Nanotechnology **12** (2001) 316–321 PII: S0957-4484(01)20733-8

Exponential assembly

George D Skidmore, Eric Parker, Matthew Ellis, Neil Sarkar and Ralph Merkle

Zyvex Corporation, 1321 North Plano Road, Richardson, TX 75081, USA

Received 31 January 2001, in final form 26 July 2001
Published 28 August 2001
Online at stacks.iop.org/Nano/12/316

Abstract
A replicative assembly methodology may be based on assembly stations, each with two degrees of rotational freedom. These stations share translational degrees of freedom in the three Cartesian axes by using a common translating mechanism. The methodology provides replication of the assembly stations, but due to the common translating mechanism and control systems, it cannot be termed self-replicating. The term 'exponential assembly' is proposed to differentiate this from self-replication. The exponential assembly architecture can use parts made from many manufacturing methods, provided that parts of considerable complexity can be produced. Because integrated circuit manufacturing methods used for micro-electromechanical systems allow large numbers of complex components to be produced, it is one possible method for pursuing such an approach to manufacturing. The methodology is scalable and therefore useful for producing assembly stations which might in turn produce other devices at ever-decreasing length scales as part of a top-down approach to nanotechnology.

(Some figures in this article are in colour only in the electronic version)

1. Introduction

Self-replication and nanotechnology are commonly discussed together as it has been suggested that no large-scale nanotechnology industry can develop without self-replication. Progress towards molecular nanotechnology is being made on several fronts, producing breakthroughs in molecular manipulation for chemical bond formation [1], molecular electronics [2] and the harnessing of bio-molecular motors [3]. The self-replicating entities of the biological arena are often used as an existence proof for the possibility and inevitability of self-replicating machinery. Simpler, man-made self-replicating machinery, however, has not seen much progress since the demonstrations of the first rudimentary self-reproducing machine [4, 5].

The first artificial, self-replicating mechanical structures were built by Penrose [6], and the first electromechanical systems were created by Jacobson [5] decades ago. The structures used by Penrose showed that it is possible to construct a very simple mechanical analogue for self-reproduction. The demonstration by Jacobson using model railroading parts showed that a self-assembly system using complex parts and one simple assembly step would also result in a self-reproducing analogue. Neither of these systems was able to perform a useful function afterwards, but these experiments did show that mechanical structures could perform self-replication in a somewhat limited way. These authors simplified self-replication by showing that it is not inherently a complex problem and that a rich environment of complex parts and simple reproduction steps makes tractable the self-replicating problem. Unfortunately, these demonstrations did not address the issues surrounding self-replicating manufacturing systems, as shown in one report still regarding such systems as complex [7].

Completely self-replicating systems may not prove necessary for nanotechnology, as simpler systems containing some replicating aspects will undoubtedly be easier to implement. Self-replicating systems can usefully be divided into the replicating component on the one hand, and the environment on the other. It is generally not possible to consider a replicating system without considering an environment also. By providing a more structured and more complex environment, the complexity of the replicating component can be reduced. The self-replicating methodology described in this paper is achieved by providing a very structured and complex environment so that certain aspects of the replication process can be greatly simplified—in particular, the number and complexity of the assembly operations.

During the approach to mature molecular manufacturing, interim systems, showing just enough replicating aspects to assist the progression to more complex mature assembly systems, is desired. In examining the evolution toward nanotechnology, it has been unclear where self-replication would become necessary or even possible. Previously, it had been thought that self-replication could only be introduced near the end of the development process. We may now have the ability to develop and introduce assembly systems containing self-replicating aspects within the foreseeable future.

Exponential assembly is an architecture starting with a single robotic arm on a surface. This first robotic arm makes a second robotic arm on a facing surface by picking parts carefully arranged in advance in precise locations. The two robotic arms then assemble two others, one on each of the two facing surfaces. These four robotic arms, two on each surface, then make four more robotic arms. This process continues with the number of robotic arms steadily increasing in the pattern 1, 2, 4, 8, 16, 32, 64, etc, until some manufacturing limit is reached (both surfaces are completely populated with assembled robotic arms, for example). This is an exponential growth rate where 2^N assembly stations are assembled after N assembly operations.

To keep things simple, we want to keep the robotic arm simple. While a general purpose arm having six degrees of free movement would be able to pick parts and position them in any orientation and position desired, such a general purpose arm is more difficult to make. Outlined below is a design using robotic arms having only two rotational degrees of freedom and a gripper. Externally provided power and computer control would make all the robotic arms on a surface operate synchronously and in parallel. Each robotic arm will have only two degrees of freedom and the surface will moved in X, Y and Z. This provides a total of five degrees of freedom for each robotic arm with the three translational degrees of freedom being shared across all the robotic arms on a surface.

2. The system components

To simplify the required assembly operations, the system shares certain degrees of freedom and instructions, as well as off-loading parts manufacture to some other method. The total assembly system consists of a translating mechanism capable of accurate positioning in the three Cartesian axes, a control system and two surfaces with pre-arranged parts ready for assembly. A schematic of this system is shown in figure 1. The parts can be manufactured and arranged by any known method, making the system accessible to many length scales.

The surfaces are divided into grids, with a set of components for a station located at each grid site. The control system knows the location of each part as well as the assembly procedure. Each site on the first surface, defined as surface A, is uniquely addressed by a right-handed Cartesian coordinate system with the convention $A(x, y)$. The first grid site has address $A(1, 1)$, the final site has $A(n, n)$ as depicted in figure 2(a). The first step is to manually assemble a single station at die site $A(1, 1)$. This is accomplished by whatever means are available, as it only needs to be done once.

At this point, the first surface has an assembled and functioning station. The functioning station is then placed into

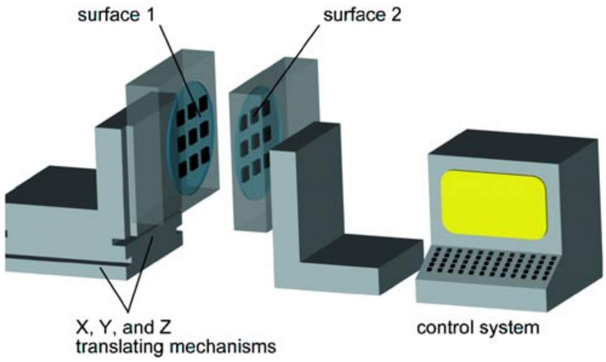

Figure 1. The above schematic shows the necessary components for an exponential assembly system. A translating mechanism capable of translating in the three Cartesian axes, a control system for controlling the stations and the shared translating mechanism, and two facing surfaces with pre-arranged arrays of unassembled components.

the system facing an opposing surface. This second surface, fabricated as a mirror image of the first, contains another set of grid sites with components ready for assembly. This second surface, defined as surface B, is addressed with a *left-handed* Cartesian coordinate system with the convention $B(x, y)$. This sites are thus labelled $B(1, 1)$ through $B(n, n)$ as shown in figure 2(a). This surface needs to be arranged to face surface A such that translations are equivalent from each of the reference frames. This is done by rotating surface B 180° about its z axis as depicted in figure 2(b).

Initially, site $A(1, 1)$ is located across from $B(1, 1)$. The assembly process is started and station $A(1, 1)$ assembles station $B(1, 1)$. The assembly process for a suggested station will be outlined in detail in the next section. The station at site $A(1, 1)$ is then returned to its as-assembled configuration and the station at $B(1, 1)$ is enabled for operation. The surfaces are then translated such that site $A(1, 1)$ is across from $B(2, 1)$, which implies that the newly constructed $B(1, 1)$ is across from $A(2, 1)$ as shown in figure 2(c). Because surface B is a mirror image copy of surface A, the same assembly sequence will work for $B(1, 1)$ in constructing $A(2, 1)$ as works for $A(1, 1)$ in constructing $B(2, 1)$. The two can perform the assembly sequence simultaneously. When the sequence is completed there are newly assembled stations at $A(2, 1)$ and at $B(2, 1)$. This iteration ends by returning the operating stations (those at $A(1, 1)$ and $B(1, 1)$) to their as-assembled configurations and enabling the newly assembled stations at $A(2, 1)$ and $B(2, 1)$ for operation. The next large translation locates $A(1, 1)$ across from $B(4, 1)$ as shown in figure 2(d) and the four operating stations can assemble another four. After some number of assembly iterations, each surface contains an entire row of assembled and functioning stations. The next translations will therefore position row $A(n, 1)$ across from row $B(n, 2)$, which also positions row $B(n, 1)$ across from row $A(n, 2)$. Entire rows are now assembled simultaneously. After iterating through the rows in a similar manner, both surfaces will be completely assembled.

3. The assembly stations

The assembly station can be configured in many ways with varying degrees of complexity. The exponential assembly

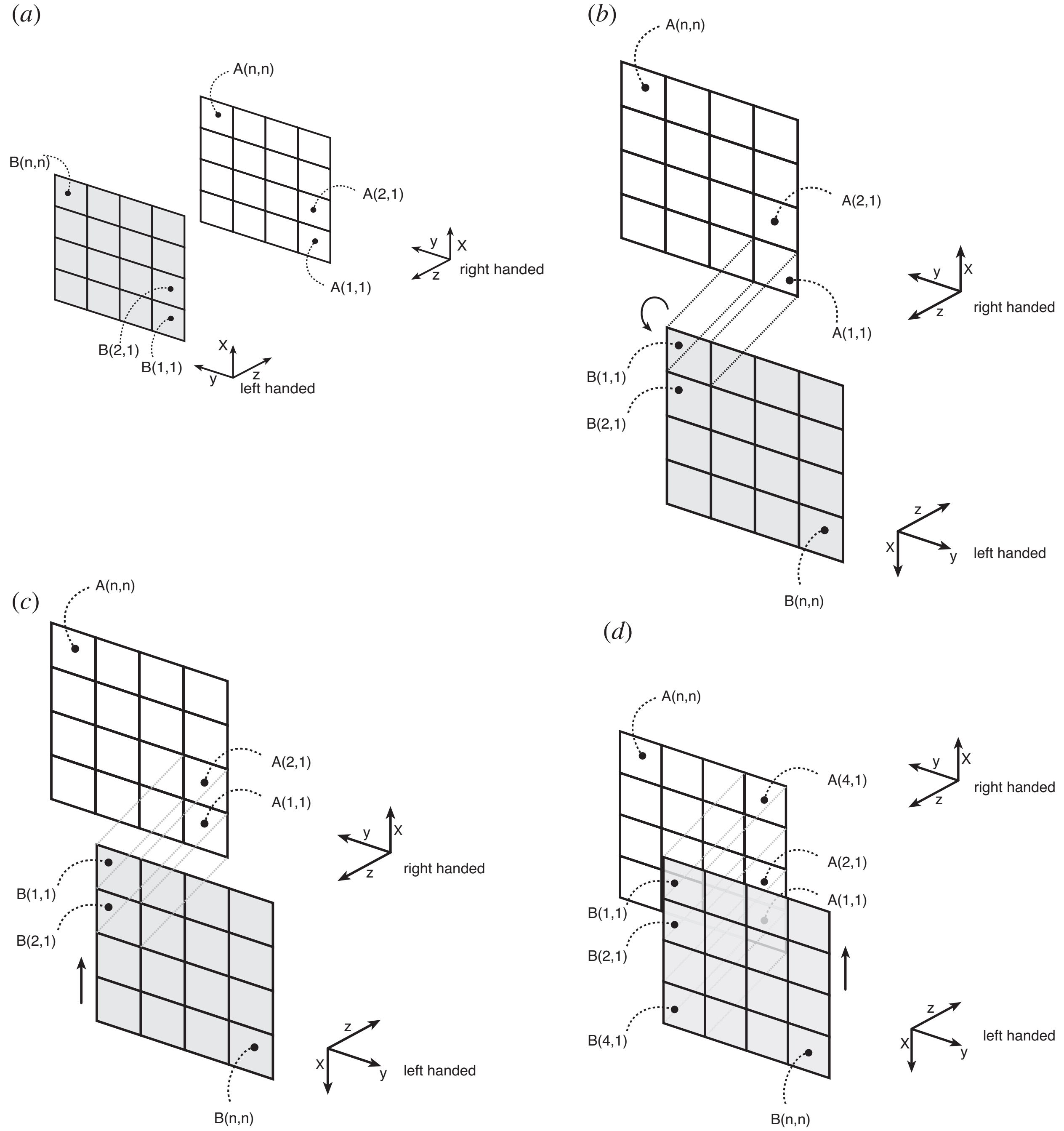

Figure 2. (*a*) Two mirror image surfaces for exponential assembly are shown in (*a*). (*b*) Surface B is then rotated 180° about its z axis and translated such that site $B(1, 1)$ is across from $A(1, 1)$. (*c*) The next translation locates $A(1, 1)$ and $B(1, 1)$ across from $B(2, 1)$ and $A(2, 1)$, respectively, as shown in (*c*). (*d*) Translating again positions $A(1, 1)$ and $B(1, 1)$ across from $B(4, 1)$ and $A(4, 1)$, respectively, as shown in (*d*).

station we will describe has the ability to rotate about two orthogonal axes, each with 90° range of motion, and the ability to grip and release parts. A schematic of this simple station is shown with the components arranged flat in figure 3(*a*). The components are quite complex, as is necessary for easy assembly. The base piece contains a rotational stage with an attachment mechanism for connecting the second component. The arm piece contains another rotation stage with an attached arm, terminating with a gripper mechanism. Also shown on the arm piece is a place for gripping the component, termed the handle, as well as the attachment mechanism for connecting to the base component. To assemble the device, the arm piece must be lifted from the surface, rotated about the y axis and attached to the base piece using the attachment points. If the attachment mechanism is secure, the entire assembly operation is complete and the device is as shown in figure 3(*b*).

This first device is now assembled. It has not been discussed with what mechanism this first assembly station is assembled, but an external gripper and set of assembly stages under manual or automated control would likely be

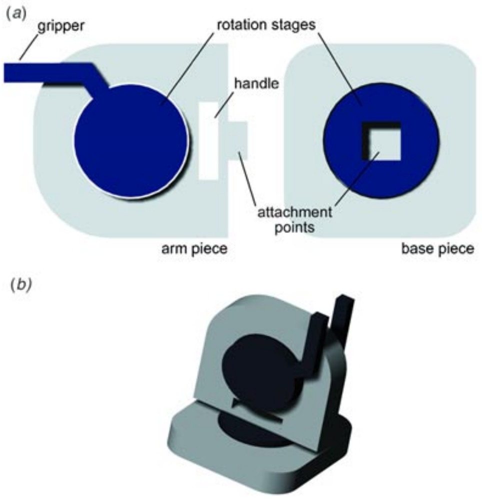

Figure 3. (*a*) The rotational stages are labelled, as are the gripper, handle and attachment points. (*b*) When the two components are assembled, the assembly station is as shown.

sufficient. The translating stages shown in figure 1, in conjunction with added rotation capability, should suffice. An animation of this first assembly sequence can be found at http://www.zyvex.com/exponential/mems.rm. Regardless of its construction method, once assembled, the first device is ready to begin the exponential assembly process.

To perform exponential assembly, two arrays of pre-arranged components are mounted facing each other. As discussed above, the second array can be a mirror image of the first. The first assembled station then performs the rotational motions and gripping while the shared translational mechanisms provide the Cartesian displacements necessary to assemble the station on the opposite surface. The assembly operation sequence showing the first component assembling the second is outlined in figure 4 and in the animation found at http://www.zyvex.com/exponential/rota.rm. Figure 4(*a*) shows the four components comprising two mirror image stations ready for assembly. Figure 4(*b*) shows the components after one has been rotated and positioned such that the two are facing each other in the necessary arrangement. Figure 4(*c*) shows the situation after the first component has been assembled by the external mechanism. The assembled station first rotates 90° about the z axis, as shown in figure 4(*d*), then it translates to grip the arm stage as shown in figure 4(*e*). The base rotation stage of the assembled station then rotates about the y axis and translates until the attachment points align as shown in figure 4(*f*). After the attachment is made, the gripper releases and translates. The assembly is now completed, as shown in figure 4(*g*). The two components are now separated a safe distance and the two rotation stages of the first station are returned to their original, as-assembled configuration. The two assembled stations are now ready to begin another assembly sequence and are translated such that each is across from a fresh set of components. The same sequence of operations is now provided to both assembled stations and they repeat the exponential assembly process until each array of components is assembled. In principle, assuming 100% yield of both components and assembly operations, the sequence produces 2^N assembly stations after N assembly sequences.

The assembly rate will suffer if individual devices cannot be enabled as assumed above. The above sequence of operations presumes that a single station, once assembled, is enabled, or 'turned-on', and ready for motion. If another approach is taken—for example, an approach where all of the stations on a surface are 'hard-wired' to move simultaneously, even when unassembled—the above operations will not work. This is because the unassembled base rotation stages are moving at an undesirable time. This problem is easily overcome by operating only one surface during an assembly sequence, leaving the other surface inactive. After one sequence, the first surface is made inactive and the second is operated. Thus, the first assembly station assembles first one on the opposing surface, and then another. There are now two assembled on the second surface and one on the first. The second surface is then enabled and used to build two more on the first while the first remains inactive. The end result is a ping-pong mode of assembly where progression follows a Fibonacci sequence instead of the aforementioned 2^N sequence. The growth rate of assembled stations has been diminished, but the control system or wiring of the stations has been greatly simplified.

We have described a conceptually simple assembly station. In practice, we expect that more complex operations and structures will prove desirable; these more evolved and complex systems evolving naturally from the original concept. For example, the attachment may be insecure and require additional bracing support, which can easily be included by adding pieces and assembly operations. More complex components could also be used, giving the stations additional degrees of freedom, larger ranges of motion, or advanced features of many kinds. Although this additional functionality is unnecessary for exponential assembly, it adds to the usefulness of the station for assembling devices other than like-copies of itself.

Once an array of these components is assembled, it can be used to do parallel assembly of additional like-stations such that complete arrays can be assembled in one sequence. It might also prove useful to assemble arrays of more complex robotic manipulators, end products, etc. A wide range of assemblies could be constructed using these or similar stations, provided that the parts are of appropriate size and arrangement, and that collisions can be avoided during assembly.

4. Implementation methods

There has been no discussion about the length scales involved in this exponential assembly methodology. In principle, the system is not confined to any specific size regime—it should be possible to construct systems having many sizes. The replicative demonstration by Jacobson mentioned above was built using model railroading parts and the exponential assembly system defined here could likely be demonstrated at such a scale. It is not yet known at what scale such a system would perform an economically useful function. Developments in micro-electromechanical systems (MEMS) [8] show that it is possible to fabricate large numbers of micro-machine components onto a silicon wafer with high precision and great complexity. MEMS technology may be an appropriate starting technology for manufacturing the arrays of components necessary for an exponential assembly demonstration.

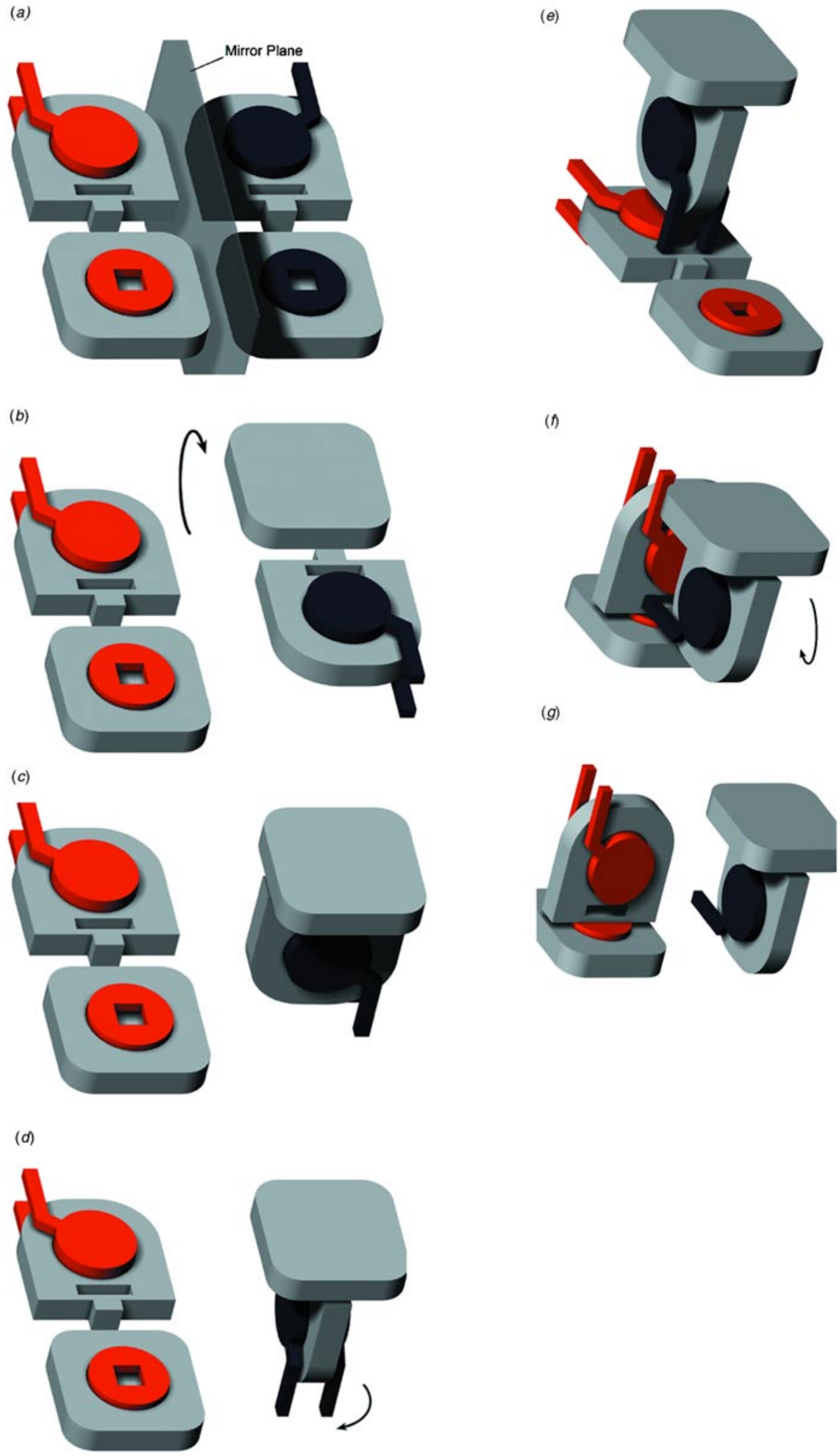

Figure 4. (*a*) The stations are first shown unassembled. (*b*) After rotating one of the stations, the two are shown facing each other. (*c*) The first is then assembled. (*d*) The assembled station first rotates about the z axis. (*e*) It translates to grip the component of the other. (*f*) The other second rotation stage of station 1 then rotates 90° and translates such that the attachment points are aligned with the translation. (*g*) After the attachment is made, the gripper releases and the assembly is completed.

5. Summary

Exponential assembly has been proposed as a possible replicative methodology for assembly-based manufacturing systems. Separate assembly stations having two rotational degrees of freedom are, in principle, able to assemble like-copies when coupled to shared translating mechanisms and a control system. The methodology should be possible to implement at different length scales, but the integrated circuit manufacturing technology used in MEMS appears to be well suited as a starting point for component production. If this methodology is scalable, it may be useful for nanotechnology, as it does allow for an exponential increase in assembly capability, making possible the production of large numbers of assembled systems.

References

[1] Hla S-W, Bartels L, Meyer G and Rieder K-H 2000 *Phys. Rev. Lett.* **85** 2777

[2] Reinerth W A, Jones L, Burgin T P, Zhou C-W, Muller C J, Deshpande M R, Reed M A and Tour J M 1998 *Nanotechnology* **9** 246–50

[3] Montemagno C, Bachand G, Stelick S and Bachand M 1999 *Nanotechnology* **10** 225–31

[4] Penrose L S 1959 *Sci. Am.* **200** 105–14

[5] Jacobson H 1958 *Am. Sci.* **46** 255–84

[6] Penrose L S and Penrose R 1957 *Nature* **179** 1183

[7] Freitas R and Gilgreath W P (ed) 1982 *Advanced Automation for Space Missions* NASA CP-2255 web page http://www.islandone.org/MMSG/aasm/

[8] Madou M 1997 *Fundamentals of Microfabrication* (Boca Raton, FL: CRC Press)