Séminaires & Congrès
4, 2000, p. 115–127

APPLICATIONS OF CURVED BERNSTEIN-GELFAND-GELFAND SEQUENCES

by

David M.J. Calderbank

Abstract. — I discuss applications of Bernstein-Gelfand-Gelfand sequences in conformal differential geometry.

***Résumé* (Applications des suites de Bernstein-Gelfand-Gelfand courbées).** — J'étudie des applications des suites de Bernstein-Gelfand-Gelfand en géométrie différentielle conforme.

One of the themes of mathematics in the twentieth century has been the growing realization that representation theory and geometry are closely related. There are at least two aspects to this. Firstly, there is the geometric study of representation theory that follows inevitably from the definition of a Lie group, where global methods of geometry and topology are applied to homogeneous spaces. Secondly, there is the increasing use of representation theory as a tool and language for the invariant analysis of geometric structure—that is to say, the local (pointwise) aspects of differential geometry. Although this second aspect also has a long history (the work of Cartan stretches back into the nineteenth century), I think it is fair to say that only in the last twenty years or so has representation theory really begun to gain ground as an alternative to the hands-on approach of local coordinate computations. One area which has motivated this shift is quaternionic geometry, the study of which only intensified relatively recently in the history of differential geometry, driven by many different forces, such as supersymmetry, the classification of metric holonomies, and the geometry of moduli spaces. Confronted by an unfamiliar geometry, geometers turned to the representation theory of $\mathbb{H}^* \cdot \mathrm{GL}(n, \mathbb{H})$ and its subgroups (such as $\mathrm{Sp}(n)$ and $\mathrm{Sp}(1)\,\mathrm{Sp}(n)$) as an efficient way to develop intuition.

2000 *Mathematics Subject Classification.* — Primary 53A55; Secondary 53A30, 53C25.
Key words and phrases. — Bernstein-Gelfand-Gelfand sequences, self-duality, Einstein metrics, Einstein-Weyl geometry, conformally invariant operators.

This paper concerns the role of representation theory in geometric structures, but, in keeping with the theme of the conference (and the author's interests), the emphasis will be on conformal and Riemannian geometry, rather than quaternionic geometry. The aim is demonstrate that even in these classical subjects, representation theoretic methods not only provide an efficient language when computations become unmanageable, but also that they lead to genuinely new insights and constructions. Nowhere has this been more true than in the study of invariant differential operators, and so I will focus on an area where much progress has been made in recent years: Bernstein-Gelfand-Gelfand sequences.

1. Parabolic geometries and the BGG sequences

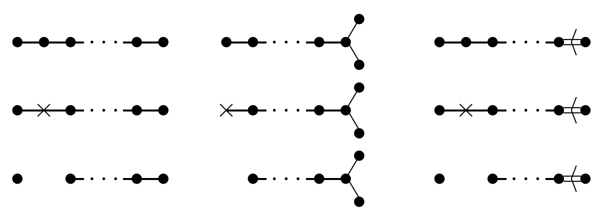
A parabolic geometry is a geometry modelled on a generalized flag variety G/P, where G is a semisimple Lie group and P is a parabolic subgroup. These compact homogeneous spaces arise naturally in representation theory as projectivized orbits of highest weight vectors in irreducible representations. A parabolic subgroup P has a decomposition $P = L \ltimes N$, where L is reductive and N is nilpotent—L is often called the Levi factor. It is convenient to describe parabolic subgroups by crossing nodes on the Dynkin diagram of G, so that if the crossed nodes (and adjoining lines) are deleted, the result is the Dynkin diagram of (the semisimple part of) L. Such diagrams will also denote the corresponding flag varieties. Strictly speaking, these diagrams describe complex geometries, whereas real flag varieties should be denoted by Satake diagrams with crosses. I shall ignore this distinction. Here are some examples, together with names for the corresponding parabolic geometries (or rather, for suitable real forms of these geometries).

Quaternionic — Even conformal — Quaternionic CR

G semisimple Lie group

P parabolic subgroup

L Levi factor



The key example for this paper is $G = \mathrm{SO}(n+1,1)$, $P = \mathrm{CO}(n) \ltimes (\mathbb{R}^n)^*$, with $G/P \cong S^n$. This is the conformal sphere, identified as the "sky" in $(n+1,1)$-dimensional spacetime. The Dynkin diagram is shown above for $n \geqslant 6$ even. For $n = 1$–5, the diagrams are ×, × ×, ×⇒•, •—×—•, ×—•⇒•. The geometries for $n = 1, 2$ are projective and Möbius geometry respectively.

Geometries "modelled on" homogeneous spaces are most simply defined as Cartan geometries, i.e., one views the homogeneous geometry on G/P as a principal P-bundle

$G \to G/P$ equipped with the parallelism $TG \cong G \times \mathfrak{g}$; then a curved analogue of this is a principal P-bundle $\mathcal{G} \to M$ equipped with a Cartan connection.

1.1. Definition. — Let M be a manifold of the same dimension as G/P.

(i) A *Cartan geometry* of type $(\mathfrak{g}, P)$ on M is a principal P-bundle $\pi\colon \mathcal{G} \to M$, together with a P-equivariant $\mathfrak{g}$-valued 1-form $\eta\colon T\mathcal{G} \to \mathfrak{g}$ such that for each $u \in \mathcal{G}$, $\eta_u\colon T_u\mathcal{G} \to \mathfrak{g}$ is an isomorphism restricting to the canonical isomorphism between $T_u(\mathcal{G}_{\pi(u)})$ and $\mathfrak{p}$.

(ii) The *curvature* $K\colon \Lambda^2 T\mathcal{G} \to \mathfrak{g}$ of a Cartan geometry is defined by
$$K(U,V) = d\eta(U,V) + [\eta(U), \eta(V)].$$
It induces a curvature function $\kappa\colon \mathcal{G} \to \Lambda^2\mathfrak{g}^* \otimes \mathfrak{g}$ via
$$\kappa(u)(\xi,\chi) = K_u\big(\eta^{-1}(\xi), \eta^{-1}(\chi)\big) = [\xi,\chi] - \eta_u[\eta^{-1}(\xi), \eta^{-1}(\chi)],$$
where $u \in \mathcal{G}$ and the brackets are the Lie bracket in $\mathfrak{g}$ and the Lie bracket of vector fields on $\mathcal{G}$.

Cartan geometries usually arise from a more familiar geometric structure on M by a process of prolongation [**10, 26**]. In the case of conformal geometry this prolongation procedure is the famous construction of the normal Cartan connection.

Associated to a P-module $\mathbb{E}$ is a vector bundle $E = \mathcal{G} \times_P \mathbb{E}$. For example, the Cartan connection η identifies the tangent bundle of M with $\mathcal{G} \times_P \mathfrak{g}/\mathfrak{p}$. An important special case of this is a G-module $\mathbb{W}$, viewed as a P-module by restriction. In this case, $W = \mathcal{G} \times_P \mathbb{W} = \tilde{\mathcal{G}} \times_G \mathbb{W}$, where $\tilde{\mathcal{G}} = \mathcal{G} \times_P G$. A Cartan connection on $\mathcal{G}$ induces a principal bundle connection of $\tilde{\mathcal{G}}$ and hence a covariant derivative on W. In various contexts, such linear representations of the Cartan connection have been called "tractor" or "local twistor" connections [**1, 2**].

When the Cartan connection is flat (e.g., on the homogeneous model), parallel sections of W can be identified, at least locally, with $\mathbb{W}$, so we have
$$0 \to \mathbb{W} \to \mathrm{C}^\infty(W) \to \mathrm{C}^\infty(T^*M \otimes W). \tag{1.1}$$
This is the beginning of a resolution, the (dualized, generalized) Bernstein-Gelfand-Gelfand resolution, of $\mathbb{W}$ (or, more accurately, of the sheaf of parallel sections of W).

A differential geometer or topologist, asked to extend (1.1), would immediately come up with a resolution by a complex of first order differential operators, the twisted de Rham complex:
$$\mathrm{C}^\infty(W) \to \mathrm{C}^\infty(T^*M \otimes W) \to \mathrm{C}^\infty(\Lambda^2 T^*M \otimes W) \to \mathrm{C}^\infty(\Lambda^3 T^*M \otimes W) \to \cdots$$
The problem with this resolution is that W is often quite a complicated bundle, and hence so are the bundles in this resolution. Bernstein, Gel'fand and Gel'fand [**5**] found a way to break up these bundles under the action of P using differential projections and hence obtain a resolution with much simpler bundles, but perhaps higher order

differentials. (In fact, they only considered the Borel case, and their construction was generalized to arbitrary parabolics by Lepowsky [**22**].)

This paper is about the curved version of this construction, which began with the work of Baston [**3**] (see also [**7, 14, 16**]), who gave a double complex construction of the BGG resolution in the case that N is abelian, and argued that his construction could be generalized to curved geometries, with one crucial difference. Namely (as one might expect from the twisted de Rham sequence above) the curved BGG sequence is no longer a complex. Unfortunately, Baston's proofs were not entirely clear, and it was not until the work of Čap, Slovák, and Souček [**11**] that the curved BGG sequences were obtained for arbitrary parabolic geometries, and with full arguments.

My goal in this paper is to try to indicate why this is a significant achievement and why curved BGG sequences might be useful in practice. In other words, my focus will be on applications rather than the theory.

For applications, the main thing needed from the theory is an efficient algorithm for computing the bundles in the BGG sequence. These bundles are associated to Lie algebra homology groups, and most algorithms to compute them are based on Kostant's theorem [**20**], which states that this homology can be obtained by applying the affine action of Weyl group reflections to the highest weight vector of $\mathbb{W}$.

I want to explain how to do this, using some notation devised a few years ago for representations of parabolic subgroups.

Let G have rank m and let $\varepsilon_1, \dots \varepsilon_m$ be an orthonormal basis for the Cartan subalgebra $\mathfrak{h} \cong \mathfrak{h}^*$ so that the roots are given in the "standard" form that one finds in any book on Lie theory (e.g. [**19**]): for type A_m and G_2 it is more convenient to identify $\mathfrak{h}$ with $\mathbb{R}^{m+1}/\langle(1,1,\dots 1)\rangle$ or $\mathbb{R}^3/\langle(1,1,1)\rangle$. In terms of this basis, all but one or two of the simple roots of $\mathfrak{g}$ are of the form $\varepsilon_i - \varepsilon_{i+1}$.

The highest weights of irreducible representations may now be described by m-tuples $(\lambda_1, \dots \lambda_m)$—or $(m+1)$-tuples for A_m and G_2—where the entries are "integral" in the sense that the inner products with the coroots are integers (this usually means the entries are all integers, or all half-integers in some cases). In order to have a representation of G, this weight has to be to be *G-dominant*, i.e., the inner product of $\sum \lambda_i \varepsilon_i$ with the simple roots should be non-negative. In practice this means that $\lambda_1 \geqslant \lambda_2 \geqslant \dots \geqslant \lambda_m$ with some additional inequalities depending on the type of the Lie group. For representations of P, only P-dominance is needed, i.e., the inequality only needs to hold for roots of the Levi factor of P. In terms of the Dynkin diagram, whose nodes are in one-to-one correspondence with simple roots, this means the inequalities corresponding to crossed nodes need not hold.

If the node corresponding to $\varepsilon_i - \varepsilon_{i+1}$ is crossed, then the highest weight of a P-representation need not satisfy $\lambda_i \geqslant \lambda_{i+1}$. This can be indicated by writing the highest weights of P-representations as $(\lambda_1, \dots \lambda_i | \lambda_{i+1}, \dots \lambda_m)$. If this highest weight happens to satisfy $\lambda_i \geqslant \lambda_{i+1}$ anyway, then the corresponding P-representation is the irreducible P-subrepresentation generated by the highest weight vector of the G-representation associated to $(\lambda_1, \dots \lambda_i, \lambda_{i+1}, \dots \lambda_m)$.

This works well for the simple examples which arise in practice. If more nodes are crossed, then more bars are needed, although in general, one would have to dream up a different notation for the exceptional roots at the right hand end of the Dynkin diagram.

The Weyl group is generated by reflections in the hyperplanes orthogonal to the simple roots. The advantage of working in an orthonormal basis is that reflections are easy to handle. In particular, a (not necessarily simple) root $\varepsilon_i - \varepsilon_j$ acts by exchanging λ_i and λ_j. The affine action is obtained by conjugating this action with translation by the half sum of the positive roots. In practice this means that $(\varepsilon_i - \varepsilon_j).\lambda = \tilde{\lambda}$ where $\tilde{\lambda}_i = \lambda_j - (j-i)$ and $\tilde{\lambda}_j = \lambda_i + (j-i)$. Note that if $\lambda_i \geqslant \lambda_j$ (for $i < j$) then $\tilde{\lambda}_i < \tilde{\lambda}_j$.

A more systematic notation involves using the basis corresponding to the fundamental weights, perhaps indicating the coefficients by writing them above the nodes on the Dynkin diagram [**4**]. The problem with this notation is that non-simple root reflections are difficult to apply, which entails a two pass procedure to compute the BGG sequence (again, see [**4**]).

The zeroth bundle in the BGG sequence of a G-representation λ is associated to the irreducible P-subrepresentation with the same highest weight. The first bundle is associated to the direct sum of all irreducible P-representations whose highest weight can be obtained by applying a simple affine root reflection to λ. The second bundle is obtained by applying a second, not necessarily simple, affine root reflection to these weights and keeping the P-dominant weights which can be obtained from λ by a composite of two simple affine root reflections. One continues applying root reflections in this way, so that the length of the element of the Weyl group (in terms of simple roots) increases by one at each step. It is usually easy to see which reflections will give P-dominant weights: in simple examples there are often only one or two reflections which work, and the they will often automatically increase the length of the element of the Weyl group by one. Furthermore, the same sequence of reflections generates the BGG sequence of any generic G-representation. In conformal geometry, the sequence of root reflections begins $\varepsilon_1 - \varepsilon_2, \varepsilon_1 - \varepsilon_3, \varepsilon_1 - \varepsilon_4, \dots$.

For example, the BGG resolution of the trivial representation in conformal geometry starts as follows:

$$0 \to (0,0,0,\dots 0) \to (0|0,0,\dots 0) \to (-1|1,0,\dots 0) \to (-2|1,1,0,\dots 0) \to \cdots$$

Here, and throughout the paper, I adopt the usual convention of denoting a representation, an associated bundle and its (sheaf of) smooth sections by the corresponding highest weight. In conventional terms, the first entry of $(\lambda_1|\lambda_2, \dots \lambda_m)$ is the conformal weight, while the remaining elements describe the representation of $\mathrm{SO}(n)$, the semisimple part of L. If $\sum_{i \geqslant 2} |\lambda_i| = k$ then the representation is a subspace