\gamma'}\} = \{-2ImH(\gamma,\gamma')\} = -2S.$$

We would like to have $c_1(L) = S$. For this we should change H to $-H/2$. However the corresponding function $\alpha_\gamma(v)' = e^{\frac{\pi}{2}H(\gamma,\gamma)+\pi H(v,\gamma)}$ is not a theta factor. It satisfies

$$\alpha_{\gamma+\gamma'}(v)' = e^{i\pi S(\gamma,\gamma')}\alpha_\gamma(v+\gamma')'\alpha_{\gamma'}(v)'.$$

We correct the definition by replacing $\alpha_\gamma(v)'$ with $\alpha_\gamma(v)'\chi(\gamma)$, where the map

$$\chi : \Lambda \to \mathbb{C}_1^*$$

has the property

$$\chi(\gamma+\gamma') = \chi(\gamma)\chi(\gamma')e^{i\pi S(\gamma,\gamma')}, \quad \forall \gamma,\gamma' \in \Lambda. \tag{10.10}$$

We call such a map a *semi-character* of Λ. An example of a semi-character is the map $\chi(\gamma) = e^{i\pi S'(\gamma,\gamma)}$, where S' is any bilinear form on Λ with values in $\mathbb{Z}$ such that

$$S'(\gamma,\gamma') - S'(\gamma',\gamma) = S(\gamma,\gamma'). \tag{10.11}$$

Now we can make the right definition of the theta factor associated to S. We set

$$\alpha_\gamma(v) = e^{\frac{\pi}{2}H(\gamma,\gamma)+\pi H(v,\gamma)}\chi(\gamma). \tag{10.12}$$

Clearly $\gamma \to \chi^2(\gamma)$ is a character of Λ (i.e., a homomorphism of abelian groups $\Lambda \to \mathbb{C}_1^*$). Obviously, any character defines a theta factor whose values are constant functions. Its first Chern class is zero. Also note that two semi-characters differ by a character and any character can be given by a formula

$$\chi(\gamma) = e^{2\pi i l(\gamma)},$$

where $l : V \to \mathbb{R}$ is a real linear form on V.

We define the line bundle $L(H,\chi)$ as the line bundle corresponding to the theta factor (10.12). It is clear now that

$$c_1(L(H,\chi)) = S. \tag{10.13}$$

10.3 Now let us interpret global sections of any line bundle constructed from a theta factor $\alpha \in Z^1(V, \mathcal{O}(V)^*)$. Recall that L is isomorphic to the line bundle obtained as the orbit space $V \times \mathbb{C}/\Lambda$ where Λ acts by formula (10.2). Let $s : V \to V \times \mathbb{C}$ be a section of the trivial bundle. It has the form $s(v) = (v, \phi(v))$ for some holomorphic function ϕ on V. So we can identify it with a holomorphic function on V. Assume that, for any $\gamma \in \Lambda$, and any $v \in V$,

$$\phi(v+\gamma) = \alpha_\gamma(v)\phi(v). \tag{10.14}$$

This means that $\gamma \cdot s(v) = s(v+\gamma)$. Thus s descends to a holomorphic section of $L = V \times \mathbb{C}/\Lambda \to V/\Lambda = T$. Conversely, every holomorphic section of L can be lifted to a holomorphic section of $V \times \mathbb{C}$ satisfying (10.14).

We denote by $\Gamma(T, L_\alpha)$ the complex vector space of global sections of the line bundle L_α defined by a theta factor α. In view of the above,

$$\Gamma(T, L_\alpha) = \{\phi \in \mathrm{Hol}(V) : \phi(z+\gamma) = \alpha_\gamma(v)\phi(v)\}. \tag{10.15}$$

Applying this to our case $L = L(H,\chi)$ we obtain

$$\Gamma(T, L(H,\chi)) \cong \{\phi \in \mathrm{Hol}(V) : \phi(v+\gamma) = e^{\frac{\pi}{2}H(\gamma,\gamma')+\pi H(v,\gamma)}\chi(\gamma)\phi(v),\ \forall v \in V, \gamma \in \Lambda\}, \tag{10.16}$$

Let $\beta_\gamma(v) = g(v+\gamma)/g(v)$ be a trivial theta factor. Then the multiplication by g defines an isomorphism

$$\Gamma(T, L_\alpha) \cong \Gamma(T, L_{\alpha\beta}).$$

We shall show that the vector space $\Gamma(T, L(H,\chi))$ is finite-dimensional and compute its dimension.

But first we need some lemmas.

Lemma 1. *Let $S : \Lambda \times \Lambda \to \mathbb{Z}$ be a non-degenerate skew-symmetric bilinear form. Then there exists a basis $\omega_1, \ldots, \omega_{2n}$ of Λ such that $S(\omega_i, \omega_j) = d_i\delta_{i+n,j}, i = 1, \ldots, n$. Moreover, we may assume that the integers d_i are positive and $d_1|d_2|\ldots|d_n$. Under this condition they are determined uniquely.*

Proof. This is well-known, nevertheless we give a proof. We use induction on the rank of Λ. The assertion is obvious for $n = 2$. For any $\gamma \in \Lambda$ the subset of integers $\{S(\gamma,\gamma'), \gamma' \in \Lambda\}$ is a cyclic subgroup of $\mathbb{Z}$. Let d_γ be its positive generator. We set d_1 to be the minimum of the d_γ's and choose ω_1, ω_{n+1} such that $S(\omega_1, \omega_{n+1}) = d_1$. Then for any $\gamma \in \Lambda$, we have $d_1|S(\gamma,\omega_1), S(\gamma,\omega_{n+1})$. This implies that

$$\gamma - \frac{S(\gamma,\omega_1)}{d_1}\omega_{n+1} - \frac{S(\gamma,\omega_{n+1})}{d_1}\omega_1 \in \Lambda' = (\mathbb{Z}\omega_1 + \mathbb{Z}\omega_{n+1})^\perp.$$

Now we use the induction assumption on Λ'. There exists a basis $\omega_2, \ldots, \omega_n, \omega_{n+2}, \ldots, \omega_{2n}$ of Λ' satisfying the properties from the statement of the lemma. Let $d_2, \ldots, d_n$ be the corresponding integers. We must have $d_1|d_2$, since otherwise $S(k\omega_1 + \omega_2, \omega_{n+1} + \omega_{n+2}) = kd_1 + d_2 < d_1$ for some integer k. This contradicts the choice of d_1. Thus $\omega_1, \ldots, \omega_{2n}$ is the desired basis of Λ.

Lemma 2. *Let H be a positive definite Hermitian form on a complex vector space V and let Λ be a lattice in V such that $S = Im(H)$ satisfies (10.5). Let $\omega_1, \ldots, \omega_{2n}$ be a basis of Λ chosen as in Lemma 1 and let Δ be the diagonal matrix $diag[d_1, \ldots, d_n]$. Then the last n vectors ω_i are linearly independent over $\mathbb{C}$ and, if we use these vectors to identify V with $\mathbb{C}^n$, the remaining vectors $\omega_{n+1}, \ldots, \omega_{2n}$ form a matrix $\Omega = X + iY \in M_n(\mathbb{C})$ such that*

(i) $\Omega\Delta^{-1}$ *is symmetric;*
(ii) $Y\Delta^{-1}$ *is positive definite.*

Proof. Let us first check that the vectors $\omega_{n+1}, \ldots, \omega_{2n}$ are linearly independent over $\mathbb{C}$. Suppose $\lambda_1\omega_{n+1} + \ldots + \lambda_n\omega_{2n} = 0$ for some $\lambda_i = x_i + iy_i$. Then

$$w = -\sum_{i=1}^{n} x_i\omega_{n+i} = i\sum_{i=1}^{n} y_i\omega_{n+i} = iv.$$

We have $S(iv, v) = S(w, v) = 0$ because the restriction of S to $\mathbb{R}\omega_{n+1} + \ldots + \mathbb{R}\omega_{2n}$ is trivial. Since $S(iv, v) = H(v, v)$, and H was assumed to be positive definite, this implies $v = w = 0$ and hence $x_i = y_i = \lambda_i = 0, i = 1, \ldots, n$.

Now let us use $\omega_{n+1}, \ldots, \omega_{2n}$ to identify V with $\mathbb{C}^n$. Under this identification, $\omega_{n+i} = e_i$, the i-th unit vector in $\mathbb{C}^n$. Write

$$\omega_j = (\omega_{1j}, \ldots, \omega_{nj}) = (x_{1j}, \ldots, x_{nj}) + i(y_{1j}, \ldots, y_{nj}) = \mathrm{Re}(\omega_j) + i\mathrm{Im}(\omega_j), \quad j = 1, \ldots, n.$$

We have

$$d_i\delta_{ij} = S(\omega_i, e_j) = \sum_{k=1}^{n} \big(x_{ki}S(e_k, e_j) + y_{ki}S(ie_k, e_j)\big) = \sum_{k=1}^{n} y_{ki}S(ie_k, e_j) = \sum_{k=1}^{n} S(ie_j, e_k)y_{ki}.$$

Let $A = (S(ie_j, e_k))_{j,k=1,\ldots,n}$ be the matrix defining $H : \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}$. Then the previous equality translates into the following matrix equality

$$\Delta = A \cdot Y. \tag{10.17}$$

Since A is symmetric and positive definite, we deduce from this property (ii). To check property (i) we use that

$$0 = S(\omega_i, \omega_j) = S(\sum_{k=1}^{n} x_{ki}e_k + y_{ki}ie_k, \sum_{k=1}^{n} x_{kj}e_k + y_{kj}ie_k) =$$

$$= \sum_{k'=1}^{n} x_{k'j}(\sum_{k=1}^{n} y_{ki}S(ie_k, e_{k'})) - \sum_{k=1}^{n} x_{ki}(\sum_{k'=1}^{n} y_{k'j}S(ie'_k, e_k)) =$$

$$= \sum_{k'=1}^{n} x_{k'j}d_i\delta_{k'i} - \sum_{k=1}^{n} x_{ki}d_j\delta_{kj} = x_{ij}d_i - x_{ji}d_j = d_id_j(x_{ij}d_j^{-1} - x_{ji}d_i^{-1}).$$

In matrix notation this gives

$$X \cdot \Delta^{-1} = (X \cdot \Delta^{-1})^t.$$

This proves the lemma.

Corollary. *Let $H : V \times V \to \mathbb{C}$ be a Hermitian positive definite form on V such that $Im(H)(\Lambda \times \Lambda) \subset \mathbb{Z}$. Let Π be a $2n \times n$-matrix whose columns form a basis of Λ with respect to some basis of V over $\mathbb{R}$. Let A be the matrix of $Im(H)$ with respect to this basis. Then the following Riemann-Frobenius relations hold*
(i) $\Pi A^{-1} \Pi^t = 0$;
(ii) $-i\Pi A^{-1} \bar{\Pi}^t > 0$.

Proof. It is easy to see that the relations do not depend on the choice of a basis. Pick a basis as in Lemma 2. Then, in block matrix notation,

$$\Pi = \begin{pmatrix} \Omega & I_n \end{pmatrix}, \quad A = \begin{pmatrix} 0_n & \Delta \\ -\Delta & 0_n \end{pmatrix}. \tag{10.18}$$

This gives

$$\Pi A^{-1} \Pi^t = \begin{pmatrix} \Omega & I_n \end{pmatrix} \begin{pmatrix} 0_n & \Delta^{-1} \\ -\Delta^{-1} & 0_n \end{pmatrix} \begin{pmatrix} \Omega^t \\ I_n \end{pmatrix} = -\Delta^{-1}\Omega^t + \Omega\Delta^{-1} = 0,$$

$$-i\Pi A^{-1} \bar{\Pi}^t = -i \begin{pmatrix} \Omega & I_n \end{pmatrix} \begin{pmatrix} 0_n & \Delta^{-1} \\ -\Delta^{-1} & 0_n \end{pmatrix} \begin{pmatrix} \bar{\Omega}^t \\ I_n \end{pmatrix} =$$

$$= -i(-\Delta^{-1}\bar{\Omega}^t + \Omega\Delta^{-1}) = -i(2iY\Delta^{-1}) = 2Y\Delta^{-1} > 0.$$

Lemma 3. *Let $\lambda : \Gamma \to \mathbb{C}_1^*$ be a character of Λ. Let $\omega_1, \ldots, \omega_{2n}$ be a basis of Λ. Define the vector c_γ by the condition:*

$$\lambda(\omega_i) = e^{2\pi i S(\omega_i, c_\lambda)}, i = 1, \ldots, 2n.$$

Note that this is possible because S is non-degenerate. Then

$$\phi(v) \to \phi(v + c_\gamma)e^{\pi H(v, c_\lambda)}$$

defines an isomorphism from $\Gamma(T, L(H, \chi))$ to $\Gamma(T, L(H, \chi \cdot \lambda))$.

Proof. Let

$$\tilde{\phi}(v) = \phi(v + c_\lambda)e^{\pi H(v, c_\lambda)}.$$

Then

$$\tilde{\phi}(v+\gamma) = e^{\pi H(v+\gamma, c_\lambda)}\phi(v+c_\lambda+\gamma) = e^{\pi H(v, c_\lambda)}\phi(v+c_\lambda)\chi(\gamma)e^{\pi(\frac{1}{2}H(\gamma,\gamma)+H(v+c_\lambda,\gamma))}e^{\pi H(\gamma, c_\lambda)}$$

$$= \tilde{\phi}(v)e^{\pi(H(\gamma,c_\lambda)-H(c_\lambda,\gamma))}e^{\pi(\frac{1}{2}H(\gamma,\gamma)+H(v,\gamma))}\chi(\gamma) = \tilde{\phi}(v)e^{2\pi i S(\gamma,c_\lambda)}e^{\pi(\frac{1}{2}H(\gamma,\gamma)+H(v,\gamma))}\chi(\gamma)$$

$$= \tilde{\phi}(v)e^{\pi(\frac{1}{2}H(\gamma,\gamma)+H(v,\gamma))}\chi(\gamma)\lambda(\gamma).$$

This shows that $\tilde{\phi} \in \Gamma(T, L(H, \chi \cdot \lambda)$. Obviously the map $\phi \to \tilde{\phi}$ is invertible.

10.4 From now on we shall use the notation of Lemma 2. In this notation V is identified with $\mathbb{C}^n$ so that we can use $\mathbf{z} = (z_1, \ldots, z_n)$ instead of v to denote elements of V. Our lattice Λ looks like

$$\Lambda = \mathbb{Z}\omega_1 + \ldots + \mathbb{Z}\omega_n + \mathbb{Z}e_1 + \ldots + \mathbb{Z}e_n.$$

The matrix

$$\Omega = [\omega_1, \ldots, \omega_n] \tag{10.19}$$

satisfies properties (i),(ii) from Lemma 2. Let

$$V_1 = \mathbb{R}e_1 + \ldots + \mathbb{R}e_n = \{\mathbf{z} \in \mathbb{C}^n : Im(\mathbf{z}) = 0\}.$$

We know that the restriction of S to V_1 is trivial. Therefore the restriction of H to V_1 is a symmetric positive definite quadratic form. Let $B : V \times V \to \mathbb{C}$ be a quadratic form such that its restriction to $V_1 \times V_1$ coincides with H (just take B to be defined by the matrix $(H(e_i, e_j))$. Then

$$\alpha'_\gamma(v) = \alpha_\gamma(v)e^{-\pi B(\mathbf{z},\gamma)-\frac{\pi}{2}B(\gamma,\gamma)} = \alpha_\gamma(v)\left(e^{-\frac{\pi}{2}B(\mathbf{z}+\gamma,\mathbf{z}+\gamma)}/e^{-\frac{\pi}{2}B(\mathbf{z},\mathbf{z})}\right) =$$

$$= \chi(\gamma)e^{\pi(H-B)(\mathbf{z},\gamma)+\frac{\pi}{2}(H-B)(\gamma,\gamma)}.$$

Since α and α' differ by a trivial theta factor, they define isomorphic line bundles.

Also, by Lemma 3, we may choose any semi-character χ since the dimension of the space of sections does not depend on its choice. Choose χ in the form

$$\chi_0(\gamma) = e^{i\pi S'(\gamma,\gamma)}, \tag{10.20}$$

where S' is defined in (10.11) and its restriction to V_1 and to $V_2 = \mathbb{R}\omega_1 + \ldots + \mathbb{R}\omega_n$ is trivial, For example, one may take S' to be defined in the basis $\omega_1, \ldots, \omega_n, e_1, \ldots, e_n$ by the matrix $\begin{pmatrix} 0_n & \Delta + I_n \\ I_n & 0 \end{pmatrix}$. We have

$$\chi_0(\gamma) = 1, \gamma \in V_1 \cup V_2.$$

So we will be computing the dimension of $\Gamma(T, L_{\alpha^\sharp})$ where

$$\alpha^\sharp_\gamma(\mathbf{z}) = \chi_0(\gamma)e^{\pi(H-B)(\mathbf{z},\gamma)+\frac{\pi}{2}(H-B)(\gamma,\gamma)}, \tag{10.21}$$

Using (10.17), we have, for any $\mathbf{z} \in V$ and $k = 1, \ldots, n$

$$(H-B)(\mathbf{z}, e_k) = \sum_{i=1}^n z_i(H-B)(e_i, e_k) = 0,$$

$$(H-B)(\mathbf{z}, \omega_k) = \mathbf{z} \cdot (H(e_i, e_j)) \cdot \bar{\omega}_k - \mathbf{z} \cdot ((B(e_i, e_j)) \cdot \omega_k = \mathbf{z}(H(e_i, e_j))(\bar{\omega}_k - \omega_k) =$$

$$-2i\mathbf{z} \cdot (S(ie_i, e_j)) \cdot Im\bar{\omega}_k = -2i\mathbf{z} \cdot (AY)e_k = -2i\mathbf{z} \cdot \Delta \cdot e_k = -2id_k z_k. \tag{10.22}$$

Theorem 2.

$$\dim_{\mathbb{C}}\Gamma(T, L(H,\chi)) = |\Delta| = d_1 \cdots d_n.$$

Proof. By (10.22), any $\phi \in \Gamma(T, L_\alpha^\sharp)$ satisfies

$$\phi(\mathbf{z} + e_k) = \alpha_{e_k}(\mathbf{z})\phi(\mathbf{z}) = \phi(\mathbf{z}), \quad k = 1, \ldots, n,$$

$$\phi(\mathbf{z} + \omega_k) = \alpha_{\omega_k}(\mathbf{z})\phi(\mathbf{z}) = e^{-2\pi i d_k(z_k + \frac{1}{2}\omega_{kk})}\phi(\mathbf{z}), \quad k = 1, \ldots, n. \tag{10.23}$$

The first equation allows us to expand $\phi(\mathbf{z})$ in Fourier series

$$\phi(\mathbf{z}) = \sum_{\mathbf{r} \in \mathbb{Z}^n} a_{\mathbf{r}} e^{2\pi i \mathbf{r} \cdot \mathbf{z}}.$$

By comparing the coefficients of the Fourier series, the second equality allows us to find the recurrence relation for the coefficients

$$a_{\mathbf{r}} = e^{-\pi i(2\mathbf{r}\cdot\omega_k + d_k\omega_{kk})} a_{\mathbf{r} - d_k e_k}. \tag{10.24}$$

Let

$$M = \{\mathbf{m} = (m_1, \ldots, m_n) \in \mathbb{Z}^n : 0 \le m_i < d_i\}.$$

Using (10.24), we are able to express each $a_{\mathbf{r}}$ in the form

$$a_{\mathbf{r}} = e^{2\pi i \lambda(\mathbf{r})} a_{\mathbf{m}},$$

where $\lambda(\mathbf{r})$ is a function in $\mathbf{r}$, $\mathbf{r} \equiv \mathbf{m} \bmod (d_1, \ldots, d_n)$, satisfying

$$\lambda(\mathbf{r} - d_k e_k) = \lambda(\mathbf{r}) - \mathbf{r}\cdot\omega_k - \frac{1}{2}d_k\omega_{kk}.$$

This means that the difference derivative of the function λ is a linear function. So we should try some quadratic function to solve for λ. We can take

$$\lambda(\mathbf{r}) = \frac{1}{2}\mathbf{r}\cdot\Omega\cdot(\Delta^{-1}\cdot\mathbf{r}).$$

We have

$$\lambda(\mathbf{r} - d_k e_k) = \frac{1}{2}(-d_k e_k + \mathbf{r})\cdot\Omega\cdot(-e_k + \Delta^{-1}\cdot\mathbf{r}) =$$

$$= \lambda(\mathbf{r}) - \frac{1}{2}\Big(-d_k e_k\cdot\Omega\Delta^{-1}\cdot\mathbf{r} - \mathbf{r}\cdot\Omega\cdot e_k + d_k e_k\cdot\Omega\cdot e_k\Big) = \lambda(\mathbf{r}) - \mathbf{r}\cdot\omega_k - \frac{1}{2}d_k\omega_{kk}.$$

This solves our recurrence. Clearly two solutions of the recurrence (10.24) differ by a constant factor. So we get

$$\phi(\mathbf{z}) = \sum_{\mathbf{m}\in M} c_{\mathbf{m}}\theta_{\mathbf{m}}(\mathbf{z}),$$

for some constants $c_{\mathbf{m}}$, and

$$\theta_{\mathbf{m}}(\mathbf{z}) = \sum_{\mathbf{r}\in\mathbb{Z}^n} e^{2\pi i\lambda(\mathbf{m}+\Delta\cdot\mathbf{r})}e^{2\pi i\mathbf{z}\cdot(\mathbf{m}+\Delta\cdot\mathbf{r})} = \sum_{\mathbf{r}\in\mathbb{Z}^n} e^{\pi i((\mathbf{m}\cdot\Delta^{-1}+\mathbf{r})\cdot(\Delta\Omega)\cdot(\Delta^{-1}\mathbf{m}+\mathbf{r})+2\mathbf{z}\cdot(\mathbf{m}+\Delta\cdot\mathbf{r}))}. \tag{10.25}$$

The series converges uniformly on any bounded subset. In fact, since $\Omega' = \Omega\cdot\Delta^{-1}$ is a positive definite symmetric matrix, we have

$$||Im(\Omega\cdot\Delta^{-1})\cdot\mathbf{r}|| \geq C||\mathbf{r}||^2,$$

where C is the minimal eigenvalue of Ω'. Thus

$$\sum_{\mathbf{r}\in\mathbb{Z}^n} |e^{2\pi i\lambda(\mathbf{m}+\Delta\cdot\mathbf{r})}||e^{2\pi i\mathbf{z}\cdot(\mathbf{m}+\Delta\cdot\mathbf{r})}| = \sum_{\mathbf{r}\in\mathbb{Z}^n} |e^{-\pi Im(\Omega')||(\mathbf{m}+\Delta\cdot\mathbf{r})||^2}|\mathbf{q}^{\mathbf{m}+\Delta\cdot\mathbf{r}}|$$

$$\leq \sum_{\mathbf{r}\in\mathbb{Z}^n} e^{-C\pi||\mathbf{m}+\Delta\cdot\mathbf{r}||^2}|\mathbf{q}|^{\mathbf{m}+\Delta\cdot\mathbf{r}},$$

where $\mathbf{q} = e^{2\pi i\mathbf{z}}$. The last series obviously converges on any bounded set.

Thus we have shown that $\dim\Gamma(T, L_\alpha) \leq \#M = |\Delta|$. One can show, using the uniqueness of Fourier coefficients for a holomorphic function, that the functions $\theta_{\mathbf{m}}$ are linearly independent. This proves the assertion.

10.5 Let us consider the special case when

$$d_1 = \ldots = d_n = d.$$

This means that $\frac{1}{d}S|\Lambda\times\Lambda$ is a unimodular bilinear form. We shall identify the set of residues M with $(\mathbb{Z}/d\mathbb{Z})^n$. One can rewrite the functions $\theta_{\mathbf{m}}$ in the following way:

$$\theta_{\mathbf{m}}(\mathbf{z}) = \sum_{\mathbf{r}\in\mathbb{Z}^n} e^{\pi i(\frac{1}{d}\mathbf{m}+\mathbf{r})\cdot(d\Omega)\cdot(\frac{1}{d}\mathbf{m}+\mathbf{r})+2\pi i d\mathbf{z}\cdot(\frac{1}{d}\mathbf{m}+\mathbf{r})}. \tag{10.26}$$

Definition. Let $(\mathbf{m},\mathbf{m}') \in (\mathbb{Z}/d\mathbb{Z})^n \oplus (\mathbb{Z}/d\mathbb{Z})^n$ and let Ω be a symmetric complex $n\times n$-matrix with positive definite imaginary part. The holomorphic function

$$\theta_{\mathbf{m},\mathbf{m}'}(\mathbf{z};\Omega) = \sum_{\mathbf{r}\in\mathbb{Z}^n} e^{\pi i\left((\frac{1}{d}\mathbf{m}+\mathbf{r})\cdot\Omega\cdot(\frac{1}{d}\mathbf{m}+\mathbf{r})+2(\mathbf{z}+\frac{1}{d}\mathbf{m}')\cdot(\frac{1}{d}\mathbf{m}+\mathbf{r})\right)}$$

is called the *Riemann theta function of order d with theta characteristic* $(\mathbf{m},\mathbf{m}')$ with respect to Ω.

A similar definition can be given in the case of arbitrary Δ. We leave it to the reader.

So, we see that $\Gamma(T, L_{\alpha^\sharp}) \cong \Gamma(T, L(H,\chi_0))$ has a basis formed by the functions

$$\theta_{\mathbf{m}}(\mathbf{z}) = \theta_{\mathbf{m},0}(d\mathbf{z}; d\Omega), \quad \mathbf{m} \in (\mathbb{Z}/d\mathbb{Z})^n. \tag{10.27}$$

Using Lemma 3, we find a basis of $\Gamma(T, L(H,\chi))$. It consists of functions

$$e^{-\frac{\pi}{2}\mathbf{z}\cdot A\cdot\mathbf{z}}\theta_{\mathbf{m},0}(d(\mathbf{z}+\mathbf{c};d\Omega)),$$

where A is the symmetric matrix $(H(e_i,e_j)) = (S(ie_i,e_j))$ and $\mathbf{c}\in V$ is defined by the condition

$$\chi(\gamma) = e^{2\pi i S(c,\gamma)}, \gamma = e_i,\omega_i, i = 1,\ldots,n.$$

Let us see how the functions $\theta_{\mathbf{m},\mathbf{m}'}(\mathbf{z},\Omega)$ change under translation of the argument by vectors from the lattice Λ. We have

$$\theta_{\mathbf{m},\mathbf{m}'}(\mathbf{z}+e_k;\Omega) = \sum_{\mathbf{r}\in\mathbb{Z}^n} e^{\pi i\left((\frac{1}{d}\mathbf{m}+\mathbf{r})\cdot\Omega\cdot(\frac{1}{d}\mathbf{m}+\mathbf{r})+2(\mathbf{z}+e_k+\frac{1}{d}\mathbf{m}')\cdot(\frac{1}{d}\mathbf{m}+\mathbf{r})\right)} =$$

$$= \sum_{\mathbf{r}\in\mathbb{Z}^n} e^{\pi i\left((\frac{1}{d}\mathbf{m}+\mathbf{r})\cdot\Omega\cdot(\frac{1}{d}\mathbf{m}+\mathbf{r})+2(\mathbf{z}+\frac{1}{d}\mathbf{m}')\cdot(\frac{1}{d}\mathbf{m}+\mathbf{r})\right)} e^{\frac{2\pi i m_i}{d}} = e^{\frac{2\pi i m_k}{d}}\theta_{\mathbf{m},\mathbf{m}'}(\mathbf{z};\Omega),$$

$$\theta_{\mathbf{m},\mathbf{m}'}(\mathbf{z}+\omega_k;\Omega) = \sum_{\mathbf{r}\in\mathbb{Z}^n} e^{\pi i\left((\frac{1}{d}\mathbf{m}+\mathbf{r}-e_k)\cdot\Omega\cdot(\frac{1}{d}\mathbf{m}+\mathbf{r}-e_k)+2(\mathbf{z}+\omega_k+\frac{1}{d}\mathbf{m}')\cdot(\frac{1}{d}\mathbf{m}+\mathbf{r}-e_k)\right)} =$$

$$= \sum_{\mathbf{r}\in\mathbb{Z}^n} e^{\pi i\left((\frac{1}{d}\mathbf{m}+\mathbf{r})\cdot\Omega\cdot(\frac{1}{d}\mathbf{m}+\mathbf{r})+2(\mathbf{z}+\omega_k+\frac{1}{d}\mathbf{m}')\cdot(\frac{1}{d}\mathbf{m}+\mathbf{r})\right)} e^{-\pi(2z_k+\omega_{kk})+2\frac{\pi i m_k}{d}} =$$

$$= e^{\frac{2\pi i m'_k}{d}} e^{-i\pi(2z_k+\omega_{kk})}\theta_{\mathbf{m},\mathbf{m}'}(\mathbf{z};\Omega). \tag{10.28}$$

Comparing this with (10.23), this shows that

$$\theta_{\mathbf{m},\mathbf{m}'}(\mathbf{z};\Omega)^d \in \Gamma(T, L_{\alpha^\sharp}) \cong \Gamma(T, L(H,\chi_0)). \tag{10.29}$$

This implies that $\theta_{\mathbf{m},\mathbf{m}'}(\mathbf{z};\Omega)$ generates a one-dimensional vector space $\Gamma(T,L)$, where $L^{\otimes d}\cong L(H,\chi_0)$. This line bundle is isomorphic to the line bundle $L(\frac{1}{d}H,\chi)$ where $\chi^n = \chi_0$. A line bundle over T with unimodular first Chern class is called a *principal polarization.* Of course, it does not necessary exist. Observe that we have d^{2n} functions $\theta_{\mathbf{m},\mathbf{m}'}(\mathbf{z};\Omega)^d$ in the space $\Gamma(T,L_{\alpha^\sharp})$ of dimension d^n. Thus there must be some linear relations between the d-th powers of theta functions with characteristic. They can be explicitly found.

Let us put $\mathbf{m}=\mathbf{m}'=0$ and consider the Riemann theta function (without characteristic)

$$\Theta(\mathbf{z};\Omega) = \sum_{\mathbf{r}\in\mathbb{Z}^n} e^{i\pi(\mathbf{r}\cdot\Omega\cdot\mathbf{r}+2\mathbf{z}\cdot\mathbf{r})}. \tag{10.30}$$

We have

$$\Theta(\mathbf{z}+\frac{1}{d}\mathbf{m}'+\frac{1}{d}\mathbf{m}\cdot\Omega;\Omega) = \sum_{\mathbf{r}\in\mathbb{Z}^n} e^{i\pi(\mathbf{r}\cdot\Omega\cdot\mathbf{r}+2(\mathbf{z}+\frac{1}{d}\mathbf{m}'+\frac{1}{d}\mathbf{m}\cdot\Omega)\cdot\mathbf{r})} =$$

$$= \sum_{\mathbf{r}\in\mathbb{Z}^n} e^{i\pi((\mathbf{r}+\frac{1}{d}\mathbf{m})\cdot\Omega\cdot(\mathbf{r}+\frac{1}{d}\mathbf{m})+2(\mathbf{z}+\frac{1}{d}\mathbf{m}')(\frac{1}{d}\mathbf{m}+\mathbf{r})-\frac{1}{d^2}(\mathbf{m}\cdot\mathbf{m}'+\mathbf{m}\cdot\Omega\frac{1}{d}\mathbf{m})} =$$

$$= e^{-i\pi \frac{1}{d^2}(\mathbf{m}\cdot\mathbf{m}'+\mathbf{m}\cdot\Omega\frac{1}{d}\mathbf{m})}\theta_{\mathbf{m},\mathbf{m}'}(\mathbf{z};\Omega). \tag{10.31}$$

Thus, up to a scalar factor, $\theta_{\mathbf{m},\mathbf{m}'}$ is obtained from Θ by translating the argument by the vector

$$\frac{1}{d}(\mathbf{m}'+\mathbf{m}\Omega) \in \frac{1}{d}\Lambda/\Lambda = T_n = \{a \in T : na = 0\}. \tag{10.32}$$

This can be interpreted as follows. For any $a \in T$ denote by t_a the translation automorphism $x \to x + a$ of the torus T. If $v \in V$ represents a, then t_a is induced by the translation map $t_v : V \to V, \mathbf{z} \to \mathbf{z}+v$. Let L_α be the line bundle defined by a theta factor α. Then $\beta : \gamma \to \alpha_\gamma(\mathbf{z}+v)$ defines a theta factor such that $L_\beta = t_a^*(L)$. Its sections are functions $\phi(\mathbf{z}+v)$ where ϕ represents a section of L. Thus the theta functions $\theta_{\mathbf{m},\mathbf{m}'}(\mathbf{z};\Omega)$ are sections of the bundle $t_a^*(L^\sharp)$ where $\Theta \in \Gamma(T, L^\sharp)$ and a is an n-torsion point of T represented by the vector (10.32).

10.6 Let L be a line bundle over T. Define the group

$$G(L) = \{a \in T : t_a^*(L) \cong L\}.$$

If L is defined by a theta factor $\alpha_\gamma(\mathbf{z})$, then $t_a^*(L)$ is defined by the theta factor $\alpha_\gamma(\mathbf{z}+v)$, where v is a representative of a in V. Thus, $t_a^*(L) \cong L$ if and only if $\alpha_\gamma(\mathbf{z}+v)/\alpha_\gamma(\mathbf{z})$ is a trivial theta factor, i.e.,

$$\alpha_\gamma(\mathbf{z}+v)/\alpha_\gamma(\mathbf{z}) = g(\mathbf{z}+\gamma)/g(v)$$

for some function $g \in \mathcal{O}(V)^*$. Take $L = L(H,\chi)$. Then

$$g(\mathbf{z}+\gamma)/g(\mathbf{z}) = e^{\frac{\pi}{2}(H(\gamma,\gamma)+2H(\mathbf{z}+v,\gamma))}/e^{\frac{\pi}{2}(H(\gamma,\gamma)+2H(\mathbf{z},\gamma))} = e^{\pi H(v,\gamma)}.$$

Multiplying $g(\mathbf{z})$ by $e^{\pi H(\mathbf{z},v)} \in \mathcal{O}(V)^*$, we get that

$$e^{\pi H(v,\gamma)}e^{\pi H(\mathbf{z}+\gamma,v)}/e^{\pi H(\mathbf{z},v)} = e^{\pi(H(v,\gamma)+H(\gamma,v))} = e^{2\pi i \mathrm{Im}H(v,\gamma)} = e^{2\pi i S(v,\gamma)} \tag{10.33}$$

is the trivial theta factor. This happens if and only if $e^{2\pi i S(v,\gamma)} = 1$. This is true for any theta factor which is given by a character $\chi : \Lambda \to \mathbb{C}_1^*$. In fact $\chi(\lambda) = g(\mathbf{z}+\gamma)/g(\mathbf{z})$ implies that $|g(\mathbf{z})|$ is periodic with respect to Λ, hence is bounded on the whole of V. By Liouville's theorem this implies that g is constant, hence $\chi(\lambda) = 1$. Now the condition $e^{2\pi i S(v,\gamma)} = 1$ is equivalent to

$$S(v,\gamma) \in \mathbb{Z}, \quad \forall \gamma \in \Lambda.$$

So

$$G(L(H,\chi)) \cong \Lambda_S := \{v \in V : S(v,\gamma) \in \mathbb{Z}, \quad \forall \gamma \in \Lambda\}/\Lambda \cong \bigoplus_{i=1}^n (\mathbb{Z}/d_i\mathbb{Z})^2. \tag{10.34}$$

If the invariants $d_1, \ldots, d_n$ are all equal to d, this implies

$$G(L(H,\chi)) = \frac{1}{d}\Lambda/\Lambda = T_n \cong (\frac{1}{d}\mathbb{Z}/\mathbb{Z})^{2n} \cong (\mathbb{Z}/d\mathbb{Z})^{2n}.$$

Now let us define the *theta group* of L by

$$\tilde{G}(L) = \{(a, \psi) : a \in G(L), \psi : t_a^*(L) \to L \quad \text{is an isomorphism}\}.$$

Here the pairs $(a, \psi), (a', \psi')$ are multiplied by the rule

$$((a, \psi) \cdot (a', \psi') = (a + a', \psi \circ t_a^*(\psi') : t_{a+a'}(L) = t_a^*(t_{a'}^*(L)) \xrightarrow{t_a^*(\psi')} t_a^*(L) \xrightarrow{\psi} L).$$

The map $\tilde{G}(L) \to G(L), (a, \psi) \to a$ is a surjective homomorphism of groups. Its kernel is the group of automorphisms of the line bundle L. This can be identified with $\mathbb{C}^*$. Thus we have an extension of groups

$$1 \to \mathbb{C}^* \to \tilde{G}(L) \to G(L) \to 1.$$

Let us take $L = L(H, \chi)$. Then the isomorphism $t_a^*(L) \to L$ is determined by the trivial theta factor $e^{\pi H(v,\gamma)}$ and an isomorphism $\psi : t_a^*(L) \to L$ by a holomorphic invertible function $g(\mathbf{z})$ such that

$$g(\mathbf{z} + \gamma)/g(\mathbf{z}) = e^{\pi H(v,\gamma)}. \tag{10.35}$$

It follows from (10.33) that $\psi(\mathbf{z}) = e^{\pi H(\mathbf{z},v)}$ satisfies (10.35). Any other solution will differ from this by a constant factor. Thus

$$\tilde{G}(L) = \{(v, \lambda e^{\pi H(\mathbf{z},v)}), v \in G(L), \lambda \in \mathbb{C}^*\}.$$

We have

$$((v, \lambda e^{\pi H(\mathbf{z},v)}) \cdot (v', \lambda' e^{\pi H(\mathbf{z},v')}) = (v + v', \lambda\lambda' e^{\pi H(\mathbf{z}+v,v')} e^{\pi H(\mathbf{z},v)}) =$$

$$= (v + v', \lambda\lambda' e^{\pi H(\mathbf{z},v+v')} e^{\pi H(v,v')}).$$

This shows that the map

$$(v, \lambda e^{\pi H(\mathbf{z},v)}) \to (v, \lambda)$$

is an isomorphism from the group $\tilde{G}(L)$ onto the group $\tilde{\Lambda}_S$ which consists of pairs $(v, \lambda), v \in G(\Lambda, S), \lambda \in \mathbb{C}$ which are multiplied by the rule

$$(v, \lambda) \cdot (v', \lambda') = (v + v', \lambda\lambda' e^{\pi H(v,v')}).$$

This group defines an extension of groups

$$1 \to \mathbb{C}^* \to \tilde{\Lambda}_S \to \Lambda_S \to 1.$$

The map $(v, \lambda) \to (v, \lambda/|\lambda|)$ is a homomorphism from $\tilde{\Lambda}_S$ into the quotient of the Heisenberg group $\tilde{V}$ by the subgroup $\tilde{\Lambda}$ of elements $(\gamma, \lambda) \in \Lambda \times \mathbb{C}_1^*$. Its image is the subgroup of elements $(v, \lambda), v \in G(L)$ modulo $\tilde{\Lambda}$. Its kernel is the subgroup $\{(0, \lambda) : \lambda \in \mathbb{R}\} \cong \mathbb{R}$.

Let us see that the group $\tilde{G}(L)$ acts naturally in the space $\Gamma(T, L)$. Let s be a section of L. Then we have a canonical identification between $t_a^*(L)_{x-a}$ and L_x. Take $(a, \psi) \in \tilde{G}(L)$.

Then $x \to \psi(s(x))$ is a section of $t_a^*(L)$, and $x \to \psi(s(x-a))$ is a section of L. It is easily checked that the formula

$$((a,\psi)\cdot s)(x) = \psi(s(x-a))$$

is a representation of $\tilde{G}(L)$ in $\Gamma(T,L)$. This is sometimes called the *Schrödinger representation of the theta group.* Take now $L_{\alpha^\sharp} \cong L(H,\chi_0)$. Identify s with a function $\phi(\mathbf{z})$ satisfying

$$\phi(\mathbf{z}+\gamma) = \chi_0(\gamma)e^{\pi(\frac{1}{2}(H-B)(\gamma,\gamma)+(H-B)(\mathbf{z},v))}\phi(\mathbf{z}).$$

Represent (a,ψ) by $(v, \lambda e^{\pi H(\mathbf{z},v)})\frac{g(\mathbf{z}+v)}{g(\mathbf{z})}$. Then

$$\tilde{\phi}(\mathbf{z}) = (a,\psi)\cdot\phi(\mathbf{z}) = \lambda e^{\pi(H-B)(\mathbf{z}-v,v)}\phi(\mathbf{z}-v). \tag{10.36}$$

We have, by using (10.22),

$$\frac{1}{d}e_i\cdot\theta_{\mathbf{m}}(\mathbf{z}) = e^{\pi(H-B)(\mathbf{z}-\frac{1}{d}e_i,\frac{1}{d}e_i)}\theta_{\mathbf{m}}(z-\frac{1}{d}e_i) = \theta_{\mathbf{m}}(z-\frac{1}{d}e_i) =$$

$$\sum_{\mathbf{r}\in\mathbb{Z}^n} e^{\pi i(\frac{1}{d}\mathbf{m}+\mathbf{r})\cdot(d\Omega)\cdot(\frac{1}{d}\mathbf{m}+\mathbf{r})+2\pi i d(\mathbf{z}-\frac{1}{d}e_i)\cdot(\frac{1}{d}\mathbf{m}+\mathbf{r})} = e^{\frac{2\pi i m_i}{d}}\theta_{\mathbf{m}}(z),$$

$$\frac{1}{d}\omega_i\cdot\theta_{\mathbf{m}}(\mathbf{z}) = e^{\pi(H-B)(\mathbf{z}-\frac{1}{d}\omega_i,\frac{1}{d}\omega_i)}\theta_{\mathbf{m}}(z-\frac{1}{d}\omega_i) =$$

$$= e^{-2\pi i(z_i+\frac{\omega_i}{2d})}\sum_{\mathbf{r}\in\mathbb{Z}^n} e^{\pi i(\frac{1}{d}\mathbf{m}+\mathbf{r})\cdot(d\Omega)\cdot(\frac{1}{d}\mathbf{m}+\mathbf{r})+2\pi i d(\mathbf{z}-\frac{1}{d}\omega_i)\cdot(\frac{1}{d}\mathbf{m}+\mathbf{r})} =$$

$$e^{-2\pi i(z_i+\frac{\omega_i}{2d})}\sum_{\mathbf{r}\in\mathbb{Z}^n} e^{\pi i(\frac{1}{d}(\mathbf{m}-e_i)+\mathbf{r})\cdot(d\Omega)\cdot(\frac{1}{d}(\mathbf{m}-e_i)+\mathbf{r})+2\pi i d(\mathbf{z}\cdot(\frac{1}{d}(\mathbf{m}-e_i)+\mathbf{r})} = \theta_{\mathbf{m}-e_i}(z). \tag{10.37}$$

10.7 Recall that given $N+1$ linearly independent sections $s_0,\ldots,s_N$ of a line bundle L over a complex manifold X they define a holomorphic map

$$f: X' \to \mathbb{P}^N, x \to (s_0(x)\ldots, s_N(x)),$$

where X' is an open subset of points at which not all s_i vanish. Applying this to our case we take $L_{\alpha^\sharp} \cong L(H,\chi_0)$ and $s_i = \theta_{\mathbf{m}}(\mathbf{z})$ (with some order in the set of indices M) and obtain a map

$$f: T' \to \mathbb{P}^N, \quad N = d_1\cdots d_n - 1.$$

where $d_1|\ldots|d_n$ are the invariants of the bilinear form $Im(H)|\Lambda\times\Lambda$.

Theorem (Lefschetz). *Assume $d_1 \geq 3$, then the map $f: T \to \mathbb{P}^N$ is defined everywhere and is an embedding of complex manifolds.*

Proof. We refer for the proof to [**Lange**].

Corollary. *Let $T = \mathbb{C}^n/\Lambda$ be a complex torus. Assume that its period matrix Π composed of a basis of Λ satisfies the Riemann-Frobenius conditions from the Corollary to Lemma 2 in section 9.3. Then T is isomorphic to a complex algebraic variety. Conversely, if T has a structure of a complex algebraic variety, then the matrix Π of Λ satisfies the Riemann-Frobenius conditions.*

This follows from a general fact (Chow's Theorem) that a closed complex submanifold of projective space is an algebraic variety. Assuming this the proof is clear. The Riemann-Frobenius conditions are equaivalent to the existence of a positive definite Hermitian form H on $\mathbb{C}^n$ such that $Im(H)|\Lambda\times\Lambda \subset \mathbb{Z}$. This forms defines a line bundle $L(H,\chi)$. Replacing H by $3H$, we may assume that the condition of the Lefschetz theorem is satified. Then we apply the Chow Theorem.

Observe that the group $G(L)$ acts on T via its action by translation. The theta group $\tilde{G}(L)$ acts in $\Gamma(T,L)^*$ by the dual to the Schrödinger representation. The map $T \to \mathbb{P}(\Gamma(T,L)^*)$ given by the line bundle L is equivariant.

Example 1. Let $n = 1$, so that $E = \mathbb{C}/\mathbb{Z}\omega_1 + \mathbb{Z}\omega_2$ is a Riemann surface of genus 1. Choose complex coordinates such that $\omega_2 = 1, \tau = \omega_1 = a + bi$. Replacing τ by $-\tau$, if needed, we may assume that $b = \mathrm{Im}(\tau) > 0$. Thus the period matrix $\Pi = [\tau, 1]$ satisfies the Riemann-Frobenius conditions when we take $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. The corresponding Hermitian form on $V = \mathbb{C}$ is given by

$$H(z,z') = z\bar{z}'H(1,1) = z\bar{z}'S(i,1) = z\bar{z}'S(\frac{1}{b}(\tau - a), 1) = \frac{z\bar{z}'}{b}S(\tau,1) = \frac{z\bar{z}'}{\mathrm{Im}\tau}. \tag{10.38}$$

The Riemann theta function of order 1 is

$$\Theta(\tau,z) = \sum_{r\in\mathbb{Z}} e^{i\pi(\tau r^2 + 2rz)}. \tag{10.39}$$

The Riemann theta functions of order d with theta characteristic are

$$\theta_{m,m'}(\tau,z) = \sum_{r\in\mathbb{Z}} e^{i\pi((r+\frac{m}{d})^2\tau + 2(r+\frac{m}{d})(z+\frac{m'}{d}))}, \tag{10.40}$$

where $(m,m') \in (\mathbb{Z}/d\mathbb{Z})^2$. Take the line bundle L with $c_1(L) = dH$ whose space of sections has a basis formed by the functions

$$\theta_m(\mathbf{z}) = \theta_{m,0}(d\tau, dz) = \sum_{r\in\mathbb{Z}} e^{i\pi((r+\frac{m}{d})^2 d\tau + 2(dr+m)z)}, \tag{10.41}$$

where $m \in \mathbb{Z}/d\mathbb{Z}$. It follows from above that $L = L_{\alpha^\sharp}$, where

$$\alpha^\sharp_{a+b\tau}(z) = e^{-2\pi i(bdz + db^2\tau)}.$$

The theta group $\tilde{G}(L)$ is isomorphic to an extension

$$0 \to \mathbb{C}^* \to \tilde{G} \to (\mathbb{Z}/d\mathbb{Z})^2 \to 1$$

Let $\sigma_1 = (\frac{1}{d} + \Lambda, 1), \sigma_2 = (\frac{\tau}{d} + \Lambda, 1)$. Then σ_1, σ_2 generate $\tilde{G}$ with relation

$$[\sigma_1, \sigma_2] = (0, e^{2\pi i d \mathrm{Im}(H)(\frac{1}{d}, \frac{\tau}{d})}) = (0, e^{-\frac{2\pi i}{d}}). \tag{10.42}$$

It acts linearly in the space $\Gamma(E, L(dH, \chi_0^d))$ by the formulae (10.36)

$$\sigma_1(\theta_m(\mathbf{z})) = e^{-\frac{2\pi i m}{d}} \theta_m(\mathbf{z}), \quad \sigma_2(\theta_m(\mathbf{z})) = \theta_{m-1}(\mathbf{z}).$$

Let us take $d = 2$ and consider the map

$$f : E \to \mathbb{P}^1, \quad z \to (\theta_{0,0}(2\tau, 2z), \theta_{1,0}(2\tau, 2z)).$$

Notice that the functions $\theta_{m,0}(2\tau, 2z)$ are even since

$$\theta_{0,0}(2\tau, 2z) = \sum_{r \in \mathbb{Z}} e^{i\pi((r^2 2\tau) + 2(-2z))} = \sum_{r \in \mathbb{Z}} e^{i\pi((-r)^2 2\tau + 2((-r)2z)} = \theta_{0,0}(2\tau, 2z)$$

$$\theta_{1,0}(2\tau, -2z) = \sum_{r \in \mathbb{Z}} e^{i\pi((-r-\frac{1}{2})^2 2\tau + 2((-r-\frac{1}{2})2z)} =$$

$$= \sum_{r \in \mathbb{Z}} e^{i\pi((-r+1-\frac{1}{2})^2 2\tau) + 2((-r+1-\frac{1}{2})2z)} = \theta_{1,0}(2\tau, 2z).$$

Thus the map f is constant on the pairs $(a, -a) \in E$. This shows that the degree of f is divisible by 2. In fact the degree is equal to 2. This can be seen by counting the number of zeros of the function $\theta_m(z) - c$ in the fundamental parallelogram $\{\lambda + \mu\tau : 0 \le \lambda, \mu \le 1\}$. Or, one can use elementary algebraic geometry by noticing that the degree of the line bundle is equal to 2.

Take $d = 3$. This time we have an embedding $f : E \to \mathbb{P}^2$. The image is a cubic curve. Let us find its equation. We use that the equation is invariant with respect to the Schrödinger representation. If we choose the coordinates (t_0, t_1, t_2) in $\mathbb{P}^2$ which correspond to the functions $\theta_0(z), \theta_1(z), \theta_2(z)$ then the action of $\tilde{G}$ is given by the formulae

$$\sigma_1 : (t_0, t_1, t_2) \to (t_0, \zeta t_1, \zeta^2 t_2), \quad \sigma_2 : (t_0, t_1, t_2) \to (t_2, t_0, t_1),$$

where $\zeta = e^{\frac{2\pi i}{3}}$. Let W be the space of homogeneous cubic polynomials. It decomposes into eigensubspaces with respect to the action of σ_1: $W = W_0 + W_2 + W_3$. Here W_i is spanned by monomials $t_0^a t_1^b t_2^c$ with $a + b + c = 3, b + 2c \equiv i \mod 3$. Solving these congruences, we find that $(a, b, c) = (c, c, c) \mod 3$. This implies that the equation of $f(E)$ is in the *Hesse form*

$$\lambda_0 t_0^3 + \lambda_1 t_1^3 + \lambda_2 t_2^3 + \lambda_4 t_0 t_1 t_2 = 0.$$

Since it is also invariant with respect to σ_2, we obtain the equation

$$F(\lambda) = t_0^3 + t_1^3 + t_2^3 + \lambda t_0 t_1 t_2 = 0. \tag{10.43}$$

Since $f(E)$ is a nonsingular cubic,

$$\lambda^3 \neq -27.$$

we can also throw in more symmetries by considering the metaplectic representation. The group $Sp(2, \mathbb{Z}/3\mathbb{Z})$ is of order 24, so that together with the group $E_3 \cong (\mathbb{Z}/3\mathbb{Z})^2$ it defines a subgroup of the group of projective transformations of projective plane whose order is 216. This group is called the *Hesse group*. The elements g of $Sp(2, \mathbb{Z}/3\mathbb{Z})$ transform the curve $F(\lambda) = 0$ to the curve $F(\lambda') = 0$ such that the transformation $g : \lambda \to \lambda'$ defines an isomorphism from the group $Sp(2, \mathbb{Z}/3\mathbb{Z})$ onto the octahedron subgroup of automorphisms of the Riemann sphere $\mathbb{P}^1(\mathbb{C})$.

Example 2. Take $n = 2$ and consider the lattice with the period matrix

$$\Pi = \begin{pmatrix} \sqrt{-2} & \sqrt{-5} & 1 & 0 \\ \sqrt{-3} & \sqrt{-7} & 0 & 1. \end{pmatrix}$$

Suppose there exists a skew-symmetric matrix

$$A = \begin{pmatrix} 0 & b_{12} & b_{13} & b_{14} \\ -b_{12} & 0 & b_{23} & b_{24} \\ -b_{13} & -b_{23} & 0 & b_{34} \\ -b_{14} & -b_{24} & -b_{34} & 0 \end{pmatrix}$$

such that $\Pi \cdot A \cdot \Pi^t = 0$. Then

$$b_{12} + b_{13}\sqrt{-3} + b_{14}\sqrt{-7} - b_{23}\sqrt{-2} + b_{34}\sqrt{-14} - b_{24}\sqrt{-5} - b_{34}\sqrt{-15} = 0.$$

Since the numbers $1, \sqrt{-3}, \sqrt{-7}, \sqrt{-14} - \sqrt{-15}, \sqrt{-2}$ are linearly independent over $\mathbb{R}$, we obtain $A = 0$. This shows that our torus is not algebraic.

Example 3. Let $n = d = 2$ and T be a torus admiting a principal polarization. Then the theta functions $\theta_{\mathbf{m},0}(2\mathbf{z}; 2\Omega), \mathbf{m} \in (\mathbb{Z}/2\mathbb{Z})^2$ define a map of degree 2 from T to a surface of degree 4 in $\mathbb{P}^3$. The surface is isomorphic to the quotient of T by the involution $a \to -a$. It has 16 singular points, the images of 16 fixed points of the involution. The surface is called a *Kummer surface*. Similar to example 1, one can write its equation:

$$\lambda_0(t_0^4 + t_1^4 + t_2^4 + t_3^4) + \lambda_1(t_0^2 t_1^2 + t_2^2 t_3^2) + \lambda_2(t_0^2 t_2^2 + t_1^2 t_3^2) + \lambda_3(t_0^2 t_3^2 + t_1^2 t_2^2) + \lambda_4 t_0 t_1 t_2 t_3 = 0.$$

Here

$$\lambda_0^3 - \lambda_0(\lambda_1^2 + \lambda_2^2 + \lambda_3^2 - \lambda_4^2) + 2\lambda_1\lambda_2\lambda_3 = 0.$$

If $n = 2, d = 3$, the map $f : T \to \mathbb{P}^8$ has its image a certain surface of degree 18. One can write its equations as the intersection of 9 quadric equations.

10.8 When we put the period matrix Π satisfying the Riemann-Frobenius relation in the form $\Pi = [\Omega \ I_n]$, the matrix $\Omega\Delta^{-1}$ belongs to the *Siegel domain*

$$\mathcal{Z}_n = \{Z = X + iY \in M_n(\mathbb{C}) : \Omega = \Omega^t, Y > 0\}.$$

It is a domain in $\mathbb{C}^{n(n+1)/2}$ which is homogeneous with respect to the action of the group

$$Sp(2n, \mathbb{R}) = \{M \in GL(2n, \mathbb{R}) : M \begin{pmatrix} 0_n & I_n \\ -I_n & 0_n \end{pmatrix} \cdot M^t = \begin{pmatrix} 0_n & I_n \\ -I_n & 0_n \end{pmatrix}\}.$$

The group $Sp(2n, \mathbb{R})$ acts on $\mathcal{Z}_n$ by the formula

$$M : Z \to M \cdot Z := (A\Omega + B) \cdot (C\Omega + D)^{-1}, \tag{10.44}$$

where we write M as a block matrix of four $n \times n$-matrices

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}.$$

(one proves that $C\Omega + D$ is always invertible). So we see that any algebraic torus defines a point in the Siegel domain $\mathcal{Z}_n$ of dimension $n(n+1)/2$. This point is defined modulo the transformaton $Z \to M \cdot Z$, where $M \in \Gamma_{n,\Delta} := Sp(2n, \mathbb{Z})_\Delta$ is the subgroup of $Sp(2n, \mathbb{R})$ of matrices with integer entries satisfying

$$M \begin{pmatrix} 0_n & \Delta \\ -\Delta & 0_n \end{pmatrix} \cdot M^t = \begin{pmatrix} 0_n & \Delta \\ -\Delta & 0_n \end{pmatrix}$$

This corresponds to changing the basis of Λ preserving the property that $S(\omega_i, \omega_{j+n}) = d_i\delta_{ij}$. The group $\Gamma_{n,\Delta}$ acts discretely on $\mathcal{Z}_n$ and the orbit space

$$\mathcal{A}_{g,d_1,\ldots,d_n} = \mathcal{Z}_n/\Gamma_{n,\Delta}$$

is the coarse moduli space of abelian varieties of dimension n together with a line bundle with $c_1(L) = S$ where S has the invariants $d_1, \ldots, d_n$. In the special case when $d_1 = \ldots = d_n = d$, we have $\Gamma_{n,\Delta} = Sp(2n, \mathbb{Z})$. It is called the *Siegel modular group* of order n. The corresponding moduli space $\mathcal{A}_{g,d,\ldots,d} \cong \mathcal{A}_{n,1,\ldots,1}$ is denoted by $\mathcal{A}_n$. It parametrizes principal polarized abelian varieties of dimension n. When $n = 1$, we obtain

$$\mathcal{Z}_1 = H := \{z = x + iy : y > 0\}, \ Sp(2, \mathbb{R}) = SL(2, \mathbb{R}).$$

The group Γ_1 is the modular group $\Gamma = SL(2, \mathbb{Z})$. It acts on the upper half-plane H by Moebius transformations $z \to az + b/cz + d$. The quotient H/Γ is isomorphic to the complex plane $\mathbb{C}$.

The theta functions $\theta_{\mathbf{m},\mathbf{m}'}(\mathbf{z}; \Omega)$ can be viewed as holomorphic function in the variable $\Omega \in \mathcal{Z}_n$. When we plug in $z = 0$, we get

$$\theta_{\mathbf{m},\mathbf{m}'}(0; \Omega) = \sum_{\mathbf{r} \in \mathbb{Z}^n} e^{\pi i \left((\frac{1}{d}\mathbf{m}+\mathbf{r}) \cdot \Omega \cdot (\frac{1}{d}\mathbf{m}+\mathbf{r}) + 2(\frac{1}{d}\mathbf{m}') \cdot (\frac{1}{d}\mathbf{m}+\mathbf{r}) \right)}$$

These functions are *modular forms* with respect to some subgroup of Γ_n. For example in the case $n = 1, d = 2k$

$$\theta_0(d\tau, 0) = \sum_{r\in\mathbb{Z}} e^{i\pi 2kr^2\tau}$$

is a modular form of weight k with respect to Γ (see [**Serre**]).

Exercises.

1. Prove that the cohomology group $H^k(\Lambda, \mathbb{Z})$ is isomorphic to $\bigwedge^k(\Lambda)^*$.

2. Let Γ be any group acting holomorphically on a complex manifold M. Define automorphy factors as 1-cocycles from $Z^1(\Gamma, \mathcal{O}(M)^*)$ where Γ acts on $\mathcal{O}(M)^*$ by translation in the argument. Show that the function $\alpha_\gamma(x) = det(d\gamma_x)^k$ is an example of an automorphy factor.

3. Show that any theta factor on $\mathbb{C}^n$ with respect to a lattice Λ is equivalent to a theta factor of the form $e^{a_\gamma\cdot\mathbf{z}+b_\gamma}$ where $a : \Gamma \to \mathbb{C}^n, b : \Lambda \to \mathbb{C}$ are certain functions.

4. Assuming that the previous is true, show that
 (i) a is a homomorphism,
 (ii) $E(\gamma, \gamma') = a_\gamma \cdot \gamma' - a_{\gamma'} \cdot \gamma$ is a skew-symmetric bilinear form,
 (iii) the imaginary part of the function $c(\gamma) = b_\gamma - \frac{1}{2}a_\gamma \cdot \gamma$ is a homomorphism.

5. Show that the group of characters $\chi : T \to \mathbb{C}^*$ is naturally isomorphic to the torus V^*/Λ^*, where $\Lambda^* = \{\phi : V \to \mathbb{R} : \phi(\Lambda) \subset \mathbb{Z}\}$. Put the complex structure on this torus by defining the complex structure on V^* by $J\phi(v) = \phi(Jv)$, where J is a complex structure on V. Prove that
 (i) the complex torus $T^* = V^*/\Lambda^*$ is algebraic if $T = V/\Lambda$ is algebraic;
 (ii) $T^* = V^*/\Lambda^* \cong T$ if T admits a principal polarization;
 (iii) T^* is naturally isomorphic to the kernel of the map $c_1 : Pic(T) \to H^2(T, \mathbb{Z})$.

6. Show that a theta function $\theta_{\mathbf{m},\mathbf{m}'}(\mathbf{z}; \Omega)$ is even if $\mathbf{m} \cdot \mathbf{m}'$ is even and odd otherwise. Compute the number of even functions and the number of odd theta functions $\theta_{\mathbf{m},\mathbf{m}'}(z; \Omega)$ of order d.

7. Find the equations of the image of an elliptic curve under the map given by theta functions $\theta_{\mathbf{m}}(\mathbf{z})$ of order 4.

Lecture 11. FIBRE G-BUNDLES

11.1 From now on we shall study *fields*, like electric, or magnetic or gravitation fields. This will be either a scalar function $\psi(\mathbf{x}, t)$ on the space-time $\mathbb{R}^4$, or, more generally, a section of some vector bundle over this space, or a connection on such a bundle. Many such functions can be interpreted as caused by a particle, like an electron in the case of an electric field. The origin of such functions could be different. For example, it could be a wave function $\psi(x)$ from quantum mechanics. Recall that its absolute value $|\psi(x)|$ can be interpreted as the distribution density for the probability of finding the particle at the point x. If we multiply ψ by a constant $e^{i\theta}$ of absolute value 1, it will not make any difference for the probability. In fact, it does not change the corresponding state, which is the projection operator P_ψ. In other words, the "phase" θ of the particle is not observable. If we write $\psi = \psi_1 + i\psi_2$ as the sum of the real and purely imaginary parts, then we may think of the values of ψ as vectors in a 2-plane $\mathbb{R}^2$ which can be thought of as the "internal space" of the particle. The fact that the phase is not observable implies that no particular direction in this plane has any special physical meaning. In quantum field theory we allow the interpretation to be that of the internal space of a particle. So we should think that the particle is moving from point to point and carrying its internal space with it. The geometrical structure which arises in this way is based on the notion of a fibre bundle. The disjoint union of the internal spaces forms a fibre bundle $\pi : S \to M$. Its fibre over a point $x \in M$ of the space-time manifold M is the internal space S_x. It could be a vector space (then we speak about a vector bundle) or a group (this leads to a principal bundle). For example, we can interpret the phase angle θ as an element of the group $U(1)$. It is important that there is no common internal space in general, so one cannot identify all internal spaces with the same space F. In other words, the fibre bundle is not trivial, in general. However, we allow ourselves to identify the fibres along paths in M. Thus, if $x(t)$ depends on time t, the internal state $\tilde{x}(t) \in S_{x(t)}$ describes a path in S lying over the original path. This leads to the notion of "parallel transport" in the fibre bundle, or equivalently, a " connection ". In general, there is no reason to expect that different paths from x to y lead to the same parallel transport of the internal state. They could differ by application of a symmetry group acting in the fibres (the structure group of the fibre bundle). Physically, this is viewed as a "phase shift". It is produced by the external field.

So, a connection is a mathematical interpretation of this field. Quantitatively, the phase shift is described by the "curvature" of the connection .

If $\phi(x) \in \mathbb{R}^n$ is a section of the trivial bundle, one can define the trivial transport by the differential operators $\phi(x) \to \phi(x + \Delta_\mu x) = \phi(x) + \frac{\partial}{\partial x_\mu}\phi\Delta_\mu x$. In general, still taking the trivial bundle, we have to replace $\partial_\mu = \frac{\partial}{\partial x_\mu}$ with

$$\nabla_\mu = \partial_\mu + A_\mu,$$

where $A_\mu(x) : S_x \to S_x$ is the operator acting in the internal state space S_x (an element of the structure group G). The phase shift is determined by the commutator of the operators ∇_μ:

$$F_{\mu\nu} = [\nabla_\mu, \nabla_\nu] = \partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu].$$

Here the group G is a Lie group, and the commutator takes its value in the Lie algebra $\mathfrak{g}$ of G. The expression $\{F_{\mu\nu}\}$ can be viewed as a differential 2-form on M with values in $\mathfrak{g}$. It is the curvature form of the connection $\{A_\mu\}$.

In general our state bundle $S \to M$ is only *locally trivial.* This means that, for any point $x \in M$, one can find a coordinate system in its neighborhood U such that $\pi^{-1}(U_i)$ can be identified with $U \times F$, and the connection is given as above. If V is another neighborhood, then we assume that the new identification $\pi^{-1}(V) = V \times F$ differs over $U \cap V$ from the old one by the "gauge transformation" $g(x) \in G$ so that $(x, s) \in U \times F$ corresponds to $(x, g(x)s)$ in $V \times F$. We shall see that the connection changes by the formula

$$A_\mu \to g^{-1}A_\mu + g^{-1}\partial_\mu g.$$

Note that we may take $V = U$ and change the trivialization by a function $g : U \to G$. The set of such functions forms a group (infinite-dimensional). It is called the gauge group.

This constitutes our introduction for this and the next lecture. Let us go to mathematics and give precise definitions of the mathematical structures involved in definitions of fields.

11.2 We shall begin with recalling the definition of a fibre bundle.

Definition. Let F be a smooth manifold and G be a subgroup of its group of diffeomorphisms. Let $\pi : S \to M$ be a smooth map. A family $\{(U_i, \phi_i)\})_{i\in I}$ of pairs (U_i, ϕ_i), where U_i is an open subset of M and $\phi_i : U_i \times F \to \pi^{-1}(U_i)$ is a diffeomorphism, is called a *trivializing family* of π if

(i) for any $i \in I$, $\pi \circ \phi_i = pr_1$;

(ii) for any $i, j \in I, \phi_j^{-1}\circ\phi_i : (U_i\cap U_j)\times F \to (U_i\cap U_j)\times F$ is given by $(x, a) \to (x, g_{ij}(x)(a))$, where $g_{ij}(x) \in G$.

The open cover $(U_i)_{i\in I}$ is called the *trivializing cover.* The corresponding diffeomorphisms ϕ_i are called the *trivializing diffeomorphisms.* The functions $g_{ij} : U_i\cap U_j \to G, x \to g_{ij}(x)$, are called the *transition functions* of the trivializing family. Two trivializing families $\{(U_i, \phi_i)\}_{i\in I}$ and $\{(V_j, \psi_j)\}_{j\in J}$ are called equivalent if for any $i \in I, j \in J$, the map $\psi_j^{-1} \circ \phi_i : (U_i \cap V_j) \times F \to \pi^{-1}(U_i \cap V_j)$ is given by a function $(x, a) \to (x, g(x)(a))$, where $g(x) \in G$.

A *fibre G-bundle with typical fibre F* is a smooth map $\pi : S \to M$ together with an equivalence class of trivializing families.

Let $\pi : S \to M$ be a fibre G-bundle with typical fibre F. For any $x \in X$ we can find a pair (U, ϕ) from some trivializing family such that $x \in U$. Let $\phi_x : F \to S_x := \pi^{-1}(x)$ be the restriction of ϕ to $\{x\} \times F$. We call such a map a *fibre marking*. If $\psi_x : F \to F$ is another fibre marking, then, by definition, $\psi_x^{-1} \circ \phi_x : F \to F$ belongs to the group G. In particular, if two markings are defined by (U_i, ϕ_i) and (U_j, ϕ_j) from the same family of trivializing diffeomorphisms, then, for any $a \in F$,

$$(\phi_i)_x(a) = (\phi_j)_x(g_{ij}(x)(a)), \tag{11.1}$$

where g_{ij} is the transition function of the trivializing family.

Fibre G-bundles form a category with respect to the following definition of morphism:

Definition. Let $\pi : S \to M$ and $\pi' : S' \to M$ be two fibre G-bundles with the same typical fibre F. A smooth map $f : S \to S'$ is called a *morphism* of fibre G-bundles if $\pi = \pi' \circ f$ and, for any $x \in M$ and fibre markings $\phi_x : F \to S_x, \psi_x : F \to S'_x$, the composition $\psi_x^{-1} \circ \phi_x : F \to F$ belongs to G.

A fibre G-bundle is called *trivial* if it is isomorphic to the fibre G-bundle $pr_1 : M \times F \to M$ with the trivializing family $\{(M, id_{M\times F})\}$.

Let $f : S \to S'$ be a morphism of two fibre G-bundles. One may always find a common trivializing open cover $(U_i)_{i\in I}$ for π and π'. Let $\{\phi_i\}_{i\in I}$ and $\{\psi_i\}_{i\in I}$ be the corresponding sets of trivializing diffeomorphisms. Then a morphism $f : S \to S'$ is defined by the maps $f_i : U_i \to G$ such that the map $(x, a) \to (x, f_i(x)(a))$ is a diffeomorphism of $U_i \times F$, and

$$f_i(x) = g'_{ij}(x)^{-1} \circ f_j(x) \circ g_{ij}(x), \quad \forall x \in U_i \cap U_j, \tag{11.2}$$

where g_{ij} and g'_{ij} are the transition functions of $\{(U_i, \phi_i)\}$ and $\{(U_i, \psi_i)\}$, respectively.

Remark 1. The group $\mathrm{Diff}(F)$ has a natural structure of an infinite dimensional Lie group. We refer for the definition to [**Pressley**]. To simplify our life we shall work with fibre G-bundles, where G is a finite-dimensional Lie subgroup of $\mathrm{Diff}(F)$. One can show that the functions g_{ij} are smooth functions from $U_i \cap U_j$ to G.

Let $(U_i)_{i\in I}$ be a trivializing cover of a fibre G-bundle $S \to M$ and g_{ij} be its transition functions. They satisfy the following obvious properties:

(i) $g_{ii} \equiv 1 \in G$,
(ii) $g_{ij} = g_{ji}^{-1}$,
(iii) $g_{ij} \circ g_{jk} = g_{ik}$ over $U_i \cap U_j \cap U_k$.

Let g'_{ij} be the set of transition functions corresponding to another trivializing family of $S \to M$ with the same trivializing cover. Then for any $x \in U_i \cap U_j$,

$$g'_{ij}(x) = h_i(x)^{-1} \circ g_{ij}(x) \circ h_j(x),$$

where

$$h_i(x) = (\psi_i)_x^{-1} \circ (\phi_i)_x : U_i \to G, i \in I.$$

We can give the following cohomological interpretation of transition functions. Let $\mathcal{O}_M(G)$ denote the sheaf associated with the pre-sheaf of groups $U \to$ {smooth functions from U to G}. Then the set of transition functions $\{g_{ij}\}_{i,j\in I}$ with respect to a trivializing family $\{(U_i, \phi_i)\}$ of a fibre G-bundle can be identified with a Čech 1-cocycle

$$\{g_{ij}\} \in Z^1(\{U_i\}_{i\in I}, \mathcal{O}_M(G)).$$

Recall that two cocycles $\{g_{ij}\}$ are called equivalent if they satisfy (11.2) for some collection of smooth functions $f_i : U_i \to G$. The set of equivalence classes of 1-cocycles is denoted by $\check{H}^1(\{U_i\}_{i\in I}, \mathcal{O}_M(G))$. It has a marked element 1 corresponding to the equivalence class of the cocycle $g_{ij} \equiv 1$. This element corresponds to a trivial G-bundle. By taking the inductive limit of the sets $\check{H}^1(\{U_i\}_{i\in I}, \mathcal{O}_M(G))$ with respect to the inductive set of open covers of M, we arrive at the set $H^1(M, \mathcal{O}_M(G))$. It follows from above that a choice of a trivializing family and the corresponding transition functions g_{ij} defines an injective map from the set $FIB_M(F;G)$ of isomorphism classes of fibre G-bundles over M with typical fibre F to the cohomology set $H^1(M, \mathcal{O}_M(G))$. Conversely, given a cocycle $g_{ij} : U_i \cap U_j \to G$ for some open cover $(U_i)_{i\in I}$ of M, one may define the fibre G-bundle as the set of equivalence classes

$$S = \coprod_{i\in I} U_i \times F/R, \tag{11.3}$$

where $(x_i, a_i) \in U_i \times F$ is equivalent to $(x_j, a_j) \in U_j \times F$ if $x_i = x_j = x \in U_i \cap U_j$ and $a_i = g_{ij}(x)(a_j)$. The properties (i),(ii),(iii) of a cocycle are translated into the definition of equivalence relation. The structure of a smooth manifold on S is defined in such a way that the factor map is a local diffeomorphism. The projection $\pi : S \to M$ is defined by the projections $U_i \times F \to U_i$. It is clear that the cover $(U_i)_{i\in I}$ is a trivializing cover of S and the trivializing diffeomorphisms are the restrictions of the factor map onto $U_i \times F$.

Summing up, we have the following:

Theorem 1. *A choice of transition functions defines a bijection*

$$FIB_M(F;G) \longleftrightarrow H^1(M, \mathcal{O}_M(G))$$

between the set of isomorphism classes of fibre G-bundles with typical fibre F and the set of Čech 1-cohomology with coefficients in the sheaf $\mathcal{O}_M(G)$. Under this map the isomorphism class of the trivial fibre G-bundle $M \times F$ corresponds to the cohomology class of the trivial cocycle.

A *cross-section* (or just a *section*) of a fibre G-bundle $\pi : S \to M$ is a smooth map $s : M \to S$ such that $p \circ s = id_M$. If $\{U_i, \phi_i\}_{i\in I}$ is a trivializing family, then a section s is defined by the sections $s_i = \phi^{-1} \circ s|U_i : U_i \to U_i \times F$ of the trivial fibre bundle $U_i \times F \to U_i$. Obviously, we can identify s_i with a smooth function $s_i : U_i \to F$ such that $s_i(x) = (x, s_i(x))$. It follows from (11.1) that, for any $x \in U_i \cap U_j$,

$$s_j(x) = g_{ij}(x) \circ s_i(x), \tag{11.4}$$

where $g_{ij} : U_i \cap U_j \to G$ are the transition functions of $\pi : S \to M$ with respect to $\{U_i, \phi_i\}_{i\in I}$.

Finally, for any smooth map $f : M' \to M$ and a fibre G-bundle $\pi : S \to M$ we define the *induced fibre bundle* (or the *inverse transform*) $f^*(S) = S \times_M M' \to M'$. It has a natural structure of a fibre G-bundle with typical fibre F. Its transition families are inverse transforms

$$f^*(\phi_i) = \phi_i \times id : f^{-1}(U_i) \times F = (U_i \times F) \times_{U_i} f^{-1}(U_i) \to \pi^{-1}(U_i) \times_{U_i} f^{-1}(U_i).$$

of the transition functions of $S \to M$. If $M' \to M$ is the identity map of an open submanifold M' of M, we denote the induced bundle by $S|M'$ and call it the *restriction* of the bundle to M'.

11.3 A sort of universal example of a fibre G-bundle is a *principal G-bundle*. Here we take for typical fibre F a Lie group G. The structure group is taken to be G considered as a subgroup of Diff(F) consisting of left translations $L_g : a \to g \cdot a$. Let $PFIB_M(G)$ denote the set of isomorphism classes of principal G-bundles. Applying Theorem 1, we obtain a natural bijection

$$PFIB_M(G) \longleftrightarrow H^1(M, \mathcal{O}_M(G)). \tag{11.5}$$

Let $s_U : U \to P$ be a section of a principal G-bundle over an open subset U of M. It defines a trivialization of $\phi_U : U \times G \to \pi^{-1}(U)$ by the formula

$$\phi_U((x, g)) = s_U(x) \cdot g.$$

Since $\phi_U((x, g \cdot g')) = s_U(x) \cdot (g \cdot g') = (s_U(x) \cdot g) \cdot g' = \phi_U((x, g)) \cdot g'$, the map ϕ_U is an isomorphism from the trivial principal G-bundle $U \times G$ to $P|U$. Let $\{(U_i, \phi_i)\}_{i\in I}$ be a trivializing family of P. If $s_i : U_i \to P$ is a section over U_i, then for any $x \in U_i \cap U_j$, we can write $s_i(x) = s_j(x) \cdot u_{ij}(x)$ for some $u_{ij} : U_i \cap U_j \to G$. The corresponding trivializations are related by

$$\phi_{U_i}((x, g)) = s_i(x) \cdot g = (s_j(x) \cdot u_{ij}(x)) \cdot g = \phi_{U_j}((x, u_{ij}(x) \cdot g)). \tag{11.6}$$

Comparing this with formula (11.1), we find that the function u_{ij} coincides with the transition function g_{ij} of the fibration P.

Let $\pi : P \to M$ be a principal G-bundle. The group G acts naturally on P by right translations along fibres $g' : (x, g) \to (x, g \cdot g')$. This definition does not depend on the local trivialization of P because left translations commute with right translations. It is clear that orbits of this action are equal to fibres of $P \to M$, and $P/G \cong M$.

Let $P \to M$ and $P' \to M$ be two principal G-bundles. A smooth map (over M) $f : P \to P'$ is a morphism of fibre G-bundles if and only if, for any $g \in G, p \in P$,

$$f(p \cdot g) = f(p) \cdot g.$$

This follows easily from the definitions. One can use this to extend the notion of a morphism of G-bundles for principal bundles with not necessarily the same structure group.

Definition. Let $\pi : P \to M$ be a principal G-bundle and $\pi' : P' \to M$ be a principal G'-bundle. A smooth map $f : P \to P'$ is a morphism of principal bundles if $\pi = \pi' \circ f$ and there exists a morphism of Lie groups $\phi : G \to G''$ such that, for any $p \in P, g \in G$,

$$f(p \cdot g) = f(p) \cdot \phi(g).$$

Any fibre G-bundle with typical fibre F can be obtained from some principal G-bundle by the following construction. Recall that a homomorphism of groups $\rho : G \to \mathrm{Diff}(F)$ is called a *smooth action* of G on F if for any $a \in F$, the map $G \to F, g \to \rho(g)(x)$ is a smooth map. A smooth action is called *faithful* if its kernel is trivial. Fix a principal G-bundle $P \to M$ and a smooth faithful action $\rho : G \to \mathrm{Diff}(F)$ and consider a fibre $\rho(G)$-bundle with typical fibre F over M defined by (11.4) using the transition functions $\rho(g_{ij})$, where g_{ij} is a set of transition functions of P. This bundle is called the fibre G-bundle associated to a principal G-bundle by means of the action $\rho : G \to \mathrm{Diff}(F)$ and transition functions g_{ij}. Clearly, a change of transition functions changes the bundle to an isomorphic G-bundle. Also, two representations with the same image define the same G-bundles. Finally, it is clear that any fibre G-bundle is associated to a principal G'-bundle, where $G' \to G$ is an isomorphism of Lie groups.

There is a canonical construction for an associated G-bundle which does not depend on the choice of transition functions.

Let $\rho : G \to \mathrm{Diff}(F)$ be a faithful smooth action of G on F. The group G acts on $P \times F$ by the formula $g : (p, a) \to (p \cdot g, \rho(g^{-1})(a))$. We introduce the orbit space

$$P \times_\rho F := P \times F/G, \tag{11.7}$$

Let $\pi' : P \times_\rho F \to M$ be defined by sending the orbit of (p, a) to $\pi(p)$. It is clear that, for any open set $U \subset M$, the subset $\pi^{-1}(U) \times F$ of $P \times F$ is invariant with respect to the action of G, and $\pi'^{-1}(U) = \pi^{-1}(U) \times F/G$. If we choose a trivializing family $\phi_i : U_i \times G \to \pi^{-1}(U_i)$, then the maps

$$(U_i \times G) \times F \to U_i \times F, ((x, g), a) \to (x, \rho(g)(x))$$

define, after passing to the orbit space, diffeomorphisms $\pi'^{-1}(U) \to U_i \times F$. Taking inverses we obtain a trivializing family of $\pi' : P \times_\rho F \to M$. This defines a structure of a fibre G-bundle on $P \times_\rho F$.

The principal bundle P acts on the associated fibre bundle $P \times_\rho F$ in the following sense. There is a map $(P \times_\rho F) \times P \to P \times_\rho F$ such that its restriction to the fibres over a point $x \in M$ defines the map $F \times P_x \to F$ which is the action ρ of $G = P_x$ on F.

Let $S \to M$ be a fibre G-bundle and G' be a subgroup of G. We say that the structure group of S can be *reduced* to G' if one can find a trivializing family such that its transition functions take values in G'. For example, the structure group can be reduced to the trivial group if and only if the bundle is trivial.

Let G' be a Lie subgroup of a Lie group G. It defines a natural map of the cohomology sets

$$i : H^1(M, \mathcal{O}_M(G')) \to H^1(M, \mathcal{O}_M(G)).$$

A principal G-bundle P defines an element in the image of this map if and only if its structure group can be reduced to G'.

11.4 We shall deal mainly with the following special case of fibre G-bundles:

Definition. A rank n *vector bundle* $\pi : E \to M$ is a fibre G-bundle with typical fibre F equal to $\mathbb{R}^n$ with its natural structure of a smooth manifold. Its structure group G is the group $GL(n, \mathbb{R})$ of linear automorphisms of $\mathbb{R}^n$.

The local trivializations of a vector bundle allow one to equip its fibres with the structure of a vector space isomorphic to $\mathbb{R}^n$. Since the transition functions take values in $GL(n, \mathbb{R})$, this definition is independent of the choice of a set of trivializing diffeomorphisms. For any section $s : M \to E$ its value $s(x) \in E_x = \pi^{-1}(x)$ cannot be considered as an element in $\mathbb{R}^n$ since the identification of E_x with $\mathbb{R}^n$ depends on the trivialization. However, the expression $s(x) = 0 \in E_x$ is well-defined, as well as the sum of the sections $s + s'$ and the scalar product λs. The section s such that $s(x) = 0$ for all $x \in M$ is called the *zero section*. The set of sections is denoted by $\Gamma(E)$. It has the structure of a vector space. It is easy to see that

$$\mathcal{E} : U \to \Gamma(E|U)$$

is a sheaf of vector spaces. We shall call it the *sheaf* of sections of E.

If we take $E = M \times \mathbb{R}$ the trivial vector bundle, then its sheaf of sections is the structure sheaf $\mathcal{O}_M$ of the manifold M.

One can define the category of vector bundles. Its morphisms are morphisms of fibre bundles such that restriction to a fibre is a linear map. This is equivalent to a morphism of fibre $GL(n, \mathbb{R})$-bundles.

We have

$$\{\text{ rank } n \text{ vector bundles over } M/\text{isomorphism}\} \longleftrightarrow H^1(M, \mathcal{O}_M(GL(n, \mathbb{R}))).$$

It is clear that any rank n vector bundle is associated to a principal $GL(n, \mathbb{R})$-bundle by means of the identity representation id. Now let G be any Lie group. We say that E is a G-bundle if E is isomorphic to a vector bundle associated with a principal G-bundle by means of some linear representation $G \to GL(n, \mathbb{R})$. In other words, the structure group of E can be reduced to a subgroup $\rho(G)$ where $\rho : G \to GL(n, \mathbb{R}) \subset \mathrm{Diff}(\mathbb{R}^n)$ is a faithful smooth (linear) action. It follows from the above discussion that

$$\{\text{rank } n \text{ vector } G\text{-bundles/isomorphism}\} \longleftrightarrow H^1(M, \mathcal{O}_M(G)) \times Rep_n(G).$$

where $\mathrm{Rep}_n(G)$ stands for the set of equivalence classes of faithful linear representations $G \to GL(n, \mathbb{R})$ ($\rho \sim \rho'$ if there exists $A \in GL(n, \mathbb{R})$ such that $\rho(g) = A\rho'(g)A^{-1}$ for all $g \in G$).

For example, we have the notion of an orthogonal vector bundle, or a unitary vector bundle.

Example 1. The tangent bundle $T(M)$ is defined by transition functions g_{ij} equal to the Jacobian matrices $J = (\frac{\partial x_\alpha}{\partial y_\beta})$, where $(x_1, \ldots, x_n), (y_1, \ldots, y_n)$ are local coordinate functions in U_i and U_j, respectively. The choice of local coordinates defines the trivialization $\phi_i : U_i \times \mathbb{R}^n \to \pi^{-1}(U_i)$ by sending $(x, (a_1, \ldots, a_n))$ to $(x, \sum_{i=1}^n a_i \frac{\partial}{\partial x_i})$. A section of the tangent bundle is a vector field on M. It is defined locally by $\eta = \sum_{i=1}^n a_i(x) \frac{\partial}{\partial x_i}$, where $a_i(x)$ are smooth functions on U_i. If $\eta = \sum_{i=1}^n b_i(y) \frac{\partial}{\partial y_i}$ represents η in U_j, we have

$$a_j(y) = \sum_{i=1}^n b_i(x) \frac{\partial x_j}{\partial y_i}.$$

One can prove (or take it as a definition) that the structure group of the tangent bundle can be reduced to $SL(n, \mathbb{R})$ if and only if M is orientable. Also, if M is a symplectic manifold of dimension $2n$, then using Darboux's theorem one checks that the structure group of $T(M)$ can be reduced to the symplectic group $Sp(2n, \mathbb{R})$.

Let $F(M)$ be the principal $GL(n, \mathbb{R})$-bundle over M with transition functions defined by the Jacobian matrices J. In other words, the tangent bundle $T(M)$ is associated to $F(M)$. Let $e_1, \ldots, e_n$ be the standard basis in $\mathbb{R}^n$. The correspondence $A \to (Ae_1, \ldots, Ae_n)$ is a bijective map from $GL(n, \mathbb{R})$ to the set of bases (or frames) in $\mathbb{R}^n$. It allows one to interpret the principal bundle $F(M)$ as the bundle of frames in the tangent spaces of M.

Example 2. Let W be a complex vector space of dimension n and $W_{\mathbb{R}}$ be the corresponding real vector space. We have $GL_{\mathbb{C}}(W) = \{g \in GL(W_{\mathbb{R}}) : g \circ J = J \circ g\}$, where $J : W \to W$ is the operator of multiplication by $i = \sqrt{-1}$. In coordinate form, if $W = \mathbb{C}^n, W_{\mathbb{R}} = \mathbb{R}^{2n} = \mathbb{R}^n + i(\mathbb{R}^n)$, $GL(W) = GL(n, \mathbb{C})$ can be identified with the subgroup of $GL(W_{\mathbb{R}}) = GL(2n, \mathbb{R})$ consisting of invertible matrices of the form $\begin{pmatrix} A & B \\ -B & A \end{pmatrix}$, where $A, B \in Mat_n(\mathbb{R})$. The identification map is

$$X = A + iB \in GL(n, \mathbb{C}) \to \begin{pmatrix} A & B \\ -B & A \end{pmatrix} \in GL(2n, \mathbb{R}).$$

One defines a rank n *complex vector bundle* over M as a real rank $2n$ vector bundle whose structure group can be reduced to $GL(n, \mathbb{C})$. Its fibres acquire a natural structure of a complex n-dimensional vector space, and its space of sections is a complex vector space.

Suppose M can be equipped with a structure of an n-dimensional complex manifold. Then at each point $x \in M$ we have a system of local complex coordinates $z_1, \ldots, z_n$. Let $x_i = (z_i + \bar{z}_i)/2, y_i = (z_i - \bar{z}_i)/2i$ be a system of local real parameters. Let $z'_1, \ldots, z'_n$ be another system of local complex parameters, and $(x'_i, y'_i), i = 1, \ldots, n$, be the corresponding system of real parameters. Since $(z'_1, \ldots, z'_n) = (f_1(z_1, \ldots, z_n), \ldots, f_n(z_1, \ldots, z_n))$ is a holomorphic coordinate change, the functions $f_i(z_1, \ldots, z_n)$ satisfy the Cauchy-Riemann conditions

$$\frac{\partial x'_i}{\partial x_j} = \frac{\partial y'_i}{\partial y_j}, \quad \frac{\partial x'_i}{\partial y_j} = -\frac{\partial y'_i}{\partial x_j},$$

and the Jacobian matrix has the form

$$\frac{\partial(x'_1, \ldots, x'_n, y'_1, \ldots, y'_n)}{\partial(x_1, \ldots, x_n, y_1, \ldots, y_n)} = \begin{pmatrix} A & B \\ -B & A \end{pmatrix},$$

where
$$A = \frac{\partial(x'_1, \dots, x'_n)}{\partial(x_1, \dots, x_n)}, \quad B = \frac{\partial(x'_1, \dots, x'_n)}{\partial(y_1, \dots, y_n)}.$$
This shows that $T(M)$ can be equipped with the structure of a complex vector bundle of dimension n.

11.5 Let E be a rank n vector bundle and $P \to M$ be the corresponding principal $GL(n, \mathbb{R})$-bundle. For any representation $\rho : GL(n, \mathbb{R}) \to GL(n, \mathbb{R})$ we can construct the associated vector bundle $E(\rho)$ with typical fibre $\mathbb{R}^n$ and the structure group $\rho(G) \subset GL(n, \mathbb{R})$. This allows us to define many operations on bundles. For example, we can define the exterior product $\bigwedge^k(E)$ as the bundle $E(\rho)$, where

$$\rho : GL(n, \mathbb{R}) \to GL(\bigwedge^k(\mathbb{R}^n)), A \to \bigwedge^k A.$$

Similarly, we can define the tensor power $E^{\otimes k}$, the symmetric power $S^k(E)$, and the dual bundle E^*. Also we can define the tensor product $E \otimes E'$ (resp. direct sum $E \oplus E'$) of two vector bundles. Its typical fibre is $\mathbb{R}^{nm}$ (resp. $\mathbb{R}^{n+m}$), where $n = \text{rank } E, m = \text{rank } E'$. Its transition functions are the tensor products (resp. direct sums) of the transition matrix functions of E and E'.

The vector bundle $E^* \otimes E$ is denoted by $\text{End}(E)$ and can be viewed as the bundle of endomorphisms of E. Its fibre over a point $x \in M$ is isomorphic (under a trivialization map) to $V^* \otimes V = \text{End}(V, V)$. Its sections are morphisms of vector bundles $E \to E$.

Example 3. Let $E = T(M)$ be the tangent bundle. Set

$$T^p_q(M) = T(M)^{\otimes p} \otimes T^*(M)^{\otimes q}. \tag{11.8}$$

A section of $T^p_q(M)$ is called a p-covariant and q-contravariant *tensor* over M (or just a tensor of type (p, q)). We have encountered them earlier in the lectures. Tensors of type $(0, 0)$ are scalar functions $f : M \to \mathbb{R}$. Tensors of type $(1, 0)$ are vector fields. Tensors of type $(0, 1)$ are differential 1-forms. They are locally represented in the form

$$\theta = \sum_{i=1}^n a_i(x) dx_i,$$

where $dx_j(\frac{\partial}{\partial x_i}) = \delta_{ij}$ and $a_i(x)$ are smooth functions in U_i. If $\theta = \sum_{i=1}^n b_i(y) dy_i$ represents θ in U_j, then
$$a_j(x) = \sum_{i=1}^n \frac{\partial y_i}{\partial x_j} b_i(y).$$

In general, a tensor of type (p, q) can be locally given by

$$\tau = \sum_{i_1, \dots, i_p = 1}^n \sum_{j_1, \dots, j_q = 1}^n a^{i_1 \dots i_p}_{j_1 \dots j_q} \frac{\partial}{\partial x_{i_1}} \otimes \dots \otimes \frac{\partial}{\partial x_{i_p}} \otimes dx_{j_1} \otimes \dots \otimes dx_{j_q}.$$

So such a tensor is determined by collections of smooth functions $a^{i_1\ldots i_p}_{j_1\ldots j_q}$ assigned to an open subset U_i with local cooordinates $x_1,\ldots,x_n$. If the same tensor is defined by a collection $b^{i_1\ldots i_p}_{j_1\ldots j_q}$ in an open subset U_j with local coordinates $y_1,\ldots,y_n$, we have

$$a^{i_1\ldots i_p}_{j_1\ldots j_q} = \frac{\partial x_{i_1}}{\partial y_{i'_1}}\cdots\frac{\partial x_{i_p}}{\partial y_{i'_p}}\frac{\partial y_{j'_1}}{\partial x_{j_1}}\cdots\frac{\partial y_{j'_q}}{\partial x_{j_q}}b^{i'_1\ldots i'_p}_{j'_1\ldots j'_q},$$

where we use the "Einstein convention" for summation over the indices $i'_1,\ldots,i'_p,j'_1,\ldots,j'_q$.

A section of $\bigwedge^k T^*(M)$ is a smooth differential k-form on M. We can view it as an anti-symmetric tensor of type $(0,k)$. It is defined locally by $\sum_{i_1,\ldots,i_k=1}^n a_{i_1\ldots i_k}dx_{i_1}\otimes\ldots\otimes dx_{i_k}$, where

$$a_{i_{\sigma(1)}\ldots i_{\sigma(k)}} = \epsilon(\sigma)a_{i_1\ldots i_k}$$

where $\epsilon(\sigma)$ is the sign of the permutation σ. This can be rewritten in the form

$$\sum_{1\le i_1<\ldots<i_k\le n} a_{i_1\ldots i_k}dx_{i_1}\wedge\ldots\wedge dx_{i_k},$$

where

$$dx_{i_1}\wedge\ldots\wedge dx_{i_k} = \sum_{1\le i_1<\ldots<i_k\le n}\epsilon(\sigma)dx_{i_{\sigma(1)}}\otimes\ldots\otimes dx_{i_{\sigma(k)}}.$$

In particular, $\bigwedge^n(T^*(M))$ is a rank 1 vector bundle. It is called the *volume bundle*. Its sections over a trivializing open set U_i can be identified with scalar functions $a_{12\ldots n}$. Over $U_i\cap U_j$ these functions transform according to the formula

$$a_{12\ldots n}(x) = \left|\frac{\partial(y_1,\ldots,y_n)}{\partial(x_1,\ldots,x_n)}\right|b_{12\ldots n}(y(x)).$$

A manifold M is *orientable* if one can find a section of its volume bundle which takes positive values at any point. Such a section is called a *volume form*. We shall always assume that our manifolds are orientable. Obviously, any two volume forms on M are obtained from each other by multiplying with a global scalar function $f : M \to \mathbb{R}_{>0}$. A volume form Ω defines a distribution on M, i.e., a continuous linear functional on the space $C_0^\infty(M)$ of smooth functions on M with compact support. Its values on $\phi\in C_0^\infty(M)$ is the integral $\int_M \phi(x)\Omega$. For example, if $M=\mathbb{R}^n$, we may take $\Omega = dx_1\wedge\ldots\wedge dx_n$. In this case, physicists use the notation

$$\int_{\mathbb{R}^n}\phi(x)dx_1\wedge\ldots\wedge dx_n = \int d^nx\phi(x).$$

Example 4. Recall that a structure of a *pseudo-Riemannian manifold* on a connected smooth manifold M is defined by a smooth function on $Q : T(M)\to\mathbb{R}$ such that its restriction to each fibre $T(M)_x$ is a non-degenerate quadratic form. The signature of $Q|T(M)_x$ is independent of x. It is called the *signature* of the pseudo-Riemannian manifold

(M, Q). We say that (M, Q) is a Riemannian manifold if the signature is $(n, 0)$, or in other words, $Q_x = Q|T(M)_x$ is positive definite. We say that (M, Q) is *Lorentzian* (or *hyperbolic*) if the signature is of type $(n-1, 1)$ (or $(1, n-1)$). For example, the space-time $\mathbb{R}^4$ is a Lorentzian pseudo-Riemannian manifold. One can define a pseudo-metric Q by the corresponding symmetric bilinear form $B : T(M) \times T(M) \to \mathbb{R}$. It can be viewed as a section of $T^*(M) \otimes T^*(M)$. It is locally defined by

$$B(x) = \sum_{i,j=1}^{n} a_{ij}(x) dx_i \otimes dx_j, \tag{11.9}$$

where $a_{ij}(x) = a_{ji}(x)$. The associated quadratic form is given by

$$Q_x = \sum_{i=1}^{n} a_{ii} dx_i^2 + 2 \sum_{1 \le i < j \le n} a_{ij}(x) dx_i dx_j.$$

Its value on a tangent vector $\eta_x = \sum_{i=1}^{n} c_i \frac{\partial}{\partial x_i} \in T(M)_x$ is equal to

$$Q_x(\eta_x) = \sum_{i=1}^{n} a_{ii} c_i^2 + 2 \sum_{1 \le i < j \le n} a_{ij}(x) c_i c_j.$$

Using a pseudo-Riemannian metric one can define a subbundle of the frame bundle of $T(M)$ whose fibre over a point $x \in M$ consists of orthonormal frames in $T(M)_x$. This allows us to reduce the structure of the frame bundle to the orthogonal group $O(k, n-k)$ of the quadratic form $x_1^2 + \ldots + x_k^2 - x_{k+1}^2 - \ldots - x_n^2$. The structure group of the tangent bundle cannot be reduced to this subgroup unless the curvature of the metric is zero (see below).

11.6 The group of automorphisms of a fibre G-bundle P is called the *gauge group* of P and is denoted by $\mathcal{G}(P)$.

Theorem 2. *Let P be a principal G-bundle and let P^c be the associated bundle with typical fibre G and representation $G \to Aut(G)$ given by the conjugation action $g \cdot h = g \cdot h \cdot g^{-1}$ of G on itself. Then*

$$\mathcal{G}(P) \cong \Gamma(P^c).$$

Proof. This is obviously true if P is trivial. In this case the both groups are isomorphic to the group of smooth functions $M \to G$. Let $(U_i)_{i \in I}$ be a trivializing cover of P with trivializing diffeomorphims $\phi_i : U_i \times G \to P|U_i$. Any element $g \in \mathcal{G}(P)$ defines, after restrictions, the automorphisms $g_i : P|U_i \to P|U_i$ and hence the automorphisms $s_i = \phi_i^{-1} \circ g \circ \phi_i$ of $U_i \times G$. By the above, s_i is a section of $\mathcal{O}_M(G)(U_i)$. If $\psi_j : V_j \times G \to P|V_j$ is another set of automorphisms, and $s'_j \in \mathcal{O}_M(V_j)$ is the corresponding automorphism of $V_j \times G$, then $s'_j = \psi_j^{-1} \circ g \circ \psi_j$. Comparing these two sections over $U_i \cap U_j$, we get

$$s_i = (\phi_i^{-1} \circ \psi_j) \circ s'_j \circ (\psi_j^{-1} \circ \psi_i) = g_{ij}^{-1} \circ s'_j \circ g_{ij}.$$

This shows that g defines a section of P^c. Conversely, every section (s_i) of P^c defines automorphisms of trivialized bundles $P|U_i$ which agree on the intersections.

Let $S \to M$ be a fibre G-bundle associated to a principal G-bundle P. Let $\rho : G \to \mathrm{Diff}(F)$ be the corresponding representation. Then it defines an isomomorphism of the gauge groups $\mathcal{G}(P) \to \mathcal{G}(S)$. Locally it is given by $g(x) \to \rho(g(x))$.

Exercises.

1. Find a non-trivial vector bundle E over circle S^1 such that $E \oplus E$ is trivial.

2. Let $\mathbb{C}^{n+1} \setminus \{0\} \to \mathbb{P}^n(\mathbb{C})$ be the quotient map from the definition of projective space. Show that it is a principal $\mathbb{C}^*$-bundle over $\mathbb{P}^n(\mathbb{C})$. It is called the Hopf bundle (in the case $n = 1$ its restriction to the sphere $S^3 = \{(z_1, z_2) \in \mathbb{C}^2 : |z_1|^2 + |z_2|^2 = 1\}$ is the usual Hopf bundle $S^3 \to S^2$). Find its transition functions with respect to the standard open cover of projective space.

3. A manifold M is called *parallelizable* if its tangent bundle is trivial. Show that S^1 is parallelizable (a much deeper result due to Milnor and Kervaire is that S^1, S^3 and S^7 are the only parallelizable spheres).

4. Show that a rank n vector bundle is trivial if and only if it admits n sections whose values at any point are linearly independent.

5. Find transition functions of the tangent bundle to $\mathbb{P}^n(\mathbb{C})$. Show that its direct sum with the trivial rank 2 vector bundle is isomorphic to $L^{\oplus n+1}$, where L is the rank 2 vector bundle associated to the Hopf bundle by means of the standard action of $\mathbb{C}^*$ in $\mathbb{R}^2$.

6. Prove that any fibre $\mathbb{R}$-bundle over a compact manifold is trivial.

7. Using the previous problem prove that the structure group of any fibre $\mathbb{C}^*$-bundle (resp. $\mathbb{R}^*$) can be reduced to $U(1)$ (resp. $\mathbb{Z}/2\mathbb{Z}$).

8. Show that any principal G-bundle with finite group G is isomorphic to a covering space $N \to M$ with the group of deck transformations isomorphic to G. Using the previous two problems, prove that any line bundle over a simply-connected compact manifold is trivial.

Lecture 12. GAUGE FIELDS

In this lecture we shall define connection (or *gauge field*) on a vector bundle.

12.1 First let us fix some notation from the theory of Lie groups. Recall that a *Lie group* is a group object G in the category of smooth manifolds. For any $g \in G$ we denote by L_g (resp. R_g) the left (resp. right) translation action of the group G on itself. The differentials of the maps R_g, L_g transform a vector field η to the vector field $(L_g)_*(\eta), (R_g)_*(\eta)$, respectively. The *Lie algebra* of G is the vector space $\mathfrak{g}$ of vector fields which are invariant with respect to left translations. Its Lie algebra structure is the bracket operation of vector fields. For any $\eta \in \mathfrak{g}$, the vector $(dL_{g^{-1}})_g(\eta_g)$ belongs to the tangent space $T(G)_1$ of G at the identity element 1. It is independent of $g \in G$. This defines a bijective linear map $\omega_G : \mathrm{Lie}(G) \to T(G)_1$. By transferring the Lie bracket, we may identify the Lie algebra of G with the tangent space of G at the identity element $1 \in G$. Let $v \in T(G)_1$ and $\tilde{v}$ be the corresponding vector field. We have $\tilde{v}_g = (dL_g)_1(v)$, so that $(dR_{g^{-1}})_g(\tilde{v}_g) = dc(g)_1(v)$, where $c(g) : G \to G$ is the conjugacy action $h \to g \cdot h \cdot g^{-1}$. The action $v \to dc(g)_1(v)$ of G on $T(G)_1$ is called the *adjoint representation* of G and is denoted by $Ad : G \to GL(\mathfrak{g})$. We have for any $v \in T(G)_1$,

$$(R_g)_*(\tilde{v}) = Ad(g^{-1})(v). \tag{12.1}$$

From now on we shall identify vectors $v \in T(G)_1$ with left-invariant vector fields $\tilde{v}$ on G. The adjoint representation preserves the Lie bracket, i.e.,

$$Ad(g)([\eta, \tau]) = [Ad(g)(\eta), Ad(g)(\tau)]. \tag{12.2}$$

Let $\eta \in \mathfrak{g}$ and $G \to \mathfrak{g}$ be the map $g \to Ad(g)(\eta)$. Its differential at $1 \in G$ is the linear map $\mathrm{Lie}(G) \to \mathfrak{g}$. It coincides with the map $\tau \to [\eta, \tau]$. The latter is a homomorphism of the Lie algebra to itself. It is also called the *adjoint representation* (of the Lie algebra) and is denoted by $ad : \mathfrak{g} \to \mathfrak{g}$.

Most of the Lie groups we shall deal with will be various subgroups of the Lie group $G = GL(n, \mathbb{R})$. Since $GL(n, \mathbb{R})$ is an open submanifold of the vector space $\mathrm{Mat}_n(\mathbb{R})$ of real $n \times n$-matrices, we can identify $T(G)_1$ with $\mathrm{Mat}_n(\mathbb{R})$. The matrix functions $x_{ij} : A =$

$(a_{ij}) \to a_{ij}$ are a system of local parameters. We can write any vector field on $\mathrm{Mat}_n(\mathbb{R})$ in the form

$$\eta = \sum_{i,j=1}^{n} \phi_{ij}(X)\frac{\partial}{\partial x_{ij}},$$

where $\phi_{ij}(X)$ are some smooth functions on $\mathrm{Mat}_n(\mathbb{R})$. The left translation map $L_A : X \to A \cdot X$ is a linear map on $\mathrm{Mat}_n(\mathbb{R})$, so it coincides with its differential. It maps $\frac{\partial}{\partial x_{ij}}$ to $\sum_{k=1}^{n} a_{ik}\frac{\partial}{\partial x_{kj}}$, where a_{ij} is the ij-th entry of A. From this we easily get

$$(L_A)_*(\eta) = \sum_{i,j=1}^{n} \phi_{ij}(A \cdot X)(\sum_{k=1}^{n} a_{ik}\frac{\partial}{\partial x_{kj}}).$$

If we assign to η the matrix function $\Phi(X) = (\phi_{ij}(X))$, then we can rewrite the previous equation in the form

$$(L_A)_*(\Phi)(X) = \Phi(AX)A^{-1}.$$

In particular, η is left-invariant, if and only if $\Phi(AX) = \Phi(X)A$ for all $A, X \in G$. By taking $X = E_n$, we obtain that $\Phi(A) = A\Phi(I_n)$. This implies that $\eta \to X_\eta = \Phi(I_n) \in \mathrm{Mat}_n(\mathbb{R})$ is a bijective linear map from $\mathrm{Lie}(G)$ to $\mathrm{Mat}_n(\mathbb{R})$. Also we verify that

$$[\eta, \tau] = [X_\eta, X_\tau] := X_\eta X_\tau - X_\tau X_\eta.$$

The Lie algebra of $GL(n, \mathbb{R})$ is denoted by $\mathfrak{gl}(n, \mathbb{R})$. We list some standard subgroups of $GL(n, \mathbb{R})$. They are

$$SL(n, \mathbb{R}) = \{A \in GL(n, \mathbb{R}) : \det(A) = 1\}$$

(*special linear group*).

Its Lie algebra is the subalgebra of $\mathrm{Mat}_n(\mathbb{R})$ of matrices with trace zero. It is denoted by $\mathfrak{sl}(n, \mathbb{R})$.

$$O(n-k, k) = \{A \in GL(n, \mathbb{R}) : A^t \cdot J_{n-k,k} \cdot A = J_{n-k,k} := \begin{pmatrix} I_{n-k} & 0 \\ 0 & -I_k \end{pmatrix}\}$$

(*orthogonal group* of type $(n-k, k)$).

Its Lie algebra is the Lie subalgebra of $\mathrm{Mat}_n(\mathbb{R})$ of matrices A satisfying $A^t J_{n-k,k} + J_{n-k,k}A^t = 0$. It is denoted by $\mathfrak{o}(n-k, k)$.

$$Sp(2n, \mathbb{R}) = \{A \in GL(2n, \mathbb{R}) : A^t \cdot J \cdot A = J := \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix}\}$$

(*symplectic group*).

Its Lie algebra is the Lie subalgebra of $\mathrm{Mat}_n(\mathbb{R})$ of matrices A such that $A^t \cdot J + J \cdot A = 0$. It is denoted by $\mathfrak{sp}(2n, \mathbb{R})$.

$$GL(n, \mathbb{C}) = \{A = \begin{pmatrix} X & Y \\ -Y & X \end{pmatrix} \in GL(2n, \mathbb{R}), X, Y \in \mathrm{Mat}_n(\mathbb{R})\}$$

(*complex general linear group*).

Its Lie algebra is isomorphic to the Lie algebra of complex $(n \times n)$-matrices and is denoted by $\mathfrak{gl}(n, \mathbb{C})$.

$$U(n) = \{A \in GL(n, \mathbb{C}) : A^t \cdot A = I_{2n}\}$$

(*unitary group*).

Its Lie algebra consists of skew-symmetric matrices A from $\mathfrak{gl}(n, \mathbb{C})$). It is denoted by $\mathfrak{u}(n)$.

Other groups are realized as subgroups of $GL(n, \mathbb{R})$ by means of a morphism of Lie groups $G \to GL(n, \mathbb{R})$ which are called linear representations. For example, $GL(n, \mathbb{C})$ is isomorphic to the Lie group of complex invertible $n \times n$-matrices $X + iY$ under the representation

$$X + iY \to \begin{pmatrix} X & Y \\ -Y & X \end{pmatrix}.$$

Under this map, the unitary group $U(n)$ is isomorphic to the group of complex invertible $n \times n$-matrices A satisfying $A \cdot \bar{A}^t = I_n$, where $\overline{X + iY} = X - iY$.

12.2 We start with a principal G-bundle $P \to M$. For any point $p = (x, g) \in P$, the tangent space $T(P)_p$ contains $T(P_x)_p$ as a linear subspace. Its elements are called *vertical tangent vectors*. We denote the subspace of vertical vectors $T(P_x)_p$ by $T(P)_p^v$.

A *connection* on P is a choice of a subspace $H_p \subset T(P)_p$ for each $p \in P$ such that

$$T(P)_p = H_p \oplus T(P)_p^v.$$

This choice must satisfy the following two properties:

(i) (smoothness) the map $T(P) \to T(P)$ given by the projection maps $T(P)_p \to T(P)_p^v \subset T(P)_p$ is smooth;

(ii) (equivariance) for any $g \in G$, considered as an automorphism of P, we have

$$(dg)_p(H_p) = H_{p \cdot g}.$$

Now observe that P_x is isomorphic to the group G although not canonically (up to left translation isomorphism of G). To any tangent vector $\xi \in T(P_x)_p$, we can associate a left invariant vector field on G. This is independent of the choice of an isomorphism $P_x \cong G$. This allows one to define an isomorphism

$$\alpha_p : T(P)_p^v \to \mathfrak{g}, \tag{12.3}$$

where $\mathfrak{g}$ is the Lie algebra of G. If $p' = g \cdot p$, then

$$\alpha_{p'} = Ad(g^{-1}) \circ \alpha_p.$$

Let $\mathfrak{g}_P = P \times \mathfrak{g}$ be the trivial vector bundle with fibre $\mathfrak{g}$. Then the projection maps $T(P)_p \to T(P)_p^v$ define a linear map of vector bundles $T(P) \to \mathfrak{g}_P$. This is equivalent to a section of the bundle $T(P)^* \otimes \mathfrak{g}_P$ and can be viewed as a differential 1-form A with values in the Lie algebra $\mathfrak{g}$. Its restriction to the vertical space $T(P)_p$ coincides with the isomorphism α_p. It is zero on the horizontal part H_p of $T(P)_p$. From this we deduce that the form A satisfies

$$A(dg_p(\xi_p)) = Ad(g^{-1})A(\xi_p), \quad \forall \xi_p \in T(P)_p, \forall g \in G. \tag{12.4}$$

Conversely, given $A \in \Gamma(T(P)^* \otimes \mathfrak{g}_P)$ which satisfies (12.4) and whose restriction to $T(P)_p^v$ is equal to the map $\alpha_p : T(P)_p^v \to \mathfrak{g}$, it defines a connection. We just set

$$H_p = \mathrm{Ker}(A_p).$$

Let $\{H_p\}_{p \in P}$ be a connection on P. Under the projection map $\pi : P \to M$, the differential map $d\pi$ maps H_p isomorphically onto the tangent space $T(M)_{\pi(p)}$. Thus we can view the connection as a linear map $A : \pi^*(T(M)) \to T(P)$ such that its composition with the differential map $T(P) \to \pi^*(T(M))$ is the identity. Now given a section ξ of $T(M)$ (a vector field), it defines a section of $\pi^*(T(M))$ by the formula $\pi^*(\xi)(p) = \xi(\pi(p))$, where we identify $\pi^*(T(M))_p$ with $T(M)_{\pi(p)}$. Using A we get a section $\xi^* = A(\pi^*(\xi))$ of $T(P)$. For any $p \in P$,

$$\xi_p^* \in H_p.$$

The latter is expressed by saying that ξ^* is a *horizontal vector field*.

Thus a connection allows us to lift any vector field on M to a unique horizontal vector field on P. In particular one can lift an integral curve $\gamma : [0,1] \to M$ of vector fields to a horizontal path $\gamma^* : [0,1] \to P$. The latter means that $\pi \circ \gamma^* = \gamma$ and the velocity vectors $\frac{d\gamma^*}{dt}$ are horizontal. If we additionally fix the initial condition $\gamma^*(0) = p \in P_{\gamma(0)}$, this path will be unique. Let us prove the existence of such a lifting for any path $\gamma : [0,1] \to M$.

First, we choose some curve $p(t)(0 \le t \le 1)$ in P lying over $\gamma(t)$. This is easy, since P is locally trivial. Now we shall try to correct it by applying to each $p(t)$ an element $g(t) \in G$ which acts on fibres of P by right translation. Consider the map $t \to q(t) = p(t) \cdot g(t)$ as the composition $[0,1] \to P \times G \to P, t \to (p(t), g(t)) \to p(t) \cdot g(t)$. Let us compute its differential. For any fixed $(p_0, g_0) \in P \times G$, we have the maps: $\mu_1 : P \to P, p \to p \cdot g_0$, and $\mu_2 : G \to P, g \to p_0 \cdot g$. The restriction of the map $\mu : P \times G \to P$ to $P \times \{g_0\}$ (resp. $\{p_0\} \times G$) coincides with μ_1 (resp. μ_2). This easily implies that

$$dq(t) = dp(t) \cdot g(t) + p(t) \cdot dg(t) = dp(t) \cdot g(t) + q(t) \cdot g(t)^{-1} \cdot dg(t). \tag{12.5}$$

Here the left-hand side is a tangent vector at the point $q(t)$, the term $dp(t) \cdot g(t)$ is the translate of the tangent vector $dp(t)$ at $p(t)$ by the right translation $R_{g(t)}$, the term $q(t) \cdot g(t)^{-1} \cdot dg(t)$ is the vertical tangent vector at $q(t)$ corresponding to the element $g(t)^{-1}dg(t)$ of the Lie algebra $\mathfrak{g}$ under the isomorphism $\alpha_{q(t)}$ from (12.3). Using (12.5), we are looking for a path $g(t) : [0,1] \to G$ satisfying

$$A_{q(t)}(dq(t)) = Ad(g(t)^{-1})A_{p(t)}(dp(t)) + g(t)^{-1}dg(t) = 0.$$

This is equivalent to the following differential equation on G:

$$A_{p(t)}(dp(t)) = -Ad(g(t))(g(t)^{-1}dg(t)) = -g(t)^{-1}dg(t).$$

This can be solved.

Thus the connection allows us to move from the fibre $P_{\gamma(0)}$ to the fibre $P_{\gamma(t)}$ along the horizontal lift of a path connecting the points $\gamma(0)$ and $\gamma(t)$.

12.3 There are other nice things that a connection can do. First let introduce some notation. For any vector bundle E over a manifold X we set

$$\Lambda^k(E) = \bigwedge^k T^*(X) \otimes E,$$

$$\mathcal{A}^k(E)(X) = \Gamma(\Lambda^k(E))$$

The elements of this space are called differential k-forms on X with values in E. In the special case when E is the trivial bundle of rank 1, we drop E from the notation. Thus

$$\mathcal{A}^k(X) = \Gamma(\bigwedge^k T^*(X))$$

is the space of smooth differential k-forms on X. Also, if $E = V_X$ is the trivial vector bundle with fibre V, then we set

$$\mathcal{A}^k(V)(X) = \mathcal{A}^k(V_X)(X).$$

Thus a connection A on a principal G-bundle is an element of the space $\mathcal{A}^1(L(G))(P)$.

Let us define the *covariant derivative*. Let P be a principal G-bundle over M. For any differential k-form ω on P with values in a vector space V (i.e. $\omega \in \mathcal{A}^k(V)(P)$) we can define its covariant derivative $d^A(\omega)$ with respect to connection A by

$$d^A(\omega)(\tau_1, \ldots, \tau_{k+1}) = d\omega(\tau_1^h, \ldots, \tau_{k+1}^h), \tag{12.6}$$

where the superscript is used to denote the horizontal part of the vector field. Here we also use the usual operation of exterior derivative satisfying the following *Cartan's formula*:

$$d\omega(\xi_1 \ldots, \xi_{k+1}) = \sum_{i=1}^{k+1}(-1)^{i+1}\xi_i(\omega(\xi_1, \ldots, \hat{\xi}_i, \ldots, \xi_{k+1}))+$$

$$+\sum_{1\le i<j\le k+1}(-1)^{i+j}\omega([\xi_i, \xi_j]), \xi_1, \ldots, \hat{\xi}_i, \ldots, \hat{\xi}_j, \ldots, \xi_{k+1}). \tag{12.7}$$

In the first part of this formula we view vector fields as derivations of functions.

We set

$$F_A = d^A(A), \tag{12.8}$$

where the connection A is viewed as an element of the space $\mathcal{A}^1(L(G))(P)$. This is a differential 2-form on P with values in $\mathfrak{g}$. It is called the *curvature form* of A. By definition

$$F_A(\xi_1, \xi_2) = \xi_1^h(A(\xi_2^h)) - \xi_1^h(A(\xi_1^h)) - A([\xi_1^h, \xi_2^h]) = -A([\xi_1^h, \xi_2^h]). \tag{12.9}$$

Here we have used that the value of A on a horizontal vector is zero. This gives a geometric interpretation of the curvature form we alluded to in the introduction.

In particular,

$$F_A = 0 \Longleftrightarrow [\xi, \tau] \quad \text{is horizontal for any horizontal vector fields } \xi, \tau.$$

Assume $\omega \in \mathcal{A}^p(L(G))(P)$ satisfies the following two conditions:

(a) $\omega(\xi_1, \ldots, \xi_k) = 0$ if at least one of the vector fields ξ_i^h is vertical.

(b) for any $g \in G$ acting on P by right translatons, $(R_g)^*(\omega)(\xi) = Ad(g^{-1})(\omega(\xi))$.

For example, the curvature form F_A satisfies these properties. If A, A' are two connections, then the difference $A - A'$ satisfies these properties (since their restrictions to vertical vector fields coincide).

Let us try to descend such ω to M by setting

$$\omega(\xi_1, \ldots, \xi_k)(x) = \omega(ds_U(\xi_1), \ldots, ds_U(\xi_k)), \tag{12.10}$$

where we choose a local section $s_U : U \to P$. A different choice of trivialization gives $s_U = s_V \cdot g_{UV}$, where g_{UV} is the transition function of P (see (11.3)). Differentiating, we have (similar to (12.5))

$$ds_U = ds_V \cdot g_{UV} + s_U \cdot g_{UV}^{-1} dg_{UV}. \tag{12.11}$$

Recall that, for any $\tau_x \in T(M)_x$, $ds_V \cdot g_{UV}(\tau_x)$ equals the image of the tangent vector $ds_V(\tau_x) \in T(P)_{s_V(x)}$ under the differential of the right translation $g_{UV} : P \to P$. The second term $s_U \cdot g_{UV}^{-1} dg_{UV}(\tau_x)$ is the image of $(dg_{UV})_x(\tau_x) \in T(G)_{g_{UV}(x)}$ in $T(P_{s_U(x)})^v$ under the isomorphism $T(G)_{g_{UV}(x)} \cong \mathfrak{g} \cong T(P_x)_{s_U(x)}$. Since ω vanishes on vertical vectors and also satisfies property (b), we obtain that the right-hand side of (12.10) changes, when we replace s_U with s_V, by applying to the value the transformation $Ad(g_{UV}^{-1})$. Thus, if we introduce the vector bundle $Ad(P)$ associated to P by means of the adjoint representation of G in $\mathfrak{g}$, the left-hand side of (12.11) is a well-defined section of the bundle $\mathcal{A}^1(Ad(P))$.

The covariant derivative operator d^A obviously preserves properties (a) and (b), hence descends to an operator

$$d^A : \mathcal{A}^k(Ad(P))(M) \to \mathcal{A}^{k+1}(Ad(P))(M).$$

Also its definition is local, so we obtain a homomorphism of sheaves

$$d^A : \mathcal{A}^k(Ad(P)) \to \mathcal{A}^{k+1}(Ad(P)). \tag{12.12}$$

Taking $k = 0$, we get a homomorphism

$$\nabla_A : Ad(P) \to \mathcal{A}^1(Ad(P)). \tag{12.13}$$

It satisfies Leibnitz's rule: for any section $s \in Ad(P)(U)$, and a function $f : U \to \mathbb{R}$,

$$\nabla_A(fs) = df \otimes s + f\nabla_A(s). \tag{12.14}$$

12.4 Let us consider the natural pairing

$$\mathcal{A}^k(\mathfrak{g})(P) \times \mathcal{A}^k(\mathfrak{g})(P) \to \mathcal{A}^{k+s}(\mathfrak{g})(P), \quad (\omega, \omega') \to [\omega, \omega'] \tag{12.15}$$

It is defined by writing any $\omega \in \mathcal{A}^k(\mathfrak{g})(P)$ as a sum of the expressions $\alpha \otimes \xi$, where $\alpha \in \mathcal{A}^k(X)$ and $\xi \in \mathfrak{g}$, This allows us to define the pairing and to set

$$[\alpha \otimes \xi, \alpha' \otimes \xi'] = (\alpha \wedge \alpha') \otimes [\xi, \xi'].$$

We leave to the reader to check the following:

Lemma 1. *For any $\phi \in \Lambda^p(X, \mathfrak{g}), \psi \in \Lambda^q(X, \mathfrak{g}), \rho \in \Lambda^r(X, \mathfrak{g})$,*
(i) $d[\phi, \beta] = [d\phi, \psi] + (-1)^p[\phi, d\psi]$,
(ii) $(-1)^{pr}[[\phi, \psi], \rho] + (-1)^{rq}[[\rho, \phi], \psi] + (-1)^{qp}[[\psi, \rho], \phi] = 0$,
(iii) $[\psi, \phi] = -(-1)^{pq}[\phi, \psi]$.

Let us prove the following

Theorem 1. *Let $A \in \Lambda^1(P, \mathfrak{g})$ be a connection on P and F_A be its curvarure form.*
(i) (The structure equation):

$$F_A = dA + \frac{1}{2}[A, A].$$

(ii) (Bianchi's identity):

$$d^A(F_A) = 0.$$

(iii) for any $\beta \in \mathcal{A}^k(Ad(P))(M)$

$$d^A(\beta) = d\beta + [A, \beta].$$

Proof. (i) We want to compare the values of each side on a pair of vector fields ξ, τ. By linearity, it is enough to consider three cases: ξ, τ are horizontal; ξ is vertical, τ is horizontal; both fields are vertical. In the first case we have, applying (12.7),

$$(dA + \frac{1}{2}[A, A])(\xi, \tau) = dA(\xi^h, \tau^h) + \frac{1}{2}([A(\xi^h), A(\tau^h)] - [A(\tau^h), A(\xi^h)]) =$$

$$= \xi^h(A(\tau^h) - \tau^h A(\xi^h) - A([\xi^h, \tau^h]) = -A([\xi^h, \tau^h]) = F_A(\xi, \tau).$$

In the second case,

$$(dA + \frac{1}{2}[A, A])(\xi, \tau)) = \xi^v(A(\tau^h)) - \tau^h(A(\xi^v)) - A([\xi^v, \tau^h]) +$$

$$+\frac{1}{2}([A(\xi^v), A(\tau^h)] - [A(\tau^h), A(\xi^v)]) = -\tau^h A(\xi^v) - A([\xi^v, \tau^h]) = 0 = F_A(\xi, \tau).$$

Here we use that at any point ξ^v can be extended to a vertical vector field $\eta^\sharp$ for some vector $\eta \in \mathfrak{g}$ (recall that G acts on P). Hence $\tau(A(\xi^v)) = \tau(A(\eta^\sharp)) = 0$ since $A(\eta^\sharp) = \eta$ is constant. Also, we use that $[\xi^v, \tau^h] = 0$. To see this, we may assume that $\xi^v = \eta^\sharp$. Then, using the property of Lie derivative from Lecture 2, we obtain that

$$[\xi^v, \tau^h] = \mathcal{L}_{\eta^\sharp}(\tau^h) = \lim_{t\to 0} \frac{R^*_{\eta^t}(\tau^h) - \tau^h}{t}.$$

The last expression shows that $[\xi^v, \tau^h]$ is a horizontal vector.

Finally, let us take $\xi = \xi^v, \tau = \tau^v$. As above we may assume that $\xi = \eta^\sharp, \tau = \theta^\sharp$. Then $F_A(\xi, \tau) = 0$, and

$$(dA + \frac{1}{2}[A,A])(\xi^v, \tau^v)) = \xi^v(A(\tau^v)) - \tau^v(A(\xi^v)) + \frac{1}{2}([A(\xi^v), A(\tau^v)] - [A(\tau^v), A(\xi^v)])-$$

$$-A([\xi^v, \tau^v]) = \xi^v(A(\tau^v)) - \tau^v(A(\xi^v)) - A([\xi^v, \tau^v]) + [A(\xi^v), A(\tau^v)] - [\eta, \theta)] + [\eta, \theta)] = 0.$$

(ii) This follows easily from Lemma 1 and (i). We leave the details to the reader.

(iii) We check only in the case $k = 1$, and leave the general case to the reader. Since $\beta(\xi) = \beta(\xi^h)$, we have

$$d^A(\beta)(\xi_1, \xi_2) = \xi_1^h(\beta(\xi_2)) - \xi_1^h(\beta(\xi_1)) - \beta([\xi_1, \xi_2]).$$

On the other hand,

$$(d\beta + [A, \beta])(\xi_1, \xi_2) = d\beta(\xi_1, \xi_2) + [A, \beta](\xi_1, \xi_2) =$$

$$= \xi_1(\beta(\xi_2)) - \xi_2(\beta(\xi_1)) - \beta([\xi_1, \xi_2] + [A(\xi_1), \beta(\xi_2)] - [A(\xi_2), \beta(\xi_1)] =$$
$$= d^A(\beta)(\xi_1, \xi_2) + \xi_1^v(\beta(\xi_2)) - \xi_2^v(\beta(\xi_1)) + [A(\xi_1), \beta(\xi_2)] - [A(\xi_2), \beta(\xi_1)].$$

As in the proof of (i), we may assume that $\xi_i^v = \eta_i^\sharp$ for some $\eta_i^\sharp \in \mathfrak{g}$. since $R_g^*(\beta) = Ad(g^{-1}(\beta))$, we obtain $\eta_i^\sharp(\beta(\xi_j)) = -[\eta_i, \beta(\xi_j)]$. Also, $[A(\xi_i), \beta(\xi_j)] = [\eta_i, \beta(\xi_j)]$. This implies that

$$\xi_1^v(\beta(\xi_2)) - \xi_2^v(\beta(\xi_1)) + [A(\xi_1), \beta(\xi_2)] - [A(\xi_2), \beta(\xi_1)] = 0$$

and proves the assertion.

Corollary 1.

$$d^A \circ d^A(\omega) = [F_A, \omega].$$

Proof. We have

$$d^A(d^A(\omega)) = d^A(d\omega + [A, \omega]) = d(d\omega + [A, \omega]) + [A, d\omega + [A, \omega]] =$$

$$= d[A, \omega] + [A, d\omega] + [A, [A, \omega]].$$

Applying Lemma 1, we get

$$d^A(d^A(\omega)) = [dA, \omega] - [A, d\omega] + [A, d\omega] + \frac{1}{2}[[A, A], \omega] = [dA + \frac{1}{2}[A, A], \omega].$$

Remark 1. If G is a subgroup of $GL(n, \mathbb{R})$, then we may identify A with a matrix whose entries are 1-forms. Recall that the Lie bracket in $\mathfrak{g}$ is the matrix bracket $[X, Y] = XY - YX$. Then $[A, A] = A \wedge A - (-1)A \wedge A = 2A \wedge A = 2A^2$. Thus

$$F_A = dA + A^2.$$

Definition. A connection is called *flat* if $F_A = 0$.

12.5 Now let E be a rank d vector G-bundle over M with typical fibre a real vector space V. We may assume that E is associated to a principal G-bundle. Let $\mathcal{E}$ be the sheaf of sections of E and $\mathcal{A}^k(E)$ be the sheaf of sections of $\Lambda^k(E)$.

Definition. A *connection on a vector bundle* E is a map of sheaves

$$\nabla : \mathcal{E} \to \mathcal{A}^1(E),$$

satisfying the Leibnitz rule: for any section $s : U \to E$ and $f \in \mathcal{O}_M(U)$,

$$\nabla(fs) = df \otimes s + f\nabla(s).$$

Let us explain how to construct a connection on E by using a connection A on the corresponding principal bundle P. Choose a local section $e : U \to P$ (which is equivalent to a trivialization $U \times G \to \pi^{-1}(U)$) and consider $\theta_e = e^*(A)$ as a 1-form on U with coefficients in $\mathfrak{g}$. Thus, for any tangent vector $\tau_x \in T(M)_x$, we have

$$e^*(A)(\tau_x) = A(de_x(\tau_x)).$$

If $e' : U' \to P$ is another section, then by (12.13), for any $x \in U \cap U'$,

$$\theta_e(\tau_x) = Ad(g)^{-1}(\theta_{e'}(\tau_x)) + g^{-1}dg,$$

where $g : U \cap U' \to G$ is a smooth map such that $e(x) = e'(x) \cdot g(x)$. Let $(U_i)_{i \in I}$ be an open cover of M trivializing P, and let $e_i : U_i \to P$ be some trivializing diffeomorphisms. The forms $\theta_i = \theta_{e_i}$ are related by

$$\theta_i = Ad(g_{ij})^{-1} \cdot \theta_j + g_{ij}^{-1}dg_{ij},$$

where g_{ij} are the transition functions. We see that, because of the second term in the right-hand side, $\{\theta_i\}_{i \in I}$ do not form a section of any vector bundle. However, the difference $A - A'$ of two connections defines a section $\{\theta_i - \theta_i'\}_{i \in I}$ of $\mathcal{A}^1(Ad(P))$. Applying the representation $d\rho : \mathfrak{g} \to \mathrm{End}(V) = \mathfrak{gl}(V))$, we obtain a set of 1-forms θ_e^ρ on U with values in $d\rho(\mathfrak{g}) \subset \mathrm{End}(V)$. They satisfy

$$\theta_e^\rho = \rho(g)^{-1} \cdot \theta_{e'}^\rho \cdot \rho(g) + \rho(g)^{-1} d\rho(g). \tag{12.16}$$

Note that the trivialization $e : U \to P$ defines a trivialization of $U \times V \to E|U$. In fact,

$$E_U = P_U \times_G V \cong (U \times G) \times_G V \cong U \times (G \times_G V) \cong U \times V.$$

Now we can define the connection ∇_A^ρ on E as follows. Using the trivialization, we can consider $s : U \to E$ as a $\mathbb{R}^n$-valued function. This means that $s = (s_1, \ldots, s_n)$, where s_i are scalar functions. Then we set

$$\nabla_A^\rho(s) = ds + \theta_e^\rho \cdot s, \tag{12.17}$$

where we view θ_e^ρ as an endomorphism in the space of sections of E which depends on a vector field. It obviously satisfies the Leibnitz rule. To check that our formula is well-defined, we compare $\nabla_A^\rho(s)$ with $\nabla_A^\rho(s')$, where $s'(x) : U' \to E$ and $s' = \rho(g)s$ for some transition function $g : U \cap U' \to G$ of the prinicipal G-bundle. We have

$$ds + \theta_e^\rho \cdot s = d(\rho(g)^{-1}s') + \rho(g)^{-1}\theta_{e'}^\rho \rho(g) \cdot s + \rho(g)^{-1}d\rho(g) \cdot s =$$

$$d\rho(g)^{-1}\rho(g) \cdot s + \rho(g)^{-1}ds' + \rho(g)^{-1}\theta_{e'}^\rho \cdot s' + \rho(g)^{-1}d\rho(g) \cdot s =$$

$$= \rho(g)^{-1}(ds' + \theta_{e'}^\rho \cdot s') + [d\rho(g)^{-1}\rho(g) + \rho(g)^{-1}d\rho(g)] \cdot s = \rho(g)^{-1}(ds' + \theta_{e'}^\rho \cdot s').$$

Here, at the last step, we have used that $0 = d(\rho(g)^{-1} \cdot \rho(g)) = d\rho(g)^{-1}\rho(g) + \rho(g)^{-1}d\rho(g)$. This shows that $\nabla_A^\rho(s)$ takes values in $\mathcal{A}^1(E)$ so that our map ∇_A^ρ is well-defined.

Remark 2. Let $\mathcal{T}_M$ be the sheaf of sections of $T(M)$. Its sections over U are vector fields on U. There is a canonical pairing

$$\mathcal{T}_M \otimes \mathcal{A}^1(E) \to \mathcal{E}.$$

Composing it with $\nabla_A^\rho \otimes 1 : \mathcal{E} \otimes \mathcal{T}_M \to \mathcal{A}^1(E) \to \mathcal{E}$, we get a map

$$\mathcal{E} \otimes \mathcal{T}_M \to \mathcal{E},$$

or, equivalently,

$$\mathcal{T}_M \to \mathcal{E}nd(E), \tag{12.18}$$

where $\mathcal{E}nd(E)$ is the sheaf of sections of $E^* \otimes E = End(E)$. It follows from our definition of ∇_A^ρ that the image of this map lies in the sheaf $Ad(E)$ which is the sheaf of sections of the image of the natural homomorphism $d\rho : Ad(P) \to \mathcal{E}nd(E)$.

We can also define the curvature of the connection ∇_A^ρ. We already know that F_A can be considered as a section of the sheaf $\mathcal{A}^2(Ad(P))$ on M. We can take its image in $\mathcal{A}^2(Ad(E))$ under the map $Ad(P) \to Ad(E)$. This section is denoted by F_A^ρ and is called the curvature form of E. Locally, with respect to the trivialization $e : U \to P$, it is given by

$$(F_A^\rho)_e = d\theta_e^\rho + \frac{1}{2}[\theta_e^\rho, \theta_e^\rho] = d\theta_e^\rho + \theta_e^\rho \wedge \theta_e^\rho. \tag{12.19}$$

Here for two 1-forms α, β with values in $\mathrm{End}(V)$,

$$[\alpha,\beta](\xi_1,\xi_2) = [\alpha(\xi_1),\beta(\xi_2)]-[\alpha(\xi_2),\beta(\xi_1)] = 2(\alpha(\xi_1)\beta(\xi_2)-\beta(\xi_2)\alpha(\xi_1)) := 2\alpha\wedge\beta(\xi_1,\xi_2).$$

Finally, we can extend ∇_A^ρ to

$$d_A^\rho : \mathcal{A}^k(E) \to \mathcal{A}^{k+1}(E)$$

by requiring $d_A^\rho = \nabla_A^\rho$ for $k=0$ and

$$d_A^\rho(\omega\wedge\alpha) = d\omega\wedge\alpha + (-1)^{deg(\omega)}\omega\wedge\nabla_A^\rho(\alpha).$$

for $\omega \in \mathcal{A}^k(U), \alpha \in \mathcal{E}(U)$. Using Theorem 1, we can easily check that

$$d_A^\rho \circ d_A^\rho(\beta) = [F_A^\rho, \beta].$$

12.6 Let us consider the special case when $G = GL(n,\mathbb{R})$ and E is a rank n vector bundle. In this case the Lie algebra $\mathfrak{g}$ is equal to the algebra of matrices $\mathrm{Mat}_n(\mathbb{R})$ with the bracket $[X,Y] = XY - YX$. We omit ρ in the notation ∇_A^ρ. A trivialization $e : U \to P$ is a choice of a basis $(e_1(x),\ldots,e_n(x))$ in E_x. This defines the trivialization of $E|U$ by the formula

$$U \times \mathbb{R}^n \to E|U, (x,(a_1,\ldots,a_n)) \to a_1e_1(x)+\ldots+a_ne_n(x).$$

A section $s : U \to E$ is given by $s(x) = a_1(x)e_1(x)+\ldots+a_n(x)e_n(x)$. The 1-form θ_e is a matrix $\theta_e = (\theta_{ij})$ with entries $\theta_{ij} \in \mathcal{A}^1(U)$. If we fix local coordinates $(x_1,\ldots,x_d)$ on U, then

$$\theta_{ij} = \sum_{k=1}^d \Gamma_{ij}^k dx_k.$$

The connection ∇_A is given by

$$\nabla_A(s) = \sum_{i=1}^n d(a_i(x)e_i(x)) = \sum_{i=1}^n da_i(x)e_i(x) + \sum_{j=1}^n a_j(x)\nabla_A(e_j(x)) =$$

$$= \sum_{i,j=1}^n (\sum_{k=1}^d (\frac{\partial a_i}{\partial x_k} + a_j(x)\Gamma_{ij}^k dx_k)e_i(x). \tag{12.20}$$

If we view ∇_A in the sense (12.18), then we can put (12.19) in the following equivalent forms:

$$\nabla_A(\frac{\partial}{\partial x_k})(s) = \sum_{i=1}^n (\frac{\partial a_i}{\partial x_k} + \sum_{j=1}^n a_j(x)\Gamma_{ij}^k)e_i(x), \tag{12.21}$$

$$\nabla_{A,k} = \nabla_A(\frac{\partial}{\partial x_k}) = \frac{\partial}{\partial x_k} + (\Gamma_{ij}^k)_{1\le i,j\le n} = \frac{\partial}{\partial x_k} + A_k^e. \tag{12.22}$$

This is an operator which acts on the section $a_1(x)e_1(x)+\ldots+a_n(x)e_n(x)$ by transforming it into the section $b_1(x)e_1(x)+\ldots+b_n(x)e_n(x)$, where

$$\begin{pmatrix} b_1(x) \\ \vdots \\ b_n(x) \end{pmatrix} = \begin{pmatrix} \frac{\partial a_1(x)}{\partial x_k} \\ \vdots \\ \frac{\partial a_n(x)}{\partial x_k} \end{pmatrix} + \begin{pmatrix} \Gamma^k_{11} & \ldots & \Gamma^k_{1n} \\ \vdots & \vdots & \vdots \\ \Gamma^k_{n1} & \ldots & \Gamma^k_{nn} \end{pmatrix} \cdot \begin{pmatrix} a_1(x) \\ \vdots \\ a_n(x) \end{pmatrix}.$$

If we change the trivialization to $e(x) = e'(x)\cdot g(x)$, then we can view $g(x)$ as a matrix function $g : U \to GL(n,\mathbb{R})$, and get

$$A_k^e = g^{-1}A_k^{e'}g + g^{-1}\frac{\partial g}{\partial x_k}. \tag{12.23}$$

By taking the Lie bracket of operators (12.22), we obtain

$$[\nabla_{A,i},\nabla_{A,j}] = [\frac{\partial}{\partial x_i} + A_i^e, \frac{\partial}{\partial x_j} + A_j^e] = \frac{\partial}{\partial x_i}A_j^e - \frac{\partial}{\partial x_j}A_i^e + [A_i^e, A_j^e]. \tag{12.24}$$

Comparing it with (12.19), we get

$$F_A = \sum_{\mu,\nu=1}^{d}[\nabla_\mu^A, \nabla_\nu^A]dx_\mu \wedge dx_\nu. \tag{12.25}$$

Example 1. Let us consider a very special case when E is a rank 1 vector bundle (a *line bundle*). In this case everything simplifies drastically. We have

$$\nabla_{A,k} = \frac{\partial}{\partial x_k} + \Gamma^k.$$

In different trivialization $e(x) = e'(x)g(x)$, we have

$$\Gamma^k = \Gamma'^k + \frac{\partial \log g(x)}{\partial x_k}.$$

If we put $\Gamma = \sum_{k=1}^d \Gamma^k dx_k$, then

$$\Gamma = \Gamma' + d\log g(x). \tag{12.26}$$

Thus a connection on a line bundle defined by transition functions g_{ij} is a section of the principal $\mathbb{R}^{\dim M}$-bundle on M with transition functions $\{d\log(g_{ij}(x))\}$ belonging to the affine group $\mathrm{Aff}(\mathbb{R}^n)$. The difference of two connections is a section of the cotangent bundle $T^*(M)$. The curvature is a section of $\bigwedge^2(T^*(M))$ given by 2-forms $d\Gamma$.

Example 2. Let $E = T(M)$ be the tangent bundle of M. The Lie derivative $\mathcal{L}_\eta$ coincides with the connection $\nabla_A(\eta)$ such that all sections (vector fields) are horizontal. This means that $\Gamma^k_{ij} \equiv 0$. We shall discuss connections on the tangent bundle in more detail later.

12.7 Let us summarize the description of the set of all connections on a principal G-bundle P. It is the set $\mathrm{Con}(P)$ of all differential 1-forms A on P with values in $\mathfrak{g} = Lie(G)$ satisfying

(i) $(R_g)^*(A) = Ad(g^{-1}) \cdot (A)$;

(ii) $A_p : T(P)_p^v \to \mathfrak{g}$ is the canonical map $\alpha_p : T(P_{\pi(p)}) \to \mathfrak{g}$.

Because of (ii), the difference of two connections is a 1-form which satisfies (i) and is identically zero on vertical vector fields. Using a horizontal lift we can descend this form to M to obtain a section of the vector bundle $T^*(M) \otimes Ad(P)$. In this way we obtain that $\mathrm{Con}(P)$ is an affine space over the linear space $\mathcal{A}^1(Ad(P))(M)$. It is not empty. This is not quite trivial. It follows from the existence of a metric on P which is invariant with respect to G. We refer to [**Nomizu**] for the proof.

The gauge group $\mathcal{G}(P)$ acts naturally on the set of sections $\Gamma(P)$ by the rule:

$$(g \cdot s)(x) = g^{-1} \cdot s(x).$$

It also acts on the set of connections on P (or on an associated vector bundle) by the rule:

$$\nabla_{g(A)}(s) = g \nabla_A (g^{-1} s).$$

Locally, if A is given by a $\mathfrak{g}$-valued 1-forms θ_e, it acts by changing the trivialization $e : U \times G \to P|U$ to $e' = g \cdot e$. As we know this changes θ_e to $Ad(g) \cdot \theta_e + g^{-1} dg$.

Exercises.

1. A section $s : U \to E$ of a vector bundle E is called *horizontal* with respect to a connection A if $\nabla_A(s) = 0$. Show that horizontal sections form a sheaf of sections of some vector bundle whose transition functions are constant functions.

2. Let $\pi : E \to M$ be a vector bundle. Show that a connection ∇ on E is obtained from a connection A on some principal G-bundle to which E is associated.

3. Show that there is a bijective correspondence between connections on a vector bundle E and linear morphisms of vector bundles $\pi^*(T(M) \to T(E)$ which are right inverses of the differential map $d\pi : T(E) \to \pi^*(T(M))$.

4. Let ∇ be a connection on a vector bundle E. A tangent vector $\xi_z \in T(E)_z$ is called horizontal if there exists a horizontal section $s : U \to E$ such that ξ_z is tangent to $s(U)$ at the point $z = s(x)$. Let $E = T(M)$ be the tangent bundle of a manifold M. Show that the natural lift $\tilde{\gamma}(t) = (\gamma(t), \frac{d\gamma}{dt})$ of paths in M is a horizontal lift for some connection on $T(M)$.

5. Let ∇ be a connection on a vector bundle. Show that it defines a canonical connection on its tensor powers, on its dual, on its exterior powers. Define the tensor product of connections.

6. Let $P \to M$ be a principal G-bundle. Fix a point $x_0 \in M$. For any closed smooth path $\gamma : [0,1] \to M$ with $\gamma(0) = \gamma(1) = x_0$ and a point $p \in P_{x_0}$ consider a horizontal lift

$\tilde{\gamma} : [0,1] \to P$ with $\tilde{\gamma}(0) = p \in P_{x_0}$. Write $\tilde{\gamma}(1) = p \cdot g$ for a unique $g \in G$. Show that g is independent of the choice of p and the map $\gamma \to g$ defines a homomorphism of groups $\rho : \pi_1(M; x_0) \to G$. It is called the *holonomy representation* of the fundamental group $\pi_1(M; x_0)$.

7. Let $E = E(\rho) = P \times_\rho V$ be the vector bundle associated to a principal G-bundle by means of a linear representation $\rho : G \to GL(V)$. Let $a : P \times V \to E$ be the canonical projection. For any $v \in V$ let $\varphi : P \to E$ be the map $p \to a(p, v)$. Use the differential of this map and a connection A on P to define, for each point $z \in E$, a unique subspace H_z of $T(E)_z$ such that $T(E_{\pi(z)}) \oplus H_z = T(E)_z$. Show that the map $\pi^*(T(M)) \to T(E)$ from Exercise 3 maps each $T(M)_x$ isomorphically onto H_z, where $z \in E_x$.

8. Let $G = U(1)$. Show that the torus $T = \mathbb{R}^4/\mathbb{Z}^4$ admits a non-flat self-dual $U(1)$ connection if and only if it has a structure of an abelian variety.

Lecture 13. KLEIN-GORDON EQUATION

This equation describes classical fields and is an analog of the Euler-Lagrange equations for classical systems with a finite number of degrees of freedom.

13.1 Let us consider a system of finitely many harmonic oscillators which we view as a finite set of masses arranged on a straight line, each connected to the next one via springs of length ϵ. The Lagrangian of this system is

$$L = \frac{1}{2}\sum_{i=1}^{N}\left[m\dot{x}_i^2 - k(x_i - x_{i+1})^2\right] = \frac{\epsilon}{2}\sum_{i=1}^{N}\left[\frac{m}{\epsilon}\dot{x}_i^2 - k\epsilon\frac{(x_i - x_{i+1})^2}{\epsilon^2}\right]. \tag{13.1}$$

Now let us take N go to infinity, or equivalently, replace ϵ with infinitely small Δx. We can write

$$\lim_{N\to\infty} L = \frac{1}{2}\int_a^b \left[\mu\left(\frac{\partial\phi(x,t)}{\partial t}\right)^2 - Y\left(\frac{\partial\phi(x,t)}{\partial x}\right)^2\right]dx, \tag{13.2}$$

where m/ϵ is replaced with the mass density μ, and the coordinate x_i is replaced with the function $\phi(x,t)$ of displacement of the particle located at position x and time t. The function Y is the limit of $k\epsilon$. It is a sort of elasticity density function for the spring, the so-called "Young's modulus". The right-hand side of equation (13.2) is a functional on the space of functions $\phi(x,t)$ with values in the space of functions on $\mathbb{R}^3$. We can generalize the Euler-Lagrange equation to find its extremum point. Let us introduce the *action functional*

$$S = \int_{t_0}^{t_1}\int_a^b L(\phi, \partial_\mu\phi)dxdt.$$

Here L is a smooth function in two variables, and

$$\partial_\mu\phi = \left(\frac{\partial\phi}{\partial t}, \frac{\partial\phi}{\partial x}\right),$$

and the expression

$$\mathcal{L}(\phi) = L(\phi, \partial_\mu \phi)$$

is a functional on the space of functions ϕ. In our case

$$\mathcal{L}(\phi) = \frac{1}{2}\mu(\frac{\partial \phi}{\partial t})^2 - \frac{1}{2}Y(\frac{\partial \phi}{\partial x})^2. \tag{13.3}$$

We shall look for the functions ϕ which are critical points of the action S.

13.2. Recall from Lecture 1 that the critical points can be found as the zeroes of the derivative of the functional S. We shall consider a more general situation. Let $L : \mathbb{R}^{n+1} \to \mathbb{R}$ be a smooth function (it will be called the *Lagrangian*). Consider the functional

$$\mathcal{L} : L_2(\mathbb{R}^n) \to L_2(\mathbb{R}^n), \quad \phi(x,t) \to L(\phi, \frac{\partial \phi}{\partial x_1}, \ldots, \frac{\partial \phi}{\partial x_n}). \tag{13.4}$$

Then the action S is the composition of $\mathcal{L}$ and the integration functional over $\mathbb{R}^n$;

$$S(\phi) = \int\limits_{\mathbb{R}^n} \mathcal{L}(\phi) d^n x = \int L(\phi, \frac{\partial \phi}{\partial x_1}, \ldots, \frac{\partial \phi}{\partial x_n}) d^n x.$$

By the chain rule, its derivative $S'(\phi)$ at the function ρ is equal to the linear functional $L_2(\mathbb{R}^n) \to \mathbb{R}$ defined by

$$S'(\phi)(h) = \int\limits_{\mathbb{R}^n} \mathcal{L}'(\phi) h d^n x.$$

The functional $\mathcal{L}$ (also often called the *Lagrangian*) is the composition of the linear operator

$$L_2(\mathbb{R}) \to L_2(\mathbb{R}^n, \mathbb{R}^{n+1}) = L_2(\mathbb{R}^n)^{n+1}, \quad \phi \to \Phi = (\phi, \frac{\partial \phi}{\partial x_1}, \ldots, \frac{\partial \phi}{\partial x_n})$$

and the functional

$$L_2(\mathbb{R}^n, \mathbb{R}^{n+1}) \to L_2(\mathbb{R}^n), \quad \Phi \to L \circ \Phi.$$

The derivative of the latter functional at $\Phi = (\Phi_1, \ldots, \Phi_{n+1})$ is equal to the linear functional $L_2(\mathbb{R}^n, \mathbb{R}^{n+1}) \to L_2(\mathbb{R}^n)$

$$(h_1, \ldots, h_{n+1}) \to \sum_{i=0}^{n} \frac{\partial L}{\partial y_i}(\Phi(x)) h_i,$$

where we use coordinates $y_0, \ldots, y_n$ in $\mathbb{R}^{n+1}$. This implies that the derivative of $\mathcal{L}$ at ϕ is equal to the linear functional

$$\mathcal{L}'(\phi) : h \to \frac{\partial L}{\partial y_0}(\phi, \frac{\partial \phi}{\partial x_1}, \ldots, \frac{\partial \phi}{\partial x_n}) h + \sum_{\mu=1}^{n} \frac{\partial L}{\partial y_\mu}(\phi, \frac{\partial \phi}{\partial x_1}, \ldots, \frac{\partial \phi}{\partial x_n}) \frac{\partial h}{\partial x_i}. \tag{13.5}$$

Let us set

$$\frac{\partial \mathcal{L}}{\partial \phi} = \frac{\partial L}{\partial y_0}(\phi, \frac{\partial \phi}{\partial x_1}, \ldots, \frac{\partial \phi}{\partial x_n}), \quad \frac{\partial \mathcal{L}}{\partial \partial_\nu \phi} = \frac{\partial L}{\partial y_\nu}(\phi, \frac{\partial \phi}{\partial x_1}, \ldots, \frac{\partial \phi}{\partial x_n}), \quad \nu = 1, \ldots, n \tag{13.6}$$

These are of course functions on $\mathbb{R}^n$. In this notation we can rewrite (13.5) as

$$\mathcal{L}'(\phi)h = \frac{\delta \mathcal{L}}{\delta \phi} h + \sum_{\nu=1}^{n} \frac{\partial \mathcal{L}}{\partial \partial_\nu \phi} \frac{\partial h}{\partial x_\nu}.$$

This easily implies that

$$S(\phi)(h) = \int_{\mathbb{R}^n} (\frac{\partial \mathcal{L}}{\partial \phi} h + \frac{\partial \mathcal{L}}{\partial \partial_1 \phi} \frac{\partial h}{\partial x_1} + \ldots + \frac{\partial \mathcal{L}}{\partial \partial_n \phi} \frac{\partial h}{\partial x_n}) d^n x.$$

In the physicists notation, we write $\delta\phi$ instead of h, and $\delta\partial_i \phi = \partial_i \delta\phi$ instead of $\frac{\partial h}{\partial x_i}$ and find that, if the action has a critical point at ϕ,

$$\int_{\mathbb{R}^n} (\frac{\partial \mathcal{L}}{\partial \phi} \delta\phi + \frac{\partial \mathcal{L}}{\partial \partial_1 \phi} \delta\partial_1 \phi + \ldots + \frac{\partial \mathcal{L}}{\partial \partial_n \phi} \delta\partial_n \phi) d^n x = 0, \tag{13.7}$$

or, using the Einstein convention (to which we shall switch from time to time without warning),

$$\int_{\mathbb{R}^n} (\frac{\partial \mathcal{L}}{\partial \phi} \delta\phi + \frac{\partial \mathcal{L}}{\partial \partial_\nu \phi} \delta\partial_\nu \phi) d^n x = 0.$$

Let us restrict the action to the affine subspace of functions ϕ with fixed boundary condition. This means that ϕ has support on some open submanifold $V \subset \mathbb{R}^n$ with boundary δV, and $\phi|\delta V$ is fixed. For example, $V = [0, \infty) \times \mathbb{R}^3 \subset \mathbb{R}^4$ with $\delta V = \{0\} \times \mathbb{R}^3$. If we are looking for the minimum of the action S on such a subspace, we have to take $\delta\phi$ equal to zero on δV. Integrating by parts, we find (using ∂_ν to denote $\frac{\partial}{\partial x_\nu}$)

$$0 = \int_{\mathbb{R}^n} (\frac{\partial \mathcal{L}}{\partial \phi} \delta\phi + \frac{\partial \mathcal{L}}{\partial \partial_\nu \phi} \delta\partial_\nu \phi) d^n x = \int_V \partial_\nu (\delta\phi \frac{\partial \mathcal{L}}{\partial \partial_\nu \phi}) d^n x + \int_V (\frac{\partial \mathcal{L}}{\partial \phi} - \partial_\nu \frac{\partial \mathcal{L}}{\partial \partial_\nu \phi}) \delta\phi d^n x =$$

$$= \int_{\delta V} (\delta\phi \frac{\partial \mathcal{L}}{\partial \partial_\nu \phi}) d^{n-1} x + \int_V (\frac{\partial \mathcal{L}}{\partial \phi} - \partial_\nu \frac{\partial \mathcal{L}}{\partial \partial_\nu \phi}) \delta\phi d^n x = \int_V (\frac{\partial \mathcal{L}}{\partial \phi} - \partial_\nu \frac{\partial \mathcal{L}}{\partial \partial_\nu \phi}) \delta\phi d^n x.$$

From this we deduce the following *Euler-Lagrange equation*:

$$\frac{\partial \mathcal{L}}{\partial \phi} - \partial_\nu \frac{\partial \mathcal{L}}{\partial \partial_\nu \phi} = \frac{\partial \mathcal{L}}{\partial \phi} - \sum_{\nu=1}^{n} \partial_\nu \frac{\partial \mathcal{L}}{\partial \partial_\nu \phi} = 0. \tag{13.8}$$

Observe the analogy with the Euler-Lagrange equations from Lecture 1:

$$\frac{\partial L}{\partial q_i}(\mathbf{x}_0, \frac{d\mathbf{x}_0}{dt}) - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i}(\mathbf{x}_0, \frac{d\mathbf{x}_0}{dt}) = 0, i = 1, \ldots, n.$$

In our case $(q_1(t),\dots,q_n(t))$ is replaced with $\phi(x,t)$, and $(\dot q_1(t),\dots,\dot q_n(t))$ is replaced with $\frac{\partial\phi(x,t)}{\partial t}$.

Returning to our situation where $\mathcal{L}$ is given by (13.3), we get from (13.8)

$$\frac{\partial^2\phi}{\partial x^2} = \left(\frac{\mu}{Y}\right)\frac{\partial^2\phi}{\partial t^2}. \tag{13.9}$$

This is the standard wave equation for a one-dimensional system traveling with velocity $\sqrt{Y/\mu}$.

Let us take $n = 4$ and consider the space-time $\mathbb{R}^4$ with coordinates (t, x_1, x_2, x_3). That is, we shall consider the functions $\phi(t,x) = \phi(t, x_1, x_2, x_3)$ defined on $\mathbb{R}^4$. We use $x^\nu, \nu = 0,1,2,3$ to denote t, x_1, x_2, x_3, respectively. Set

$$\partial_\nu = \frac{\partial}{\partial x^\nu}, \nu = 0,1,2,3,$$

$$\partial^\nu = \frac{\partial}{\partial x^\nu}, \nu = 0, \quad \partial^\nu = -\frac{\partial}{\partial x^\nu}, \nu = 1,2,3.$$

The operator

$$\Box = \sum_{\nu=0}^{3}\partial_\nu\partial^\nu = \frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x_1^2} - \frac{\partial^2}{\partial x_2^2} - \frac{\partial^2}{\partial x_3^2}.$$

is called the *D'Alembertian*, or *relativistic Laplacian.*

If we take

$$\mathcal{L} = \frac{1}{2}\left(\sum_{\nu=0}^{3}\partial_\nu\phi\partial^\nu\phi - m^2\phi^2\right), \tag{13.10}$$

the Euler-Lagrange equation gives

$$\Box\phi + m^2\phi = 0. \tag{13.11}$$

This is called the *Klein-Gordon equation.*

13.3 We can generalize the Euler-Lagrange equation to fields which are more general than scalar functions on $\mathbb{R}^n$, namely sections of some vector bundle E over a smooth n-dimensional manifold M with some volume form $d\mu$ equipped with a connection $\nabla : \Gamma(E) \to \mathcal{A}^1(E)(M)$. For this we have to consider the Lagrangian as a function $L : E \oplus \Lambda^1(E) \to \mathbb{R}$. In a trivialization $E \cong U \times \mathbb{R}^r$, our field is represented by a vector function $\phi = (\phi_1,\dots,\phi_r)$, and the Lagrangian by a scalar function on $U \times \mathbb{R}^r \times \mathbb{R}^{nr}$. Let $\Gamma = \Gamma(E) \oplus \mathcal{A}^1(E)(M)$ be the space of sections of the vector bundle $E \oplus \Lambda^1(E)$. We have a canonical map $\Gamma(E) \to \Gamma$ which sends ϕ to $(\phi, \nabla(\phi))$. Composing it with L, we get the map

$$\mathcal{L} : \Gamma(E) \to C^\infty(M), \phi \to \mathcal{L}(\phi) = L(\phi(x), \nabla\phi(x)).$$

Then we can define the action functional on X by the formula

$$S(\phi) = \int_M \mathcal{L}(\phi)d\mu = \int_M L(\phi, d\phi)d\mu.$$

The condition for ϕ to be an extremum point is

$$\int_M \left(\frac{\partial \mathcal{L}}{\partial \phi}\delta\phi + \frac{\partial \mathcal{L}}{\partial \nabla\phi}\nabla\delta\phi\right)d^n x = 0 \tag{13.12}$$

Here $\delta\phi$ is any function in $L_2(M,\mu)$ with $||\delta\phi|| < \epsilon$ for sufficiently small ϵ. The partial derivative $\frac{\partial \mathcal{L}}{\partial \phi}$ is the partial derivative of $\mathcal{L}$ with respect to the summand $\Gamma(E)$ of Γ at the point $(\phi, \nabla\phi)$. The partial derivative $\frac{\partial \mathcal{L}}{\partial \nabla\phi}$ is the partial derivative of $\mathcal{L}$ with respect to the summand $\mathcal{A}^1(E)(M)$ of Γ computed at the point $(\phi, \nabla\phi)$. This partial derivative is a linear map $\mathcal{A}^1(E)(M) \to C^\infty(M)$ which we apply at the point $\nabla\delta\phi$. We leave it to the reader to deduce the corresponding Euler-Lagrange equation.

The Klein-Gordon Lagrangian can be extended to the general situation we have just described. Let us fix some function $q : E \to \mathbb{R}$ whose restriction to each fibre E_x is a quadratic form on E_x. Let g be a pseudo-Riemannian metric on M, considered as a function $g : T(M) \to \mathbb{R}$ whose restriction to each fibre is a non-degenerate quadratic form on $T(M)_x$. If we view g as an isomorphism of vector bundles $g : T(M) \to T^*(M)$, then its inverse is an isomorphism $g^{-1} : T^*(M) \to T(M)$ which can be viewed as a metric $g^{-1} : T^*(M) \to \mathbb{R}$ on $T^*(M)$. We define the Klein-Gordon Lagrangian

$$L = -u \circ q + \lambda g^{-1} \otimes q : E \oplus \Lambda^1(E) \to \mathbb{R},$$

Its restriction to the fibre over a point $x \in M$ is given by the formula

$$(a_x, \omega_x \otimes b_x) \to -u(x)q(a_x) + \lambda(x)q(b_x)g^{-1}(\omega_x) \tag{13.13}$$

for some appropriate non-negative valued scalar functions λ and u. Usually q is taken to be a positive definite form. It is an analog of potential energy. The term $\lambda q g^{-1}$ is the analog of kinetic energy.

The simplest case we have considered above corresponds to the situation when $M = \mathbb{R}^n, E = M \times \mathbb{R}$ is the trivial bundle with the trivial connection $\phi \to d\phi$. The Lagrangian $\mathcal{L} : M \times \mathbb{R}^{n+1} \to \mathbb{R}$ does not depend on $x \in M$. The Lagrangian from (13.10) corresponds to taking g to be the Lorentz metric, the quadratic function q equals $(x,z) \to -m^2 z^2$, and the scalar λ is equal to $1/2$.

A little more general situation frequently considered in physics is the case when E is the trivial vector bundle of rank r with trivial connection. Then ϕ is a vector function $(\phi_1, \ldots, \phi_n)$, $d\phi = (d\phi_1, \ldots, d\phi_n)$ and the Euler-Lagrange equation looks as follows:

$$\frac{\partial \mathcal{L}}{\partial \phi_i} = \sum_{\mu=1}^{n} \partial_\mu \frac{\partial \mathcal{L}}{\partial \partial_\mu \phi_i}, \quad i = 1, \ldots, r.$$

13.4 The Lagrangian (13.10) and the corresponding Euler-Lagrange equation (13.11) admits a large group of symmetries. Let us explain what this means. Let G be any Lie group which acts smoothly on a manifold M (see Lecture 2).

Let $\pi : E \to M$ be a vector bundle on M. A *lift* of the action $G \times M \to M$ is a smooth action $G \times E \to E$ such that, for any $z \in E_x$, $g \cdot z \in E_{g\cdot x}$. Given such a lift, G acts naturally on the space of sections of E by the formula

$$(g \cdot s)(x) = g \cdot s(g^{-1} \cdot x). \tag{13.14}$$

If, additionally, the restriction of any $g \in G$ to E_x is a linear map $E_x \to E_{g\cdot x}$ (such a lift is called a *linear lift*) then the previous formula defines a linear representation of G in the space $\Gamma(E)$.

Suppose, for example, that $E = M \times \mathbb{R}^n$ is the trivial bundle. Let G be a Lie group of diffeomorphisms of M and $\rho : G \to GL(n, \mathbb{R})$ be its linear representation. Then the formula

$$g \cdot (x, a) = (g \cdot x, \rho(g)(a))$$

defines a linear lift of the action of G on M to E. In particular, if $n = 1$ and ρ is trivial, then $g \in G$ transforms a scalar function ϕ into a functon $\phi' = g^{-1} \circ \phi$. It satisfies

$$\phi'(g \cdot x) = \phi(x). \tag{13.15}$$

If we identify a section of E with a smooth vector function $\underline{\phi}(x) = (\phi_1(x), \ldots, \phi_n(x))$ on M, then g acts on $\Gamma(E)$ by the formula

$$g \cdot \underline{\phi}(x) = (\psi_1(x), \ldots, \psi_n(x)),$$

where

$$\psi_i(x) = \sum_{j=1}^n \alpha_{ij}\phi_j(g^{-1} \cdot x), \quad \rho(g) = (\alpha_{ij}) \in GL(n, \mathbb{R}), \quad i = 1, \ldots, n.$$

The tangent bundle $T(M)$ admits the *natural lift* of any action of G on M. For any $\tau_x \in T(M)_x$ we set

$$g \cdot \tau_x = (dg)_x(\tau_x) \in T(M)_{g\cdot x}.$$

A section of $T(M)$ is a vector field $\tau : x \to \tau_x$. We have

$$(g \cdot \tau)_x = g \cdot \tau_{g^{-1}\cdot x} = (dg)_{g^{-1}\cdot x}(\tau_{g^{-1}\cdot x}).$$

If we view a vector field τ as a derivation $C^\infty(M) \to C^\infty(M)$, this translates to

$$(g \cdot \tau)_x(\phi) = \tau_{g^{-1}\cdot x}(g^*(\phi)).$$

Let $g \in G$ define a diffeomorphism from an open set $U \subset M$ with coordinate functions $\underline{x} = (x_1, \ldots, x_n) : U \to \mathbb{R}^n$ to an open set $V = g(U)$ with coordinate functions $\underline{y} = (y_1, \ldots, y_n) : V \to \mathbb{R}^n$. For any $a \in U$ we have

$$y_i(g(a)) = \psi_i(x_1(a), \ldots, x_n(a)), i = 1, \ldots, n, \tag{13.16}$$

for some smooth function $\psi : \underline{x}(U) \to \mathbb{R}^n$. Thus

$$(g \cdot \frac{\partial}{\partial x_j})(y_i) = \frac{\partial}{\partial x_j}(\psi_i(x_1, \ldots, x_n)) = \frac{\partial \psi_i}{\partial x_j}.$$

This implies

$$g_*(\frac{\partial}{\partial x_j})_{g\cdot x} := (g \cdot \frac{\partial}{\partial x_j})_{g\cdot x} = \sum_{i=1}^{n} \frac{\partial \psi_i}{\partial x_j}(x) \frac{\partial}{\partial y_i}, \tag{13.17}$$

$$g \cdot \sum_{j=1}^{n} a^j(x_1, \ldots, x_n) \frac{\partial}{\partial x_j} = \sum_{i=1}^{n} b^i(y) \frac{\partial}{\partial y_i},$$

where

$$b^i(y) = \sum_{j=1}^{n} a^j(g^{-1} \cdot y) \frac{\partial \psi_i}{\partial x_j}(g^{-1} \cdot y).$$

The cotangent bundle $T^*(M)$ also admits the *natural lift.* For any $\omega_x \in T^*(M)_x = T(M)^*_x$ and $\tau_{g\cdot x} \in T(M)_{g\cdot x}$,

$$(g \cdot \omega_x)(\tau_{g\cdot x}) = \omega_x(g^{-1} \cdot \tau_{g\cdot x}).$$

A section of $T^*(M)$ is a differential 1-form $\omega : x \to \omega_x$. We have

$$(g \cdot \omega)_x(\tau_x) = g \cdot \omega_{g^{-1}\cdot x}(\tau_x) = \omega_{g^{-1}\cdot x}(g^{-1} \cdot \tau_x).$$

On the other hand, if $f : M \to N$ is any smooth map of manifolds, one defines the inverse image of a differential 1-form ω on N by

$$f^*(\omega)_x(\tau) = \omega_{f(x)}(df_x(\tau)).$$

If $f = g : M \to M$, we have

$$g \cdot \omega = (g^{-1})^*(\omega).$$

If $g : U \to V$ as above, then

$$g^{-1} \cdot dy_i(\frac{\partial}{\partial x_j}) = g^*(dy_i)(\frac{\partial}{\partial x_j}) = dy_i(g_*(\frac{\partial}{\partial x_j})) = dy_i(\sum_{k=1}^{n} \frac{\partial \psi_k}{\partial x_j} \frac{\partial}{\partial y_k}) = \frac{\partial \psi_i}{\partial x_j}.$$

Therefore

$$g^{-1} \cdot (dy_i) := g^*(dy_i) = \sum_{j=1}^{n} \frac{\partial \psi_i}{\partial x_j} dx_j = d(g^*(y_i)).$$

$$g^{-1} \cdot (\sum_{i=1}^{n} b_i(y) dy_i) = \sum_{i=1}^{n} b_i(g \cdot x) \frac{\partial \psi_i}{\partial x_j} dx_j.$$

Now assume that G consists of linear transformations of $M = \mathbb{R}^n$. Fix coordinate functions $x_1, \ldots, x_n$ such that $\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_n}$ gives a basis in each $T(M)_{\mathbf{a}}$, and $dx_1, \ldots, dx_n$ gives a basis in $T^*(M)_{\mathbf{a}}$. Let $g \in G$ be identified with a matrix $A = (\alpha_{ij})$. It acts on $\mathbf{a} = (a_1, \ldots, a_n)$ by transforming it to

$$\mathbf{b} = A \cdot \mathbf{a}.$$

We have

$$x_i(A \cdot \mathbf{a}) = b_i = \sum_{j=1}^{n} \alpha_{ij} a_j = \sum_{j=1}^{n} \alpha_{ij} x_j(\mathbf{a}).$$

Thus (13.16) takes the form

$$g^*(x_i) = \phi_i(x_1, \ldots, x_n) = \sum_{j=1}^{n} \alpha_{ij} x_j, i = 1, \ldots, n,$$

and

$$g \cdot \frac{\partial}{\partial x_j} = g_*(\frac{\partial}{\partial x_j}) = \sum_{i=1}^{n} \alpha_{ij} \frac{\partial}{\partial x_i},$$

$$g^{-1} \cdot dx_i = g^*(dx_i) = \sum_{j=1}^{n} \alpha^{ij} dx_j,$$

where

$$A^{-1} = (\alpha^{ij}).$$

13.5 Let us identify the Lagrangian $\mathcal{L}$ with a map $\mathcal{L} : C^\infty(M) \to C^\infty(M), \phi(x) \to \mathcal{L}(\phi)(x)$. We use the unknown x to emphasize that the values of $\mathcal{L}$ are smooth functions on M. Let $g : M \to M$ be a diffeomorphism. It acts on the source space and on the target space of $\mathcal{L}$, transforming $\mathcal{L}$ to $g\mathcal{L}$, where

$$(g \cdot \mathcal{L})(\phi) = g \cdot \mathcal{L}(g^{-1} \cdot \phi) = \mathcal{L}(g^*(\phi)((g^{-1})^* \cdot x).$$

We say that the *Lagrangian is invariant* with respect to g if $g \cdot \mathcal{L} = \mathcal{L}$, or, equivalently,

$$\mathcal{L}(g^*(\phi))(x) = \mathcal{L}(\phi)(g \cdot x). \tag{13.18}$$

Example. Let $M = \mathbb{R}$ and $\mathcal{L}(\phi) = \phi'(x)$. Take $g : x \to 2x$. Then

$$\mathcal{L}(g^*(\phi)) = \phi(2x)' = 2\phi'(2x) = 2\mathcal{L}(\phi)(2x) = 2\mathcal{L}(\phi)(g \cdot x).$$

So, $\mathcal{L}$ is not invariant with respect to g. However, if we take $h : x \to x + c$, then

$$\mathcal{L}(h^*(\phi)) = \phi(x + c)' = \phi(x + c) = \mathcal{L}(\phi)(x + c) = \mathcal{L}(\phi)(h \cdot x).$$

So, $\mathcal{L}$ is invariant with respect to h.

Assume $\mathcal{L}$ is invariant with respect to g and g leaves the volume form d^nx invariant. Then

$$S(g^*(\phi)) = \int_M \mathcal{L}(g^*(\phi))d^nx = \int_M \mathcal{L}(\phi))(g\cdot x)g^*(d^nx) = \int_M \mathcal{L}(\phi)(x)d^nx = S(\phi).$$

This shows that g leaves the level sets of the action S invariant. In particular, it transforms critical fields to critical fields. Or, equivalently, it leaves invariant the set of solutions of the Euler-Lagrange equation.

Let η be a vector field on M and g_η^t be the associated one-parameter group of diffeomorphisms of M. We say that η is an *infinitesimal symmetry* of the Lagrangian $\mathcal{L}$ if $\mathcal{L}$ is invariant with respect to the diffeomorphisms g_η^t.

Let L_η be the Lie derivative associated to the vector field η (see Lecture 2). Assume η is an infinitesimal symmetry of $\mathcal{L}$. Then $\mathcal{L}(\phi)(g_\eta^t\cdot x) = \mathcal{L}(\phi(g_\eta^t\cdot x))$. This implies

$$L_\eta(\mathcal{L}(\phi,d\phi)(x)) = \lim_{t\to 0}\frac{\mathcal{L}(\phi)(g_\eta^t\cdot x)-\mathcal{L}(\phi)(x)}{t} = \frac{\partial\mathcal{L}}{\partial\phi}L_\eta(\phi)+\frac{\partial\mathcal{L}}{\partial d\phi}L_\eta(d\phi), \tag{13.19}$$

where L_η denotes the Lie derivative with respect to the vector field η. By Cartan's formula,

$$L_\eta(\phi) = \eta(\phi), \quad L_\eta(d\phi) = d\eta(\phi).$$

Let us assume $M=\mathbb{R}^n$ and $\mathcal{L}(\phi,d\phi) = F(\phi,\partial_1\phi,\ldots,\partial_n\phi)$ for some smooth function F in $n+1$ variables. Let $\eta = \sum_i a^i(x)\frac{\partial}{\partial x_i}$, then

$$\eta(\phi) = \sum_i a^i(x)\frac{\partial\phi}{\partial x_i},$$

$$d\eta(\phi) = d(\sum_i a^i\frac{\partial\phi}{\partial x_i}) = \sum_{i,j=1}^n(\frac{\partial a^i}{\partial x_j}\frac{\partial\phi}{\partial x_i}+a^i(x)\frac{\partial^2\phi_i}{\partial x_j\partial x_i})dx_j.$$

We can now rewrite (13.19) as

$$a^\mu\partial_\mu\mathcal{L} = \frac{\partial\mathcal{L}}{\partial\phi}a^\nu(x)\partial_\nu\phi+\frac{\partial\mathcal{L}}{\partial\partial_\mu\phi}(\partial_\mu a^\nu\partial_\nu\phi+a^\nu(x)\partial_\nu\partial_\mu\phi).$$

Assume now that ϕ satisfies the Euler-Lagrange equation. Then we can rewrite the previous equality in the form

$$a^\mu\partial_\mu\mathcal{L} = \partial_\mu\frac{\partial\mathcal{L}}{\partial\partial_\mu\phi}a^\nu(x)\partial_\nu\phi+\frac{\partial\mathcal{L}}{\partial\partial_\mu\phi}(\partial_\mu a^\nu\partial_\nu\phi+a^\nu(x)\partial_\nu\partial_\mu\phi).$$

Since we assume that g_η^t preserves the volume form $\omega = dx_1\wedge\ldots\wedge dx_n$, we have

$$L_\eta(\omega) = d\langle\eta,\omega\rangle = (\sum_{i=1}^n\frac{\partial a^i}{\partial x_i})\omega = 0.$$

This shows that

$$a^{\mu}\frac{\partial \mathcal{L}}{\partial x_{\mu}} = \partial_{\mu}(a^{\mu}\mathcal{L}) - \mathcal{L}\partial_{\mu}(a^{\mu}) = \partial_{\mu}(a^{\mu}\mathcal{L}).$$

Now combine everything under ∂_{μ} to obtain

$$0 = -\partial_{\mu}(a^{\mu}\mathcal{L}) + \partial_{\mu}\frac{\partial \mathcal{L}}{\partial\partial_{\mu}\phi}a^{\nu}(x)\partial_{\nu}\phi + \frac{\partial \mathcal{L}}{\partial\partial_{\mu}\phi}(\partial_{\mu}a^{\nu}\partial_{\nu}\phi + a^{\nu}(x)\partial_{\nu}\partial_{\mu}\phi) =$$

$$= \partial_{\mu}(\frac{\partial \mathcal{L}}{\partial\partial_{\mu}\phi}a^{\nu}\partial_{\nu}\phi - \delta^{\mu}_{\nu}a^{\nu}\mathcal{L}) = \sum_{\nu,\mu=1}^{n}\partial_{\mu}(\frac{\partial \mathcal{L}}{\partial\partial_{\mu}\phi}a^{\nu}\partial_{\nu}\phi - \delta^{\nu}_{\mu}a^{\nu}\mathcal{L}), \tag{13.20}$$

where $\delta^{\nu}_{\mu} = \delta_{\nu\mu}$ is the Kronecker symbol.

Set

$$J^{\mu} = \sum_{\nu=1}^{n}(a^{\nu}(\frac{\partial \mathcal{L}}{\partial\partial_{\mu}\phi}\partial_{\nu}\phi - \delta_{\mu\nu}\mathcal{L}), \quad \mu = 1, \ldots, n, \tag{13.21}$$

The vector-function

$$\mathbf{J} = (J^1, \ldots, J^n)$$

is called the *current* of the field ϕ with respect to the vector field η. From (13.21) we now obtain

$$\text{div}\mathbf{J} = \partial_{\mu}J^{\mu} = \sum_{\mu=1}^{n}\frac{\partial J^{\mu}}{\partial x_{\mu}} = 0. \tag{13.22}$$

For example, let $n = 4, (x_1, x_2, x_3, x_4) = (t, x_1, x_2, x_3)$. We can write $(J^1, J^2, J^3, J^4) = (j^0, \mathbf{j})$ and rewrite (13.21) in the form

$$\frac{\partial j^0}{\partial t} = -\text{div}(\mathbf{j}) := -(\sum_{i=1}^{3}\frac{\partial j^i}{\partial x_i}).$$

Thus, if we introduce the *charge*

$$Q(t) = \int_{\mathbb{R}^3} j^0(t, x)d^3x,$$

and assume that $\mathbf{j}$ vanishes fast enough at infinity, then the divergence theorem implies the conservation of the charge

$$\frac{dQ(t)}{dt} = 0.$$

Remarks. 1. If we do not assume that the infinitesimal transformation of M preserves the volume form, we have to require that the action function is invariant but not the Lagrangian. Then, we can deduce equation (13.21) for any field ϕ satisfying the Euler-Lagrange equation.

2. We can generalize equation (13.22) to the case of fields more general than scalar fields. In the notation of 13.3, we assume that E is equipped with a lift with respect to some

one-parameter group of diffeomorphisms g_η^t of M. This will define a canonical lift of g_η^t to $E \oplus \Lambda^1(E)$. At this point we have to further assume that the connection on E is invariant with respect to this lift. Now we consider the Lagrangian $\mathcal{L} : E \oplus (\Lambda^1(E)) \to \mathbb{R}$ as the operator $\mathcal{L} : \Gamma(E) \to \Gamma(E), \phi \to \mathcal{L}(\phi, \nabla\phi)$. We say that η is an infinitesimal symmetry of $\mathcal{L}$ if $g_\eta^t \cdot \mathcal{L}(g_\eta^{-t} \cdot \phi) = \mathcal{L}(\phi)$ for any $\phi \in \Gamma(E)$. Here we consider the action of g_η^t on sections as defined in 13.5. Now equation (13.19) becomes

$$\nabla(\tau)(\mathcal{L}(\phi)) = \frac{\partial\mathcal{L}}{\partial\phi}\nabla(\tau)(\phi) + \frac{\partial\mathcal{L}}{\partial\nabla(\phi)}\nabla(\nabla(\tau)(\phi)).$$

Here we consider the connection ∇ both as a map $\Gamma(E) \to \Gamma(\Lambda^1(E))$ and as a map $\Gamma(T(M)) \to \Gamma(End(E))$. We leave to the reader to find its explicit form analogous to equation (13.20), when we trivialize E and choose a connection matrix Γ_{ij}^k.

13.6 Now we shall consider various special cases of equation (13.20). Let us assume that $M = \mathbb{R}^n$ and $G = \mathbb{R}^n$ acts on itself by translations. Since ∂_ν commute with translations, any Lagrangian $\mathcal{L}(\phi, d\phi)$ is invariant with respect to translations. Also translations preserve the standard volume in $\mathbb{R}^n$. We can identify the Lie algebra of G with $\mathbb{R}^n$. For any $\eta = (a^1, \dots, a^n) \in \mathbb{R}^n$, the vector field $\eta^\sharp$ is the constant vector field $\sum_i a^i \frac{\partial}{\partial x_i}$. Thus we get from (13.20) that

$$\sum_{\nu,\mu=1}^n a^\nu \partial_\mu \big(\sum_{\nu=1}^n \frac{\partial\mathcal{L}}{\partial\partial_\mu\phi}\partial_\nu\phi - \delta_{\nu\mu}\mathcal{L}\big) = 0.$$

Since this is true for all η, we get

$$\sum_{\mu=1}^n \partial_\mu \big(\frac{\partial\mathcal{L}}{\partial\partial_\mu\phi}\partial_\nu\phi - \delta_\nu^\mu\mathcal{L}\big) = 0, \quad \nu = 1, \dots, n.$$

The tensor

$$(T_\nu^\mu) = (\frac{\partial\mathcal{L}}{\partial\partial_\mu\phi}\partial_\nu\phi - \delta_\nu^\mu\mathcal{L}) := \sum_{\nu,\mu}^n (\frac{\partial\mathcal{L}}{\partial\partial_\mu\phi}\partial_\nu\phi - \delta_\nu^\mu\mathcal{L})\frac{\partial}{\partial x_\mu} \otimes dx_\nu \tag{13.23}$$

is called the *energy-momentum tensor*. For example, consider the Klein-Gordon action where

$$\mathcal{L} = \frac{1}{2}(\partial_\mu\phi\partial^\mu\phi - m^2\phi^2).$$

We have

$$(T_\nu^\mu) = \begin{pmatrix} (\partial_0\phi)^2 - \mathcal{L} & \partial_0\phi\partial_1\phi & \partial_0\phi\partial_2\phi & \partial_0\phi\partial_3\phi \\ -\partial_0\phi\partial_1\phi & -(\partial_1\phi)^2 - \mathcal{L} & -\partial_1\phi\partial_2\phi & -\partial_1\phi\partial_3\phi \\ -\partial_0\phi\partial_2\phi & -\partial_1\phi\partial_2\phi & -(\partial_2\phi)^2 - \mathcal{L} & -\partial_2\phi\partial_3\phi \\ -\partial_0\phi\partial_3\phi & -\partial_1\phi\partial_3\phi & -\partial_2\phi\partial_3\phi & -(\partial_3\phi)^2 - \mathcal{L} \end{pmatrix}.$$

Since ϕ satisfies Klein-Gordon equation (13.11), we easily check that each row of the matrix T is a vector whose divergence is equal to zero. the frequency

We have

$$T_0^0 = \frac{1}{2}(m^2\phi^2 + \sum_{i=0}^{3}(\partial_i\phi)^2).$$

If we integrate T_0^0 over $\mathbb{R}^3$ with coordinates (x_1, x_2, x_3), we obtain the analog of the total energy function

$$\frac{1}{2}\sum_{i=1}^{n}(m^2q_i^2 + \dot{q}_i^2).$$

We also have $T^{0k} = \partial_0\phi\partial_k\phi = \frac{\partial\mathcal{L}}{\partial\partial_0\phi}\partial_k\phi, k = 1, 2, 3$. In classical mechanics $p_i = \frac{\partial\mathcal{L}}{\partial\dot{q}_i}$ is the momentum coordinate. Thus the integral over $\mathbb{R}^3$ of the vector function (T^{01}, T^{02}, T^{03}) represents the analog of the momentum.

13.7 Consider the orthogonal group $G = O(1, 3)$ of linear transformations $A : \mathbb{R}^4 \to \mathbb{R}^4$ which preserve the quadratic form $x_0^2 - x_1^2 - x_2^2 - x_3^2$. Take the Klein-Gordon Lagrangian $\mathcal{L}$. Since $(\partial_0, \partial_1, \partial_2, \partial_3)$ transforms in the same way as the vector (x_0, x_1, x_2, x_3), we obtain that $\partial_\nu\partial^\nu(\phi(A \cdot x)) = \partial_\nu\partial^\nu(\phi)(A \cdot x)$. Thus

$$\mathcal{L}(\phi(A \cdot x)) = \partial_\nu\partial^\nu(\phi(A \cdot x)) - m^2\phi(A \cdot x)^2 = \partial_\nu\partial^\nu(\phi)(A \cdot x) - m^2\phi^2(A \cdot x) = \mathcal{L}(\phi)(A \cdot x).$$

This shows that the Lagrangian is invariant with respect to all transformations from the group $O(1, 3)$. Let us find the corresponding currents.

As we saw in Lecture 12, the Lie algebra of $O(1, 3)$ consists of matrices $A \in Mat_4(\mathbb{R})$ satisfying $A \cdot J + J \cdot A^t = 0$, where

$$J = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

If $A = (a_i^j)$, this means that

$$a_0^0 = 0, a_0^i = a_i^0, \quad i \neq 0, \quad a_j^i = -a_i^j, i, j \neq 0.$$

This can also be expresed as follows. Let us write

$$J = (g_{ij}),$$

and $A \cdot J = (a_{ij})$, where

$$a_{ij} = \sum_{k=0}^{3} a_i^k g_{kj}.$$

Then $A = (a_i^j)$ belongs to the Lie algebra of $O(1, 3)$ if and only if the matrix $A \cdot J = (a_{ij})$ is a skew-symmetric matrix. In fact, this generalizes to any orthogonal group. Let $O(f)$

be the group of linear transformations preserving the quadratic form $f = g_{\nu\mu}x^\nu x^\mu$. Here we switched to superscript indices for the coordinate functions x_i. Then $A = (a_i^j)$ belongs to the Lie algebra of $O(f)$ if and only if $A \cdot (g_{\nu\mu}) = (a_i^\nu g_{\nu j})$ is skew-symmetric. The reason that we change the usual notation for the entries of A is that we consider A as an endomorphism of the linear space V, i.e., a tensor of type $(1,1)$. On the other hand, the quadratic form f is considered as a tensor of type $(0,2)$ (x_i forming a basis of the dual space). The contraction map $(V^* \otimes V) \otimes (V^* \otimes V^*) \to V^* \otimes V^*$ corresponds to the product $A \cdot (g_{ij})$ and has to be considered as a quadratic form, hence the notation a_{ij}.

Under the natural action of $GL(n,\mathbb{R}) \to \mathbb{R}^n, A \to A \cdot \mathbf{v}$, the matrix $A = (a_i^j)$ goes to

$$(v_1, \ldots, v_n) \cdot A^t = (\sum_{j=1}^n a_1^j v_j, \ldots, \sum_{j=1}^n a_n^j v_j).$$

Under this map the coordinate function x^i on $\mathbb{R}^n$ is pulled back to the function $\sum_{j=1}^n x_i^j c_j$ on $\mathrm{Mat}_n(\mathbb{R})$, where x_i^j is the coordinate function on $Mat_n(\mathbb{R})$ whose value on the matrix (δ_{ij}^{kl}) is equal to δ_{ij}^{kl}. Using (13.17), we see that the vector field $\frac{\partial}{\partial x_i^j}$ on $GL_n(\mathbb{R})$ goes to the vector field $x^j \frac{\partial}{\partial x_i}$ (no summation here!). Therefore, the matrix A^t, identified with the vector field $\eta = \sum_{i,j=1}^n a_j^i \frac{\partial}{\partial x_i^j}$, goes to the vector field

$$\eta^\sharp = \sum_{i=1}^n (\sum_{j=1}^n a_j^i x^j) \frac{\partial}{\partial x_i} = a_\mu^\nu x^\mu \partial_\nu = a^\nu \partial_\nu,$$

where $a^\nu = a_\rho^\nu x^\rho$, $a_{\rho\nu} = -a_{\nu\rho}$.

Note that

$$\sum_{\nu=0}^3 \partial_\nu (a^\nu) = \sum_{\rho=0}^3 a_\rho^\rho = 0,$$

so that the group $O(1,3)$ preserves the standard volume in $\mathbb{R}^3$.

We can rewrite equation (13.20) as follows

$$0 = \partial_\mu \big(\frac{\partial \mathcal{L}}{\partial \partial_\mu \phi} x^\alpha a_\alpha^\nu \partial_\nu \phi - \delta_\nu^\mu x^\alpha a_\alpha^\nu \mathcal{L}\big) = a_\alpha^\nu \partial_\mu \big(\frac{\partial \mathcal{L}}{\partial \partial_\mu \phi} x^\alpha \partial_\nu \phi - \delta_\nu^\mu x^\alpha \mathcal{L}\big) =$$

$$= g_{\beta\nu} a_\alpha^\nu \partial_\mu \big(\frac{\partial \mathcal{L}}{\partial \partial_\mu \phi} g^{\beta\nu} x^\alpha \partial_\nu \phi - g^{\beta\nu} \delta_\nu^\mu x^\alpha \mathcal{L}\big) = a_{\beta\alpha} \partial_\mu \big(\frac{\partial \mathcal{L}}{\partial \partial_\mu \phi} x^\alpha \partial^\beta \phi - \delta^{\beta\mu} x^\alpha \mathcal{L}\big) =$$

$$= \sum_{0 \le \beta < \alpha \le 3}^n a_{\beta\alpha} \partial_\mu \big(\frac{\partial \mathcal{L}}{\partial \partial_\mu \phi} (x^\alpha \partial^\beta \phi - x^\beta \partial^\alpha \phi) - (\delta^{\beta\mu} x^\alpha - \delta^{\alpha\mu} x^\beta) \mathcal{L}\big), \tag{13.24}$$

where $(g^{\mu\nu}) = (g_{\mu\nu})^{-1}$. Since the coefficients $a_{\beta\alpha}, \beta < \alpha$, are arbitrary numbers, we deduce that

$$\partial_\mu \big(\frac{\partial \mathcal{L}}{\partial \partial_\mu \phi} (x^\alpha \partial^\beta \phi - x^\beta \partial^\alpha \phi) - (\delta^{\beta\mu} x^\alpha - \delta^{\alpha\mu} x^\beta) \mathcal{L}\big) = 0.$$

Let

$$T^{\beta\mu} = g^{\beta\nu} T^\mu_\nu = \frac{\partial \mathcal{L}}{\partial \partial_\mu \phi} \partial^\beta \phi - \delta^{\beta\mu} \mathcal{L}$$

where (T^μ_ν) is the energy-momentum tensor (13.23). Set

$$\mathcal{M}^{\mu,\alpha\beta} = T^{\beta\mu} x^\alpha - T^{\alpha\mu} x^\beta. \tag{13.25}$$

Now the equation $\operatorname{div}(\mathbf{J}) = 0$ is equiavlent to

$$\partial_\mu \mathcal{M}^{\mu,\alpha\beta} = \sum_{\mu=1}^n \partial_\mu \mathcal{M}^{\mu,\alpha\beta} = 0.$$

We can introduce the charge

$$M^{\alpha\beta} = \int_{\mathbb{R}^3} \mathcal{M}^{0,\alpha\beta} d^3 x.$$

Then it is conserved

$$\frac{dM^{\alpha\beta}}{dt} = 0.$$

When we restrict the values α, β to $1, 2, 3$, we obtain the conservation of angular momentum. For this reason, the tensor $(\mathcal{M}^{\mu,\alpha\beta})$ is called the *angular momentum tensor.*

13.8 The Klein-Gordon equation was proposed as a relativistic version of Schrödinger equation by E. Schrödinger, Gordon and O. Klein in 1926-1927. The original idea was to use it to construct a relativistic 1-particle theory. However, it was abandoned very soon due to several difficulties.

The first difficulty is that the equation admits obvious solutions

$$\phi(x,t) = e^{i(k\cdot x + Et)}$$

with

$$|k|^2 + m^2 = k_1^2 + k_2 + k_3^2 + m^2 - E^2 = 0. \tag{13.26}$$

Since the Schrödinger equation (5.6) should give us

$$i\frac{\partial \phi}{\partial t} = -E\phi = H\phi,$$

this implies that $-E$ is the eigenvalue of the Hamiltonian. Thus we have to interpret $-E$ as the energy. However, equation (13.26) admits positive solutions for E, so this leads to the negative energy of a system consisting of one particle. We, of course, have seen this already in the case of a particle in central field, for example, the electron in the hydrogen atom. However in that case, the energy was bounded from below. In our case, the energy spectrum is not bounded from below. This means we can extract an arbitrarily large amount of energy from our system. This is difficult to assume. In the next Lectures we shall show how to overcome this difficulty by applying second quantization.

Exercises.

1. Let $\tilde{g} : E \to E$ be a lift of a diffeomorphism $g : M \to M$ to a vector bundle over M. Show that the dual bundle (as well as the exterior and tensor powers) admits a natural lift too.

2. Generalize the Euler-Lagrange equation (13.8) to the case when the Lagrangian depends on $x \in \mathbb{R}^n$.

3. Let $\mathbf{J} = (\frac{\partial \mathcal{L}}{\partial \partial_t \phi}\partial_t\phi - \mathcal{L}, \frac{\partial \mathcal{L}}{\partial \partial_{x_1}\phi}\partial_t\phi, \frac{\partial \mathcal{L}}{\partial \partial_{x_2}\phi}\partial_t\phi, \frac{\partial \mathcal{L}}{\partial \partial_{x_3}\phi}\partial_t\phi)$, where $\mathcal{L}$ is the Klein-Gordon Lagrangian (13.10). Verify directly that $\mathrm{div}(\mathbf{J}) = 0$. What symmetry transformation in the Lagrangian gives us this conserved current?

4. Let $M = \mathbb{R}$ and $\mathcal{L}(\phi)(x) = \phi(x) + 2\log|\phi'(x)| - \log|\phi''(x)|$. Find the Euler-Lagrange equation for this Lagrangian. Note that it depends on the second derivative so this case was discussed in the lecture. Observing that $\mathcal{L}$ is invariant with respect to transformations $x \to \lambda x$, where $\lambda \in \mathbb{R}^*$, find the corresponding conservation law.

5. The Klein-Gordon Lagrangian for complex scalar fields $\phi : \mathbb{R}^4 \to \mathbb{C}$ has the form $\mathcal{L}(\phi) = \partial_\mu\phi\partial^\mu\bar{\phi} - m^2|\phi|^2$. Obviously it is invariant with respect to transformations $\phi \to \lambda\phi$, where $\lambda \in \mathbb{C}$ with $|\lambda| = 1$. Find the corresponding current vector $\mathbf{J}$ and the conserved charge.